Informed source extraction from a mixture of sources exploiting second order temporal structure

Informed source extraction from a mixture of sources exploiting second order temporal structure

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus, prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op donderdag 17 januari 2019 om 16.00 uur

door

Brian Brand Antonius Johannes Bloemendal

geboren te Deurne

Dit proefschrift is goedgekeurd door de promotor en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof. dr. ir J.H. Blom
promotor:	prof. dr. ir. J.W.M. Bergmans
copromotoren:	dr. ir. P.C.W. Sommen
	dr. ir. J. van de Laar (Philips Research)
leden:	prof. dr. ir. M. Moonen (KU Leuven)
	dr. ir Emmanuël Habets (International Audio Laboratories Erlangen)
	prof. dr. S. Weiland
adviseur:	ing. C.P. Janse (Philips Research)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Informed source extraction from a mixture of sources exploiting second order temporal structure

Brian Brand Antonius Johannes Bloemendal

This research was kindly supported by Philips Research, The Netherlands.

A catalogue record is available from the Eindhoven University of Technology Library. ISBN: 978-90-386-4661-9.

Cover design by: Frank Gijpmans - Libelnet
Reproduction by: Gildeprint

Summary

Informed source extraction from a mixture of sources exploiting second order temporal structure

This work considers scenarios where one desired source and at least one interfering source are mixed by a linear, time-invariant, multiple-input multiple-output (MIMO) mixing system. The sources are mutually statistically independent and only their mixtures are observed, i.e., both the sources and the MIMO system are unknown. The desired source is to be recovered from the observations via a multiple-input single-output (MISO) source extraction (SE) filter. This thesis deals with the development of SE algorithms that identify the MISO filter that extracts the desired source. In the literature many blind and informed SE algorithms are available. Blind algorithms use only the observations and assumptions on the statistical and temporal structure of the sources to perform the task at hand; however, they extract the sources randomly and cannot guarantee extraction of the desired source. The rationale behind informed algorithms is to exploit additional a priori information with respect to blind algorithms in order to improve their performance. Informed SE algorithms are available that extract potentially 'interesting' sources by exploiting statistical measures that are assumed to be carried only by 'interesting' sources. Although potentially interesting sources are extracted, the a priori information is not used to guarantee immediate extraction of the desired source. Consequently, an iterative approach may be used to extract a source and subsequently classify it. If not the desired source is extracted, then the source is removed from the observations and a new source is extracted. These iterative SE algorithms are inefficient since several iterations may be required until the desired source is extracted. Informed SE algorithms that use a priori information in order to guarantee immediate extraction of only the desired source are missing. Furthermore, many of the available blind and informed algorithms consider instantaneous mixtures of sources. Convolutive mixtures are considered to be difficult to handle and a time-frequency transformation is often used to transform convolutive mixtures into complex-valued instantaneous mixtures per frequency bin. This time-frequency transformation sets high standards on the available a priori information in order to ensure extraction of the desired source per frequency bin.

This thesis focuses on SE algorithms that utilize a priori information in order to guarantee immediate extraction of the desired source. We exploit the second order temporal structure (SOTS), i.e., non-stationarity, temporal correlatedness, and spatial uncorrelatedness, of the sources. Many blind algorithms in literature exploit this SOTS and work by applying subspace techniques to different arrangements of correlation data from observations of MIMO instantaneous mixtures. These subspace techniques can be used to solve blind SE problems as mathematical problems where systems of homogeneous polynomial equations have to be solved. Each root of such a system corresponds uniquely to the extraction filter for one of the sources. For SE the key problem is to identify the root that corresponds to the filter that extracts the desired source. In this thesis the design of informed SE algorithms, i.e., algorithms that identify the desired extraction filter, is formulated as the design of optimization problems where polynomial equations ensure that the desired extraction filter is identified. The objective functions of the optimization problems are constructed in such a way that they allow for the incorporation of a priori information about the mixing and the correlation parameters of the sources. Another important step in this work is that the numerical solutions are reformulated as smallest and largest eigenvalue problems, for which efficient and numerically stable algorithms exist. We discuss the design of objective functions and derive conditions that have to hold on the available a priori information such that extraction of the desired source is guaranteed. Specifically for complex instantaneous mixtures we have developed design techniques that allow for the incorporation of a priori information about physical parameters in the form of direction of arrival information. Finally, we have applied subspace techniques to correlation data from observations of convolutive mixtures. By applying these subspace techniques we are able to formulate new systems of polynomial equations of which the roots correspond to finite impulse response (FIR) filters from MIMO mixing systems that consist of short FIR filters. This is considered to be a first step towards the design of convolutive SE algorithms that are more flexible in the use of a priori information than SE algorithms based on a time-frequency transformation.

Contents

Glossary

List of acronyms

ASP	Array signal processing
BSP	Blind signal processing
BSS	Blind source separation
BSE	Blind source extraction
BW	Bandwidth
cICA	Constrained independent component analysis
DOA	Direction of arrival
ECG	Electrocardiography
EEG	Electroencephalography
EMG	Electromyography
EVD	Eigenvalue decomposition
GEVD	Generalized eigenvalue decomposition
GSC	Generalized sidelobe canceller
HOS	Higher order statistics
ICA	Independent component analysis
ICA-r	Independent component analysis with reference
LCMV	Linear constrained minimum variance
LTI	Linear time-invariant
MIMO	Multiple-input multiple-output
MISO	Multiple-input single-output
MMSE	Minimum mean square error
MVDR	Minimum variance distortionless response
MWF	Multichannel Wiener filter
NF-ROS	Noise-free region of support
SINR	Signal to interference and noise ratio
SIR	Signal to interference ratio
SNR	Signal to noise ratio
SOS	Second order statistics

SOTS	Second order temporal structure
SVD	Singular value decomposition
VAD	Voice activity detection

List of textual abbreviations

e.g.	exempli gratia: for example
i.e.	id est: that means, in other words
s.t.	subject to
viz.	videlicet: namely
w.r.t.	with respect to

List of mathematical symbols and variables

a_i^j	Mixing element from the j'th source to the i'th sensor
$\mathbf{A}$	Mixing matrix
$\mathbf{a}^i$	Mixing column vector, i.e., i'th column of mixing matrix $\mathbf{A}$
$\tilde{\mathbf{a}}_j$	Row j of mixing matrix $\mathbf{A}$
$\mathbf{B}$	Reference mixing matrix
c	Symbol indicating type of conjugation
D	Number of sensors
$\mathbf{e}^i$	Column vector with a one at the i'th row and zeros elsewhere
$\tilde{\mathbf{e}}_i$	Row vector with a one at the i'th column and zeros elsewhere
γ	Shape parameter for selection beamformers or objective functions
$\mathbf{h}$	Extraction filter representing the mixing column of the desired source
Ω	Symbol representing the noise-free region of support (NF-ROS)
Ω_κ	Time-lag pair κ from the NF-ROS
K	Number of time-lag pairs in NF-ROS
L	Number of vectors used to incorporate a priori information
λ	Eigenvalue of a matrix
$\mathbf{\Lambda}$	Diagonal matrix containing eigenvalues
M	Matrix from which the eigenvectors are extraction filters
$\boldsymbol{\mu}$	Eigenvector of a matrix
p	Order of a polynomial
$\mathbf{P}$	Permutation matrix
$\mathbf{P}_0$	Projection matrix
q	Number of conjugations in a polynomial
$\mathbf{Q}$	Whitening matrix
$\mathbf{s}$	Vector containing source signals
S	Number of sources
$\mathbf{T}$	Transformation matrix for the mixing matrix
$\boldsymbol{\upsilon}$	Vector of observations of reference signals
$\mathbf{v}$	Vector of variables used in polynomial functions

$\tilde{\mathbf{w}}$	Extraction filter that maximizes SIR
$\mathbf{W}$	Source separation matrix
$\mathbf{x}$	Vector containing observed sensor signals
$\mathbf{\Gamma}$	Matrix containing linear combinations of correlation matrices
$\mathbf{z}$	Vector of variables or indeterminates used in polynomials
$\boldsymbol{\xi}^l$	Column vector designed based on a priori information
$\tilde{\boldsymbol{\xi}}_l$	Row vector designed based on a priori information
$\mathbf{\Xi}$	Matrix designed based on a priori information
η	Arbitrary scaling value unequal zero
σ^2	Representing the power of a signal
θ	Symbol for direction of arrival or angle
$\mathbf{\Phi}$	Polynomial coefficient matrix
$(\cdot)^T$	Transpose of a matrix or vector
$(\cdot)^H$	Hermitian transpose of a matrix or vector
$(\cdot)^*$	Conjugate of a matrix, vector, or scalar value
$(\cdot)^{-1}$	Inverse of a square matrix
$(\cdot)^{-T}$	Inverse and transpose of a square matrix
$(\cdot)^\dagger$	Pseudo-inverse of a matrix
$\mathrm{tr}\,(\cdot)$	Trace of a matrix
$\mathrm{diag}(\mathbf{v})$	Matrix with the elements of vector $\mathbf{v}$ on its diagonal
$\mathrm{diag}(\mathbf{V})$	Vector containing the elements on the diagonal of matrix $\mathbf{V}$
$\lvert\cdot\rvert$	Absolute value
$\lVert\cdot\rVert_2$	Euclidean norm
$\lVert\cdot\rVert_F$	Frobenius norm
$\langle\cdot,\cdot\rangle$	Euclidean inner product
$\diamond$	Khatri-Rao product
$\otimes$	Kronecker product
$\bigoplus$	Direct sum

1 Introduction

1.1 Scope and motivation

The work in this thesis is motivated by the increase in availability and usage of systems and devices with multiple sensors. Nowadays, systems exist that have multiple sensors measuring in different domains such as a telephone equipped with both a camera and a microphone. The multi-sensor systems considered in this work are equipped with multiple sensors that measure in the same domain such as a teleconferencing system with multiple microphones. Applications of these multi-sensor systems and devices can be found in areas such as audio and acoustics, biomedical engineering, and telecommunication. Furthermore, efforts in energy harvesting and wireless telecommunication contribute to the deployment of wireless sensor networks [1–3]. Using multiple sensors instead of only one sensor is beneficial due to several reasons. It can lead for example to better performance in terms of intelligibility for speech enhancement and to a better bit error ratio in communication systems due to a better spatial coverage on the one hand and combining the signals in a smart way, such as weighted averaging, on the other hand. Furthermore, using multiple sensors allows for new use cases such as source localization. Techniques for processing of the signal captured with such multi-sensor systems are researched and developed in the field of *array signal processing* (ASP) [4–6] and the related research fields *independent component analysis* (ICA) and *blind source separation* (BSS) [7–11].

In this work the focus is on the source extraction problem [8, 12–18]. Here one source has to be extracted or recovered from a mixture of multiple sources. In many multi-sensor applications a specific, desired source has to be extracted; however, in the literature mainly source extraction algorithms that extract randomly one of the sources from the mixture are found. The main objective of this thesis is to provide insight in the source extraction problem and to develop and design source extraction

algorithms for which immediate extraction of the desired source is guaranteed. To elucidate the problem we provide three examples in different application areas, i.e., audio and acoustics, biomedical engineering, and telecommunication. These examples are used to discuss, clarify, and underpin the problems and assumptions considered in this thesis.

- Typical *audio and acoustics* devices that contain multiple sensors are mobile phones, tablet computers, smart televisions, and hearing aids. For each of these devices we observe an increased demand on its performance in order to serve new applications. Mobile phones have te be operated hands-free and from a distance, a smart television has to be operated and communicated with by using only voice while sitting a few meters from the television set; and, hearing aids have to work comfortably in crowded environments. These modern applications put pressure on the currently available signal processing algorithms. Ideally, sensors are placed close to the desired source such that the observations contain this desired source with high signal to noise ratio (SNR). In modern applications more flexibility and comfort for the user is desired; therefore, these applications do not allow in general for sensor placement nearby the desired source. As a result, mixtures of the desired and interfering sources are observed where the interfering sources potentially have a similar or even higher sound level than the desired source. Consequently, new algorithms are required in order to extract the desired source.
- Another field of application where the source extraction problem is relevant is in the field of *biomedical engineering*. Many biomedical engineering applications are aimed at obtaining information about a process in the body based on non-invasive measurements. For such measurements, multiple sensors are placed on the surface of the body or at a distance from the body. Examples of such measurements are electrocardiography (ECG), electroencephalography (EEG), and electromyography (EMG) measurements. These sensors measure signals arising from multiple active processes while one is often interested in the signal related to only one of these processes. For example, an ECG measurement taken from a pregnant women contains both the fetal and maternal ECG while often only the fetal ECG is desired. Another example is neurofeedback. In such an application the EEG measurements are used to in order to measure the response of the brain to an externally applied stimulus. These measurements contain the response to the stimulus and the default activity of the brain. For neurofeedback extraction of the response to the external activity is desired.
- In the field of *telecommunications* the number of transceivers is growing rapidly while the available bandwidth is not increasing at the same pace. Consequently, simultaneous transmission in the same frequency band cannot be avoided. Communication bandwidth can then only be increased by exploiting spatial diversity. An example application where simultaneous transmission is likely to appear is for example in densely populated areas such as festival sites where many people use their mobile phone simultaneously; and, extraction of a specific signal is desired. Another application is found in wireless sensor networks. Nowadays many wireless sensor networks are being deployed. In such a network each

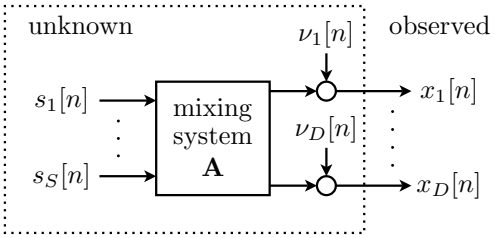


Figure 1.1: Overview of a multi-sensor scenario in discrete time domain where S sources are mixed by an LTI, MIMO mixing system $\mathbf{A}$. The mixtures are corrupted by additive noise and only the D noisy mixtures $x_1[n], \cdots, x_D[n]$ are observed.

wireless sensor node is equipped with a transmitter or transceiver and transmits its signal either to its neighbors or to a base station. In dense wireless sensor networks multiple sensor nodes transmit simultaneously and extraction of the signal transmitted by a specific node is a relevant problem. Moreover, the objective of many wireless sensor networks is to gather information from the environment by collaboration of the nodes, which can also be considered a multi-sensor source extraction problem.

1.2 Problem formulation

As discussed in the previous section, multi-sensor systems and devices can be used for a wide range of applications. In this work we focus on applications where one desired source has to be extracted from a linear mixture of multiple sources. In this section we discuss commonly used models for mixing systems, objectives for source extraction filters, and methods to identify source extraction filter parameters.

1.2.1 Mixing models

In the considered source extraction scenario a number of sources are mixed by a linear, time-invariant (LTI), multiple-input multiple-output (MIMO) mixing system. In practice the outputs of the mixing system are measured by multiple sensors. These sensors typically have an analog to digital converter (ADC) in order to obtain a discrete time representation of the output signals. In most literature all sensors are assumed to be the same and only corrupt the mixtures with additive sensor noise. Following these assumptions, the mixtures are assumed to be corrupted by additive sensor noise and only the additive noisy measurements are observed, i.e., both the original sources and the mixing system are unknown. A graphical interpretation of this model for S sources and D sensors is depicted in Figure 1.1. Mathematically, the discrete time observations from an LTI, MIMO mixing system have the following

structure:

$$x_i[n] = \sum_{j=1}^{S} a_i^j \star s_j[n] + \nu_i[n] \qquad \forall\, i \in \mathcal{D} \tag{1.1}$$

where $x_i[n]$ represents the signal measured by the i'th sensor at discrete time instance n, $s_j[n]$ represents the j'th source at discrete time n, $\nu_i[n]$ represents the noise measured by the i'th sensor, a_i^j is the element in the mixing system $\mathbf{A}$ that represents the relationship between the j'th source and i'th sensor, $\star$ is the operator corresponding to the type of mixing system, and $\mathcal{D} = [1, \cdots, D]$ is a set of integers from 1 up to and including D.

In the literature mainly three types of mixing models have been considered for LTI, MIMO mixing systems. These mixing models are closely related to each other and each model is specifically relevant for one of the application areas. The first model is the real-valued instantaneous mixing model. In this model the relationships between the sources and the sensors are considered to be real-valued gains. This model is mostly applicable in biological processes that work with electrical charges. The electrical activity inside a human body is observed instantaneously at the surface of a body. The second mixing model is the complex-valued instantaneous mixing model. This model is particularly useful for handling mixing systems where narrowband sources are delayed before they are observed. Such a mixing model holds for example in narrowband communication systems. In such a system the path from the source to a sensor results in a delay. Due to different path lengths between a source and the sensors, the sources are observed at the sensors with different delays. Finally, the third model is the convolutive mixing model. In this model the relationship between the sources and the sensors is modeled by filters. The most widely used type of filter to model convolutive mixtures is the finite impulse response (FIR) filter. Convolutive mixing models are applicable in applications such as modern communication systems and acoustical applications. In these applications the sources arrive at the sensors via multiple paths. In many algorithms convolutive mixtures are dealt with by transforming the observations from time domain to time-frequency domain. Under some conditions, the time-frequency transformation allows for the interpretation of the convolutive mixing system as a complex-valued instantaneous mixing system per frequency bin.

Another relevant property of the mixing systems are its dimensions, i.e., the number of sources that are active and the number of sensors that are available. Typically three classes are considered with respect to the dimensionality of mixing systems, i.e., the square, over-determined, and under-determined mixing system. A square mixing system has the same number of sensors as there are active sources and is the most widely considered mixing system in the literature. The over-determined and under-determined mixing systems have respectively more and fewer sensors than sources. In general, the under-determined mixing case is considered to be more difficult than the square or over-determined mixing scenario.

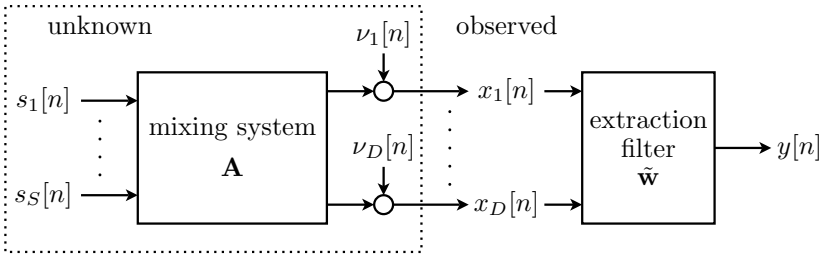


Figure 1.2: Overview of a MIMO mixing system followed by an LTI, MISO extraction filter $\tilde{\mathbf{w}}$.

1.2.2 Extraction filters

Extraction of the desired source from MIMO mixtures of sources is accomplished by applying a multiple-input single-output (MISO) extraction filter to the observations. An overview of the mixing system followed by an extraction filter is depicted in Figure 1.2. Mathematically, the LTI extraction filter works as follows:

$$y[n] = \sum_{i=1}^{D} w^i \star x_i[n] \tag{1.2}$$

where $y[n]$ is the output of the MISO extraction filter $\tilde{\mathbf{w}}$. Dependent on the type of mixing model, each extraction filter element w^i is either a real-valued gain, complex-valued gain, or an FIR filter that is applied to the i'th sensor signal. The goal of a source extraction algorithm is to recover the desired source at the output of the extraction filter. Due to the linear characteristic of both the mixing system and the extraction filter, the output of the extraction filter generally consists of three components, i.e., the desired source, the undesired sources, and noise. It follows immediately from linear algebra that in the presence of noise it is impossible to obtain only the desired source at the output of the linear extraction filter. Consequently, an objective function that accomplishes source extraction has to be chosen for the extraction filter. Examples of objective functions that lead to source extraction are to maximize the desired source to interference plus noise ratio (SINR), to maximize the desired source to noise ratio (SNR) under the constraint that the interfering sources are suppressed, and to minimize the mean squared difference between the extracted signal and a reference that represents the desired source.

In applications where the sources are non-stationary, the extraction filter can be followed by a single-channel, time-varying post-processor. Such a post-processor is often used in speech enhancement applications and uses desired source, interference, and noise power estimates to suppress the compound signal when noise and interference are dominant in the signal.

1.2.3 Identification of extraction filter parameters

In many applications only mixtures of the sources are observed, i.e., both the original sources and the mixing parameters are unavailable. Consequently, the source extraction filter parameters have to be identified from these observations. Many studies have shown that source extraction filter parameters can be identified by exploiting the observed mixtures and some assumptions on statistical parameters of the sources. The most basic assumptions made on many LTI, MIMO mixtures are that the sources have zero mean and are mutually statistically independent. The assumption of mutually statistically independent sources has led to the research field called *independent component analysis (ICA)* [7] and means that sources do not provide information about each other. Mutual statistical independence implies that the sources have zero crosscorrelation and zero crosscumulant for any order. Besides being mutually statistically independent, the sources must have sufficiently different statistical features. Without exploiting this diversity in statistical parameters source extraction filter identification is in general not feasible. Later we return to this property and discuss the assumptions on the statistical features of the sources in a more formal and detailed way.

In the past, a wide variety of assumptions on the (statistical properties of) sources has been made as a basis for development of source extraction algorithms [7, 8, 11]. Typically, these assumptions are grouped as being assumptions on second order statistics (SOS) or higher (than second) order statistics (HOS). Algorithms based on SOS assume and exploit non-whiteness or coloredness of stationary sources, i.e., the presence of temporal structure. If the sources are non-stationary then they are assumed to be stationary for a certain amount of time before their statistical properties change. These assumptions allow for the consideration of the SOS of the sources in terms of autocorrelation functions with two parameters, i.e., a time and autocorrelation lag parameter. Since the statistics may vary over time, the autocorrelation function varies over time. In order to be able to extract a source from a mixture we require that sufficiently many time-lag pairs are available for which the source autocorrelation functions are non-zero. Another regularly exploited property of the temporal structure of the sources is sparseness. This is a very specific form of non-stationarity, which is often exploited in speech enhancement algorithms. A sparse source is only active for a limited amount of time. Exploiting this property leads to simpler mixtures. Finally, a widely used assumption on the temporal structure of the sources in the field of telecommunication is cyclostationarity. This assumption means that the signals are non-stationary; however, the statistical features of the sources repeat themselves periodically. Algorithms based on assumptions on the HOS of the sources typically exploit diversity in the distribution of the sources. A widely used HOS property is non-Gaussianity, which is the deviation of a signal from having its HOS parameters equal to zero. This immediately addresses a potential problem when exploiting HOS. Since Gaussian sources cannot be observed with HOS at most one source with a Gaussian distribution is allowed in the mixture. Signals with non-zero HOS are split into two groups. A group having sub-Gaussian distribution and a group having super-Gaussian distribution, i.e., distributions having smaller and larger peaks than the Gaussian distribution, respectively. Each group has its own sets of specialized al-

gorithms to deal with the signals. Finally, algorithms exist that exploit assumptions on the higher order temporal structure of sources in order to identify source extraction filter parameters, i.e., coloredness and non-stationarity for HOS parameters.

As discussed before, a subset of the assumptions on the statistical properties of the sources is required in order to be able to extract a source from a mixture of multiple sources. However, a source extraction algorithm based on only these assumptions extracts randomly one of the sources and is not able to guarantee that the desired source is extracted. Additional a priori information is required in order to guarantee extraction of the desired source. This additional a priori information can be available in many different forms. In many applications in the fields of audio and acoustics and telecommunication some information about the mixing or the statistical parameters of the sources is available, for example information about the direction of arrival (DOA) of the desired source. In biomedical applications often a reference signal is available such as the on-off characteristic of the stimulus applied during EEG measurements or an estimate of the fetal heartbeat with respect to the maternal heartbeat. Extraction algorithms that exploit this additional a priori information in order to ensure extraction of the desired source are called *informed source extraction algorithms*.

1.3 Related work

Traditionally, knowledge about structure in the mixing parameters, for example DOA information and a free-field acoustic model, is used to design or calculate source extraction filters [4,5] in the form of spatial filters or beamformers. Nowadays, in many applications source extraction is desired in more challenging environments and for new applications where little knowledge about the mixing parameters or the original sources is available and the source extraction filter parameters have to be identified with a source extraction algorithm using only the observations [8,11]. Many different assumptions have been exploited in order to develop source extraction algorithms. In the majority of algorithms the assumption that all sources are mutually statistically independent is exploited. Apart from that, the algorithms strongly depend on the properties of the mixtures, the available a priori information, and the objective function for the source extraction filter.

One way to perform source extraction is to utilize a blind source separation (BSS) algorithm [8, 10, 11]. A BSS algorithm identifies the parameters of a MIMO source separation filter that has the original sources separated at its outputs. The source separation filter can be viewed as a set of parallel extraction filters for all the sources in the mixture. Most BSS algorithms perform source separation by cancelling interfering sources. When additional noise reduction is applied, source extraction filters based on BSS typically extract the sources according to the source extraction objective function where the SNR is maximized under the constraint that the interfering sources are suppressed. The filter parameters are calculated by exploiting only the observations and some assumptions on statistical properties of the sources as we have discussed in Section 1.2.3. Since a limited amount of a priori information is used by a BSS algorithm, the source separation filter has the sources in a random order at the output.

In order to deal with this permutation of the sources, a classifier is required that uses additional a priori information in order to select the output corresponding to the desired source. This strategy implies redundancy since first all sources are separated and subsequently only one of the separated sources is used. Furthermore, the source separation filter exploits more resources than strictly necessary since the separation filter consists of many individual extraction filters.

In the field of acoustics, source extraction algorithms are typically speech enhancement algorithms. Most state of the art speech enhancement algorithms heavily depend on sparseness of the desired (speech) source and estimate the noise and interference components when the desired source is inactive. Such speech enhancement algorithms have been presented for example in [12] and [19]. In [12] a speech enhancement algorithm based on the multichannel Wiener filter (MWF) is presented. The objective for the algorithm is to minimize the mean squared error (MMSE) between the filter output and a reference. Ideally, this reference is chosen as the desired source; however, the desired source is not observed. In order to overcome this problem, in [12], the statistical properties of the desired source are estimated by exploiting sparsity of the desired (speech) source. The statistical parameters of the noise and interference are estimated during inactive periods of the desired source; subsequently, the interference and noise are assumed to be stationary and are subtracted from the statistics of observations that contain the desired source, interfering sources, and noise. Such a method has several disadvantages, e.g., voice activity detection (VAD) is required and the desired source has to be sufficiently sparse. Furthermore, the mixing parameters are not allowed to change over time; and the interfering sources and noise are required to be sufficiently stationary over silent and non-silent periods of the desired source. The speech enhancement algorithm presented in [19] is based on the generalized sidelobe canceler (GSC) [20,21]. A GSC consists of three components, namely: a beamformer, a blocking matrix, and an adaptive filter. The beamformer is a spatial filter that extracts the desired source from the observations and can be used to apply additional constraints on the spatial behavior of the GSC, e.g., to attenuate sound coming from the direction of an interfering source. In parallel, the blocking matrix is applied to the observations in order to generate signals consisting of only interference and noise. The adaptive filter is applied to the signals from the blocking matrix in order to enhance the output of the beamformer. The most difficult aspect of the GSC is to identify the filter parameters of the beamformer and the blocking matrix. Leakage of the desired signal into the output of the blocking matrix leads to degradation of the desired source [22]. In [19] the linear subspaces corresponding to the desired source and interfering sources are estimated during respectively non-silent and silent periods of the desired source. These subspaces are used to build a beamformer that enhances the desired source and cancels the interfering signals. The blocking matrix is orthogonal to the subspaces of the desired and interfering sources. Overall, the beamformer cancels the interfering sources and the blocking matrix and adaptive filter are used to maximize the SNR. In the literature this GSC is called a linear constraint minimum variance (LCMV) filter. An advantage of the GSC approach over the MWF approach in [12] is that the interfering sources and noise are not required to be stationary over silent and non-silent periods of the desired speech as long as the mixing system is not changing; however, this method also requires a VAD and the desired source

has to be sufficiently sparse. Speech enhancement algorithms that do not depend on sparseness of the source signals exist as well. One class of algorithms are geometrically constrained source extraction algorithms [23,24]. In these algorithms geometric constraints, typically in the form of DOA information, are used in order to steer extraction filters in a certain direction. In [23] geometric constraints were merged for the first time with source separation algorithms. The geometric constraints are used in order to solve the permutation problem, which occurs in a convolutive mixture in each frequency bin. In [24] geometric constraints are combined with a source separation algorithm called independent vector analysis (IVA) [25,26] in order to extract the desired source at a predefined output. In the IVA algorithm the permutation per frequency bin is solved by exploiting dependency among frequencies in the source signals. This dependency is exploited by considering the frequency domain source signals as multivariate random variables. A disadvantage of these algorithms is that geometric constraints are enforced by adding penalty terms to a cost function, which in turn leads to a bias or regularization in the identified extraction filter, especially for ill-conditioned mixing systems.

Recently another category of speech enhancement algorithms has been introduced that exploit non-stationarity of the source signals by relying on the assumption that in each time-frequency bin at most one source signal is dominant [27–29]. Different techniques have been applied in order to exploit this property. In [27] a DOA based speech detector is designed in order to classify if either the desired source or interference is dominant in each time-frequency bin. Subsequently, statistical properties of the desired and undesired signals are estimated from which data-dependent spatial filters are computed. Instead of a DOA based speech detector in [27], the method in [28] considers the use of a speech presence probability (SPP) estimator in order to estimate the statistical properties of non-stationary noise. An additional assumption that the spatial coherence of the noise is less than the spatial coherence of the desired speech is used in order to develop a necessary signal dependent a priori speech absence probability estimator. The estimated noise statistics are used in order to compute either an MWF or a conditional minimum mean-squared error filter [30]. This latter filter is obtained by multiplying the output of the MWF filter by the estimated a posteriori SPP. In [29] two model based expectation maximization (EM) algorithms are presented that allow for the incorporation of a priori information about either the mixing parameters or the statistical properties of the desired source. In multiple EM iterations the extraction filters, statistical features of the speech and noise, and speech presence probability are estimated, under the assumption that at most one of the source signals is dominant in each time-frequency bin. It is shown that for a deterministic clean speech model the EM algorithm results in a minimum variance distortionless beamformer (MVDR), whereas for a complex Gaussian model for the clean speech the EM algorithm results in the MWF filter, i.e., an MVDR followed by a single channel Wiener filter.

Source extraction algorithms that extract only one source from a mixture of sources without requiring sparsity are also found in other fields than speech enhancement [8,14,31]. Many of these algorithms are blind source extraction (BSE) algorithms [8] that extract a single source without incorporating a priori information re-

lated to a specific source. Consequently, such an algorithm extracts randomly a source. Although some BSE algorithms extract potentially interesting sources based on properties that are likely to be carried by interesting sources, such as non-Gaussianity and linear predictability [8, 14], no guarantees can be given on which source is extracted. In the last decade, several source extraction algorithms have been developed that incorporate reference signals. In the field of telecommunications, a popular way to extract sources is to incorporate reference signals in a contrast function. Contrast functions are cost functions based on higher order statistical features of the sources that attain a local maximum if a source is extracted [11, 32]. These contrast functions have been combined with reference signals in [17, 33] in order to increase the performance of the algorithms in terms of complexity, robustness, and convergence. Although performance is increased by the use of reference signals, immediate extraction of the desired source cannot be guaranteed with these algorithms. Finally, source extraction algorithms exist in the field of biomedical engineering that are derived from the constrained ICA (cICA) framework. The cICA framework allows for the incorporation of equality and inequality constraints in the optimization problem and has led to a reference based source extraction algorithm that includes an upper bound on the distance between the extracted signal and a reference as an inequality constraint [13, 34]. The upper bound on the distance is controlled by means of a threshold parameter and can lead to undesired behavior of the algorithm. If the threshold parameter is chosen too low then no source will be extracted, while a too large threshold parameter allows for multiple solutions such that an undesired source might be extracted. If a priori information that can be used to determine a suitable threshold parameter is available, then immediate extraction of the desired source can be guaranteed. Otherwise, the threshold parameter has to be found experimentally, which leads to a suboptimal algorithm.

1.4 Objectives of this thesis

In this thesis the focus is on source extraction algorithms for which immediate extraction of the desired source can be guaranteed.

From the previous sections it follows that many different properties and assumptions have been exploited in the past in order to develop source extraction algorithms. Mainly three MIMO mixing models have been considered in the literature, i.e., the real-valued instantaneous mixing model, the complex-valued instantaneous mixing model, and the convolutive mixing model. Furthermore, either second order statistical properties of the sources such as non-whiteness or non-stationarity or higher order statistical properties such as non-Gaussianity have been exploited in order to identify source extraction filters. Finally, the presence of noise makes that an objective function has to be chosen for the source extraction filter.

Despite the large number of source separation and source extraction algorithms in the literature, only a few source extraction algorithms are available that deal with immediate extraction of the desired source. These informed algorithms such as speech enhancement algorithms [12, 19, 27–29] or the reference based or constrained ICA

based algorithms [13, 34], which are concerned with immediate extraction of the desired source, either have very specific requirements such as knowledge about sparsity of the desired source or they are not sufficiently reliable in order to guarantee extraction of the desired source. Furthermore, many source extraction algorithms identify source extraction filter parameters based on a single objective function. Typically, source extraction algorithms are designed in such a way that interfering sources are annihilated without considering the gain potentially applied to the noise.

The objectives of this thesis are to provide insight into the source extraction problem and to provide tools that support the design and development of application specific informed source extraction algorithms. These source extraction algorithms should work for a wide variety of applications and have to guarantee immediate extraction of the desired source. Therefore, the source extraction problem is considered for various mixing models, different types of a priori information, and multiple objectives for the source extraction filter.

1.5 Thesis outline and contribution

In Chapter 2 we discuss several algorithms found in the literature that are relevant for the work in this thesis. Both source separation and source extraction algorithms are considered since the assumptions and approaches for these problems form the basis of the work in this thesis. For each approach we discuss assumptions on the source and noise signals and we consider the fundamental steps in the algorithms to perform the task at hand. Furthermore, we discuss two generic approaches for solving the closely related blind system identification problem. These approaches form the basis for the informed source extracted algorithms developed in this thesis.

In Chapter 3 we show that immediate extraction of the desired source can be guaranteed for real-valued instantaneous mixtures if a rough estimate of the autocorrelation function or the mixing parameters corresponding to the desired source is available. Based on this rough estimate, we take a combination of sensor correlation matrices in such a way that the eigenvectors corresponding to the smallest eigenvalue of this combination define extraction filters for the desired source. We derive conditions that have to hold on the available a priori information such that extraction of the desired source is guaranteed. The eigenvectors are used to define extraction filters and an additional noise reduction step is proposed such that the desired source is either extracted by maximizing the SINR or by maximizing the SNR under the constraint that the interference is canceled.

Expanding the results from Chapter 3 to complex-valued mixtures is not trivial and in order to guarantee extraction of the desired source from complex-valued mixtures a different combination of sensor correlation matrices is required. In Chapter 4 we show how to combine sensor correlation matrices for complex-valued mixtures in order to incorporate a priori information about the autocorrelation function or the mixing parameters of the desired source. Moreover, we present design tools in the form of

selection beamformers and selection patterns that can be used to design informed source extraction algorithms based on DOA information. Furthermore, we present an algorithm to choose between the two objective functions for the source extraction filters.

In Chapter 5 we show that the source extraction algorithms from Chapters 3 and 4 can be viewed and implemented as reference based source extraction algorithms. We show how to design or choose the reference systems based on available information about the mixing parameters or the autocorrelation function of the desired source such that immediate extraction of the desired source is guaranteed. Furthermore, we show that reference based source extraction using two non-overlapping subsets of sensors can be beneficial in the presence of noise that is uncorrelated among the two subsets.

In Chapter 6 we combine the algorithms and insights from the previous chapters into a single framework where the informed source extraction problem is formulated as an optimization problem with polynomial constraints. In this optimization problem the polynomial constraints ensure that the solution set consists of only extraction filters and the objective function is used to identify the solution that represents the extraction filter for the desired source. This formulation of the source extraction problem as an optimization problem allows for a more fundamental analysis of the source extraction problem. Selection of the desired solution is based on the objective function that is chosen for the optimization problem. The polynomial optimization problem is represented and solved as a largest eigenvalue problem. Furthermore, we show that the source extraction problem can be solved for certain under-determined mixtures, i.e., mixtures with more sources than sensors, we consider reference based source extraction algorithms in the form of polynomial optimization problems, and we show a natural connection to using higher order statistics.

In Chapter 7 our goal is to translate our insights on the instantaneous mixing scenario and the informed source extraction problem towards informed source extraction from convolutive mixtures. Although convolutive mixtures exist in many different forms, we focus in this chapter on a specific form that is based on finite impulse responses (FIR). We apply our subspace based approach to second order statistical features obtained from the observations of such mixtures and formulate systems of homogeneous polynomial equations from which the roots correspond to the coefficients of the mixing system. Although a rather specific mixing system is discussed, these results are considered a first step towards the development of informed source extraction algorithm design for convolutive mixtures.

Finally, Chapter 8 concludes this thesis and considerations for future research are presented.

1.6 Publications by the author

Journals

- B.B.A.J. Bloemendal, J. van de Laar, & P.C.W. Sommen. A single stage approach to blind source extraction. Signal Processing, Volume 93, Issue 2, February 2013, (pp. 432-444).
- B.B.A.J. Bloemendal, J. van de Laar, & P.C.W. Sommen. Design of signal extraction algorithms based on second order statistics exploiting beamforming techniques. Signal Processing, Volume 96, Part B, March 2014, (pp. 125-137).

Proceedings

- B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Multichannel signal enhancement using a remote wireless microphone. Proceedings of the International Workshop on Acoustic Signal Enhancement (IWAENC), 4-6 September 2012, Aachen, Germany.
- B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Blind source extraction for a combined fixed and wireless sensor network. Proceedings of the 20th European Signal Processing Conference (EUSIPCO), 27-31 August 2012, Bucharest, Romania, (pp. 1264-1268).
- Bloemendal, B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Beamformer design exploiting blind source extraction techniques. Proceedings of the 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Kyoto, Japan, 25-30 March 2012 (pp. 2589-2592).
- R.G.M. Hilkens & B.B.A.J. Bloemendal. An educational DSP platform based on a TMS320C5505 eZDSP. Proceedings of the 5th European DSP Education and Research Conference (EDERC), 13-14 September 2012, Amsterdam, The Netherlands.
- B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Blind extraction algorithm with direct desired signal selection. Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), 4-7 October 2010, Jerusalem, Israel. (pp. 13-16).
- B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Robust blind extraction of a signal with the best match to a prescribed autocorrelation. Proceedings of the 18th European Signal Processing Conference (EUSIPCO), 23-27 August 2010, Aalborg, Denmark. (pp. 1811-1815).
- B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Overdetermined blind source extraction exploiting a generalized sidelobe canceller structure. Proceedings of the International Workshop for Acoustic Echo and Noise Control (IWAENC), 30 August - 2 September 2010, Tel Aviv, Israel.
- B.B.A.J. Bloemendal, A.C.P. van Meer & P.C.W. Sommen. Realtime digital signal processing environment for educational exercises and lab experiments.

Proceedings of the Annual Workshop on Circuits, Systems and Signal Processing (ProRISC), 18-19 November 2010, Veldhoven, The Netherlands.

- B.B.A.J. Bloemendal, J. van de Laar & P.C.W. Sommen. Instantaneous blind signal extraction using second order statistics. Proceedings of the International Workshop for Acoustic Echo and Noise Control (IWAENC), 14-17 September 2008, Seattle, Washington USA.

Patent applications

- R.M.M. Derkx, B.B.A.J. Bloemendal & V.A.R. Aarts. Processing apparatus and method for determining an ambulation motion of a subject. Patent application WO 2018/029275 A1, February 2018.

2

Relevant blind signal processing methodologies and algorithms

Blind signal processing (BSP) algorithms such as blind source separation and blind source extraction algorithms can be used in order to separate or extract sources from a mixture of source signals. Such blind algorithms exploit only the observations and certain assumptions on statistical features of the sources to perform the task at hand, i.e., both the mixing system and the source signals are unknown. In order to guarantee extraction of the desired source, BSP algorithms must be followed by a classifier that selects the desired source. The objective in this thesis is to incorporate the a priori information used by the classifier into BSP algorithms, leading to informed source extraction algorithms.

In this chapter we consider several BSP algorithms found in the literature that can form the basis for the development of informed source extraction algorithms. First we present assumptions on the mixing model and assumptions on the source and noise signals that are often exploited in BSP algorithms. Subsequently, we discuss a variety of algorithms that perform BSP tasks. First we distinguish explicitly between algorithms that exploit non-Gaussianity of the source signals and algorithms based on temporal structure in the source signals. Finally, we discuss the two generic approaches for solving BSP problems that can exploit both non-Gaussianity and temporal structure. For each algorithm we present the fundamental steps towards source extraction and we highlight the pros and cons of the different methodologies.

2.1 Introduction

In a source extraction problem, one desired source signal has to be extracted from an observed mixture of several source signals. Many blind signal processing (BSP) algorithms exist that separate or extract sources. In the literature these algorithms are known as blind source separation (BSS) and blind source extraction (BSE) algorithms, respectively. Furthermore, blind system identification (BSI) algorithms, used to blindly identify or estimate the mixing system, exist as well [7,8,11,35]. BSP algorithms use only the observations and assumptions on statistical features of the sources to perform the task at hand, i.e., without knowing the mixing system and the source signals. A consequence of this blindness is that sources can only be separated or extracted in a random order due to a permutation indeterminacy. A straightforward way to deal with this permutation indeterminacy is via a two-stage approach where a BSS or BSE algorithm is followed by a classifier. This classifier uses additional a priori information in order to select the desired source. Instead of using a BSS algorithm, the mixing system identified with a BSI algorithms can be inverted, such that a source separation filter is obtained. Disadvantages of such two-stage approaches are that the algorithms are inefficient, utilizing more resources than strictly required. Furthermore, such two-stage approaches based on BSE algorithms can only be used in batch mode.

Source extraction algorithms that extract potentially interesting signals based on measures such as non-Gaussianity and linear predictability are available in the literature. However, when using these algorithms extraction of the desired source cannot be guaranteed, i.e., these algorithms also suffer from the permutation problem. The existing source extraction algorithms for which extraction of the desired source can be guaranteed are typically tailored to a specific application and exploit specific a priori information about the mixing parameters or the source signals. In this thesis, algorithms are developed for which extraction of the desired source is guaranteed by incorporating the a priori information used by a classifier into BSP algorithms, leading to so-called *informed source extraction* algorithms. In the current chapter we present an overview of BSP algorithms that are relevant for the remainder of this thesis and that can form the basis for the development of informed source extraction algorithms. More extensive overviews and detailed discussions of BSP algorithms can be found in the following books [7,8,11,35,36].

In general, two underlying actions can be distinguished in a BSP algorithm. One activity is spatial whitening of the source signal components in the observations. The other activity is to resolve an unknown rotation of the spatially whitened source signal components. In a BSS algorithm this rotation is identified for all sources, while in a BSE algorithm this rotation is identified for only one of the sources. One of the most prominent differences between BSP algorithms is in the way in which the rotation is identified; however, BSP algorithms that do not require spatial whitening in a preprocessing step exist as well. BSP algorithms can be divided into two main categories. Algorithms belonging to the first category exploit differences in probability density functions (pdf) of mutually statistically independent sources by considering higher than second order statistical features. These algorithms arise from the research field

of independent component analysis (ICA) and are often called algorithms based on higher order statistics (HOS) [7]. The algorithms from the second category exploit differences in temporal structure of the sources, typically through the use of second order statistical features. These algorithms use second order statistics (SOS) in order to exploit coloredness and non-stationarity of mutually uncorrelated sources [8]. Differences between the algorithms within a category are often related to implementation, complexity, and convergence.

Methods that do not require a spatial whitening pre-processing step and that are flexible in their use of second order and higher order statistical features to perform BSP tasks exist as well [36]. In [36] the main focus is on the blind system identification (BSI) problem where the objective is to recover the individual mixing columns, up to the well known permutation and arbitrary scaling, instead of the source signals. By structuring the sensor correlation data in a specific form, the BSI problem is projected onto two mathematical problems, i.e., on a root finding problem for a system of homogeneous polynomials and on a multi-matrix generalized eigenvalue decomposition problem. Although BSI can be used in a two-stage approach to perform source extraction, the derivation and formulation of two dual mathematical problems are considered more relevant for the derivation of new informed source extraction algorithms in the remainder of this thesis.

The outline of this chapter is as follows. First we introduce and discuss the mixing model and widely used assumptions on the mixing system and source and noise signals in Section 2.2. In Section 2.3, BSS and BSE algorithms are discussed that assume non-Gaussianity of the sources and exploit higher than second order statistical features to deal with these activities. BSS and BSE algorithms that exploit temporal structure in the source signals and second order statistical features are explained in Section 2.4. The methods to formulate the BSI problem as two dual mathematical problems are reviewed in Section 2.5. Finally, this chapter is concluded in Section 2.6.

2.2 Model and assumptions

In order to facilitate the detailed discussions on the source extraction and separation algorithms, we first present a model for the observed mixtures that is typically considered in the literature. Subsequently, we present commonly used assumptions on the mixing system and the source and noise signals.

2.2.1 The mixing model

As discussed in (1.1) of Chapter 1, in a source extraction scenario with D sensors the observations are assumed to have the following structure in terms of mixing parameters, source signals, and noise signals:

$$x_i[n] = \sum_{j \in \mathcal{S}} a_i^j \star s_j[n] + \nu_i[n] \quad \forall i \in \mathcal{D} \tag{2.1}$$

where the operator $\star$ depends on the type of mixing system. For instantaneous mixing systems, the operation $\star$ represents multiplication and the mixing parameters a_i^j are, potentially complex-valued, gains. For convolutive mixtures the operation $\star$ represents convolution and the mixing parameters a_i^j are impulse responses or filters.

Although direct methods exist [17,37,38], a filter bank or Fourier transform is often used to simplify the source extraction or separation problem for these convolutive mixtures [9]. When the frequency resolution of the filter bank or Fourier transform is sufficiently large, a convolutive mixing problem can be approximated as independent complex-valued instantaneous mixing problems per frequency bin. In this thesis we mainly focus on source extraction for real-valued and complex-valued instantaneous mixtures and in the remainder of this chapter the methodologies are discussed for instantaneous mixing scenarios, unless explicitly mentioned otherwise.

For instantaneous mixing systems consisting of S sources and D sensors, the following vector-matrix representation of the sensor signals can be derived from the mixing model in (2.1):

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n] \tag{2.2}$$

where the signals are collected in the vectors

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x_1[n] \\ \vdots \\ x_D[n] \end{bmatrix}, \ \mathbf{s}[n] \triangleq \begin{bmatrix} s_1[n] \\ \vdots \\ s_S[n] \end{bmatrix}, \text{ and } \boldsymbol{\nu}[n] \triangleq \begin{bmatrix} \nu_1[n] \\ \vdots \\ \nu_D[n] \end{bmatrix} \tag{2.3}$$

and for the mixing system we have the matrix

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{a}^1 & \cdots & \mathbf{a}^S \end{bmatrix} = \begin{bmatrix} a_1^1 & \cdots & a_1^S \\ \vdots & \ddots & \vdots \\ a_D^1 & \cdots & a_D^S \end{bmatrix} \tag{2.4}$$

where the mixing coefficients a_i^j for $(i, j) \in \mathcal{D} \times \mathcal{S}$ are either real-valued or complex-valued scalars.

Whether sources can be extracted or separated depends on properties of and knowledge about the mixing system and the source and noise signals. In the remainder of this section we discuss assumptions on mixing systems and source and noise signals frequently made in the field of blind signal processing (BSP). Furthermore, we discuss two widely known model indeterminacies. Later in this chapter we show how these assumptions are used and how the indeterminacies are dealt with in different algorithms.

2.2.2 Assumptions on the mixing system and signals

In the field of BSP only the observations and assumptions on the source and noise signals are used in order to perform the task at hand, i.e., extraction or separation of

sources. Assumptions shared among almost all literature on BSP are that the source signals have zero mean and that the sources are ergodic, stationary and *mutually statistically independent* [7,8,11]. The latter assumption can be expressed mathematically in the following way:

$$p_s(s_1, s_2, \cdots, s_S) = p_1(s_1)p_2(s_2)\cdots p_S(s_S) \tag{2.5}$$

where $p_s(s_1, s_2, \cdots, s_S)$ is the joint probability density function of the source signals and $p_j(s_j)$ for $j \in \mathcal{S}$ are marginal probability density functions for the individual sources. This independence assumption means that for any order statistics the cross-cumulants of the source signals are zero. In practice, this property is typically exploited by considering up to at most fourth order statistical features. However, higher than fourth order statistical features can be considered as well. Next to the assumption of statistical independence, the source signals are assumed to be sufficiently rich and diverse. These latter assumptions mean that the considered statistical features are non-zero and different from each other. When considering higher than second order statistical features this implies that the sources are considered to have a non-Gaussian distribution.

Another widely used assumption is that the source signals contain *temporal structure*. This temporal structure can occur in the form of coloration and non-stationarity of the signals. Also when exploiting statistical features based on temporal structure of the sources, these features have to be sufficiently diverse. Note that temporal structure is assumed not to exist among different source signals, i.e., the mutual statistical independence assumption includes that a delayed source signal is also mutually independent from other, potentially delayed, source signals. Later in this chapter we discuss several source extraction methods that exploit some form of the assumptions on the source signals presented in this section; however, first we introduce assumptions on the noise signals.

In many papers the noise signals are assumed to have a relatively simple temporal and spatial structure such that statistical parameters of the noise signals can either can be estimated or are considered to be zero for the considered statistical parameters. For example, when the noise signals are spatially and temporally white Gaussian signals, then the powers can be estimated during noise only segments such that the noise can be removed from the sensor correlation data. For an over-determined mixing system the information from the extra sensors can be used to estimate the noise powers [8]. In order to focus on the source extraction and separation methodologies, we neglect noise in the observations in the remainder of this chapter, unless specifically mentioned otherwise. This leads to the following noise-free mixing model for the observations $\mathbf{x}[n]$:

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] \tag{2.6}$$

Next we use this noise-free model in order to discuss two model indeterminacies that are characteristic for BSP problems.

2.2.3 Indeterminacies in blind signal processing mixing models

The mixing model and the corresponding assumptions on the source and noise signals in the previous section are sufficient to allow for tackling BSP problems, up to a certain extent. With these assumptions, identification of the mixing matrix and extraction or separation of the sources can be accomplished up to two model indeterminacies. These two model indeterminacies manifest themselves in an unknown scaling and permutation of the mixing columns and the sources. The indeterminacies cannot be identified without exploiting additional a priori information; however, different heuristic approaches exist to handle them [7, 8, 11]. In this section we discuss the basics of these two indeterminacies.

The model indeterminacies originate from the multiplicative structure of the mixing parameters and the source signals in the mixing model. This multiplicative structure allows for the following decomposition of the observations according to:

$$\mathbf{x}[n] = \mathbf{A}\left(\mathbf{V}\right)^{-1}\mathbf{V}\mathbf{s}[n] = \hat{\mathbf{A}}\hat{\mathbf{s}}[n] \tag{2.7}$$

where $\mathbf{V}$ is an invertible matrix of size $S \times S$, $\hat{\mathbf{s}}[n] = \mathbf{V}\mathbf{s}[n]$ are virtual source signals and $\hat{\mathbf{A}} = \mathbf{A}\left(\mathbf{V}\right)^{-1}$ is a virtual mixing matrix.

Enforcing the assumption of mutually statistically independent source signals restricts the permissible structure of transformation matrix $\mathbf{V}$. In case the virtual source signals need to be mutually statistically independent, then the product $\mathbf{V}\mathbf{s}[n]$ cannot result in linear combinations of the source signals. The only operations that do not reduce the number of sources and do not destroy statistical independence are scaling and permutation of the observations. Consequently, if $\mathbf{V}$ is different from scaled permutation, then the virtual sources are linear combinations of the independent sources, which means that at least two sources are mutually statistically dependent. There is a special case where multiple sources are merged and can be considered as a single source. In this special case the mixing system is not full rank. As long as the desired source is not one of the merged sources, this specific problem can be dealt with in a source extraction problem by considering the multiple sources as a single source.

Since only a scaling and permutation of the mixing columns and the source signals is allowed, the following decomposition of the observations can be considered:

$$\mathbf{x}[n] = \mathbf{A}\left(\mathbf{P}\mathbf{D}\right)^{-1}\mathbf{P}\mathbf{D}\mathbf{s}[n] \tag{2.8}$$

where $\mathbf{P}$ and $\mathbf{D}$ are a permutation and scaling matrix of size $S \times S$, respectively. A permutation matrix has exactly one non-zero value in each row and column, being a one, and zeros elsewhere, whereas a scaling matrix has non-zero scaling values on its diagonal and zeros elsewhere. Blind identification of the mixing system is accomplished in case the mixing matrix $\mathbf{A}$ is identified up to the unknown scaling and permutation matrix. Similarly, BSS is accomplished in case the sources $\mathbf{s}[n]$ are found in an arbitrary order with arbitrary variance and a BSE is successful in case one of the sources with arbitrary variance is found.

In the remainder of this chapter we discuss several source separation and extraction algorithms that exploit different assumptions on the mixing system and signals.

Methods that exploit higher than second order statistical features are discussed in Section 2.3. In Section 2.4 methods that exploit temporal structure in the form of coloredness and non-stationarity are discussed. Most of these methods work for both real-valued and complex-valued mixing scenarios. However, in order to focus on the methodologies instead of mathematical details we discuss the algorithms in the context of real-valued mixing scenarios.

2.3 BSP methods exploiting non-Gaussianity

As discussed in Section 2.1, BSP algorithms employ two activities, i.e., spatial whitening and identification of an unknown rotation. In many BSP algorithms these two activities are implemented in consecutive steps. Therefore, we first discuss the principle of de-correlation and spatial whitening of the data. Subsequently, we discuss several BSS and BSE algorithms that rely on non-Gaussianity of the source signals and exploit HOS in order to identify the remaining rotation.

2.3.1 De-correlation and spatial whitening

De-correlation and spatial whitening are two related signal processing techniques that are widely used in BSP. De-correlation and spatial whitening do not induce separation of the sources; however, the observations are transformed into mutually uncorrelated signals. A consequence of spatial whitening is that the mixing system becomes an orthogonal mixing system.

A technique for performing de-correlation is principal component analysis (PCA). PCA can be applied via the eigenvalue decomposition of the following correlation matrix of noise-free observations $\mathbf{x}[n]$, also called sensor correlation matrix:

$$\mathbf{R}^x = \begin{bmatrix} r^x_{11} & \cdots & r^x_{1D} \\ \vdots & \ddots & \vdots \\ r^x_{D1} & \cdots & r^x_{DD} \end{bmatrix} \triangleq \mathbb{E}\{\mathbf{x}[n]\,(\mathbf{x}[n])^T\} \tag{2.9}$$

where $r^x_{i_1 i_2}$ is defined as the correlation between two signals $x_{i_1}[n]$ and $x_{i_2}[n]$, i.e.,

$$r^x_{i_1 i_2} \triangleq \mathbb{E}\{x_{i_1}[n] x_{i_2}[n]\} \quad \forall\,(i,j) \in \mathcal{D} \times \mathcal{D}. \tag{2.10}$$

From the structure in the noise-free mixing model in (2.6) it follows that the sensor correlation matrix has the following structure:

$$\mathbf{R}^x = \mathbf{A}\mathbf{R}^s(\mathbf{A})^T \tag{2.11}$$

where the source correlation matrix $\mathbf{R}^s$ is a diagonal matrix with the following elements on its diagonal:

$$r^x_{jj} \triangleq \mathbb{E}\{s_j[n] s_j[n]\} \quad \forall\, j \in \mathcal{S}. \tag{2.12}$$

Note that the source correlation matrix $\mathbf{R}^s$ is diagonal due to the assumption of mutually statistically independent sources. Generalizations of these correlations are discussed later in Sections 2.3 and 2.4.

The eigenvalue decomposition of the sensor correlation matrix is given as follows:

$$\mathbf{R}^x \triangleq \mathbf{U}\mathbf{\Lambda}\left(\mathbf{U}\right)^{-1} \tag{2.13}$$

where $\mathbf{U}$ contains the eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of the sensor correlation matrix. Applying the matrix containing the eigenvectors to the observations leads to de-correlation of the observations, i.e, the correlation matrix of the de-correlated observations $(\mathbf{U})^T\mathbf{x}[n]$ is a diagonal matrix.

Next to de-correlation, PCA can be used to perform dimensionality reduction. Dimensionality reduction is applied in case the number of sensors is larger than the number of sources. Ideally, D observations of S sources, with $D > S$, are mapped to S observations without loosing information about the sources. With PCA, this dimensionality reduction is accomplished by using only the eigenvectors corresponding to the S largest eigenvalues.

A generalization of de-correlation that is often applied in the field of BSP is spatial whitening. Via spatial whitening the observations are transformed into a set of uncorrelated signals having normalized power. Whitening of the data is accomplished by applying a whitening transform $\mathbf{Q}$ onto the observations such that the correlation matrix of the transformed sensor data is the identity matrix.

The spatial whitening matrix $\mathbf{Q}$ can also be calculated from the eigenvalue decomposition of the sensor correlation matrix $\mathbf{R}^x$. Given the eigenvalue decomposition in (2.13), the whitening matrix is calculated as:

$$\mathbf{Q} = \mathbf{\Lambda}^{-\frac{1}{2}}\left(\mathbf{U}\right)^T \tag{2.14}$$

Again, dimensionality reduction can be incorporated by selecting only the eigenvectors corresponding to the S largest eigenvalues. An alternative approach to calculate a spatial whitening matrix is by performing a Cholesky decomposition on the sensor correlation matrix, i.e., $\mathbf{R}^x = \mathbf{L}(\mathbf{L})^T$, where $\mathbf{L}$ is a lower triangular matrix with real-valued and positive diagonal elements. The whitening matrix is then given as $\mathbf{Q} = (\mathbf{L})^{-1}$.

A consequence of spatial whitening in the instantaneous mixing scenario is that the mixing column vectors become orthogonal to each other. However, although the whitened signals are mutually uncorrelated, they are not necessarily independent. The reason for this is that whitened signals are rotation invariant, i.e., if we apply an arbitrary orthogonal matrix to the spatially whitened observations the correlation matrix remains the identity matrix [7, 8, 11]. Consequently, in order to perform source extraction or separation the rotation leading to independent signals has to be identified. In the remainder of this section we discuss well established BSS and BSE algorithms that identify this rotation using higher order statistical features.

2.3.2 Eigenvalue decomposition of a fourth order cumulant matrix

Cumulants are used extensively in statistics and serve to approximate joint probability density functions by considering an expansion of the probability density function up to a certain order. In the field of BSP, cumulants up to the fourth order are typically considered due to computational complexity and for robustness. In this section we also focus on fourth order cumulants.

A cumulant tensor can be seen as a generalization of the sensor correlation matrix by considering more than two signals. In fact, the sensor correlation matrix is a second order cumulant tensor, whereas the fourth order sensor cumulant tensor can be considered a four dimensional matrix built from the following collection of fourth order cumulants of the observations:

$$\kappa^x_{i_1 i_2 i_3 i_4} \triangleq \mathrm{cum}(x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}) \quad \forall\, (i_1, i_2, i_3, i_4) \in \mathcal{D} \times \mathcal{D} \times \mathcal{D} \times \mathcal{D} \tag{2.15}$$

where we have omitted the time index n. For zero mean signals x_{i_1}, x_{i_2}, x_{i_3} and x_{i_4}, such a fourth order cumulant $\kappa^x_{i_1 i_2 i_3 i_4}$ can be estimated as follows:

$$\begin{aligned}\kappa^x_{i_1 i_2 i_3 i_4} = \mathbb{E}\{x_{i_1} x_{i_2} x_{i_3} x_{i_4}\} - \mathbb{E}\{x_{i_1} x_{i_2}\}\mathbb{E}\{x_{i_3} x_{i_4}\} \\ - \mathbb{E}\{x_{i_1} x_{i_3}\}\mathbb{E}\{x_{i_2} x_{i_4}\} - \mathbb{E}\{x_{i_1} x_{i_4}\}\mathbb{E}\{x_{i_2} x_{i_3}\}\end{aligned} \tag{2.16}$$

Cumulants carry certain useful properties that can be exploited when performing BSP. The first property is that any crosscumulant of mutually independent signals is zero, i.e., a cumulant is zero if at least one of the signals used to calculate a crosscumulant is independent from another signal. When a signal is sufficiently rich, the autocumulant of that signal will be non-zero. In summary, this leads to the following conditions for fourth order cumulants of *sufficiently rich*, independent source signals:

$$\kappa^s_{j_1 j_2 j_3 j_4} \triangleq \mathrm{cum}(s_{j_1}, s_{j_2}, s_{j_3}, s_{j_4}) \quad \begin{cases} \neq 0 & \text{if} \quad j_1 = j_2 = j_3 = j_4 \\ = 0 & \text{otherwise} \end{cases} \tag{2.17}$$

For fourth order cumulants, the richness condition is typically fulfilled if the source signals are non-Gaussian. This condition immediately indicates an important limitation of using HOS for BSP. At most one of the source signals is allowed to have a Gaussian probability density function. If multiple sources have a Gaussian probability density function, then these sources cannot be separated using HOS since all their higher than second order statistical features are zero.

The second property exploited in BSP algorithms is that cumulants are multilinear. A consequence of this multi-linearity property in combinations with the assumption of mutually statistically independent sources is that the noise-free sensor cumulant data from (2.6) has the following structure:

$$\kappa^x_{i_1 i_2 i_3 i_4} = \sum_{j=1}^{S} a^j_{i_1} a^j_{i_2} a^j_{i_3} a^j_{i_4} \kappa^s_{jjjj} \quad \forall\, (i_1, i_2, i_3, i_4) \in \mathcal{D} \times \mathcal{D} \times \mathcal{D} \times \mathcal{D}. \tag{2.18}$$

A two dimensional cross section of this fourth order sensor cumulant tensor can be obtained by fixing two of the indices and evaluating the tensor over the remaining indices. Due to symmetry of the fourth order sensor cumulant tensor we can fix the indices i_1 and i_2 and evaluate over the indices i_3 and i_4 without loss of generality. The structure of the resulting sensor cumulant matrix $\mathbf{F}^x_{i_1 i_2}$ of size $D \times D$ in terms of the mixing parameters and the fourth order source autocumulants is as follows:

$$\mathbf{F}^x_{i_1 i_2} = \begin{bmatrix} \kappa^x_{i_1 i_2 11} & \cdots & \kappa^x_{i_1 i_2 1D} \\ \vdots & \ddots & \vdots \\ \kappa^x_{i_1 i_2 D1} & \cdots & \kappa^x_{i_1 i_2 DD} \end{bmatrix} = \mathbf{A}\mathbf{F}^s_{i_1 i_2} \left(\mathbf{A}\right)^T \tag{2.19}$$

where $\mathbf{F}^s_{i_1 i_2}$ is the following diagonal matrix of size $S \times S$ containing weighted source autocumulants:

$$\mathbf{F}^s_{i_1 i_2} \triangleq \operatorname{diag}\left(\begin{bmatrix} a^1_{i_1} a^1_{i_2} \kappa^s_{1111} & \cdots & a^S_{i_1} a^S_{i_2} \kappa^s_{SSSS} \end{bmatrix}\right). \tag{2.20}$$

Instead of a two dimensional cross section of the fourth order sensor cumulant tensor, the fourth order sensor cumulant tensor is often reduced to a sensor cumulant matrix by taking linear combinations of the two dimensional cross sections. Such linear combinations are typically denoted as follows:

$$\mathbf{F}^x(\mathbf{M}) = \sum_{i_1, i_2} m^{i_2}_{i_1} \mathbf{F}^x_{i_1 i_2} = \mathbf{A}\mathbf{F}^s(\mathbf{M}) \left(\mathbf{A}\right)^T \tag{2.21}$$

where $\mathbf{M}$ is a real-valued matrix of size $D \times D$ and $\mathbf{F}^s(\mathbf{M})$ is the following diagonal matrix of size $S \times S$:

$$\mathbf{F}^s(\mathbf{M}) \triangleq \operatorname{diag}\left(\begin{bmatrix} (\mathbf{a}^1)^T \mathbf{M} \mathbf{a}^1 \kappa^s_{1111} & \cdots & (\mathbf{a}^S)^T \mathbf{M} \mathbf{a}^S \kappa^s_{SSSS} \end{bmatrix}\right). \tag{2.22}$$

Note that $\mathbf{F}^x(\mathbf{M}) = \mathbf{F}^x_{i_1 i_2}$ holds when choosing $\mathbf{M} = \mathbf{e}_{i_1} \tilde{\mathbf{e}}^{i_2}$, with $\mathbf{e}_{i_1}$ and $\tilde{\mathbf{e}}^{i_2}$ a column and row vector of length D, with a one at the i_1'th row and i_2'th column, respectively, and zeros elsewhere. We mainly focus on the structure of $\mathbf{F}^x(\mathbf{M})$. Later on, we discuss certain properties that have to hold for the matrix $\mathbf{M}$, but for now $\mathbf{M}$ can be an arbitrary matrix of size $D \times D$.

When spatially whitened sensor data $\breve{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$ instead of the original sensor data $\mathbf{x}$ is used to calculate sensor cumulant matrices such as (2.21), then the mixing matrix is transformed into an orthogonal matrix $\breve{\mathbf{A}} = \mathbf{Q}\mathbf{A}$ and the structure of the sensor cumulant tensor for whitened data becomes an eigenvalue decomposition, i.e.,

$$\mathbf{F}^{\breve{x}}(\mathbf{M}) = \breve{\mathbf{A}}\mathbf{F}^{\breve{s}}(\mathbf{M})(\breve{\mathbf{A}})^T \tag{2.23}$$

where $\mathbf{F}^{\breve{s}}$ is based on the columns of the whitened mixing matrix $\breve{\mathbf{A}}$, i.e.,

$$\mathbf{F}^{\breve{s}}(\mathbf{M}) \triangleq \operatorname{diag}\left(\begin{bmatrix} (\breve{\mathbf{a}}^1)^T \mathbf{M} \breve{\mathbf{a}}^1 \kappa^s_{1111} & \cdots & (\breve{\mathbf{a}}^S)^T \mathbf{M} \breve{\mathbf{a}}^S \kappa^s_{SSSS} \end{bmatrix}\right). \tag{2.24}$$

Consequently, for whitened sensor data the eigenvectors $\boldsymbol{\mu}$ that solve eigenvalue problem $\mathbf{F}^{\breve{x}}(\mathbf{M})\boldsymbol{\mu} = \lambda\boldsymbol{\mu}$ correspond to the columns of the orthogonal mixing matrix

$\breve{\mathbf{A}}$, up to an arbitrary permutation and scaling as discussed in Section 2.2.3. Note that when dimensionality reduction is applied the virtual number of sensors is S, leading to a matrix $\mathbf{M}$ of size $S \times S$.

In combination with the spatial whitening matrix $\mathbf{Q}$, an estimate $\hat{\mathbf{A}}$ of the original mixing matrix $\mathbf{A}$ and an estimate $\hat{\mathbf{W}}$ of the source separation filter $\mathbf{W}$ can be calculated, up to the unknown permutation and scaling, in the following way:

$$\hat{\mathbf{A}} = (\mathbf{Q})^{-1}\,\hat{\breve{\mathbf{A}}} \tag{2.25}$$

$$\hat{\mathbf{W}} = (\hat{\breve{\mathbf{A}}})^T\mathbf{Q} \tag{2.26}$$

where $\hat{\breve{\mathbf{A}}}$ is the estimated orthogonal mixing matrix built from the collection of eigenvectors of $\mathbf{F}^{\breve{x}}(\mathbf{M})$.

A potential problem of this eigenvalue decomposition method is that it only works if all eigenvalues are distinct. If two or more eigenvalues are the same, then only the linear subspace spanned by the corresponding eigenvectors can be identified. A straightforward approach to deal with this problem is to choose different sensor cumulant matrices, i.e., matrices $\mathbf{M}$, until unique eigenvalues are found. If $\mathbf{M}$ is a randomly chosen matrix, it is in practice unlikely that eigenvalues are equal; however, for certain matrices $\mathbf{M}$ the sensor cumulant matrix $\mathbf{F}^{\breve{x}}(\mathbf{M})$ may become ill-conditioned.

A summary of the method for identifying the mixing system or source separation matrix with arbitrary scaling and permutation as presented in this section is as follows:

- Compute the sensor correlation matrix $\mathbf{R}^x = \mathbb{E}\{\mathbf{x}[n]\,(\mathbf{x}[n])^T\}$
- Compute the spatial whitening matrix $\mathbf{Q}$ from the sensor correlation matrix
- Compute whitened observations $\breve{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$
- Compute fourth order cumulants $\kappa^{\breve{x}}_{i_1 i_2 i_3 i_4}$ of the whitened data
- Construct $\mathbf{F}^{\breve{x}}(\mathbf{M})$ from the fourth order cumulants and a random matrix $\mathbf{M}$ of size $S \times S$
- Compute the eigenvalue decomposition $\mathbf{F}^{\breve{x}}(\mathbf{M}) = \breve{\mathbf{A}}\mathbf{\Lambda}(\breve{\mathbf{A}})^T$
- Compute the estimated source signals by $\hat{\mathbf{s}}[n] = (\breve{\mathbf{A}})^T\breve{\mathbf{x}}[n]$

In the following section the eigenvalue decomposition based approach is extended to the simultaneous diagonalization of multiple sensor cumulant matrices.

2.3.3 Simultaneous diagonalization of cumulant matrices

The BSS algorithm from the previous section can be extended to the use of multiple fourth order cumulant matrices [39]. The idea is to increase robustness and to evade the problem of similar eigenvalues by simultaneous diagonalization of multiple sensor cumulant matrices. In other words, the objective is to identify a single matrix $\mathbf{V}$ such that the following matrices $(\mathbf{V})^T\mathbf{F}^x(\mathbf{M})\mathbf{V}$ for several matrices $\mathbf{M}$ are diagonal. When $\mathbf{V}$ diagonalizes all considered sensor cumulant matrices $\mathbf{F}^x(\mathbf{M})$ then its columns correspond again to the columns of the mixing matrix $\mathbf{A}$, up to an arbitrary permutation and scaling factor.

Simultaneous diagonalization is only feasible for series of matrices with very specific structure. The cumulant matrices that follow the model in (2.6) ideally contain this specific structure; however, in practice the cumulant matrices suffer from modelling and estimation errors. Consequently, simultaneous diagonalization has to be approximated. An algorithm found in the literature that performs this approximation is the joint diagonalization of eigenmatrices (JADE) algorithm [39].

The JADE algorithm requires the observations to be spatially whitened. As we have seen in the previous section, diagonalization of a single sensor cumulant matrix based on whitened observations $\breve{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$ can be accomplished with an orthogonal matrix $\mathbf{V}$.

A natural way to measure diagonality of a matrix is the ratio of sum of squares of diagonal elements over the sum of squares of off-diagonal elements. Since the sum of squares of a matrix remains constant when orthogonal transformations are applied, the following two approaches can be considered as well to diagonalize multiple sensor cumulant matrices:

- Minimize the sum of squares of off-diagonal elements of the transformed sensor cumulant matrices $(\mathbf{V})^T \mathbf{F}^{\breve{x}}(\mathbf{M})\mathbf{V}$
- Maximize the sum of squares of diagonal elements of the transformed sensor cumulant matrices $(\mathbf{V})^T \mathbf{F}^{\breve{x}}(\mathbf{M})\mathbf{V}$

Note that these objectives are only equivalent if the observations are whitened. This equivalence leads to the following objective function used by the JADE algorithm:

$$J_{JADE}(\mathbf{V}, \mathbf{F}^{\breve{x}}) = \sum_i \left|\left|\text{diag}\left((\mathbf{V})^T \mathbf{F}^{\breve{x}}(\mathbf{M}_i)\mathbf{V}\right)\right|\right|_2^2 \tag{2.27}$$

for several matrices $\mathbf{M}_i$. Maximization of this objective function can be accomplished using the Jacobi technique extended for diagonalization of multiple matrices. The Jacobi technique works by successively applying Givens rotations, which makes the JADE algorithm an iterative algorithm.

What remains is the selection of a set of matrices $\mathbf{M}_i$. In Section 2.3.2 we already discussed the use of a single (arbitrary) matrix $\mathbf{M}$. A natural choice for multiple matrices $\mathbf{M}_i$ that utilize all available cumulant data is to select these matrices as $\mathbf{e}_{i_1}\tilde{\mathbf{e}}^{i_2}$ for all $(i_1, i_2) \in \mathcal{S} \times \mathcal{S}$, which leads to a total of S^2 matrices $\mathbf{M}_i$; however, this approach with S^2 matrices is computationally unattractive for mixing systems with many sources. Furthermore, the method is redundant since a total of S properly chosen matrices $\mathbf{M}_i$ are sufficient in order to consider all available cumulant data. These S properly chosen matrices can be obtained by selecting the eigen-matrices that correspond to the S largest eigenvalues of $\mathbf{F}^{\breve{x}}(\mathbf{M}_i) = \lambda\mathbf{M}_i$. In the literature many alternatives for obtaining a suitable set of matrices $\mathbf{M}_i$ are available, always making a trade-off between robustness and computational complexity.

2.3.4 BSE by maximizing non-Gaussianity

The methods from the previous sections where sensor cumulant matrices are diagonalized are typically used to identify the entire mixing system or separate all sources simultaneously. In [40] a well known family of algorithms has been introduced that can be used for both blind extraction of one or a few sources and for blindly separating all sources in a sequential fashion. In this section we focus on the extraction of only a single source signal. Deflation techniques for sequential extraction of additional source signals are discussed in for example [7, 8].

BSS was first studied in the field of independent component analysis (ICA). The fundament of ICA comes from the Central Limit Theorem, which states that under certain conditions the distribution of a sum of independent random variables tends towards a Gaussian distribution [8]. Consequently, identification of an extraction filter can be accomplished by maximizing a measure of non-Gaussianity at the output. Often used measures of non-Gaussianity are kurtosis and negentropy; however, many different measures have been proposed in the literature [7, 8, 11]. In this section we discuss gradient and fixed-point iteration BSE algorithm based on kurtosis and negentropy. A more appropriate name for the fixed-point iteration algorithms is to describe them as Newton methods; however, in the original publication the term fixed-point was already used [40].

2.3.4.1 BSE algorithms based on kurtosis

Kurtosis is the fourth-order cumulant of a signal and for zero mean signals the kurtosis is defined as follows:

$$\text{kurt}(y) = \mathbb{E}\{(y[n])^4\} - 3(\mathbb{E}\{(y[n])^2\})^2 \tag{2.28}$$

where $y[n] = \tilde{\mathbf{w}}\mathbf{x}[n]$. When the observations are spatially whitened and the extraction filters have unit norm, then we have $\mathbb{E}\{(y[n])^2\} = 1$, leading to $\text{kurt}(y[n]) = \mathbb{E}\{(y[n])^4\} - 3$. For a Gaussian signal we have $\mathbb{E}\{(y[n])^4\} = 3(\mathbb{E}\{(y[n])^2\})^2$ such that the kurtosis of a Gaussian signal is zero. Consequently, kurtosis can be considered a measure of non-Gaussianity since a non-zero kurtosis indicates that the signal does not have a Gaussian distribution.

Source extraction can be accomplished by maximizing the absolute kurtosis of the extracted signal, under the constraints of a normalized extraction filter and spatially whitened observations, i.e.,

$$\tilde{\mathbf{w}}_{opt} = \max_{\tilde{\mathbf{w}}} \left| (\tilde{\mathbf{w}}\check{\mathbf{x}}[n])^4 - 3 \right| \tag{2.29}$$

$$\text{s. t. } ||\tilde{\mathbf{w}}||_2^2 = 1 \tag{2.30}$$

where $\check{\mathbf{x}}[n]$ are spatially whitened observations.

Next we present two optimization algorithms for finding the filter $\tilde{\mathbf{w}}$ that extracts one of the independent components from a spatially whitened mixture. The first

algorithm is a gradient algorithm and the second algorithm is a fixed-point iteration algorithm.

The gradient of the objective function in (2.29), i.e., the absolute kurtosis of the extracted signal, is given as follows:

$$\frac{\partial \left|\text{kurt}(\tilde{\mathbf{w}}\breve{\mathbf{x}}[n])\right|}{\partial \mathbf{w}} = 4\,\text{sign}(\text{kurt}(\tilde{\mathbf{w}}\breve{\mathbf{x}}[n])\left[\mathbb{E}\{\breve{\mathbf{x}}(\tilde{\mathbf{w}}\breve{\mathbf{x}})^3\} - 3\mathbf{w}\,||\tilde{\mathbf{w}}||_2^2\right]. \tag{2.31}$$

Note that we take the gradient with respect to $\mathbf{w}$ instead of $\tilde{\mathbf{w}}$ for notational convenience. In the remainder of this section we assume that $\tilde{\mathbf{w}} = (\mathbf{w})^T$.

This gradient can be used in an iterative or adaptive algorithm in order to identify a source extraction filter. In order to solve the optimization problem appropriately, the extraction filter coefficients $\tilde{\mathbf{w}}$ have to be normalized after every step. Due to this normalization, only the direction of the extraction filter is of interest when calculating an update of the extraction filter based on the gradient. Therefore it is justified to remove the second term in the brackets of the gradient since it mainly scales the resulting filter.

These observations lead to the following gradient algorithm for finding a source extraction filter by maximizing the absolute kurtosis:

- Compute the sensor correlation matrix $\mathbf{R}^x = \mathbb{E}\{\mathbf{x}[n]\,(\mathbf{x}[n])^T\}$
- Compute the spatial whitening matrix $\mathbf{Q}$ from the sensor correlation matrix
- Compute whitened observations $\breve{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$
- Iterate the following steps until convergence:
  - $\mathbf{w} \leftarrow \mathbf{w} + \gamma\,\text{sign}\{\text{kurt}(\tilde{\mathbf{w}}\breve{\mathbf{x}}[n])\}\mathbb{E}\{\breve{\mathbf{x}}[n](\tilde{\mathbf{w}}\breve{\mathbf{x}}[n])^3\}$
  - $\tilde{\mathbf{w}} \leftarrow (\mathbf{w})^T / \,||\mathbf{w}||_2$
- Compute the extracted signal $y[n] = \tilde{\mathbf{w}}\mathbf{Q}\mathbf{x}[n]$.

where $\gamma > 0$ is a learning rate parameter. One way to detect convergence is to calculate the inner product between two consecutive extraction filters. In case this inner product is close to one, the iteration step provides only a small update indicating that the algorithm hs converged. Extensions of this fixed-point iteration algorithm, for example for real-time implementation can be found in [7].

A disadvantage of gradient based methods is the dependence on a step size or learning rate parameter. Inappropriate selection of this parameter leads to slow convergence or instability of the gradient algorithm. In [40] fixed-point iteration source extraction algorithms with a fast convergence and no learning rate parameter were presented. These properties make the algorithms both fast and reliable [7]. Next we discuss a fixed-point iteration algorithm that maximizes the absolute kurtosis of the extracted signal.

In [7] it is observed for the gradient algorithm that a source extraction filter is identified in case the gradient of the absolute kurtosis in (2.31) points in the direction of the extraction filter. In other words, at an optimal point the gradient is given as $\alpha\tilde{\mathbf{w}}$, with α a non-zero scalar and $\tilde{\mathbf{w}}$ an extraction filter with unit norm. The fixed-point iteration algorithm is an iterative method that works by fixing the new extraction filter in the direction of the normalized gradient of the absolute kurtosis in (2.31), which is calculated with using the extraction filter from the previous iteration.

In a two step approach first the gradient in (2.31) is calculated, followed by a normalization step. In such a case, we are again only interested in the direction of the gradient and not its magnitude. Consequently, we are able to remove all scaling related parts in (2.31), leading to the following estimate of the gradient:

$$\mathbf{w} \leftarrow \mathbb{E}\{\breve{\mathbf{x}}[n](\tilde{\mathbf{w}}\breve{\mathbf{x}}[n])^3\} - 3\mathbf{w} \tag{2.32}$$

Note that we omitted the term $||\tilde{\mathbf{w}}||_2^2$ since it equals one for normalized extraction filters.

The fixed-point iteration algorithm that maximizes absolute kurtosis is summarized as follows:

- Compute the sensor correlation matrix $\mathbf{R}^x = \mathbb{E}\{\mathbf{x}[n]\,(\mathbf{x}[n])^T\}$
- Compute the spatial whitening matrix $\mathbf{Q}$ from the sensor correlation matrix
- Compute whitened observations $\breve{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$
- Iterate the following steps until convergence:
  - $\mathbf{w} \leftarrow \mathbb{E}\{\breve{\mathbf{x}}[n](\tilde{\mathbf{w}}\breve{\mathbf{x}}[n])^3\} - 3\mathbf{w}$
  - $\tilde{\mathbf{w}} \leftarrow \mathbf{w}/\,||\mathbf{w}||_2$
- Compute the extracted signal $y[n] = \tilde{\mathbf{w}}\mathbf{Q}\mathbf{x}[n]$.

Again the inner product between two consecutive extraction filters can be used to in order to test for convergence. A more formal algorithm derivation of this algorithm using Newton's method can be found in [7].

BSE algorithms based on kurtosis are relatively easy to compute; however, a drawback of using kurtosis in practice is its sensitivity to outliers [7]. In order to obtain robust algorithms that produce reliable extraction filters, a lot of (stationary) data is required. By using alternative measures of non-Gaussianity more robust BSE algorithms can be derived. One of such measures is negentropy. Negentropy is a more robust measure of non-Gaussianity than kurtosis; however, it is also more computationally demanding.

2.3.4.2 BSE algorithms based on negentropy

A general formulation of ICA based on the concept of mutual information between signals instead of a data model such as the often used instantaneous mixing model is introduced in [32]. By formulating mutual information as a function of negentropy it was shown that maximization of negentropy leads to extraction of independent source signals. First we discuss the concept of mutual information, followed by an overview of fast and robust source extraction algorithms.

Negentropy is a measure of non-Gaussianity based on the information-theoretic quantity of differential entropy. Differential entropy is a measure for the unpredictability of a random variable and is defined as follows:

$$H(y) = -\int f(y)\log f(y)dy \tag{2.33}$$

where $f(y)$ represents the probability density function of the random variable y. Given a certain variance, differential entropy is maximal when the random variable y has

a Gaussian distribution [32]. This property can be used to normalize differential entropy, leading to the following definition of negentropy, i.e.,

$$J(y) = H(y_{\text{gauss}}) - H(y) \tag{2.34}$$

where y_{gauss} is a Gaussian random variable with the same variance as y. Normalization of differential entropy using a Gaussian random variable with the same variance as y ensures that the negentropy is zero for a Gaussian signal and positive otherwise [32], which makes negentropy a natural measure for non-Gaussianity.

Negentropy can be used in order to express the dependence between random variables in the form of mutual information. For mutually uncorrelated signals, mutual information in terms of negentropy is defined as follows:

$$I(y_1, y_2, \cdots, y_S) \triangleq J(y_1, y_2, \cdots, y_S) - \sum_{j=1}^{S} J(y_j) \tag{2.35}$$

Minimizing mutual information means that mutual independence of the individual components y_j is maximized. Minimization of mutual information is accomplished in case the negentropy of the individual components y_j is maximized. When choosing the components y_j as the outputs of a spatial filter applied to spatially whitened observations, then maximizing the negentropy of the output leads to extraction of one of the independent source signals. A drawback of negentropy is that it is computationally demanding and it requires the estimation of the probability density functions of the signals [7]. Therefore, approximations of negentropy are typically used in order to obtain robust and efficient algorithms for performing BSE [7, 8, 40]. In the remainder of this section we derive a family of algorithms based on maximization of approximations of negentropy.

In [40] several approximations of negentropy are considered that all share the following common form:

$$J(y) = [\mathbb{E}\{G(y)\} - \mathbb{E}\{G(v)\}]^2 \tag{2.36}$$

where v is a Gaussian variable with zero mean and unit variance, and $G(\cdot)$ is a suitably chosen non-quadratic function that approximates negentropy. How to choose a suitable approximation $G(\cdot)$ is explained later in this section.

Using this approximation, the following constrained optimization problem can be formulated in order to extract an arbitrary independent source signal from spatially whitened observations [40]:

$$\begin{aligned} \tilde{\mathbf{w}} &= \arg\max_{\tilde{\mathbf{w}}} |\mathbb{E}\{G(y[n])\} - \mathbb{E}\{G(v)\}|^2 \\ &\text{s. t. } ||\tilde{\mathbf{w}}||_2^2 = 1 \end{aligned} \tag{2.37}$$

where $y[n] = \tilde{\mathbf{w}}\breve{\mathbf{x}}[n]$, $\breve{\mathbf{x}}[n]$ are spatially whitened observations, and $G(y)$ is a suitably chosen convex function that approximates negentropy.

Different algorithms can be derived that solve the optimization problem in (2.37). In the previous section a gradient algorithm and a fixed-point iteration algorithm are introduced for maximizing absolute kurtosis. Such algorithms can also be derived for maximizing approximations of negentropy. In this section we focus on the fixed-point iteration algorithm for solving the optimization problem in (2.37). Details of a gradient algorithm approach can be found for example in [7].

The fixed-point iteration approach from the previous section can be used in order to derive a fast method for finding a maximum of the optimization problem in (2.37). This fast method is well known in literature as the FastICA algorithm [40]. The gradient of (2.37) is given as follows:

$$\frac{\partial \left|\mathbb{E}\{G(y[n])\} - \mathbb{E}\{G(v)\}\right|^2}{\partial \mathbf{w}} = 2\left[\mathbb{E}\{G(y[n])\} - \mathbb{E}\{G(v)\}\right]\mathbb{E}\{g(y[n])\check{\mathbf{x}}[n]\} \tag{2.38}$$

where $g(\cdot)$ is the derivative of $G(\cdot)$. Note that we again calculate the gradient with respect to $\mathbf{w}$ instead of $\tilde{\mathbf{w}}$. Also in the remainder of this section we assume that $\tilde{\mathbf{w}} = (\mathbf{w})^T$.

Combining the gradient with a normalization step for the extraction filter allows us to ignore the scaling term $2\left[\mathbb{E}\{G(y[n])\} - \mathbb{E}\{G(v)\}\right]$, which leads to the following fixed-point iteration update for identifying the source extraction filter:

$$\mathbf{w} \leftarrow \mathbb{E}\{g(y[n])\check{\mathbf{x}}[n]\}. \tag{2.39}$$

Due to differences in algebraic properties, convergence properties of this update step, complemented with normalization, for approximations of negentropy are not as good as the convergence properties of the fixed-point iteration algorithm based on kurtosis. In order to improve the convergence properties an approximative Newton method was introduced in [40] that has an update of the following form:

$$\mathbf{w} \leftarrow \mathbb{E}\{g(y[n])\check{\mathbf{x}}[n]\} - \mathbb{E}\{g'(y[n])\}\mathbf{w} \tag{2.40}$$

where $g'(\cdot)$ is the derivative of $g(\cdot)$ and the extraction filter is normalized after each iteration.

In order to further increase robustness of the fixed-point iteration source extraction algorithm a stabilized version of the fixed-point iteration algorithm was introduced in [40]. This stabilized version uses a step size parameter μ that can be decreased with the iteration count. Choosing small μ increases certainty of converge at the cost of convergence speed. The stabilized fixed-point iteration algorithm has the following update rule:

$$\mathbf{w} \leftarrow \mathbf{w} - \mu\frac{\mathbb{E}\{\check{\mathbf{x}}[n]g(y[n])\} - \beta\mathbf{w}}{\mathbb{E}\{g'(y[n])\} - \beta} \tag{2.41}$$

where $\beta = \mathbb{E}\{\tilde{\mathbf{w}}\check{\mathbf{x}}[n]g(y[n])\}$ and the extraction filter is normalized after each iteration.

What remains is the selection of an appropriate approximation for negentropy.

The following approximations were considered in the original paper [40]:

$$G_1(y) = \frac{1}{a_1} \log \cosh(a_1 y) \tag{2.42}$$

$$G_2(y) = -\frac{1}{a_2} \exp(-\frac{a_2 y^2}{2}) \tag{2.43}$$

$$G_3(y) = \frac{1}{4} y^4 \tag{2.44}$$

where $1 \leq a_1 \leq 2, a_2 \approx 1$ are constants. Function $G_1(y)$ is good for general-purpose source extraction. For extracting highly super-Gaussian signals or for high robustness $G_2(y)$ should be used. Using kurtosis, i.e., $G_3(y)$, should only be considered for extracting sub-Gaussian signals since this function is very sensitive to outliers [40].

The FastICA algorithm [40] can be summarized as follows:

- Compute the sensor correlation matrix $\mathbf{R}^x = \mathbb{E}\{\mathbf{x}[n](\mathbf{x}[n])^T\}$
- Compute the spatial whitening matrix $\mathbf{Q}$ from the sensor correlation matrix
- Compute whitened observations $\breve{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$
- Iterate the following steps until convergence:
  1. $\mathbf{w} \leftarrow \mathbb{E}\{g(y[n])\breve{\mathbf{x}}[n]\} - \mathbb{E}\{g'(y[n])\}\mathbf{w}$
  2. $\tilde{\mathbf{w}} \leftarrow (\mathbf{w})^T / ||\mathbf{w}||_2$
- Compute the extracted signal $y[n] = \tilde{\mathbf{w}}\mathbf{Q}\mathbf{x}[n]$

where $g(\cdot)$ and $g'(\cdot)$ are the first and second order derivative of the function $G(\cdot)$, which is used to approximate negentropy. Alternatively, the stabilized update rule from (2.41) with a suitable choice for $\boldsymbol{\mu}$ can be used instead of the update rule from (2.40).

Although these algorithms can be used to extract one or a few sources, it is impossible to extract a specific source or to know which source will be extracted. In [40] it is suggested to choose the approximation of negentropy depending on the attraction pool. This attraction pool is the region in which an algorithm will be attracted to a certain local optimum. By choosing a proper function and initialization, one can increase the chance that immediately the desired source is extracted; however, immediate extraction of the desired source cannot be guaranteed with these methods.

2.3.5 ICA based source extraction using reference signals

In applications where one or more desired sources have to be extracted from a mixture of multiple sources, it is necessary to require a priori information that indicates these desired sources. In [34, 41, 42] it is assumed that this kind of a priori information is available in the form of reference signals that are related to the desired sources. By formulating inequality constraints based on the reference signals, the optimization problem in (2.37) can be extended to become a constrained ICA problem [34]. Such constrained ICA problems can be solved using Newton-like algorithms as presented in [34, 41, 42], leading to source extraction algorithms for which extraction of the desired source is guaranteed.

In this section we discuss the principles of the source extraction algorithms in [34, 41, 42]. First we introduce the constrained ICA framework in a generalized form that allows for both equality and inequality constraints [34]. Subsequently, we show how the reference signals can be incorporated into the constrained ICA framework such that extraction of a desired source signal is guaranteed. Although the constrained ICA framework can be used to extract multiple source signals, we focus on the extraction of a single desired source signal.

As discussed in the previous sections, source extraction or separation can be accomplished by maximizing a certain measure of non-Gaussianity. By adding constraints to this optimization problem, a priori information can be incorporated in order to ensure extraction of a desired source signal. The constrained ICA framework from [34] allows for both equality and inequality constraints.

The general form of the ICA optimization problem including constraints has the following form:

$$\begin{aligned} \max_{\tilde{\mathbf{w}}} \quad & J(y) \\ \text{s.t.} \quad & g_i(\tilde{\mathbf{w}}) \leq 0 \\ & h_j(\tilde{\mathbf{w}}) = 0 \end{aligned} \tag{2.45}$$

where $J(y)$ is a suitable ICA objective function for extracting source signals, $y = \tilde{\mathbf{w}}\mathbf{x}$ is the extracted signal, and $g_i(\tilde{\mathbf{w}})$ for $i \in \mathcal{I}$ and $h_j(\tilde{\mathbf{w}})$ for $j \in \mathcal{J}$ are an unspecified number of inequality and equality constraints.

The optimal solution of the constrained ICA optimization problem can be searched for using the method of Lagrange multipliers [43]. First, inequality constraints $g_i(\tilde{\mathbf{w}}) \leq 0$ are transformed into equality constraints by introducing slack variables z_i, i.e., $\hat{g}_i(\tilde{\mathbf{w}}) = g_i(\tilde{\mathbf{w}}) + (z_i)^2 = 0$. Next, the augmented Lagrangian function corresponding to the optimization problem in (2.45) is given as follow:

$$\mathcal{L}(\tilde{\mathbf{w}}) = J(y) + (\boldsymbol{\mu})^T \hat{\mathbf{g}}(\tilde{\mathbf{w}}) + \frac{1}{2}\gamma \, ||\hat{\mathbf{g}}(\tilde{\mathbf{w}})||_2^2 + (\boldsymbol{\lambda})^T \mathbf{h}(\tilde{\mathbf{w}}) + \frac{1}{2}\gamma \, ||\mathbf{h}(\tilde{\mathbf{w}})||_2^2 \tag{2.46}$$

where $\hat{\mathbf{g}}(\tilde{\mathbf{w}})$ and $\mathbf{h}(\tilde{\mathbf{w}})$ are vectors containing the equality constraints $\hat{g}_i(\tilde{\mathbf{w}})$ and $h_i(\tilde{\mathbf{w}})$, respectively. The vectors $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are vectors containing positive Lagrange multipliers, γ is a positive penalty parameter, and the terms $\frac{1}{2}\gamma\|\cdot\|_2^2$ ensure local convexity [43].

Through several mathematical manipulations and approximations, as discussed in [34], a Newton-like algorithm can be obtained that has the following form:

$$\Delta\tilde{\mathbf{w}} = -\eta \left(\nabla^2_{\tilde{\mathbf{w}}}\mathcal{L}\right)^{-1} \nabla_{\tilde{\mathbf{w}}}\mathcal{L} \tag{2.47}$$

where η equals one or can be decreased in order to ensure stable convergence. The Lagrange multipliers can be learned for example using the following gradient ascent method:

$$\nabla_{\boldsymbol{\mu}} = \max\{-\boldsymbol{\mu}, \eta_\mu \mathbf{g}(\tilde{\mathbf{w}})\} \tag{2.48}$$

$$\nabla_{\boldsymbol{\lambda}} = \eta_\lambda \mathbf{h}(\tilde{\mathbf{w}}) \tag{2.49}$$

where η_μ and η_λ are learning rate parameters.

By carefully selecting a suitable ICA objective function for extracting source signals and constraints based on a priori information in the form of reference signals, the constrained ICA framework can be used in order to develop source extraction algorithms for which extraction of the desired source is guaranteed. In the remainder of this section we discuss how to tailor the constrained ICA framework for this purpose.

In [34, 41, 42] the assumption is made that a reference signal $r[n]$ is available that contains a priori information about the desired source. It is assumed that the reference signal is closer to the desired source than any of the other sources in the mixture, given a certain distance measure. If we denote the distance between reference signal $r[n]$ and output signal $y[n]$ as $\varepsilon(r[n], y[n])$, then the condition on the reference signal and its corresponding distance measure can be formulated as follows:

$$\varepsilon(r[n], s_d[n]) \leq \varepsilon(r[n], s_j[n]) \quad \forall\, j \in \mathcal{S} \tag{2.50}$$

where $s_d[n]$ represents the desired source signal and $s_j[n]$ is the j'th source. Furthermore, equality is assumed to occur only when the extracted signal is the desired source signal, i.e., when $j = d$.

The condition in (2.50) can be transformed into an inequality constraint, which in turn can be applied in the constrained ICA problem. By introducing a threshold parameter ξ, the following inequality constraint can be formulated that allows for the extraction of only the desired source:

$$g(\tilde{\mathbf{w}}) = \varepsilon(r[n], y[n]) - \xi \leq 0 \tag{2.51}$$

From Section 2.3.4 we know that the optimization problem in (2.37) only has S essentially unique solutions, corresponding to the extraction of the S individual source signals. Consequently, extraction of the source closest to the reference signal is guaranteed when selecting the threshold parameter ξ as follows:

$$\varepsilon(r[n], s_d[n]) \;\leq\; \xi \;<\; \min_{j \neq d} \varepsilon(r[n], s_j[n]) \quad \forall\, j \in \mathcal{S} \tag{2.52}$$

In many practical applications the terms $\varepsilon(r[n], s_j[n])$ for $j \in \mathcal{S}$ are unknown. Furthermore, due to noise, modelling mismatches, and a natural distance between the reference signal and the actual source signal, a suitable value for the threshold parameter ξ may be difficult to obtain in practice. Therefore it is suggested [34, 41, 42] to start with a relatively low value for the threshold parameter and slowly increasing the parameter value until convergence.

Suitable choices for the distance measure that can be used in the Newton like algorithm to solve the constrianed ICA problem [34, 41, 42] are mean squared error (MSE) and correlation (COR), i.e.,

$$\varepsilon_{\text{MSE}}(r[n], y[n]) = \mathbb{E}\{(y[n] - r[n])^2\} \quad \text{and} \quad \varepsilon_{\text{COR}} = \frac{1}{(\mathbb{E}\{y[n]r[n]\})^2} \tag{2.53}$$

Both these distance measures require that the reference signal and the extracted signal are normalized, otherwise the measures are unbounded. Normalization of the

extracted signal $y[n]$ can be accomplished if we replace the unit norm constraint for the extraction filter in (2.37) by the following equality constraint:

$$h(\tilde{\mathbf{w}}) = \mathbb{E}\{(y[n])^2\} - 1 = 0. \tag{2.54}$$

Note that the constrained ICA problem is not limited to the presented distance measures.

Combining the results from the previous and the current section leads to the following constrained ICA optimization problem that ensures extraction of a desired source signal based on a reference signal:

$$\begin{aligned} \max_{\tilde{\mathbf{w}}} \quad & J(y[n]) \approx [\mathbb{E}\{G(\tilde{\mathbf{w}}\mathbf{x})\} - \mathbb{E}\{G(\nu)\}]^2 \\ \text{s.t.} \quad & g(\tilde{\mathbf{w}}) = \varepsilon(r[n], y[n]) - \xi \leq 0 \\ & h(\tilde{\mathbf{w}}) = \mathbb{E}\{(y[n])^2\} - 1 = 0 \end{aligned} \tag{2.55}$$

where $J(y[n])$ is the approximation of negentropy from (2.36), $g(\tilde{\mathbf{w}})$ is the inequality constraint from (2.51) that ensures that the desired source is extracted, and $h(\tilde{\mathbf{w}})$ is the equality constraint for normalizing the extracted signal.

2.3.6 Discussion

In many BSP algorithms two steps can be distinguished, as discussed in Section 2.1. The first step is spatial whitening and the second step is identification of an orthogonal matrix to resolve an unknown rotation. In Section 2.3.1 we discussed the processes of spatial whitening and de-correlation. In the remaining sections we presented methods found in the literature that exploit higher than second order statistical features of signals in order to identify the rotation that remains after spatial whitening. Identification of the full rotation matrix allows for separation of all sources while identification of a single column of this rotation matrix allows for extraction of a single source signal.

The algorithms in Sections 2.3.2, 2.3.3 and 2.3.4 perform either blind separation of all sources or extract randomly one of the sources. When using these algorithms, immediate extraction of a desired source cannot be guaranteed. The constrained ICA based method described in Section 2.3.5 is the only BSP based approach found in literature that guarantees immediate extraction of the desired source. Using this approach, immediate extraction of the desired source is guaranteed if certain conditions hold on a priori information in the form of reference signals. Furthermore, by adding additional constraints to the objective function, the process of spatial whitening can be included to obtain a single source extraction process. A disadvantage of this approach is that the algorithms to solve the optimization problems contain many parameters that require tuning or a priori information.

Using higher order statistical features of signals requires these signals to be sufficiently rich and sufficiently stationary. Furthermore, methods based on HOS can only deal with mixtures where at most one of the sources has a Gaussian distribution.

In many practical applications these stationarity and non-Gaussianity requirements are violated, leading to errors in estimated source extraction filters. An alternative approach to solve BSP problems is to exploit temporal structure in the signals. BSP algorithms that exploit this temporal structure are discussed in the next section.

2.4 BSP methods exploiting temporal structure

The BSP methods in the previous section require that mixed source signals are mutually independent, stationary and have different but non-Gaussian distributions. These properties are exploited by considering higher than second order statistical features in order to perform the BSP task at hand. A different category of BSP methods exists that require the independent source signals to have, mutually different, temporal structure. This temporal structure can be coloredness, non-stationarity, or a combination of both [7, 8, 11].

Most BSP methods that exploit temporal structure in the source and sensor signals use second order statistical features; however, some works exist that exploit higher order temporal structure [36]. In this section we focus on methods that exploit second order temporal structure (SOTS). First we present a robust spatial whitening procedure that exploits SOTS in order to deal with noisy observations. Subsequently, we present BSS methods based on joint diagonalization of sensor correlation matrices. Finally, we discuss a method to perform BSE using a linear predictor.

2.4.1 Robust spatial whitening

In Section 2.3.1 we discussed methods to perform spatial whitening for mixing scenarios where a noise-free sensor correlation matrix can be obtained, for example by compensating for additive noise. For applications where it is difficult or even impossible to obtain or estimate a noise-free sensor correlation matrix we discuss an alternative spatial whitening method where SOTS in the source signals is exploited in order to be robust to noise. First we discuss estimation of and assumptions on SOTS in signals. Subsequently, we discuss the robust spatial whitening procedure.

The SOTS in signals is revealed by evaluating auto- and crosscorrelations of mutually delayed signals and by considering auto- and crosscorrelations at different time instances. This leads to the following generalized and natural definition of correlation between two, potentially delayed, signals such as $x_{i_1}[n]$ and $x_{i_2}[n]$:

$$r^x_{i_1 i_2}[n,k] \triangleq \mathbb{E}\{x_{i_1}[n]x_{i_2}[n-k]\} \tag{2.56}$$

where n is a time index and k is a delay. Considering different lags k exposes coloredness of a signal while non-stationarity of signals is revealed by considering different time indices n. For stationary signals the time index n can be omitted, which we did implicitly in Section 2.3.1.

Combining correlation data for all available sensor pairs for a given time-lag pair

(n, k) into a matrix leads to the following sensor correlation matrix:

$$\mathbf{R}^x[n,k] \triangleq \mathbb{E}\{\mathbf{x}[n](\mathbf{x}[n-k])^T\} = \begin{bmatrix} r^x_{11}[n,k] & \cdots & r^x_{1D}[n,k] \\ \vdots & \ddots & \vdots \\ r^x_{D1}[n,k] & \cdots & r^x_{DD}[n,k] \end{bmatrix} \tag{2.57}$$

where D is the total number of sensors. For noise-free observations as in (2.6), the structure in such a sensor correlation matrix is very similar to the structure in (2.11), and is given as follows:

$$\mathbf{R}^x[n,k] \equiv \mathbf{A}\mathbf{R}^s[n,k](\mathbf{A})^T \tag{2.58}$$

where $\mathbf{R}^s[n,k]$ is a source correlation matrix for time-lag pair (n,k), i.e.,

$$\mathbf{R}^s[n,k] \triangleq \mathbb{E}\{\mathbf{s}[n](\mathbf{s}[n-k])^T\} = \begin{bmatrix} r^s_{11}[n,k] & \cdots & r^s_{1S}[n,k] \\ \vdots & \ddots & \vdots \\ r^s_{S1}[n,k] & \cdots & r^s_{SS}[n,k] \end{bmatrix} \tag{2.59}$$

with S the total number of sources.

The assumption of mutually statistically independent source signals, presented in Section 2.2.2, implies that for all source signal pairs the crosscorrelations are zero for all time-lag pairs. Consequently, this assumption leads to diagonal source correlation matrices $\mathbf{R}^s[n,k]$. On the other hand, the presence of sufficient and diverse temporal structure in the source signals guarantees that the diagonal elements are mostly non-zero and different when considering several time-lag pairs. Consequently, sensor correlation matrices have full rank and vary for different time-lag pairs.

For the robust spatial whitening algorithm to work we require that the noise signals have less temporal structure than the source signals in such a way that a set of time-lag pairs exist for which the sensor correlation matrices are free of noise. Furthermore, we assume to know for which time-lag pairs these assumptions hold. Since the lag zero sensor correlation matrix contains the energy of signals, this matrix will in general not be free of noise and therefore not considered in this approach.

Given a set of sensor correlation matrices for which the assumptions on the mixing system and the source and noise signals hold, spatial whitening can be performed in case a positive definite sensor correlation matrix is available or can be constructed. Zero-lag sensor correlation matrices are guaranteed to be positive definite but are also corrupted by noise. Positive definiteness is not a natural property of arbitrary-lag sensor correlation matrices and cannot be guaranteed without additional assumptions. However, in some applications a positive definite sensor correlation matrix can be constructed from a linear combination of sensor correlation matrices.

The following linear combination of sensor correlation matrices for several time-lag pairs is considered in the robust spatial whitening algorithm:

$$\mathbf{\Gamma} = \sum_{\kappa=1}^{K} \xi_\kappa \mathbf{R}^x[\Omega_\kappa] \tag{2.60}$$

where Ω_κ for $1 \leq \kappa \leq K$ is the set of time-lag pairs $(n,k)_\kappa$ for which the sensor correlation matrices are free of noise and where the coefficients ξ_κ for $1 \leq \kappa \leq K$ need to be properly selected such that $\mathbf{\Gamma}$ is a positive define matrix. Proper coefficients and the spatial whitening matrix can be found using the following algorithm [8]:

1. Calculate K noise-free sensor correlation matrices for time-lag pairs Ω_κ
2. Select arbitrary, non-zero coefficients ξ_κ
3. Calculate the linear combination $\mathbf{\Gamma} = \sum_{\kappa=1}^{K} \xi_\kappa \mathbf{R}^x[\Omega_\kappa]$
4. Calculate the eigenvalue decomposition $\mathbf{U}\mathbf{\Lambda}\left(\mathbf{U}\right)^{-1} = \mathbf{\Gamma}$
5. If $\mathbf{\Gamma}$ is positive definite, then use the eigenvectors to construct a spatial whitening matrix as in Section 2.3.1, i.e., $\mathbf{Q} = \mathbf{\Lambda}^{-\frac{1}{2}}\left(\mathbf{U}\right)^{-1}$
6. Otherwise, select the eigenvector $\mathbf{u}$ corresponding to the smallest eigenvalue and repeat from step 3 with the following updated coefficients:

$$\xi_\kappa \leftarrow \xi_\kappa + \frac{(\mathbf{u})^T \mathbf{R}[\Omega_\kappa]\mathbf{u}}{\sqrt{\sum_{l=1}^{K} (\mathbf{u})^T \mathbf{R}^x[\Omega_l]\mathbf{u}}} \quad \forall\, 1 \leq \kappa \leq K \tag{2.61}$$

In general the source autocorrelation functions have to be linearly independent for the considered time-lag pairs in order to guarantee successful spatial whitening with this algorithm. This requirement also implies that the number of time-lag pairs should be at least equal to the number of sources, i.e., $K \geq S$. In conclusion, robust spatial whitening can only be performed if the source signals contain sufficient and diverse temporal structure.

2.4.2 BSS by exploiting second order temporal structure

The basic principle of BSS methods that exploit temporal structure of signals through SOS are based on the principle of joint diagonalization of at least two sensor correlation matrices [8,11,44,45]. Three different strategies to perform this joint diagonalization can be distinguished in the literature, viz., the BSS problem is formulated as an eigenvalue decomposition problem, a generalized eigenvalue decomposition problem, or as an approximate joint diagonalization problem. In this section we discuss these different strategies and corresponding algorithms.

2.4.2.1 Eigenvalue decomposition based BSS

Eigenvalue decomposition based BSS methods are two-stage approaches that require spatial whitening as a pre-processing step. As discussed in Section 2.3.1, spatial whitening leads to whitened observations that are an orthogonal mixture of normalized source signals. Note that the normalization is an essential part of the spatial whitening process.

When considering a correlation matrix of spatially whitened observations for time-

lag pairs where the noise is not correlated we obtain the following noise-free structure:

$$\mathbf{R}^{\check{x}}[n,k] = \check{\mathbf{A}}\mathbf{R}^{\check{s}}[n,k](\check{\mathbf{A}})^T \tag{2.62}$$

where $\check{\mathbf{A}}$ is an orthogonal mixing matrix and $\check{\mathbf{s}}[n]$ corresponds to the normalized source signals. It follows immediately that (2.62) constitutes an eigenvalue decomposition when we realize that $(\check{\mathbf{A}})^T = (\check{\mathbf{A}})^{-1}$ holds for an orthogonal mixing matrix.

The eigenvectors of (2.62) correspond to the extraction filters for the sources and the collection of all eigenvectors corresponds to the source separation solution. However, in order to obtain a unique solution, and solve the source separation problem up to the unresolvable scaling and permutation, this sensor correlation matrix of spatially whitened observations has to have unique diagonal elements. It follows immediately that the matrix for performing spatial whitening cannot be used since that leads to the identity matrix. Candidate sensor correlation matrices for colored signals are those having different lags and in case of non-stationarity sensor correlation matrices different time instances should be used. Alternatively, a linear combination of sensor correlation matrices as in (2.60) can be used. The parameters for creating this linear combination of correlation matrices should be different from the parameters used in the robust spatial whitening algorithm from Section 2.4.1.

In the literature this eigenvalue decomposition based BSS approach is called Algorithm for Multiple Unknown Signals Extraction (AMUSE) [44]. In this algorithm the following steps allow for blind separation or extraction of the sources:

1. Apply spatial whitening as in Section 2.3.1 or Section 2.4.1 to the observations, i.e., $\check{\mathbf{x}}[n] = \mathbf{Q}\mathbf{x}[n]$
2. For suitable time-lag pairs, calculate a linear combination of correlation matrices of the spatially whitened observations, indicated as $\mathbf{\Gamma}$
3. Calculate the eigenvalue decomposition $\mathbf{\Gamma} = \mathbf{U}\mathbf{\Lambda}\left(\mathbf{U}\right)^{-1}$
4. If zero or equal eigenvalues are obtained, then repeat steps 2 and 3 until all eigenvalues are non-zero and distinct from each other
5. Use the eigenvector matrix as source separation filter, i.e., $\mathbf{W} = (\mathbf{U})^T$, and calculate the separated sources as $\hat{\mathbf{s}}[n] = \mathbf{W}\check{\mathbf{x}}[n]$ or $\hat{\mathbf{s}}[n] = \mathbf{W}\mathbf{Q}\mathbf{x}[n]$.

2.4.2.2 Generalized eigenvalue decomposition based BSS

Instead of a two-stage approach, the BSS problem can be solved in a single-stage approach by performing a joint diagonalization of two noise-free sensor correlation matrices. The advantage of this joint diagonalization approach is that a pre-processing stage in the form of spatial whitening is not required.

Joint diagonalization of two noise-free sensor correlation matrices $\mathbf{R}^x[\Omega_1]$ and $\mathbf{R}^x[\Omega_2]$ for two different time-lag pairs can be accomplished via the following generalized eigenvalue decomposition (GEVD):

$$\tilde{\boldsymbol{\mu}}\mathbf{R}^x[\Omega_1] = \lambda\tilde{\boldsymbol{\mu}}\mathbf{R}^x[\Omega_2]. \tag{2.63}$$

By considering the structure in the noise-free sensor correlation matrices, as given in (2.58), it follows that the generalized eigenvectors correspond to rows from the

inverse of the mixing matrix. Consequently, the source separation filters can be found by collecting all the generalized eigenvectors.

Like with the eigenvalue decomposition based approach in Section 2.4.2.1, a unique solution, up to an arbitrary scaling factor and permutation, can only be found in case all eigenvalues are distinct. If not all eigenvalues are distinct then noise-free sensor correlation matrices for different time-lag pairs or linear combinations of noise-free sensor correlation matrices as in (2.60) can be considered.

Several numerical algorithms for solving GEVD problems are available [46]. Naturally, each algorithm has its own properties such as complexity and accuracy. Furthermore, the choice for a particular algorithm depends on algebraic properties of the sensor correlation matrices. Next we discuss a few of these properties and corresponding approaches for solving the resulting GEVD problems.

- If the noise-free sensor correlation matrix $\mathbf{R}^x[\Omega_2]$ is symmetric and positive definite, for example obtained through the method as described in Section 2.4.1, then the eigenvalues are guaranteed to be real-valued. In this case the GEVD problem can be solved through a Cholesky decomposition [46] of $\mathbf{R}^x[\Omega_2]$.
- For an invertible matrix $\mathbf{R}^x[\Omega_2]$ the GEVD in (2.63) can be solved, and thus the source separation filters can be found, via the following eigenvalue decomposition [47]:

$$\tilde{\boldsymbol{\mu}}\mathbf{R}^x[\Omega_1](\mathbf{R}^x[\Omega_2])^{-1} = \lambda\tilde{\boldsymbol{\mu}}. \tag{2.64}$$

- In the general case where both $\mathbf{R}^x[\Omega_1]$ and $\mathbf{R}^x[\Omega_2]$ are two arbitrary full rank matrices the GEVD can be computed via a generalized Schur or QZ decomposition [46].

When comparing the GEVD based approach with the eigenvalue decomposition based approach from the previous section two relevant differences can be observed. First, the current approach does not require spatial whitening as a pre-processing step. Second, the mixing system is not required to be orthogonal. Consequently, the source separation filter identified via the GEVD approach is in general a non-orthogonal matrix that is applied directly to the original observations.

2.4.2.3 Joint approximate diagonalization based BSS

The BSS approaches from the previous sections allow for the diagonalization of at most two, linear combinations of, noise-free sensor correlation matrices. Since these approaches are sensitive to modelling and estimation errors, we discuss in this section a BSS approach where multiple noise-free sensor correlation matrices are jointly diagonalized. Ideally, multiple sensor correlation matrices that obey the instantaneous mixing model in (2.6) such that their structure is as in (2.58) can be diagonalized exactly. However, for more than two sensor correlation matrices the system is overdetermined and due to modelling and estimation errors joint diagonalization can only be approximated.

The idea of joint approximate diagonalization (JAD) is to decompose multiple noise-free sensor correlation matrices according to the noise-free mixing model in (2.58). A natural and commonly used criterion for performing this decomposition or JAD of noise-free correlation matrices from spatially whitened observations is as follows:

$$J(\mathbf{W}, \mathbf{\Lambda}_\kappa) = \sum_{\kappa=1}^{K} \|\mathbf{R}^{\check{\mathbf{x}}}[\Omega_\kappa] - \mathbf{U}\mathbf{\Lambda}_\kappa(\mathbf{U})^T\|_F^2 \tag{2.65}$$

where $\|\cdot\|_F$ is the Frobenius norm [46], $K \geq 1$ is the number of correlation matrices considered and where $\mathbf{\Lambda}_\kappa$ are non-zero diagonal matrices that have to be identified together with the orthogonal matrix $\mathbf{U}$. The source separation matrix is subsequently found as $\mathbf{W} = (\mathbf{U})^T$. In the literature several algorithms can be found for solving the JAD problem in (2.65), e.g., Jacobi techniques, Alternating Least squares (ALS), and PARAFAC [8, 11, 45, 48]. Note that performing this decomposition for one or two noise-free sensor correlation matrices leads to the eigenvalue decomposition and GEVD problems in the previous sections.

A summary of the JAD based BSS algorithm based on whitened observations is given as follows [45, 49]:

1. Apply spatial whitening as in Section 2.3.1 or Section 2.4.1 to the observations $\mathbf{x}[n]$
2. Estimate K sensor correlation matrices $\mathbf{R}^{\check{\mathbf{x}}}[\Omega_\kappa]$ from the whitened observations $\check{\mathbf{x}}[n]$
3. Calculate for $1 \leq \kappa \leq K$ the JAD of $\mathbf{R}^{\check{\mathbf{x}}}[\Omega_\kappa]$ by minimizing the cost function in (2.65), where $\mathbf{U}$ is an orthogonal matrix and $\mathbf{\Lambda}_\kappa$ are non-zero diagonal matrices
4. Use $\mathbf{W} = (\mathbf{U})^T$ to calculate the separated sources as $\hat{\mathbf{s}}[n] = \mathbf{W}\check{\mathbf{x}}[n]$

Also for JAD linear combinations of noise-free correlation matrices can be used instead of the noise-free correlation matrices directly. Furthermore, a JAD cost function that does not require spatial whitening exists and can be found in [50]. A drawback of this approach is that symmetric and positive definite sensor correlation matrices are required [8]. Such matrices can be found for example by taking appropriate linear combinations of sensor correlation matrices such as described in the robust spatial whitening algorithm in Section 2.4.1.

2.4.3 BSE using a linear predictor

Source extraction in a single-stage algorithm can be accomplished with an on line BSE algorithm based on a linear predictor [14, 15, 51]. In this on line algorithm, an adaptive source extraction filter $\tilde{\mathbf{w}}$ is followed by a fixed linear prediction filter $\mathbf{b}$, as is depicted in Figure 2.1. Note that this filter can have arbitrary structures such as finite impulse response (FIR) and infinte impulse response (IIR). Design of the linear prediction filter is application specific and depends on the temporal structure of the source signals. We assume that the delay required for performing prediction is embedded in the filter.

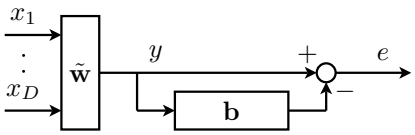


Figure 2.1: Overview of a linear prediction based BSE algorithm.

For a prediction filter $\mathbf{b}$ with impulse response $b[n]$, the extracted signal and the linear prediction error are calculated as follows, respectively:

$$y[n] = \tilde{\mathbf{w}}\mathbf{x}[n] \tag{2.66}$$
$$e[n] = y[n] - b[n] * y[n] \tag{2.67}$$

where $*$ represents convolution. In order to extract one of the sources, the extraction filter coefficients $\tilde{\mathbf{w}}$ are adapted based on the output of the linear prediction filter. Spatial whitening is a required procedure in the linear prediction based BSE algorithm. Partly due to this spatial whitening process, the source signal with the smallest normalized mean square prediction error is extracted by the algorithm. Extraction this source is accomplished by alternating the following update rules [14]:

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \mu e[n]\hat{\mathbf{x}}[n] \tag{2.68}$$
$$\mathbf{w}[n+1] \longleftarrow \mathbf{w}[n+1]/\|\mathbf{w}[n+1]\|_2 \tag{2.69}$$

where $\hat{\mathbf{x}}[n]$ is the linear prediction error per spatially whitened sensor signal, i.e., for prediction filter $\mathbf{b}$ these signals are given as $\hat{x}_i[n] = \breve{x}_i[n] - b[n] * \breve{x}_i[n]$ for $1 \leq i \leq D$.

The extraction filter that is found with this linear prediction based source extraction algorithm corresponds to the eigenvector that corresponds to the smallest eigenvalue of the positive definite 'prediction error' matrix $\mathbf{R}^{\hat{x}} \triangleq \mathbb{E}\{\hat{\mathbf{x}}[n](\hat{\mathbf{x}}[n])^T\}$. In [52] we showed that the linear prediction based source extraction algorithm including the spatial whitening process can be formulated as a GEVD problem. Consequently, the linear prediction based BSE algorithm can be considered a special case and implementation of the eigenvalue decomposition and GEVD based source separation algorithms presented in Section 2.4.2.1 and Section 2.4.2.2, respectively.

In [14] alternative linear prediction based algorithms are presented that perform spatial whitening in the update rule instead of in a pre-processing stage. Furthermore, in [15, 51] the linear prediction based source extraction algorithms are extended in order to deal with noise. For these modifications it is assumed that the noise has a relatively simple structure and that the noise is measurable such that it can be compensated for. In these approaches the contribution of the noise to the cost function is compensated for in order to obtain again the noise-free cost function from [14].

It follows that the linear prediction based BSE algorithms are relatively flexible in their configurations. Furthermore, designing prediction filters based on a priori information of the desired source allows to design algorithms for which extraction of the desired source can be guaranteed [52]. A generalization of the linear prediction based BSE algorithm for complex mixtures can be found in [53, 54].

2.4.4 Discussion

In this section we discussed BSP algorithms that exploit SOTS in the source signals in order to solve the task at hand. In Section 2.4.1 we discussed a robust spatial whitening method that exploits SOTS to perform spatial whitening on observations that are corrupted by noise that cannot be estimated or compensated. In the remaining sections we discussed BSS and BSE algorithms that exploit SOTS in the source signals. Each of these algorithms can be used in a multi-stage source extraction algorithm in order to extract a desired source. Source separation algorithms have to be followed by a classifier that selects the desired output, while source extraction algorithms should be followed by a classifier and a deflation technique in order to guarantee extraction of the desired source. These multi-stage algorithms are evidently inefficient and expensive in resources.

Exploiting SOTS has the advantage over exploiting HOS that mixtures with multiple Gaussian sources can be dealt with. Furthermore, the amount of data required to obtain robust estimates of statistical features and the computational complexity is much smaller for second order statistical features than for higher order statistical features. However, the choice for using a certain set of features always depends on the problem at hand.

2.5 BSI projected onto two dual mathematical problems

In [36] the BSI problem is projected onto two dual mathematical problems using subspace techniques. The projections apply for both second order and higher order statistical features such that both non-Gaussianity and temporal structure in the source signals can be considered. For brevity, in this thesis we will discuss the methods in the context of second order temporal structure of real-valued signals and we only present the main steps to form the relevant systems of equations. For further details and derivations of the methods and generalizations of the methods towards the use of complex signals and higher order statistical features we refer to [36].

2.5.1 BSI as a root finding problem for a system of homogeneous polynomials

The approach of [36] starts by collecting the sensor correlation data, which we introduced in Sections 2.3.1 and 2.4.1, for all $D \times D$ sensor pairs and all K time-lag pairs into a single matrix $\mathbf{C}^x$ of size $D^2 \times K$:

$$\mathbf{C}^x = \begin{bmatrix} r^x_{11}[\Omega_1] & \cdots & r^x_{11}[\Omega_K] \\ r^x_{12}[\Omega_1] & \ddots & r^x_{12}[\Omega_K] \\ \vdots & \ddots & \vdots \\ r^x_{DD}[\Omega_1] & \cdots & r^x_{DD}[\Omega_K] \end{bmatrix} \tag{2.70}$$

where Ω_κ are time-lag pairs for which the sensor auto- and crosscorrelation data is free of noise. Based on the mixing model and assumptions introduced in Section 2.2, the structure in each column in the matrix $\mathbf{C}^x$ can be represented with a Kronecker product ($\otimes$) in the following way:

$$\mathbb{E}\{\mathbf{x}[n] \otimes \mathbf{x}[n-k]\} = \sum_{j=1}^{S} (\mathbf{a}^j \otimes \mathbf{a}^j) \cdot \mathbb{E}\{s_j[n] s_j[n-k]\} \tag{2.71}$$

where the time-lag pairs $(n, k) \in \Omega_k$ differ among the columns. This structure in the sensor correlation data leads to the following structure in the sensor correlation matrix:

$$\mathbf{C}^x = (\mathbf{A} \diamond \mathbf{A}) \cdot \mathbf{C}^s = \begin{bmatrix} \mathbf{a}^1 \otimes \mathbf{a}^1 & \cdots & \mathbf{a}^S \otimes \mathbf{a}^S \end{bmatrix} \mathbf{C}^s \tag{2.72}$$

where $\diamond$ is the Khatri-Rao product, i.e., the column-wise Kronecker product [55] and the matrix $\mathbf{C}^s$ is a source autocorrelation matrix of size $S \times K$ with the following structure:

$$\mathbf{C}^s = \begin{bmatrix} r^s_{11}[\Omega_1] & \cdots & r^s_{11}[\Omega_K] \\ r^s_{22}[\Omega_1] & \ddots & r^s_{22}[\Omega_K] \\ \vdots & \ddots & \vdots \\ r^s_{SS}[\Omega_1] & \cdots & r^s_{SS}[\Omega_K] \end{bmatrix}. \tag{2.73}$$

A further look into the structure of the sensor correlation matrix shows that the rank of the matrix is at most S, i.e., the number of sources. This follows from the observation that the Khatri-Rao product $\mathbf{A} \diamond \mathbf{A}$ results in a matrix of size $D^2 \times S$ and $\mathbf{C}^s$ has a size of $S \times K$. Another remarkable property of the sensor correlation matrix $\mathbf{C}^x$ is that for real-valued mixtures, only $D(D+1)/2$ out of the D^2 rows are unique. This follows from the fact that for real-valued signals the following equality holds: $r^x_{i_1 i_2}[\Omega_\kappa] = r^x_{i_2 i_1}[\Omega_\kappa]$. Taking only the unique rows into account leads to a uniquefied sensor correlation matrix $\mathbf{C}^x_u$ of size $D(D+1)/2 \times K$ with its rank at most equal to S [36]. Consequently, the sensor correlation matrix $\mathbf{C}^x$ or its uniquefied counterpart $\mathbf{C}^x_u$ can have rank S in case both the mixing system and the source autocorrelation matrix are well conditioned and if $D(D+1)/2 \geq S$ and $K \geq S$ hold. In the remainder of this chapter we assume that these conditions are met in such a way that the rank of the sensor correlation matrix equals S.

For mixtures where the number of sensors is such that $D(D+1)/2 > S$ holds, a left nullspace exists in the uniquefied sensor correlation matrix $\mathbf{C}^x_u$, i.e., a vector or matrix $\mathbf{U}^\nu$ exists such that the following relation holds:

$$(\mathbf{U}^\nu)^T \mathbf{C}^x_u = \mathbf{0}. \tag{2.74}$$

The matrix $\mathbf{U}^\nu$ has a size $D(D+1)/2 \times D(D+1)/2 - S$ and by exploiting the structure in the sensor correlation matrix from (2.72) the following relation can be formulated for the left nullspace:

$$(\mathbf{U}^\nu)^T (\mathbf{A} \diamond_u \mathbf{A}) \mathbf{C}^s = \mathbf{0} \tag{2.75}$$

where $\diamond_u$ stands for the Khatri-Rao product combined with uniquiefication.

If we now introduce a vector $\mathbf{z}$ of length D, then the following equations form a system of homogenous polynomial equations of which the roots correspond to the mixing columns.

$$f(\mathbf{z}) = (\mathbf{U}^{\nu})^T(\mathbf{z} \otimes_u \mathbf{z}) = \mathbf{0} \tag{2.76}$$

where $\otimes_u$ is the uniquiefied Kronecker product. Consequently, solving this system of homogeneous polynomial equations leads to identification of the mixing system, up to the arbitrary scaling factor and permutation.

In [36] such systems of polynomial equations are constructed based on both SOS and HOS and for both real-valued and complex-valued mixing systems using a variety of conjugation schemes. It is also shown there that the system can be identified for certain under-determined mixtures where $D < S$. Furthermore, it is proposed to solve such systems of homogenous polynomial equations by exploiting a homotopy method. In such a method a system of polynomial equations with known solutions is transformed into the system with unknown solutions and the roots are constantly tracked.

From a source extraction point of view, one of the strong advantages of the homotopy method to solve a system of homogeneous polynomial equations is that the method can be parallelized such that each root is found by a different processor. However, it is difficult or even impossible to predict which mixing column will be identified, which is a drawback in case only one, desired mixing column or source has to be extracted.

2.5.2 BSI as a multi-matrix Generalized Eigenvalue Decomposition problem

In this section we show that for real-valued instantaneous mixtures a Multi-Matrix Generalized Eigenvalue Decomposition (MMGEVD) problem can be formulated from which the mixing system can be estimated, up to the arbitrary scaling factor and permutation.

For the sensor correlation matrix $\mathbf{C}^x$ from (2.70) in Section 2.5.1 it follows that a vector $\boldsymbol{\mu}$ must exist such that the following relation holds:

$$\mathbf{z} \otimes \mathbf{z} = \mathbf{C}^x \boldsymbol{\mu} \tag{2.77}$$

where $\mathbf{z}$ is a vector of length D. Based on the structure in the sensor correlation matrix, as introduced in (2.72), it can be shown that only S solutions exist and that the S solutions of (2.77) correspond to the mixing column vectors, up to an arbitrary scaling factor [36]. This relation is easily observed if we choose $\boldsymbol{\mu}$ as one of the columns from the inverse of the sensor correlation matrix $\mathbf{C}^s$ since then $\mathbf{z} \otimes \mathbf{z}$ equals $\mathbf{a}^j \otimes \mathbf{a}^j$. However, the sensor correlation matrix or its inverse is not directly available and therefore it has to be estimated.

Splitting the matrix $\mathbf{C}^x$ into D consecutive submatrices $\mathbf{C}^x_i$ of size $D \times K$ leads

to the following relation for each submatrix:

$$z_i \mathbf{z} = \mathbf{C}_i^x \boldsymbol{\mu} \quad \forall\, i \in \mathcal{D} \tag{2.78}$$

which follows from the Kronecker product in (2.77). When we choose $\boldsymbol{\mu}$ properly, all products $\mathbf{C}_i^x \boldsymbol{\mu}$ in (2.78) form a vector in the same direction, i.e., direction $\mathbf{z}$, but with different lengths, i.e., z_i. Consequently, the system in (2.78) describes a Multi-Matrix Generalized Eigenvalue Decomposition (MMGEVD) with generalized eigenvectors $\boldsymbol{\mu}$ and generalized eigenvalues $1/z_i$. In case $z_i \neq 0$ we can rewrite the MMGEVD problem from (2.78) in the following, better recognizable, form:

$$\mathbf{z} = \frac{1}{z_1} \mathbf{C}_1^x \boldsymbol{\mu} = \cdots = \frac{1}{z_D} \mathbf{C}_D^x \boldsymbol{\mu}. \tag{2.79}$$

Once the eigenvectors $\boldsymbol{\mu}$ are known, the mixing columns can be computed from the matrix products $\mathbf{C}_i^x \boldsymbol{\mu}$. Note that the order and scaling of the eigenvectors is arbitrary, which is a manifestation of the scaling and permutation indeterminacy.

For a system with only two submatices $\mathbf{C}_1^x$ and $\mathbf{C}_2^x$ and $D \geq S$ the MMGEVD problem reduces to a standard GEVD problem. For a GEVD problem the eigenvectors $\boldsymbol{\mu}$ can be computed for example as the columns of the following matrix product $(\mathbf{C}_1^x)^{-1} \mathbf{C}_2^x$, in case the submatrix $\mathbf{C}_1^x$ has full rank. Methods to solve the generic MMGEVD problems still need to be developed; however, the joint diagonalization approaches such as the ones presented in Sections 2.3.3 and 2.4.2.3 are closely related to solving the MMGEVD problem.

For a more detailed derivation of the method we refer to [36]. Furthermore, it is shown in [36] that MMGEVD problems can also be formulated for complex-valued mixtures and for exploiting higher order statistical features.

2.5.3 Discussion

We discussed two approaches for solving the BSI problem for instantaneous mixtures in this section, which were originally introduced in [36]. These approaches are flexible in their use of real-valued and complex-valued data and in the use of second and higher order statistical features. Furthermore, in [36] fundamental limits are derived on the number of sources that can be dealt with, given the number of sensors and the considered statistical features, and it is shown that even under-determined mixing systems with $D < S$ can be identified for certain scenarios. Consequently, the approach and insights from the work in [36] are considered very relevant for the development of generic and flexible informed source extraction algorithms.

The methods to perform BSI can be used directly in a multi-stage approach to solve the source extraction problem. However, the derivations and formulation of BSI as two dual mathematical problems are considered more relevant for the derivation and design of new informed source extraction algorithms. The challenge remaining is to incorporate a priori information into the approaches and to identify source extraction filters instead of the mixing columns.

2.6 Conclusions

In this chapter we presented an overview of BSP algorithms that are relevant for the remainder of this thesis. We discussed only a subset of the large number of BSP algorithms that are available in the literature and more extensive overviews and detailed discussions of BSP algorithms can be found for example in the following books [7, 8, 11, 35].

The BSP algorithms that we discussed in this chapter can be applied to observations of a mixture of multiple sources in order to separate all sources simultaneously or to extract one of the sources. The most widely shared assumption among these algorithms is that the sources are mutually statistically independent or mutually uncorrelated. Based on additional assumptions on the source signals that are exploited by the algorithms we distinguished two classes of BSP algorithms. Algorithms from the first class exploit higher order statistical features and are based on the assumption that the source signals have a non-Gaussian distribution. The second class of algorithms exploit second order statistical features based on the assumption that the source signals have temporal structure. Using second order statistical features has several practical advantages over using higher order statistical features; however, the choice for using a certain set of features always depends on the problem at hand. Finally, we discussed algorithms that are flexible in their use of statistical features.

Due to their blindness, BSP algorithms suffer from a scaling and permutation indeterminacy. Although in many practical applications heuristics can be used in order to deal with these indeterminacies, we found only one family of source extraction algorithms that deal with the permutation indeterminacy with a generic approach. In these algorithms, which we discussed in Section 2.3.5, a priori information about the desired source in the form of reference signals is incorporated in a constrained optimization problem. By selecting or searching for a proper threshold, the source that is closest to the reference signal is extracted by the algorithm.

In the remainder of this thesis we develop source extraction algorithms and design procedures for source extraction algorithms for which extraction of the desired source can be guaranteed. In these algorithms it is guaranteed that the source with the best match to the available a priori information is extracted without searching for or selecting a threshold parameter. Furthermore, we focus on exploiting SOTS in order to extract the desired signals; however, our approach is not limited to using only SOTS.

3

A single-stage approach to source extraction based on second order statistics

The work in this chapter is published as a journal article [56]; however, we made some modifications and additions to improve the flow and consistency in this thesis.

A single-stage approach to extract one desired source signal from a real-valued instantaneous mixture of several signals is presented in this chapter. In contrast to multi-stage or iterative methods, for example based on blind source extraction (BSE) or blind source separation (BSS) followed by a classifier, we guarantee immediate extraction of a desired source. The approach presented in this chapter is applicable to real-valued mixing scenarios where correlation matrices can be jointly diagonalized, i.e., we exploit second order temporal structure (SOTS) in real-valued signals.

Our objective is to provide insight in the source extraction problem and to derive design tools and techniques to solve the source extraction problem. We show for real-valued mixtures that prior information about the desired source's mixing column or autocorrelation function can be incorporated in a single-stage source extraction algorithm. Furthermore, we show that both types of prior information may be combined in order to obtain better results in selecting the desired source. Above that, the same algorithm may be used for calculating source extraction filters with different objective functions. Finally, a performance comparison with BSS algorithms that are followed by a classifier shows that the proposed approach is more robust to noisy observations.

3.1 Introduction

In a blind source extraction (BSE) problem, one desired source signal has to be extracted from observed mixtures of several source signals. BSE is performed by filtering the observed sensor signals with a multiple-input single-output extraction system that is identified with the least possible amount of a priori information. However, some prior information is required to distinguish between the desired and undesired sources.

As we have discussed in Chapter 1 and Chapter 2, a wide variety of blind signal processing (BSP) algorithms exist that use assumptions on different signal properties of the source, noise and sensor signals [7, 8, 10, 11, 35, 45], e.g., assumptions on the Second Order Statistics (SOS) such as stationarity, whiteness, and cyclo-stationarity; or, assumptions on sparseness or higher order statistics (HOS). There are three major advantages of using SOS over HOS: 1) SOS is less sensitive to noise and outliers, 2) SOS requires less data for their estimation, and 3) SOS can allow for more than one Gaussian signal [9]. Next to different assumptions on the signals, BSP problems are solved in different domains such as time, frequency, and time-frequency. In this thesis we focus on the SOS of the signals and solve the source extraction problem in the time domain by joint diagonalization of correlation matrices, which is a proven BSP technology [8, 11].

A straightforward, two-stage approach to perform BSE is depicted in Figure 3.1a. In the first stage a blind source separation (BSS) algorithm is used to separate all source signals. One major problem in the field of BSP is the fact that sources can be separated in a random order only, which is called the permutation indeterminacy. Therefore, a classifier is required in a second stage that uses prior information to select the desired signal. Closely related to the BSS approach are BSE methods or sequential BSS methods. These algorithms extract a possibly 'interesting' source signal from the measurements and determine, using additional a priori information, whether the desired source is extracted. If the desired source is not extracted, the signal is removed from the measurements and a new, potentially interesting, source is extracted.

In this chapter we present a single-stage procedure to perform BSE, as depicted in Figure 3.1b, for real-valued instantaneous mixtures. The permutation indeterminacy is dealt with by incorporating prior information about the desired source directly into the algorithm that identifies the desired extraction filter. In this way, the method remains as blind as the previously described two-stage approaches.

The method is based on the approach we discussed in Section 2.4.2.2, where a noise-free region of support (NF-ROS) [36] is used to find two noise-free linear combinations of correlation matrices and a generalized eigenvalue decomposition (GEVD) is applied to identify all extraction filters simultaneously [11, 36]. We show that the generalized eigenvalues contain a very specific structure, which was exploited for the first time in [52]. By taking linear combinations of noise-free correlation matrices based on a priori knowledge about the autocorrelation function of the desired source, the desired extraction filter is identified without first extracting undesired signals. In parallel we identify the desired extraction filter using a different type of a prior information,

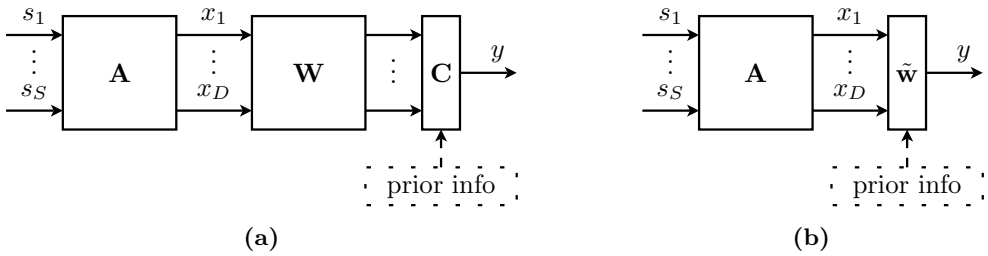


Figure 3.1: Two configurations to perform blind source extraction. In (a) a straightforward two-stage scheme is depicted where in the first stage all source signals are separated and in the second stage the desired signal is selected. In (b) the proposed single-stage scheme is depicted where the desired extraction filter is identified in a single stage

namely, information about the mixing column of the desired source together with a set of differently structured correlation matrices. We also show that combinations of available prior information can be used to increase flexibility and robustness in the selection of the desired source. Subsequently, we show that the same procedure can be used to solve the extraction problem for an objective that balances between noise and interference reduction rather than the typical objective that focusses on interference reduction. The identification of the extraction filter is combined with a noise reduction technique for the over-determined case with more sensors than sources. Here, filtered signals from a noise-only subspace are subtracted from the output of the extraction filter such that the output power is minimized. The prior information nicely connects to the structure in the observed SOS. Finally, we compare the performance of proposed algorithm with BSS methods that we equip with a classifier in the second stage.

The algorithm proposed in this chapter is called the unified instantaneous BSE (UIBSE) algorithm. Later on we realized that the algorithm is better called an informed instead of a blind source extraction algorithm. Since the work in this chapter is published as a journal paper we keep using the name UIBSE here; however, in the next chapter we refer to it as an informed source extraction algorithm.

The outline of this chapter is as follows. In Section 3.2 we give an overview of the model and assumptions. In Section 3.3 we present the basis of the UIBSE algorithm of which the additional functionality is discussed in Section 3.4. We give a summary of the UIBSE algorithm in Section 3.5. In Section 3.6 we present simulation results for combining a priori knowledge and for using different objectives in the UIBSE algorithm; additionally, in a performance analysis we compare the presented work with the SOBI [45] and SOBI-RO [49] algorithms. In Section 3.7 we discuss the results. Finally, in Section 3.8 we conclude this chapter and suggest future research.

Notation of matrices, vectors, and their elements is as follows. Matrices are denoted by bold uppercase letters and vectors by bold lowercase letters. Row vectors carry an additional tilde symbol to indicate their orientation, e.g., $\tilde{\mathbf{w}}$. Elements of a column vector are indicated by the vector letter with a subscript index such as $x_1[n]$

as an element of the vector $\mathbf{x}[n]$, with n the discrete time index. Similarly, elements of a row vector are indicated by the vector letter with a superscript index such that w^1 is the first element of the row vector $\tilde{\mathbf{w}}$. Finally, matrices are decomposed into three different elements. First, columns of a matrix are indicated as column vectors with a superscript index, i.e. $\mathbf{a}^j$ represents the j'th column of the matrix $\mathbf{A}$. Second, rows of a matrix are indicated by row vectors with a subscript index. The i'th row of the matrix $\mathbf{A}$ is thus denoted by $\tilde{\mathbf{a}}_i$. Third, scalars in a matrix are denoted by both superscript and subscript indices. The scalar matrix element that is located in the second column at the third row of $\mathbf{A}$ is thus denoted by a_3^2. Further notational issues are addressed later in this chapter.

3.2 Model and assumptions

First we introduce the mixing model and the assumptions on the SOS of the signals. Next we introduce the structure in correlation matrices that is exploited by the UIBSE algorithm.

3.2.1 Instantaneous mixing model

A model of the BSE problem is depicted in Figure 3.1b. The D sensor signals $x_1[n], \cdots, x_D[n]$, with $n \in \mathbb{Z}$, are discrete time samples of the continuous time signals $x_1(t), \cdots, x_D(t)$, where $t = nT_s$ and T_s is the sampling time such that no aliasing occurs. These sensor signals contain instantaneous mixtures of S source signals $s_1[n], \cdots, s_S[n]$ contaminated by D additive noise signals $\nu_1[n], \cdots, \nu_D[n]$. The mixing system is modeled by a real-valued matrix $\mathbf{A} \triangleq [\mathbf{a}^1, \cdots, \mathbf{a}^S] \in \mathbb{R}^{D \times S}$. Each column $\mathbf{a}^j \in \mathbb{R}^D$ of the mixing matrix contains D weights a_i^j for $i \in \{1, \cdots, D\}$ that correspond to the gain between the j'th source and the i'th sensor. The mutual relation between the sensor, source, and noise signals is then mathematically described by

$$\mathbf{x}[n] = \sum_{j=1}^{S} \mathbf{a}^j s_j[n] + \boldsymbol{\nu}[n] = \mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n] \tag{3.1}$$

where the column vectors $\mathbf{x}[n]$, $\mathbf{s}[n]$ and $\boldsymbol{\nu}[n]$ contain a collection of the sensor, source and noise signals, respectively, i.e.,

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x_1[n] \\ \vdots \\ x_D[n] \end{bmatrix}, \ \mathbf{s}[n] \triangleq \begin{bmatrix} s_1[n] \\ \vdots \\ s_S[n] \end{bmatrix} \text{ and } \boldsymbol{\nu}[n] \triangleq \begin{bmatrix} \nu_1[n] \\ \vdots \\ \nu_D[n] \end{bmatrix}$$

The extraction filter is represented by a real-valued row vector $\tilde{\mathbf{w}}$ of length D and produces the following signal $y[n]$:

$$y[n] = \sum_{i=1}^{D} w^i x_i[n] = \tilde{\mathbf{w}}\mathbf{x}[n] = \tilde{\mathbf{w}}\mathbf{A}\mathbf{s}[n] + \tilde{\mathbf{w}}\boldsymbol{\nu}[n]. \tag{3.2}$$

Using this model combined with the following assumptions on the source and noise signals and the mixing system we derive the UIBSE algorithm.

3.2.2 Assumptions on the mixing system and signals

In the previous section some assumptions were already made on the mixing system. First, the mixing system is a square or over-determined real-valued mixing system, thus of size $D \times S$ with $D \geq S$. Furthermore, the mixing system has to be an invertible mixing system.

Assumptions on the source and noise signals are made on the SOS of these signals. The SOS of signals are revealed by means of correlation functions. The correlation function of a signal pair $(p_{i_1}[n], q_{i_2}[n])$ is called the autocorrelation function if $p_{i_1}[n] = q_{i_2}[n]$ for all $n \in \mathbb{Z}$ and crosscorrelation function otherwise.

Definition 3.2.1 (Correlation functions)**:** The (auto)correlation function of a signal $p_{i_1}[n]$ and a signal $q_{i_2}[n-k]$ for all available i_1, i_2 at time $n \in \mathbb{Z}$ and with lag $k \in \mathbb{Z}$ is defined as follows:

$$r^{pq}_{i_1 i_2}[n,k] \triangleq \mathbb{E}\{p_{i_1}[n] q_{i_2}[n-k]\}.$$

If both signals are described by the same variable, then we use the short hand notation $r^{p}_{i_1 i_2}[n,k] \equiv r^{pp}_{i_1 i_2}[n,k]$.

We assume that the mathematical expectation operator can be approximated by averaging the signal product $p_{i_1}[n] q_{i_2}[n-k]$ over a certain period close to the evaluated time instance n_0, e.g., $r^{pq}_{i_1 i_2}[n_0,k] \approx 1/(N_0) \sum_{n=n_0}^{N_0-1} p_{i_1}[n] q_{i_2}[n-k]$. This assumption implies that the SOS of the signals should either be slowly varying over time or be approximately constant for a certain time and then rapidly change to a new, approximately constant value. Most conventional BSP methods use the same assumption; however, the non-stationarity of the signals is often not used explicitly. In Definition 3.2.1 the time index n is incorporated such that if the statistics of the signals change over time, then a new measurement of the statistics is obtained that can be used explicitly by the presented algorithm. This property creates flexibility by allowing for different types of signals such as stationary, non-white signals, and non-stationary signals.

By replacing the signal pair (p_{i_1}, q_{i_2}) in Definition 3.2.1 by the sensor, source, and noise signal pairs we obtain the following sets of correlation functions:

$$\begin{aligned} r^{x}_{i_1 i_2}[n,k] \quad & \forall\, 1 \leq i_1, i_2 \leq D \\ r^{s}_{i_1 i_2}[n,k] \quad & \forall\, 1 \leq i_1, i_2 \leq S \\ r^{\nu}_{i_1 i_2}[n,k] \quad & \forall\, 1 \leq i_1, i_2 \leq D \end{aligned}$$

respectively. Notice that these sets of functions consist of both auto- and crosscorrelation functions. Finally, we consider the crosscorrelation functions $r^{s\nu}_{i_1 i_2}[n,k]$ and $r^{\nu s}_{i_2 i_1}[n,k]$ that correspond to the source and noise signal pairs (s_{i_1}, ν_{i_2}) for $1 \leq i_1 \leq S$ and $1 \leq i_2 \leq D$.

Many BSP algorithms require statistically independent source signals, for example in the field of independent component analysis [7]. We require that all source signals are mutually uncorrelated, which is a less restrictive assumption. Our assumptions on the SOS of the source and noise signals are introduced in the following definition of a NF-ROS, where the assumption of uncorrelated source signals is incorporated as the first condition in this definition.

Definition 3.2.2 (Noise-free region of support)**:** The noise-free region of support (NF-ROS), also denoted by Ω, consists of a set of time-lag pairs (n, k) for which the following assumptions hold:

$$\forall\,(n,k) \in \Omega : \begin{cases} r^{s}_{i_1 i_2}[n,k] = 0\;\forall\, 1 \le i_1 \neq i_2 \le S \\ r^{\nu}_{i_1 i_2}[n,k] = 0\;\forall\, 1 \le i_1, i_2 \le D \\ r^{s\nu}_{i_1 i_2}[n,k] = 0\;\forall\, 1 \le i_1 \le S,\ 1 \le i_2 \le D \\ r^{\nu s}_{i_1 i_2}[n,k] = 0\;\forall\, 1 \le i_1 \le D,\ 1 \le i_2 \le S \end{cases}.$$

Additionally, the source autocorrelation functions are assumed to be linearly independent in the NF-ROS [36]. The total number of time-lag pairs in the set is denoted by K, thus $|\Omega| = K$. In our notation we use the symbol Ω_κ to indicate the κ'th time-lag pair $\Omega_\kappa = (n, k)_\kappa$ from the NF-ROS Ω.

In summary, the NF-ROS consists of a set of time-lag pairs for which only the source autocorrelation functions are non-zero and linearly independent [36]. The following example is given to explain the use of a NF-ROS.

Example 3.2.1: Suppose that D sensors measure a mixture of S stationary, differently colored source signals that are each contaminated by additive noise signals with a moving average 1 (MA1) temporal structure. In that case, the time index n can be ignored because the signals are stationary signals. Furthermore, lags $k = 0$ and $k = 1$ should not be taken into account because for these lags the MA1 noise contributes to the SOS of the sensor signals. Therefore, the NF-ROS may be chosen as the first $K \ge S$ lags starting from $k = 2$, thus: $\Omega = \{(n, 2), \cdots, (n, K + 1)\}$ for an arbitrary $n \in \mathbb{Z}$.

3.2.3 Structure in the Second Order Statistics

Using the model in (3.1) and the knowledge of a NF-ROS as in Definition 3.2.2, the following noise-free relation between the sensor correlation and the source autocorrelation functions is obtained:

$$r^{x}_{i_1 i_2}[\Omega_\kappa] = \sum_{j=1}^{S} a^{j}_{i_1} a^{j}_{i_2} r^{s}_{jj}[\Omega_\kappa] \quad \forall\, 1 \le i_1, i_2 \le D, \quad \forall\, \Omega_\kappa \in \Omega. \tag{3.3}$$

Notice that the j'th element in the summation of (3.3) depends only on the autocorrelation function of source j and the weighting factors of the source-sensor pairs for source j. This structure is already discussed in the literature where it is referred to as a parallel factorization (PARAFAC) model [48].

The PARAFAC model consists of a three-way tensor, which can be seen as a cube of data. Here, the PARAFAC model is a cube of correlation data $r^x_{i_1 i_2}[\Omega]$ with a size of $D \times D \times K$ and a structure as in (3.3) that contains a symmetry because the mixing matrix elements appear twice. Because of this symmetry, only two matrix slices of the PARAFAC model with a unique structure exist. First, if we fix the index for the time-lag pair Ω then we obtain the following set of correlation matrices:

$$\mathbf{R}^x_\kappa \triangleq \begin{bmatrix} r^x_{11}[\Omega_\kappa] & \cdots & r^x_{1D}[\Omega_\kappa] \\ r^x_{21}[\Omega_\kappa] & \cdots & r^x_{2D}[\Omega_\kappa] \\ \vdots & \ddots & \ddots \\ r^x_{D1}[\Omega_\kappa] & \ddots & r^x_{DD}[\Omega_\kappa] \end{bmatrix} \quad \text{for } 1 \leq \kappa \leq K. \tag{3.4}$$

Second, if we fix either the mixing column element index i_1 or i_2 we obtain the following set of correlation matrices:

$$\mathbf{C}^x_i \triangleq \begin{bmatrix} r^x_{i1}[\Omega_1] & \cdots & r^x_{i1}[\Omega_K] \\ r^x_{i2}[\Omega_1] & \cdots & r^x_{i2}[\Omega_K] \\ \vdots & \ddots & \ddots \\ r^x_{iD}[\Omega_1] & \ddots & r^x_{iD}[\Omega_K] \end{bmatrix} \quad \text{for } 1 \leq i \leq D. \tag{3.5}$$

The K correlation matrices $\mathbf{R}^x_\kappa$ have size of $D \times D$ and the D correlation matrices $\mathbf{C}^x_i$ have size $D \times K$. By collecting either K matrices $\mathbf{R}^x_\kappa$ or D matrices $\mathbf{C}^x_i$, the whole PARAFAC model is obtained.

Both types of correlation matrices can be used to identify the extraction filters, as we show in Section 3.3. The desired extraction filter is identified by exploiting the structure in the correlation matrices. Which correlation matrix structure is used, depends on the type of available a priori information.

The structure of the correlation matrices is obtained from (3.3) and is as follows (see Appendix 3.A for a derivation):

$$\mathbf{R}^x_\kappa \equiv \mathbf{A}\, \mathrm{diag}\left(\mathbf{r}^s[\Omega_\kappa]\right) (\mathbf{A})^T \tag{3.6}$$

$$\mathbf{C}^x_i \equiv \mathbf{A}\, \mathrm{diag}\left(\tilde{\mathbf{a}}_i\right) \mathbf{C}^s \tag{3.7}$$

where $\mathbf{r}^s[\Omega_\kappa]$ is a column vector that contains the source autocorrelation function values for time-lag pair κ of the NF-ROS, and $\mathrm{diag}\left(\mathbf{r}^s[\Omega_\kappa]\right)$ puts the elements of the vector $\mathbf{r}^s[\Omega_\kappa]$ on the diagonal of a matrix. The vector $\tilde{\mathbf{a}}_i$ is the i'th row of the mixing system, and $\mathbf{C}^s$ is the source autocorrelation matrix with the following structure:

$$\mathbf{C}^s \triangleq \begin{bmatrix} \mathbf{r}^s[\Omega_1] & \cdots & \mathbf{r}^s[\Omega_K] \end{bmatrix}. \tag{3.8}$$

Given the assumptions on the mixing system and the SOS of the signals in Section 3.2.2 it follows that the rank of the noise-free correlation matrices is at most equal to S. This means that we should be able to transform each correlation matrix to a matrix of size $S \times S$ without loss of information. Furthermore, all correlation matrices of a single type can be reduced in size by the same transforms. In order

to identify the transformation matrices we exploit subspace techniques to separate the null/noise-only subspace from the signal plus noise subspace. We present two methods to identify the subspaces.

1) If we perform a singular value decomposition (SVD) on one of the correlation matrices $\mathbf{C}_i^x$ with rank S, we obtain the following decomposition:

$$\text{SVD}\left(\mathbf{C}_i^x\right) = \begin{bmatrix}\mathbf{U}^s & \mathbf{U}^\nu\end{bmatrix} \begin{bmatrix}\mathbf{\Sigma}^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{V}_s \\ \mathbf{V}_\nu\end{bmatrix} \tag{3.9}$$

where $\mathbf{U}^s$ and $\mathbf{V}_s$ span the S dimensional left and right signal subspace, respectively, $\mathbf{U}^\nu$ spans the $D-S$ dimensional left noise-only subspace, and $\mathbf{V}_\nu$ spans the $K-S$ dimensional right noise-only subspace.

2) In order to improve the robustness of these methods we can also perform SVDs on the following two matrices:

$$\mathbf{C}^x \triangleq \begin{bmatrix}\mathbf{C}_1^x \\ \vdots \\ \mathbf{C}_D^x\end{bmatrix} = (\mathbf{A} \diamond \mathbf{A}) \cdot \mathbf{C}^s \tag{3.10}$$

$$\check{\mathbf{C}}^x = \begin{bmatrix}\mathbf{C}_1^x & \cdots & \mathbf{C}_D^x\end{bmatrix} = \mathbf{A} \cdot \left(\mathbf{A} \diamond (\mathbf{C}^s)^T\right)^T \tag{3.11}$$

which are constructed from a set of correlation matrices, where $\diamond$ is the Khatri-Rao product [55]. The SVD of (3.10) can be used to identify the subspace matrices $\mathbf{V}_s$ and $\mathbf{V}_\nu$ in a more robust manner compared to method 1), while the SVD of (3.11) can be used to identify the subspace matrices $\mathbf{U}^s$ and $\mathbf{U}^\nu$ more robustly. Instead of stacking all sensor correlation matrices it is also possible to use a subset of correlation matrices, e.g., only $\mathbf{C}_1^x$ and $\mathbf{C}_D^x$.

In order to perform the reduction to size $S \times S$ we use the reduction matrices $\mathbf{P}_s \triangleq (\mathbf{U}^s)^T$ and $\mathbf{Q}^s \triangleq (\mathbf{V}_s)^T$ as follows:

$$\bar{\mathbf{R}}_\kappa^x \triangleq \mathbf{P}_s \mathbf{R}_\kappa^x (\mathbf{P}_s)^T = \bar{\mathbf{A}}\, \text{diag}\left(\mathbf{r}^s[\Omega_\kappa]\right) (\bar{\mathbf{A}})^T \tag{3.12}$$

$$\bar{\mathbf{C}}_i^x \triangleq \mathbf{P}_s \mathbf{C}_i^x \mathbf{Q}^s = \bar{\mathbf{A}}\, \text{diag}\left(\tilde{\mathbf{a}}_i\right) \bar{\mathbf{C}}^s \tag{3.13}$$

with $\bar{\mathbf{A}} = \mathbf{P}_s \mathbf{A}$ and $\bar{\mathbf{C}}^s = \mathbf{C}^s \mathbf{Q}^s$. Later we also need the matrix $\mathbf{P}_\nu \triangleq (\mathbf{U}^\nu)^T$ in order to perform noise reduction.

From now on, we assume that both sets of correlation matrices are reduced to a size of $S \times S$. Based on this assumption we derive a method to identify the desired extraction filter. Later we show that the extra degrees of freedom may be used to perform noise reduction.

3.3 Extraction filter identification

The remaining filter identification procedure consists of two steps, which are visualized in Figure 3.2. First we show that the eigenvalues of a GEVD of two linear

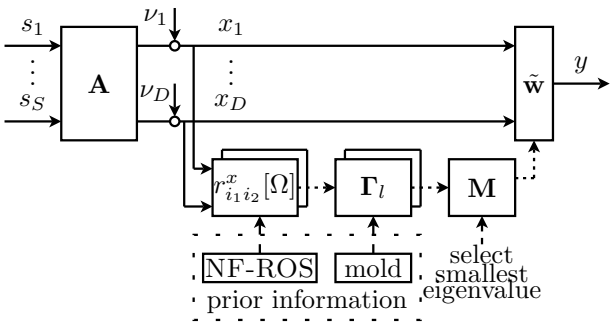


Figure 3.2: Based on the noise-free correlation data $r^x_{i_1 i_2}[\Omega]$ and a mold we take linear combinations of correlation matrices. From these linear combinations $\mathbf{\Gamma}_l$ we construct a matrix $\mathbf{M}$ whose eigenvector that corresponds to the smallest eigenvalue is the desired extraction filter $\tilde{\mathbf{w}}$.

combinations of sensor correlation matrices contain information about the sources. Subsequently, we show that a design of the weights for taking linear combinations based on a priori information can be used for characterizing the eigenvalues. Finally, we show how to create a matrix $\mathbf{M}$ from multiple linear combinations of sensor correlation matrices and how to design weight vectors such that immediate extraction of the desired source is guaranteed.

3.3.1 Identification of the desired extraction filter from two linear combinations of sensor correlation matrices

The rationale behind the source extraction algorithm comes from the joint diagonalization or generalized eigenvalue decomposition (GEVD) of two (linear combinations of) sensor correlation matrices as discussed in Section 2.4.2.2. In this section we show how generalized eigenvalues can be characterized based on a priori information such that the extraction filter corresponding to the desired source can be identified immediately.

Linear combinations of reduced sensor correlation matrices for $\bar{\mathbf{R}}^x_\kappa$ and $\bar{\mathbf{C}}^x_i$ are defined as follows, respectively:

$$\mathbf{\Gamma}_l \triangleq \sum_{\kappa=1}^{K} \xi_l^\kappa \bar{\mathbf{R}}^x_\kappa \qquad \text{and} \qquad \mathbf{\Gamma}_l \triangleq \sum_{i=1}^{D} \xi_i^l \bar{\mathbf{C}}^x_i \tag{3.14}$$

where the weights ξ_l^κ and ξ_i^l are elements of the weight vectors $\tilde{\boldsymbol{\xi}}_l$ and $\boldsymbol{\xi}^l$, respectively. Later we show design techniques for these weight vectors based on a priori information. The GEVD of two linear combinations of reduced sensor correlation matrices $\mathbf{\Gamma}_{l_1}$ and $\mathbf{\Gamma}_{l_2}$ is the set of all eigenvectors and eigenvalues $\{\tilde{\boldsymbol{\mu}}, \lambda\}$ that solve the system: $\lambda \tilde{\boldsymbol{\mu}} \mathbf{\Gamma}_{l_1} = \tilde{\boldsymbol{\mu}} \mathbf{\Gamma}_{l_2}$ [36]. If the matrix $\mathbf{\Gamma}_{l_1}$ is invertible, then the eigenvectors and eigenvalues of this

GEVD can be calculated from the following eigenvalue decomposition:

$$\lambda\tilde{\boldsymbol{\mu}} = \tilde{\boldsymbol{\mu}}\boldsymbol{\Gamma}_{l_2}\left(\boldsymbol{\Gamma}_{l_1}\right)^{-1}. \tag{3.15}$$

By exploiting the structure in the sensor correlation matrices as shown in (3.12) and (3.13) it follows that the S eigenvectors are found as $\tilde{\boldsymbol{\mu}}_j = \tilde{\mathbf{e}}_j(\bar{\mathbf{A}})^{-1}$ for both types of sensor correlation matrices. The eigenvalues corresponding to these eigenvectors for the sensor correlation matrices $\mathbf{R}^x_\kappa$ and $\mathbf{C}^x_i$, respectively, have the following structure (see Appendix 3.A for a derivation):

$$\lambda^j_{l_1 l_2} = \frac{\langle\tilde{\boldsymbol{\xi}}_{l_2}, \tilde{\mathbf{r}}^s_{jj}\rangle}{\langle\tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\mathbf{r}}^s_{jj}\rangle} \qquad \text{and} \qquad \lambda^j_{l_1 l_2} = \frac{\langle\boldsymbol{\xi}^{l_2}, \mathbf{a}^j\rangle}{\langle\boldsymbol{\xi}^{l_1}, \mathbf{a}^j\rangle} \tag{3.16}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidian inner product and $\tilde{\mathbf{r}}^s_{jj}$ is the j'th row from the source autocorrelation matrix $\mathbf{C}^s$, i.e., the autocorrelation function values in the NF-ROS of the j'th source.

Notice that each eigenvalue depends only on the weight vectors $\tilde{\boldsymbol{\xi}}_l$ or $\boldsymbol{\xi}^l$ and the autocorrelation function $\tilde{\mathbf{r}}^s_{jj}$ or the mixing column vector $\mathbf{a}^j$ of the source that is extracted by the corresponding eigenvector $\tilde{\boldsymbol{\mu}}_j$. Consequently, by designing the weight vectors $\tilde{\boldsymbol{\xi}}_l$ or $\boldsymbol{\xi}^l$ based on a priori information about the autocorrelation functions or the mixing columns of the sources we are able to characterize the generalized eigenvalues. For example, if we make sure that the following condition holds:

$$\left|\frac{\langle\boldsymbol{\xi}^{l_2}, \mathbf{a}^d\rangle}{\langle\boldsymbol{\xi}^{l_1}, \mathbf{a}^d\rangle}\right| < \left|\frac{\langle\boldsymbol{\xi}^{l_2}, \mathbf{a}^j\rangle}{\langle\boldsymbol{\xi}^{l_1}, \mathbf{a}^j\rangle}\right| \quad \forall\, j \in \mathcal{S}\backslash d \tag{3.17}$$

where d is the index for the desired source, then the absolute smallest eigenvalue of $\boldsymbol{\Gamma}_{l_2}\left(\boldsymbol{\Gamma}_{l_1}\right)^{-1}$ corresponds to the desired source, which is extracted by the corresponding eigenvector.

Disadvantages of using the matrix $\boldsymbol{\Gamma}_{l_2}\left(\boldsymbol{\Gamma}_{l_1}\right)^{-1}$ are that the eigenvalues can be both positive and negative. Furthermore, only two weight vectors, i.e., $\tilde{\boldsymbol{\xi}}_{l_1}$ and $\tilde{\boldsymbol{\xi}}_{l_2}$ or $\boldsymbol{\xi}^{l_1}$ and $\boldsymbol{\xi}^{l_2}$ can be used to incorporate a priori information. This means that only information about the two dimensional subspace spanned by these vectors is taken into account. Therefore, in the following section we combine multiple linear combinations of sensor correlation matrices into a single matrix $\mathbf{M}$. This matrix has the properties that its left eigenvectors are extraction filters and its eigenvalues are positive and depend on multiple weight vectors.

3.3.2 The desired extraction filter as a specific eigenvector using multiple linear combinations of sensor correlation matrices

By taking the square of the matrix $\boldsymbol{\Gamma}_{l_2}\left(\boldsymbol{\Gamma}_{l_1}\right)^{-1}$ its eigenvalues are squared while its eigenvectors remain the same. Consequently, all the eigenvalues are positive. By setting $l_1 = 1$ and by summing over $S-1$ of such matrices for varying l_2, we obtain

the following matrix $\mathbf{M}$:

$$\mathbf{M} = \sum_{l=2}^{S} \{\mathbf{\Gamma}_l \left(\mathbf{\Gamma}_1\right)^{-1} \mathbf{\Gamma}_l \left(\mathbf{\Gamma}_1\right)^{-1}\}. \tag{3.18}$$

By exploiting the structure in the correlation matrices as in (3.6) and (3.7) or by considering the matrices in terms of eigenvectors and eigenvalues we find the following expression for the eigenvalues of the matrix $\mathbf{M}$ (see Appendix 3.A for a derivation):

$$\lambda^j = \sum_{l=2}^{S} \frac{|\langle \tilde{\boldsymbol{\xi}}_l, \tilde{\mathbf{r}}^s_{jj} \rangle|^2}{|\langle \tilde{\boldsymbol{\xi}}_1, \tilde{\mathbf{r}}^s_{jj} \rangle|^2} \qquad \text{or} \qquad \lambda^j = \sum_{l=2}^{S} \frac{|\langle \boldsymbol{\xi}^l, \mathbf{a}^j \rangle|^2}{|\langle \boldsymbol{\xi}^1, \mathbf{a}^j \rangle|^2}. \tag{3.19}$$

Again, the left eigenvectors $\tilde{\boldsymbol{\mu}}_j$ are the extraction filters and the eigenvalues depend only on the weight vectors and either the source autocorrelation vectors or the mixing column vectors.

By designing the weight vectors based on a priori information we are able to guarantee that the extraction filter for the desired source corresponds to the smallest eigenvalue of $\mathbf{M}$. For this design we assume to have one out of two types of a priori information available in the form of a mold. The first type of mold, $\tilde{\mathbf{r}}_0$, is an estimate of the autocorrelation function of the desired source in the NF-ROS. The second type of mold, $\mathbf{a}^0$, is an estimate of the mixing column that corresponds to the desired source. Conditions for these molds are given later. For the mold $\tilde{\mathbf{r}}_0$ the S weight vectors $\tilde{\boldsymbol{\xi}}_1, \cdots, \tilde{\boldsymbol{\xi}}_S$ are designed as follows. These weight vectors are chosen as an orthonormal basis for the S dimensional signal subspace of the source autocorrelation functions $\mathbf{C}^s$, with the first vector being the normalized mold, i.e., $\tilde{\boldsymbol{\xi}}_1 = \tilde{\mathbf{r}}_0 / ||\tilde{\mathbf{r}}_0||_2$. For the mold $\mathbf{a}^0$, the S weight vectors $\boldsymbol{\xi}^1, \cdots, \boldsymbol{\xi}^S$ for creating linear combinations of correlation matrices are designed in a similar way. These weight vectors are chosen as an orthonormal basis for the S dimensional signal subspace of the mixing system $\mathbf{A}$, with the first vector being the normalized mold, i.e., $\boldsymbol{\xi}^1 = \mathbf{a}^0 / \left|\left|\mathbf{a}^0\right|\right|_2$, where $||\cdot||_2$ is the Euclidian norm. The remaining $S-1$ vectors $\boldsymbol{\xi}^2, \cdots, \boldsymbol{\xi}^S$ or $\tilde{\boldsymbol{\xi}}_2, \cdots, \tilde{\boldsymbol{\xi}}_S$ can be calculated as respectively the first $S-1$ right or left singular vectors of the signal subspace matrix $\mathbf{P}_s$ or $\mathbf{Q}^s$ projected onto the space orthogonal to the mold, i.e.,

$$\left(\mathbf{I} - \frac{(\tilde{\mathbf{r}}_0)^T \tilde{\mathbf{r}}_0}{\tilde{\mathbf{r}}_0 (\tilde{\mathbf{r}}_0)^T}\right) \mathbf{Q}^s \qquad \text{or} \qquad \mathbf{P}_s \left(\mathbf{I} - \frac{\mathbf{a}^0 (\mathbf{a}^0)^T}{(\mathbf{a}^0)^T \mathbf{a}^0}\right). \tag{3.20}$$

Using the fact that the first weight vector is the mold and that the other vectors are orthonormal, the structure in the eigenvalues can be simplified to either:

$$\lambda^j = \frac{||\tilde{\mathbf{r}}_{jj}||_2^2 - \langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}_{jj} \rangle^2}{\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}_{jj} \rangle^2} \qquad \text{or} \qquad \lambda^j = \frac{||\mathbf{a}^j||_2^2 - \langle \mathbf{a}^0, \mathbf{a}^j \rangle^2}{\langle \mathbf{a}^0, \mathbf{a}^j \rangle^2} \tag{3.21}$$

depending on the mold that is available, as is shown in [52] and [57], respectively. Exploiting the geometric interpretation that the inner product measures the angle between two normalized vectors we obtain the following eigenvalues:

$$\lambda^j = \frac{1 - (\cos \phi_j)^2}{(\cos \phi_j)^2} = (\tan \phi_j)^2 \tag{3.22}$$

where ϕ_j is either the angle between the row vectors $\tilde{\mathbf{r}}_0$ and $\tilde{\mathbf{r}}_j$ or the column vectors $\mathbf{a}^0$ and $\mathbf{a}^j$.

By performing an eigenvalue decomposition on $\mathbf{M}$ the ordering of the eigenvalues is completely arbitrary, which is a manifestation of the permutation problem that is known from BSP. However, we have shown that each eigenvalue only depends on the mold and source parameters of a single source. Therefore, by putting a constraint on the mold w.r.t. the information it represents, we are able to characterize the eigenvalue that corresponds to the desired source.

If we require the angle between the mold and the mixing columns, or source autocorrelation functions, to be smallest for the desired source, then we know that the eigenvalue that corresponds to the desired source is the smallest eigenvalue. This holds because $(\tan\phi)^2$ is symmetric and monotonically increasing when starting from $\phi = 0$. The mathematical form of this requirement is as follows:

$$\frac{|\langle \mathbf{a}^0, \mathbf{a}^d \rangle|}{||\mathbf{a}^d||} > \frac{|\langle \mathbf{a}^0, \mathbf{a}^j \rangle|}{||\mathbf{a}^j||} \quad \text{or} \quad \frac{|\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}_d \rangle|}{||\tilde{\mathbf{r}}_d||_2} > \frac{|\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}_j \rangle|}{||\tilde{\mathbf{r}}_j||_2} \tag{3.23}$$

for all $j \neq d$. This requirement characterizes the eigenvalues in such a way that the desired extraction filter is given by $\tilde{\mathbf{w}} = \tilde{\boldsymbol{\mu}}\mathbf{P}_s$, where $\tilde{\boldsymbol{\mu}}$ is the left eigenvector that corresponds to the smallest eigenvalue of the matrix $\mathbf{M}$.

We have presented an overview of the extraction filter identification algorithm. A strength of this procedure is that the desired extraction filter is identified in a single stage and the performance of the extraction filter is independent of errors in the mold, as long as the desired source is selected.

3.4 The UIBSE algorithm

The extraction filter identification procedure leaves room for additional functionality. Therefore, we show that knowledge about two different types of molds can be combined in order to select the desired source. Second, we show that the identification algorithm can be extended with a noise reduction step to identify a LCMV filter. Finally, we show that with a similar procedure the MVDR filter can be found.

3.4.1 Combining different types of prior information

In the UIBSE algorithm a mold $\mathbf{a}^0$ or $\tilde{\mathbf{r}}_0$ is used to condition the correlation data such that the desired extraction filter is selected. Here we show that knowledge about both molds can be merged in order to obtain a more robust selection mechanism, which is especially valuable if the reliability of the molds is low and/or if mixtures of many sources are observed. We present two methods for combining the information and interpret constraints that have to hold for the molds.

Suppose we create two matrices $\mathbf{M}$ as in (3.18) based on two different molds, i.e., $\mathbf{M}_a$ is related to the mold $\mathbf{a}^0$ and $\mathbf{M}_r$ is related to the mold $\tilde{\mathbf{r}}_0$. If we make the

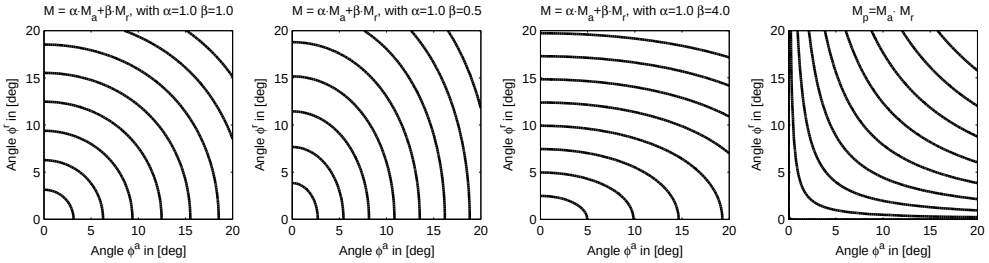


Figure 3.3: Contour maps for different combinations of prior information in terms of the molds. The contour maps are plotted in terms of the angle between the mold and the actual information that they represent. The source that is located in a trajectory closest to the point (0,0) is selected by the UIBSE algorithm. The three maps on the left are from different linear combinations of the matrices $\mathbf{M}_a$ and $\mathbf{M}_r$. The map on the right is from the product of these two matrices.

following linear combination (lc) of these two matrices: $\mathbf{M}_{lc} = \alpha\mathbf{M}_a + \beta\mathbf{M}_r$ for $\alpha \geq 0$ and $\beta \geq 0$, then the eigenvectors of matrix $\mathbf{M}_{lc}$ are again extraction filters and the corresponding eigenvalues have the following structure:

$$\lambda_{lc}^j = \alpha\lambda_a^j + \beta\lambda_r^j = \alpha(\tan\phi_j^a)^2 + \beta(\tan\phi_j^r)^2 \tag{3.24}$$

where the eigenvalues λ^j with subscript symbols lc, a, and r are the eigenvalues from the matrices $\mathbf{M}$ with the same subscript symbol, ϕ_j^a is the angle between the mold $\mathbf{a}^0$ and the j'th mixing column vector and ϕ_j^r is the angle between the mold $\tilde{\mathbf{r}}_0$ and the j'th source autocorrelation function in the NF-ROS $\tilde{\mathbf{r}}_j$.

This combination of matrices results in a non-linear geometric interpretation in terms of the angles ϕ_j^a and ϕ_j^r. If we require the smallest eigenvalue of the matrix $\mathbf{M}_{lc}$ to correspond to the desired source, then we obtain the following conditions in terms of the angles:

$$\alpha(\tan\phi_d^a)^2 + \beta(\tan\phi_d^r)^2 < \alpha(\tan\phi_j^a)^2 + \beta(\tan\phi_j^r)^2 \tag{3.25}$$

for all $j \neq d$. In order to give an interpretation to these constraints we visualized the constraints for three (α, β) pairs by means of the three contour maps on the left in Figure 3.3. Suppose we observe a mixture of two sources, one source for which the molds lead to the angle pair $(\phi_1^a, \phi_1^r) = (2, 5)$ and one source for which the molds lead to the angle pair $(\phi_1^a, \phi_1^r) = (5, 2)$, these pairs can be interpreted as coordinates on the contour maps. The values on the contour map are the values that the combined eigenvalues λ_{lc}^j take, which increase when moving away from point (0, 0). For the first contour map, this means that the eigenvalues for both sources are the same. If we select other values for α and β, as is depicted in the second and third contour map, then a separation of the eigenvalues is obtained. In the second contour map the source at (2, 5) has the lowest eigenvalue, while in the third contour map the source at (5, 2) has the smallest eigenvalue.

From this figure we observe that the weights α and β can be interpreted as a penalty weight for the corresponding angle ϕ_j^a or ϕ_j^r. If the value of α increases w.r.t.

β, then an angle ϕ_j^a is more penalized than an angle ϕ_j^r. On the other hand, if β increases w.r.t. α, then the angle ϕ_j^r is more penalized than the angle ϕ_j^a. Ultimately, if $\alpha \gg \beta$ or $\beta = 0$, then $\mathbf{M}_{lc} \simeq \mathbf{M}_a$ and if $\beta \gg \alpha$ or $\alpha = 0$, then $\mathbf{M}_{lc} \simeq \mathbf{M}_r$. Overall, a smooth trade-off is made between a larger angle in one domain or the other.

We understand that the example scenario is a very simplified scenario and in fact only $\mathbf{M}_a$ or $\mathbf{M}_r$ should be used to select the desired source in this example. However, in more complex scenarios with more sources and noisy data we believe that this approach may help to increase the selection performance of the extraction algorithm.

The second type of combination is created by the following product (p): $\mathbf{M}_p = \mathbf{M}_a\mathbf{M}_r$. Because the eigenvectors of the matrices $\mathbf{M}_a$ and $\mathbf{M}_r$ are the same, the eigenvalues of the matrix $\mathbf{M}_p$ are the products of the eigenvalues of the individual matrices, i.e.,

$$\lambda_p^j = \lambda_a^j \cdot \lambda_r^j = (\tan\phi_j^a)^2 \cdot (\tan\phi_j^r)^2 \tag{3.26}$$

where the index j corresponds to the same eigenvectors. This product of eigenvalues results again in non-linear constraints with respect to the errors in terms of angles ϕ_j^a and ϕ_j^r, i.e,

$$(\tan\phi_d^a)^2 \cdot (\tan\phi_d^r)^2 < (\tan\phi_j^a)^2 \cdot (\tan\phi_j^r)^2 \quad \forall\, j \neq d. \tag{3.27}$$

The contour map for these constraints is depicted on the right in Figure 3.3. In contrast to the linear combination method with parameters α and β, no parameters can be selected for this product based method. For the example scenario with sources at (2, 5) and (5, 2), this method has no advantage because both eigenvalues are the same. However, this contour map shows that a very high selectivity is obtained for a mold that is very close to the actual data it represents. If the error in one domain (ϕ_j^a or ϕ_j^r) increases, then the error in the other domain becomes more and more important in the selection procedure.

Depending on prior information that is available in an application, the desired form of combination, and if necessary the desired values for α and β, should be chosen. Alternatively, methods for learning optimal values for α and β could be a topic for future research.

3.4.2 UIBSE approach to identify the LCMV filter

A noise reduction technique based on a sidelobe canceller structure is already proposed in [58]. In the field of array signal processing this filter is called a linear constrained minimum variance (LCMV) filter, which has the following underlying optimization problem:

$$\tilde{\mathbf{w}}_{\text{LCMV}} = \underset{\tilde{\mathbf{w}}}{\text{argmin}}\ \mathbb{E}\{y^2[n]\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{A} = \tilde{\mathbf{e}}_d \tag{3.28}$$

where y is the output of the extraction filter and $\tilde{\mathbf{e}}_d$ is a row vector with a one at the d'th location and zeros elsewhere.

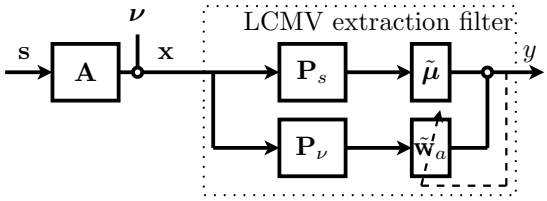


Figure 3.4: Overview of the structure for the LCMV optimization problem. In the upper branch the interferers are suppressed and an estimation of the desired source is created. The lower branch is solely used for noise reduction.

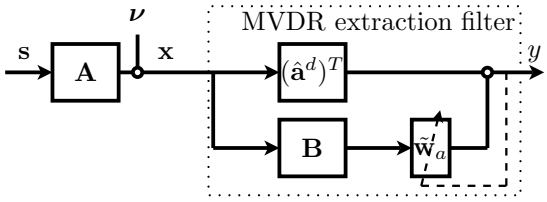


Figure 3.5: Overview of the structure for the MVDR optimization problem. In the upper branch an estimation of the desired signal is created and the lower branch is used for both noise and interference reduction.

The extraction filter $\tilde{\mathbf{w}} = \tilde{\boldsymbol{\mu}}\mathbf{P}_s$ from Section 3.3 fulfills the constraints in this optimization problem, i.e., the contribution of the undesired sources is zero and the contribution of the desired source is non-zero. By applying noise reduction, using the orthogonal noise-only subspace $\mathbf{P}_\nu$, we obtain the following overall extraction filter:

$$\boxed{\tilde{\mathbf{w}}_{\text{LCMV}} = \tilde{\boldsymbol{\mu}}\mathbf{P}_s - \tilde{\mathbf{w}}_a\mathbf{P}_\nu} \tag{3.29}$$

where the noise reduction filter $\tilde{\mathbf{w}}_a$ is identified by minimizing the output power over the $D-S$ filter coefficients of $\tilde{\mathbf{w}}_a$. Several batch and adaptive algorithms are available in the literature to solve this problem, e.g., (N)LMS and RLS.

An implementation to solve this optimization problem is depicted in Figure 3.4. The S constraints are solved in the upper branch and the $D-S$ dimensional unconstrained optimization problem is solved in the lower branch. Notice that if $D = S$ the only task of the extraction filter is to perform interference reduction.

3.4.3 UIBSE approach to identify the MVDR filter

When using BSP techniques, typically the LCMV filter is applied; however, from the field of array signal processing another filter is known. The minimum variance distortionless response (MVDR) filter has the following objective:

$$\tilde{\mathbf{w}}_{\text{MVDR}} = \underset{\tilde{\mathbf{w}}}{\text{argmin}}\ \mathbb{E}\{y^2[n]\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{a}^d = 1 \tag{3.30}$$

and the goal is to maximize the signal to interference plus noise ratio (SINR).

The optimal MVDR filter has the following structure:

$$\boxed{\tilde{\mathbf{w}}_{\text{MVDR}} = (\mathbf{a}^d)^T - \tilde{\mathbf{w}}_a \mathbf{B}} \tag{3.31}$$

where the blocking matrix $\mathbf{B}$ of size $(D-1) \times D$ is orthogonal to the transposed mixing column of the desired source $(\mathbf{a}^d)^T$ and the filter $\tilde{\mathbf{w}}_a$ performs the noise and interference reduction.

This MVDR filter requires the identification of the desired source mixing column vector $\mathbf{a}^d$, which can be done with a procedure that is very similar to the extraction filter identification procedure from Section 3.3. The set of right eigenvectors is the inverse of the set of left eigenvectors, which are in turn the row vectors from the inverse of the size-reduced mixing system. Therefore, the right eigenvector that corresponds to the smallest eigenvalue of the matrix $\mathbf{M}$ is the desired column from the size-reduced mixing system, i.e., $\boldsymbol{\mu} = \mathbf{P}_s \mathbf{a}^d$.

An implementation for the MVDR optimization problem in (3.30) is depicted in Figure 3.5. In the upper branch, the mixing column that corresponds to the desired source is used to obtain an estimate of the desired source signal. In the lower branch a blocking matrix $\mathbf{B}$ of size $(D-1) \times D$ is used that is orthogonal to the estimated mixing column of the desired source $(\mathbf{a}^d)^T = (\boldsymbol{\mu})^T \mathbf{P}_s$. As a result, the $D-1$ lower branch signals do not contain the desired source signal. The filter $\tilde{\mathbf{w}}_a$ uses these signals to reduce the noise and undesired sources in the upper branch signal. Also this filter may be identified adaptively using for example the (N)LMS or RLS algorithm.

We have shown that both filters can be obtained by following a very similar procedure; however, the choice for one of these BSE techniques to perform noise and interference reduction depends on the application.

3.5 Summary of unified instantaneous BSE (UIBSE)

Based on the following types of a priori information:

- The number of sources S
- Knowledge about at least S time-lag pairs in a NF-ROS
- An estimation of the desired source autocorrelation function $\tilde{\mathbf{r}}_0$ and/or mixing parameters $\mathbf{a}^0$

The UIBSE algorithm works as follows:

- Calculate correlation matrices based on the NF-ROS
  - $\mathbf{R}_i^x$ as in (3.6) if mold $\tilde{\mathbf{r}}_0$ is available
  - $\mathbf{C}_i^x$ as in (3.7) if mold $\mathbf{a}^0$ is available
- Calculate the SVD of either a matrix $\mathbf{C}_i^x$ or matrices $\mathbf{C}^x$ and $\check{\mathbf{C}}^x$ from (3.10) and (3.11) to identify $\mathbf{P}_s, \mathbf{P}_\nu$, and $\mathbf{Q}^s$
- Reduce correlation matrices $\mathbf{R}_i^x$ and/or $\mathbf{C}_i^x$ to size $S \times S$ matrices $\bar{\mathbf{R}}_i^x$ and/or $\bar{\mathbf{C}}_i^x$ using (3.12) and/or (3.13)

- Construct orthonormal basis $\tilde{\boldsymbol{\xi}}_1 \cdots \tilde{\boldsymbol{\xi}}_S$ and/or $\boldsymbol{\xi}^1 \cdots \boldsymbol{\xi}^S$ that span the signal subspace, with $\tilde{\boldsymbol{\xi}}_1 = \tilde{\mathbf{r}}_0$ and $\boldsymbol{\xi}^1 = \mathbf{a}^0$
- Create S linear combinations from the reduced size correlation matrices: $\boldsymbol{\Gamma}_l = \sum_{i=1}^{N} \xi_l^i \bar{\mathbf{R}}_i^x$ and/or $\boldsymbol{\Gamma}_l = \sum_{i=1}^{D} \xi_i^l \bar{\mathbf{C}}_i^x$
- Construct (a combination of) $\mathbf{M}_r$ and/or $\mathbf{M}_a$ from the matrices $\boldsymbol{\Gamma}_l$ and find the left or right eigenvector that corresponds to the smallest eigenvalue
- Construct the LCMV or MVDR extraction filter and perform noise reduction with an adaptive or batch algorithm

3.6 Simulation results

We verify the UIBSE algorithm by computer simulations. First we evaluate the desired source selection performance when different types of prior information are combined. Second, we compare the performance of the identified extraction filter, LCMV, and MVDR configurations. Finally, we compare overall performance of the presented approach with the SOBI and SOBI-RO algorithm.

3.6.1 Combining different types of prior information

We evaluate the usage of a combination of the molds $\mathbf{a}^0$ and $\tilde{\mathbf{r}}_0$ for a better performance in desired source selection with respect to using only one type of mold. For this evaluation we use three autoregressive sources, each with a single parameter being 0.9, 0.5, and 0.7, respectively. The source signals have unit power and are mixed with the following mixing system:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \mathbf{a}^3 \end{bmatrix} = \begin{bmatrix} 0.0676 & 0.0308 & -0.4075 \\ 0.9380 & 0.9604 & -0.6359 \\ -0.3400 & 0.2770 & -0.6554 \end{bmatrix}. \tag{3.32}$$

Because the observations are free of noise, the NF-ROS is chosen as the lags 0, 1, and 2. We assume to have the mold $\tilde{\mathbf{r}}_0 = \begin{bmatrix} 0.6985 & 0.5526 & 0.4442 \end{bmatrix}$, which is such that the angles ϕ_j^r between $\tilde{\mathbf{r}}_0$ and the autocorrelation functions of the sources in the NF-ROS are 5.55, 17.77, and 5.55 degrees, respectively. Additionally, we assume to have a mold $\mathbf{a}^0$. We choose the mold $\mathbf{a}^0$ in the subspace spanned by the vectors $\mathbf{a}^1$ and $\mathbf{a}^2$. Initially we choose the mold as the first mixing column vector, i.e., $\mathbf{a}^0 = \mathbf{a}^1$. Subsequently, we increase the angle ϕ_1^a between the mold $\mathbf{a}^0$ and $\mathbf{a}^1$, up to 90 degrees. We performed a Monte Carlo simulation of 100 iterations per degree for six different UIBSE algorithms. Each algorithm uses 10 000 samples of observations, the molds, and the NF-ROS to identify the desired extraction filter. The selection performance of each algorithm is measured by the Extraction Index, which is one if the desired source s_1 is extracted and otherwise zero. A different combination of the matrices $\mathbf{M}_a$ and $\mathbf{M}_r$ is taken for each algorithm. The first five algorithms are based on linear combinations $\alpha\mathbf{M}_a + \beta\mathbf{M}_r$ with (α, β) pairs being $(1, 0), (0, 1), (1, 0.5), (1, 1)$, and $(1, 4)$, respectively. The sixth algorithm uses the product $\mathbf{M}_a\mathbf{M}_r$. Note that the

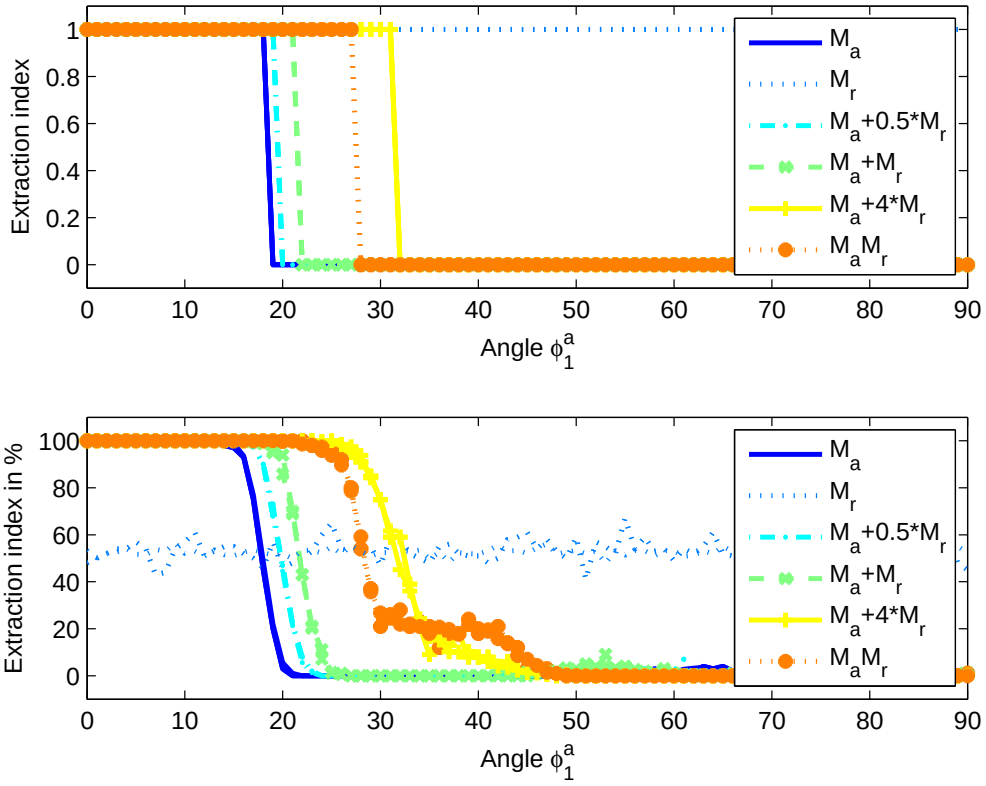


Figure 3.6: For a scenario where the angle between the mold and the actual mixing column vector of the desired source is increased, we compare the desired source selection performance of UIBSE algorithms that use different combinations of prior knowledge. The theoretical boundary per algorithm is depicted in the top of the figure, where a one indicates the extraction of the desired source and a zero indicates the extraction of an undesired source. The corresponding results from a Monte Carlo simulation are depicted in the bottom by the amount of desired source extractions per algorithm per degree.

values for α and β, which should both be positive, are randomly chosen in order to investigate their influence.

From Section 3.4.1 it follows that we can calculate for each algorithm for which angles ϕ_1^a the smallest eigenvalue corresponds to desired source, which is depicted in the top of Figure 3.6. The results of the Monte Carlo simulation are depicted in the bottom of Figure 3.6.

From Figure 3.6 we observe that the transition from desired to undesired source extraction in the Monte Carlo simulation occurs at the theoretically expected angles. Additionally, the transition in the Monte Carlo simulation goes smoothly, which is probably due to the fact that the smallest eigenvalues become rather similar near the

transition area such that small estimation errors lead to the selection of an undesired source. A special case is the algorithm that uses only $\mathbf{M}_r$. This algorithm is independent of variations in ϕ_1^a and therefore a constant value for the Extraction Index is expected. The algorithm is designed to extract that source that has the smallest value for ϕ_j^r for $j \in [1, 2, 3]$. In this example the ϕ_1^r is very similar to ϕ_3^r, but slightly smaller due to rounding. Therefore the theoretical Extraction Index is always one; however, due to estimation inaccuracies the desired extraction filter is selected only in about 50% of the cases such that the simulated Extraction Index is around 50% for each angle ϕ_1^a. As expected, we observe that the Extraction Index is independent of the angle ϕ_1^a.

We conclude that the desired source is extracted if the molds are close enough to the information that they represent. The use of a combination of prior knowledge can help to increase the performance of the system.

3.6.2 Comparing UIBSE objectives: extraction, LCMV, and MVDR

We compare different configurations of the UIBSE algorithm and show that signal extraction can be performed in a similar way for different objectives. In these simulations we show that it may be advantageous to be able to use different objectives.

We mix two speech sources, which have unit variance and are sampled at 8 kHz. The mixtures are contaminated by temporally white, spatially correlated Gaussian noise. The noise signals are generated by mixing three white Gaussian sequences with a variance of 1/30, i.e., $\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] + \mathbf{H}\boldsymbol{\nu}[n]$, where the mixing systems are given as follows:

$$\mathbf{A} = \begin{bmatrix} -0.6406 & 0.1007 \\ 0.7358 & -0.9410 \\ -0.2197 & 0.3231 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 0.1803 & 0.7249 & 0.0128 \\ 0.0195 & 0.4202 & 0.4874 \\ 0.4632 & 0.4854 & 0.9418 \end{bmatrix},$$

and where $\boldsymbol{\nu}[n]$ are temporally and spatially white noise signals.

We compare the following three configurations of the UIBSE algorithm.

1. The Extraction algorithm from Section 3.3, i.e., the extraction filter is a time-invariant filter that is identified in batch mode using all sensor data.
2. The LCMV based UIBSE algorithm from Section 3.4.2 and Figure 3.4. The filters $\mathbf{P}_s$, $\mathbf{P}_\nu$, and $\tilde{\boldsymbol{\mu}}$ are identified in batch using the methods from Section 3.3. The filter $\tilde{\mathbf{w}}_a$ is identified by applying the adaptive normalized least mean squares (N-LMS) algorithm, which has the following update function:

$$\tilde{\mathbf{w}}_a[n+1] = \tilde{\mathbf{w}}_a[n] + \frac{\mu}{\bar{\sigma}_x[n]} y\,[n]\,(\mathbf{P}_v \cdot \mathbf{x}[n])^T \tag{3.33}$$

   where $\bar{\sigma}_x[n+1] = \beta\bar{\sigma}_x[n] + (1-\beta)(\mathbf{x}[n])^T\mathbf{x}[n]$ with $\mu = 0.1$, $\beta = 0.99$, and $\tilde{\mathbf{w}}_a[0] = 0$.
3. The MVDR based UIBSE algorithm from Section 3.4.3 and Figure 3.5. Again, the filters $\bar{\mathbf{a}}^d$ and $\mathbf{B}$ are calculated in batch modus and the filter $\tilde{\mathbf{w}}_a$ is identified

by applying the adaptive N-LMS algorithm, with the following update rule:

$$\tilde{\mathbf{w}}_a[n+1] = \tilde{\mathbf{w}}_a[n] + \frac{\mu}{\bar{\sigma}_x[n]} y[n] \left(\mathbf{B} \cdot \mathbf{x}[n]\right)^T \tag{3.34}$$

where $\bar{\sigma}_x[n+1] = \beta\bar{\sigma}_x[n] + (1-\beta)(\mathbf{x}[n])^T\mathbf{x}[n]$ with $\mu = 0.1$, $\beta = 0.95$, and $\tilde{\mathbf{w}}_a[0] = [0\ 0]$.

The observed signals are split into six non-overlapping segments of 300 ms. In each segment, correlations are calculated for lags 1, 2, and 3, which leads to $N = 18$ time-lag pairs in the NF-ROS. Furthermore, the mold $\tilde{\mathbf{r}}_0^s$ is a guess of the autocorrelation function of the first source $\tilde{\mathbf{r}}_1^s$, which is calculated from the clean speech signal.

We use a method to evaluate the UIBSE algorithms that is inspired by [59]. The extracted signal obtained from the different configurations of the UIBSE algorithm is decomposed into three components: $y = s_d + e_i + e_\nu$, where the signal components s_d, e_i, and e_ν are due to the desired source, interference, and noise, respectively. These components are calculated by exciting the time variant source extraction filters obtained with the UIBSE algorithms with the observations that only contain the corresponding signals, which is possible since we have the signals and mixing systems from the simulation.

Based on the decomposition of the extracted signal we measure the performance by means of the signal to interference plus noise (SINR), signal to interference (SIR), and signal to noise (SNR) ratios, which are defined as follows:

$$\underset{10\log_{10}\frac{\mathbb{E}\{{s_d}^2\}}{\mathbb{E}\{{e_i}^2+{e_\nu}^2\}}}{\text{SINR}} \quad \underset{10\log_{10}\frac{\mathbb{E}\{{s_d}^2\}}{\mathbb{E}\{{e_i}^2\}}}{\text{SIR}} \quad \underset{10\log_{10}\frac{\mathbb{E}\{{s_d}^2\}}{\mathbb{E}\{{e_\nu}^2\}}}{\text{SNR}} .$$

The speech signals have the property that signal powers change over time; therefore, the expected signal powers are calculated by filtering the squared instantaneous signal components with a 50 ms Hanning window. In the upper subplot of Figure 3.7 the estimated source signal powers are depicted for a period of 437 ms. The differences in performance of the different configurations of the UIBSE algorithm are indicated by showing the performance improvements with respect to the arbitrarily chosen reference extraction filter that sums sensor signals. The time variant performance improvements with respect to the performance when the sensor signals are summed are depicted in the other subplots of Figure 3.7 for the same period. Note that the choice of reference filter influences the absolute performance improvements; however, the signal power dependencies remain similar for the different configurations of the UIBSE algorithm.

First, we observe that all three methods improve all performance measures with respect to summing the sensor signals. Both the extraction and LCMV algorithm have an excellent performance on the SIR measure, which means that the constraints to suppress interference are fulfilled. The difference between the Extraction and LCMV algorithm follows from the SNR measure. Mainly because the noise has spatial structure, the LCMV algorithm has a better noise reduction performance than the

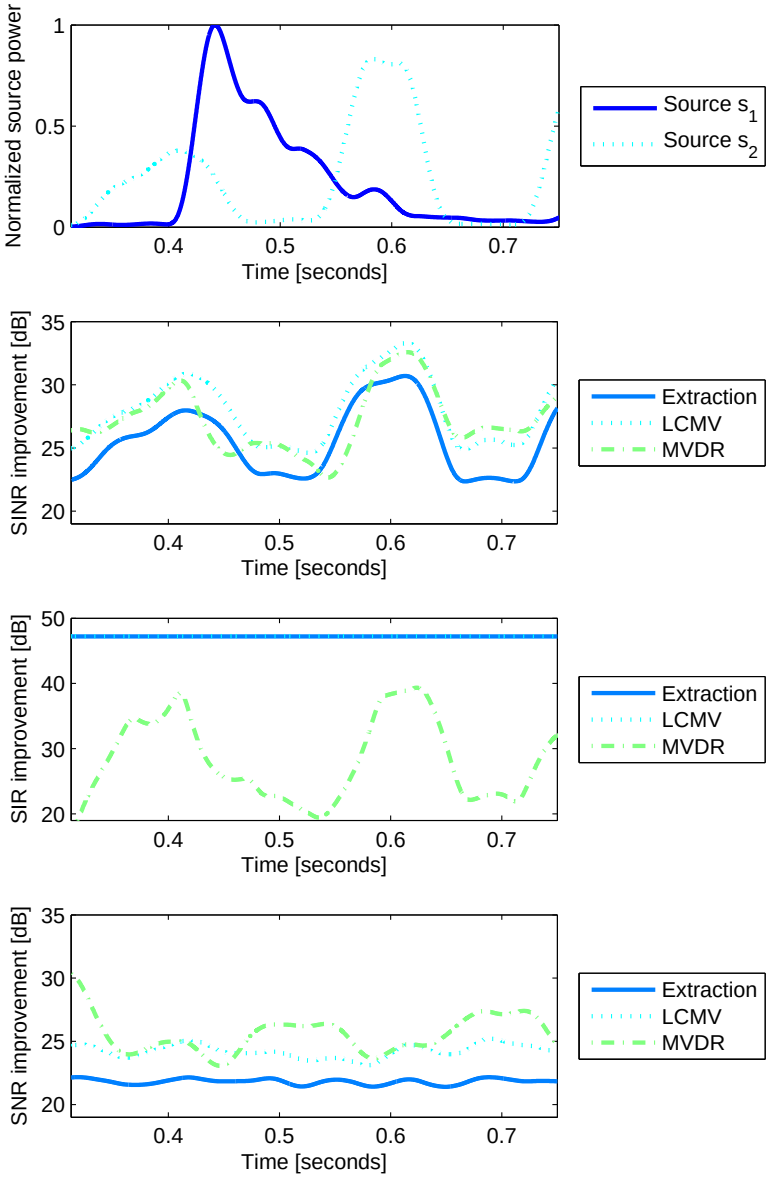


Figure 3.7: Source signal power estimates are depicted in the upper subplot for a period of 437 ms starting after 312 ms. In the other subplots the SINR, SIR, and SNR improvements are depicted for the following three algorithms: UIBSE extraction, LCMV based UIBSE, and MVDR based UIBSE. The improvement is calculated with respect to the performance of summing the sensor signals.

Extraction algorithm. The noise reduction performance of the LCMV algorithm with respect to the Extraction algorithm is also visible in the SINR measure, where a

constant improvement is obtained.

The MVDR algorithm shows a worse performance on SIR than the Extraction and LCMV performance, as was expected. Furthermore, we observe that the SIR improvement for the MVDR is highly depending on the input signal composition. In terms of SNR improvement the MVDR algorithm shows a performance that is at least similar to the LCMV algorithm and is even better at low input source and/or interference power. Finally, in terms of SINR improvement we observe that for low signal powers, i.e., relatively high noise powers the MVDR algorithm performs best.

3.6.3 Comparing performance: UIBSE, SOBI, and SOBI-RO

We compare the difference in performance for the extraction of one desired source signal from a mixture of three source signals by means of a Monte Carlo simulation. We equipped the second order blind identification (SOBI) [45] and SOBI with robust orthogonalization (SOBI-RO) [49] algorithms with a classifier that selects the desired source signal based on prior information about the mixing system. This classifier calculates the inner product of the mold with all identified and normalized mixing columns and selects the mixing column with the highest absolute value.

For the performance evaluation we mixed three autoregressive sources consisting of 10 000 samples with their pole pairs as $-0.5 \pm 0.49i$, $0.1 \pm 0.8i$, and $0.9 \pm 0.15i$, respectively. The mixed source signals are observed by three sensors. Each Monte Carlo iteration a new mixing system is chosen randomly and the actual first mixing column was used as the mold. Both the source signal powers and mixing columns are normalized such that the same source signal power is observed by the sensors for each simulation. We use the total noise power (TNP), which is the sum over the variances of the noise signals, to control the noise level at the observed signals. For these simulations the noise signals are both spatially and temporally white; furthermore, the noise power is chosen equal for each sensor.

The NF-ROS consists of every lag except lag 0; however, for very low values of the TNP this lag becomes almost noise-free. Therefore, we choose two configurations for the NF-ROS. The first NF-ROS consists of the lags 1 up to 20, which is used by UIBSE-NF, UIBSE-MVDR, and SOBI-RO. The second NF-ROS consists of the lags 0 up to 19, which is used by UIBSE-N and SOBI. Notice that UIBSE-NF (noise-free) and UIBSE-N (noisy) are the same algorithm from Section 3.3 except for the NF-ROS. The UIBSE-MVDR algorithm is the algorithm from Section 3.4.3 where the filter $\tilde{\mathbf{w}}_a$ is calculated as the optimal Wiener solution using $\mathbf{R}_x = \mathbb{E}\{(\mathbf{x}[n] \cdot \mathbf{x}[n])^T\}$, which is calculated by taking the mean over all samples.

The extracted signal is decomposed into four components using the projection method from [59], i.e., $y = s_d + e_i + e_\nu + e_a$, where s_d is the contribution from the desired source, e_i from the interference, e_ν from noise, and e_a from artifacts. Since we are using a time invariant, linear filter the artifacts component equals zero, which was also observed during simulations, and therefore omitted. This reduces the performance measures from [59] to signal to distortion ratio (SDR), which equals SINR, and the SIR and SNR measures. During the simulations we also measured

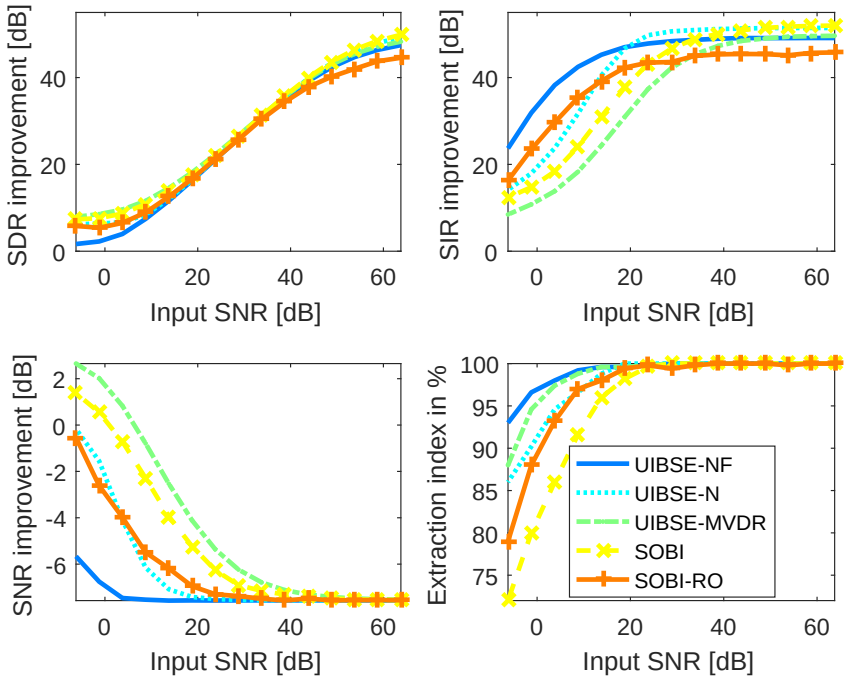


Figure 3.8: Performance comparison of SOBI, SOBI-RO and three configurations of the UIBSE algorithm for the performance measures SDR, SIR, and SNR in a Monte Carlo simulation of 500 runs. The performance improvement with respect to summing the sensor signals is depicted. In the bottom right figure the Extraction Index (EI) is depicted, which measures how often the desired source is extracted.

the performance of the classifier. We calculated the Extraction Index (EI), which indicates if the desired source was extracted.

The simulation results are depicted in Figure 3.8, where we depict the improvement of SDR, SIR, and SNR with respect to summing the sensor signals. We observe that all methods improve the SDR and SIR performance with respect to summing the sensor signals; however, the performance in terms of SNR improvement is worse for each algorithm except for very low input SNR.

The primary objective of the different algorithms is to improve the SIR. It is observed that under noisy conditions the UIBSE-NF algorithm performs best, followed by the SOBI-RO algorithm. As expected, the UIBSE-N and SOBI algorithms suffer from the additional noise. For high SNR values the SIR improvement of the UIBSE-N and SOBI algorithm raises above the SIR improvement of the UIBSE-NF algorithm. This observation can be explained by the fact that the lag zero correlation data contains all energy in the source signals leading to a good extraction filter estimate if

the data is not corrupted by noise. The UIBSE-MVDR has a different objective from the other algorithms. The otjective of the UIBSE-MVDR algorithm is to maximize the SINR whereas the objective of the other algorithms is to maximize the SIR. The effect of this difference in objective is observed in the SIR improvement in the sense that the performance is continuously below the SIR improvement of the UIBSE-NF algorihtm, except for high SNR values. For high SNR values the objective of the UIBSE-NF and UIBSE-MVDR algorithms become the same, i.e., to perform cancellation of the interfering source.

As expected, SNR improvement mainly shows for low input SNR values for all algorithms. Therefore, we focus on the low input SNR use case when evaluating SNR improvement. The UIBSE-NF algorithm shows the lowest SNR improvement, whereas the UIBSE-MVDR algorithm shows the highest SNR improvement. The UIBSE-NF algorithm does not take noise into account and it focusses fully on SIR improvement. The UIBSE-MVDR algorthm balances between SIR and SNR improvement and for low input SNR values this leads to a reduction in SIR improvement while providing a large SNR improvement. Whereas the SIR improvement of the UIBSE-N, SOBI and SOBI-RO algorithms suffers from the presence of noise, the presence of noise leads to a form of regularization in such a way that the lack of SIR improvement leads to an increase SNR improvement with respect to the UIBSE-NF algorithm. However, a drawback of these methods is that is impossible to regulate or balance the amount of noise versus interference reduction.

The SDR improvement of the algorithms is as expected from the results for SIR and SNR improvement. The UIBSE-MVDR algorithm balances between noise reduction and interference cancellation, leading to the high SDR improvement, especially for low input SNR. On the contrary, the SDR improvement of the UIBSE-NF algorithm for low input SNR is dominated by the poor SNR improvement of that algorithm for that use case. Finally, in terms of extraction index the UIBSE algorithms outperform the SOBI and SOBI-RO algorithms. As expected, best performance is obtained with the UIBSE-NF and UIBSE-MVDR algorithms, since these algorithms properly use the NF-ROS.

In order to have a general feeling about complexity we used the Matlab profiler to measure the time spent per algorithm. In these experiments, both SOBI algorithms took more than twice as much time as the UIBSE algorithms.

3.7 Discussion

The main goal of this work is to obtain insight in the BSE problem in order to derive a flexible and robust UIBSE algorithm.

We have shown that different types of prior information about the desired source can be used and combined as a mold to extract the desired source in a single-stage procedure. We have linked eigenvalues to angles between the mold and the information that they represent. This paves the way to a detection system that measures the presence of a source with a specific mixing and/or correlation structure. An example

application where we expect the prior information to be available is in fetal ECG extraction from maternal ECG. The periodicity of the heart rate could be used as an autocorrelation mold, which is already done in [60]. Alternatively, when expanding the presented methodology to complex or convolutive mixtures then a direction of arrival estimate could give an estimate of the mixing of the desired source. Another way of expanding the presented approach is to include HOS. In such a case new types of prior information can be used to select the desired source.

Furthermore, we have shown that different objectives can be used to extract the desired source. The performance with respect to these objectives depends strongly on the input SNR; and, depending on the type of application the desired extraction objective should be chosen. The LCMV filter focusses on interference reduction, while the MVDR filter optimally weights noise and interference reduction. For a wide range of applications, different configurations of the UIBSE algorithm can be created. The main choices that have to be made are the prior information in terms of the mold and NF-ROS, objective function, and noise reduction implementation.

It follows from the diverse use of prior information and the choice to solve different optimization criteria with a single method that the UIBSE method is very flexible. Furthermore, the performance comparison with the SOBI and SOBI-RO algorithms in Section 3.6 shows that the method is more robust to noise than the SOBI and SOBI-RO algorithms. Additionally, in [52] it was shown that the method is more robust to noise than the linear prediction based BSE method [51]. Finally, we have seen from the Matlab profiler that the complexity of the UIBSE algorithm is competitive with respect to SOBI and SOBI-RO.

Overall, we have presented a flexible and robust method to deal with the BSE problem. Most importantly, we have gained insight in the problem, which may lead to more efficient, robust algorithms.

3.8 Conclusion

We developed a unified, single-stage method to identify extraction filters from the second order statistics of observed sensor signals that is flexible in the use of prior information. The main objective was to exploit the full potential of the second order statistics, which lead to a method that is able to identify three types of extraction filters that extract the desired source in a single stage. These extraction filters are based on well known optimization criteria from different fields of applications and may be used for systems where mutually uncorrelated sources are instantaneously mixed and corrupted by (structured) noise. We deal with noise by the prior knowledge of a noise-free region of support (NF-ROS) [36].

By scaling the filter coefficients or the output signal power we deal with the scaling indeterminacy. The permutation problem is solved by using prior knowledge about the desired source in terms of a *mold*. With a mathematical analysis we have proven that the desired source is extracted, given that a certain condition holds for the *mold*. This condition implies that the prior information has to be a reasonable estimate of

the information it represents. We have shown that two types of prior information can be combined in the same approach to obtain a more flexible and robust desired source selection mechanism. Furthermore, we have shown that the algorithm improves the SINR, SIR, and SNR and therefore they may be used to extract a source signal. Finally, the performance of the presented work is similar or even better than the well-known SOBI and SOBI-RO methods.

We conclude that the presented work contains a flexible, robust, and unified method to deal with the BSE problem and that insight is obtained in the BSE problem, which may lead to more efficient and robust algorithms. Finally, suggestions for future research are the development of more efficient and realtime algorithms and the generalization of these concepts of signal extraction from instantaneous to convolutive mixing systems.

3.A Derivations of selected equations

Derivation of (3.6) The sensor correlation matrix $\mathbf{R}_\kappa^x$ is obtained in the following way:

$$\mathbf{R}_\kappa^x = \mathbb{E}\{\mathbf{x}[n](\mathbf{x}[n-k])^T\} = \begin{bmatrix} r_{11}^x[\Omega_\kappa] & \cdots & r_{1D}^x[\Omega_\kappa] \\ r_{21}^x[\Omega_\kappa] & \cdots & r_{2D}^x[\Omega_\kappa] \\ \vdots & \ddots & \ddots \\ r_{D1}^x[\Omega_\kappa] & \ddots & r_{DD}^x[\Omega_\kappa] \end{bmatrix} \tag{3.35}$$

where $\Omega_\kappa = (n,k)_\kappa$ is a time-lag pair in the NF-ROS. By substituting the vector with observations $\mathbf{x}[n]$ by the model in (3.1) we obtain the following structure:

$$\begin{aligned} \mathbf{R}_\kappa^x &= \mathbb{E}\{(\mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n])(\mathbf{A}\mathbf{s}[n-k] + \boldsymbol{\nu}[n-k])^T\} \\ &= \mathbf{A}\mathbb{E}\{\mathbf{s}[n](\mathbf{s}[n-k])^T\}(\mathbf{A})^T + \mathbb{E}\{\boldsymbol{\nu}[n](\boldsymbol{\nu}[n-k])^T\} \\ &\quad + \mathbf{A}\mathbb{E}\{\mathbf{s}[n](\boldsymbol{\nu}[n-k])^T\} + \mathbb{E}\{\boldsymbol{\nu}[n](\mathbf{s}[n-k])^T\}(\mathbf{A})^T. \end{aligned}$$

Applying the assumptions on the source and noise signals that hold for time-lag pairs in the NF-ROS, as given in Definition 3.2.2, leads to the following structure in the sensor correlation matrix:

$$\mathbf{R}_\kappa^x = \mathbf{A}\mathbb{E}\{\mathbf{s}[n](\mathbf{s}[n-k])^T\}(\mathbf{A})^T = \mathbf{A}\mathbf{R}_\kappa^s(\mathbf{A})^T \tag{3.36}$$

with the source autocorrelation matrix a diagonal matrix, i.e., $\mathbf{R}_\kappa^s = \operatorname{diag}(\mathbf{r}^s[\Omega_\kappa])$.

Derivation of (3.7) The sensor correlation matrix $\mathbf{C}_i^x$ is build from the concatenation of sensor correlation vectors with the following structure:

$$\mathbb{E}\{\mathbf{x}[n]x_i[n-k]\} = \begin{bmatrix} r_{1i}^x[\Omega_\kappa] \\ \vdots \\ r_{Di}^x[\Omega_\kappa] \end{bmatrix}. \tag{3.37}$$

By substituting the observations $x_i[n]$ with the model in (3.1) and immediately taking into account that the noise terms do not contribute to the sensor correlation data for time-lag pairs taken from the NF-ROS, the following structure in the sensor correlation vectors is obtained:

$$\mathbb{E}\{\mathbf{x}[n]x_i[n-k]\} = \mathbb{E}\{\mathbf{A}\mathbf{s}[n]\sum_{j=1}^{S} a_i^j s_j[n-k]\} = \mathbf{A}\sum_{j=1}^{S} a_i^j \mathbb{E}\{\mathbf{s}[n]s_j[n-k]\} \tag{3.38}$$

where $(n,k)_\kappa \in \Omega$.

Due to the assumption that the source signals are mutually uncorrelated, the vector $\mathbb{E}\{\mathbf{s}[n]s_j[n-k]\}$ is only non-zero for the j'th row. Consequently, the sensor correlation vectors can be represented in the following way:

$$\mathbb{E}\{\mathbf{x}[n]x_i[n-k]\} = \mathbf{A}\begin{bmatrix} a_i^1 r_{11}^s[\Omega_\kappa] \\ \vdots \\ a_i^S r_{SS}^s[\Omega_\kappa] \end{bmatrix} = \mathbf{A}\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i\right)\mathbf{r}^s[\Omega_\kappa] \tag{3.39}$$

where $\tilde{\mathbf{a}}_i$ is the i'th row of the mixing matrix $\mathbf{A}$. Concatenating such vectors for all time-lag pairs in the NF-ROS leads to the matrix $\mathbf{C}_i^x$ with the following structure:

$$\mathbf{C}_i^x = \mathbf{A}\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i\right)\mathbf{C}^s \tag{3.40}$$

where $\mathbf{C}^s = \begin{bmatrix}\mathbf{r}^s[\Omega_1] & \cdots & \mathbf{r}^s[\Omega_K]\end{bmatrix}$.

Derivation of (3.16) The structure in linear combinations of sensor correlation matrices for the two different forms of sensor correlation matrices can be formulated as follows:

$$\mathbf{\Gamma}_l = \sum_{\kappa=1}^{K} \xi_l^\kappa \bar{\mathbf{A}}\,\mathrm{diag}\left(\mathbf{r}^s[\Omega_\kappa]\right)\left(\bar{\mathbf{A}}\right)^T = \bar{\mathbf{A}}\left(\sum_{\kappa=1}^{K}\xi_l^\kappa\,\mathrm{diag}\left(\mathbf{r}^s[\Omega_\kappa]\right)\right)\left(\bar{\mathbf{A}}\right)^T \triangleq \bar{\mathbf{A}}\mathbf{\Lambda}_l\left(\bar{\mathbf{A}}\right)^T$$

and

$$\mathbf{\Gamma}_l = \sum_{i=1}^{D} \xi_i^l \bar{\mathbf{A}}\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i\right)\bar{\mathbf{C}}^s = \bar{\mathbf{A}}\left(\sum_{i=1}^{D}\xi_i^l\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i\right)\right)\bar{\mathbf{C}}^s \triangleq \bar{\mathbf{A}}\mathbf{\Lambda}_l\bar{\mathbf{C}}^s.$$

The matrices $\mathbf{\Lambda}_l$ are diagonal matrices, with the j'th element having the following structure for the two different sensor correlation matrix structures:

$$\sum_{\kappa=1}^{K}\xi_l^\kappa r_{jj}^s[\Omega_\kappa] = \langle\tilde{\boldsymbol{\xi}}_l, \tilde{\mathbf{r}}_{jj}^s\rangle \quad \text{and} \quad \sum_{i=1}^{D}\xi_i^l a_i^j = \langle\boldsymbol{\xi}^l, \mathbf{a}^j\rangle \tag{3.41}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidian inner product and $\tilde{\mathbf{r}}_{jj}^s$ is the j'th row from the source autocorrelation matrix $\mathbf{C}^s$, i.e., the autocorrelation function values in the NF-ROS of the j'th source.

Computing for two arbitrary linear combinations of sensor correlation matrices $\mathbf{\Gamma}_{l_1}$ and $\mathbf{\Gamma}_{l_2}$ the product $\mathbf{\Gamma}_{l_2}\left(\mathbf{\Gamma}_{l_1}\right)^{-1}$ leads to the following:

$$\mathbf{\Gamma}_{l_2}\left(\mathbf{\Gamma}_{l_1}\right)^{-1} = \bar{\mathbf{A}}\mathbf{\Lambda}_{l_2}\left(\bar{\mathbf{A}}\right)^{T}\left(\bar{\mathbf{A}}\mathbf{\Lambda}_{l_1}(\bar{\mathbf{A}})^{T}\right)^{-1} = \bar{\mathbf{A}}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\left(\bar{\mathbf{A}}\right)^{-1} \tag{3.42}$$

and

$$\mathbf{\Gamma}_{l_2}\left(\mathbf{\Gamma}_{l_1}\right)^{-1} = \bar{\mathbf{A}}\mathbf{\Lambda}_{l_2}\bar{\mathbf{C}}^{s}\left(\bar{\mathbf{A}}\mathbf{\Lambda}_{l_1}\bar{\mathbf{C}}^{s}\right)^{-1} = \bar{\mathbf{A}}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\left(\bar{\mathbf{A}}\right)^{-1}. \tag{3.43}$$

Note that the products of diagonal matrices $\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}$ form again diagonal matrices, such that (3.42) and (3.43) form eigenvalue decompositions. Finally, the eigenvalues, i.e., the diagonal elements of the matrix products $\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}$, have the following structure:

$$\frac{\langle\tilde{\boldsymbol{\xi}}_{l_2}, \tilde{\mathbf{r}}^{s}_{jj}\rangle}{\langle\tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\mathbf{r}}^{s}_{jj}\rangle} \quad \text{and} \quad \frac{\langle\boldsymbol{\xi}^{l_2}, \mathbf{a}^{j}\rangle}{\langle\boldsymbol{\xi}^{l_1}, \mathbf{a}^{j}\rangle}. \tag{3.44}$$

Derivation of (3.19) From the derivation of (3.16) we know the structure of the eigenvalue decomposition of a matrix product $\mathbf{\Gamma}_{l_2}\left(\mathbf{\Gamma}_{l_1}\right)^{-1}$. Furthermore, we exploit the property that multiplication of two matrices that have the same eigenvectors leads to a matrix with an eigenstructure where the eigenvectors are the same as the eigenvectors of the original two matrices and the eigenvalues are the products of the eigenvalues of the two matrices. More specifically, if a matrix is multiplied by itself a new matrix is obtained with the same eigenvectors and squared eigenvalues. This leads to the following result:

$$\mathbf{\Gamma}_{l_2}\left(\mathbf{\Gamma}_{l_1}\right)^{-1}\mathbf{\Gamma}_{l_2}\left(\mathbf{\Gamma}_{l_1}\right)^{-1} = \bar{\mathbf{A}}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\underbrace{\left(\bar{\mathbf{A}}\right)^{-1}\bar{\mathbf{A}}}_{=\mathbf{I}}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\left(\bar{\mathbf{A}}\right)^{-1} \tag{3.45}$$

$$= \bar{\mathbf{A}}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\left(\bar{\mathbf{A}}\right)^{-1} \tag{3.46}$$

where $\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}\mathbf{\Lambda}_{l_2}\left(\mathbf{\Lambda}_{l_1}\right)^{-1}$ are diagonal matrices with the following eigenvalues on the diagonals:

$$\left|\frac{\langle\tilde{\boldsymbol{\xi}}_{l_2}, \tilde{\mathbf{r}}^{s}_{jj}\rangle}{\langle\tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\mathbf{r}}^{s}_{jj}\rangle}\right|^{2} \quad \text{and} \quad \left|\frac{\langle\boldsymbol{\xi}^{l_2}, \mathbf{a}^{j}\rangle}{\langle\boldsymbol{\xi}^{l_1}, \mathbf{a}^{j}\rangle}\right|^{2}. \tag{3.47}$$

Finally, summing two matrices that have the same eigenvectors leads to a summation of eigenvalues, i.e.,

$$\begin{aligned}\mathbf{U}\mathbf{\Lambda}_1\left(\mathbf{U}\right)^{-1} + \mathbf{U}\mathbf{\Lambda}_2\left(\mathbf{U}\right)^{-1} \\ = \sum_j \lambda_1^j \mathbf{u}^j\tilde{\mathbf{v}}^j + \sum_j \lambda_2^j \mathbf{u}^j\tilde{\mathbf{v}}^j = \sum_j (\lambda_1^j + \lambda_2^j)\mathbf{u}^j\tilde{\mathbf{v}}^j \\ = \mathbf{U}(\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2)\left(\mathbf{U}\right)^{-1}\end{aligned}$$

where we have taken λ_i^j to be the elements on the diagonal of $\mathbf{\Lambda}_i$ and we used the substitution $\mathbf{V} = (\mathbf{U})^{-1}$.

Consequently, if we choose $l_1 = 1$ and for $l_2 = l$ with $2 \leq l \leq S$, we can formulate the following sum of matrices:

$$\mathbf{M} = \sum_{l=2}^{S} \mathbf{\Gamma}_l \left(\mathbf{\Gamma}_1\right)^{-1} \mathbf{\Gamma}_l \left(\mathbf{\Gamma}_1\right)^{-1}. \tag{3.48}$$

Applying the same sum to the corresponding eigenvalues, leads to the following eigenvalues for the two different types of sensor correlation matrices:

$$\sum_{l=2}^{S} \left| \frac{\langle \tilde{\boldsymbol{\xi}}_l, \tilde{\mathbf{r}}_{jj}^s \rangle}{\langle \tilde{\boldsymbol{\xi}}_1, \tilde{\mathbf{r}}_{jj}^s \rangle} \right|^2 \quad \text{and} \quad \sum_{l=2}^{S} \left| \frac{\langle \boldsymbol{\xi}^l, \mathbf{a}^j \rangle}{\langle \boldsymbol{\xi}^1, \mathbf{a}^j \rangle} \right|^2. \tag{3.49}$$

4

Design of signal extraction algorithms based on second order statistics exploiting beamforming techniques

The work in this chapter is published as a journal article [61]; however, we made some modifications and additions to improve the flow and consistency in this thesis.

In this chapter we extend the work from Chapter 3 and present a design strategy for efficient signal extraction algorithms that work for instantaneous mixtures of complex-valued source signals. This extension is not straightforward and it requires a different combination of sensor correlation matrices in order to guarantee immediate extraction of the desired source. Furthermore, for a parameterized mixing system new tools for the design and evaluation of signal extraction algorithms have been developed. These tools are used to ensure immediate extraction of the desired signal by exploiting knowledge on physical parameters.

First we show how to combine sensor correlation matrices for complex-valued mixtures in order to incorporate a priori information about the autocorrelation function or the mixing parameters of the desired source. Subsequently, we present design tools in the form of selection beamformers and selection patterns that can be used to design informed source extraction algorithms for a parameterized mixing model based on direction of arrival information. Additionally, we present an algorithm to choose between two objective functions for the source extraction filters. Finally, the design procedure and the properties of the extraction algorithms are evaluated via examples and experiments.

4.1 Introduction

Traditionally, in the fields of acoustics and telecommunication mainly beamformers were used to extract the desired source signal from multiple observations. These spatial filters exploit a parameterized modeling of a designed sensor array and its environment [5]. Nowadays blind signal processing (BSP) techniques are more and more applied in the fields of acoustics, telecommunication, and biomedical engineering [8,11]. A BSP algorithm separates or extracts source signals without using information about the mixing system. Each of these methods, i.e., beamforming and BSP, has their own imperfections when applied for signal extraction. Beamformers heavily depend on the array configuration at hand and are sensitive to modeling errors and false prior information. Additionally, with the rise of wireless sensor networks [3], modeling of the entire sensor array becomes practically impossible. Contrarily, BSP algorithms suffer from a scaling and permutation indeterminacy; therefore, they require a classifier that selects the desired signal.

Despite the research conducted in the last decades [8, 10, 11], BSS algorithms yield insufficient quality for many applications [9, 62]. The lack of performance of BSS algorithms has lead to a new research field called informed source separation (ISS) [62, 63]. The objective of ISS is to build source separation algorithms that exploit all relevant prior information such that the separation performance increases. Furthermore, it is unrealistic to develop a single algorithm that fits any scenario. "Instead, we must focus our efforts on developing a methodology for designing robust algorithms that are specific to the application at hand." [63]. In this context, we develop a methodology that can be used to design signal extraction algorithms based on available prior information. The main considerations for the development of such a framework can be brought back into three elements. The framework must be *flexible* in the use of available prior information; moreover, the algorithms derived from the framework have to be *robust* and *efficient*.

In Chapter 3 we presented a single-stage signal extraction method that guarantees immediate extraction of the desired signal if certain conditions hold on a mold, which consists of some amount of prior information. For robustness and efficiency purposes, the method calculates extraction filters exploiting the second order temporal structure in the data instead of higher order statistics.

In the current chapter we expand the work from the previous chapter and our goal is to provide insight into the design of signal extraction algorithms that exploit all available prior information. The main contributions in this chapter are that we convert the method from Chapter 3, based on a mold, such that it works for both complex and real-valued mixtures of signals. We also present a method to choose between the two types of source extraction filters that can be identified with the presented algorithm. Furthermore, we introduce new design and evaluation tools for developing signal extraction algorithms and we present example designs based on different types and amounts of a priori information about physical parameters. Finally, we show that the presented methodology has strong advantages over beamforming.

Part of this work has already been presented in [64], [65], and [66]. It was shown in

[64] that extraction filters with different objectives on noise and interference reduction can be obtained with the presented method. In [65], a signal extraction algorithm is applied to a combined fixed and wireless sensor network and an example signal extraction algorithm based on prior information about autocorrelation functions of sources is presented in [66].

The outline of this chapter is as follows. In Section 4.2 we introduce the mixing model and assumptions on the signals. In Section 4.3 we present the signal extraction algorithms for real and complex mixtures based on a mold that contains some amount of prior information. Furthermore, we present the method to select between the two extraction filters. The new design tools and example algorithm designs are derived in Section 4.4 and experimental results are presented in Section 4.5. Finally, a discussion and conclusions are given in Section 4.6.

Notation of vectors and matrices is with bold face, lowercase and bold face, uppercase letters, respectively. Subscript and superscript indices are used to denote respectively row and column elements of vectors and matrices. Row vectors carry an additional tilde symbol $\sim$ in order to distinguish between column and row vectors. For example, $\mathbf{a}^j$ is the j'th column vector and $\tilde{\mathbf{a}}_i$ is the i'th row vector of the matrix $\mathbf{A}$. Throughout this paper we reserve the index symbols i and j for components related to the sensors and sources, respectively. Index sets are denoted by calligraphic symbols such as $\mathcal{S} = [1, S]$, which contains all integers from 1 until S. Square brackets are used for denoting the time index n. Conjugate, transpose, and conjugate transpose are denoted by the superscript symbols $*$, T, and H, respectively, e.g., $\mathbf{a}^*$, $(\mathbf{a})^T$, and $(\mathbf{a})^H$. The Euclidean inner products for column and row vectors are defined as $\langle \mathbf{a}, \mathbf{b} \rangle \triangleq (\mathbf{a})^H \mathbf{b}$ and $\langle \tilde{\mathbf{a}}, \tilde{\mathbf{b}} \rangle \triangleq \tilde{\mathbf{a}}^* (\tilde{\mathbf{b}})^T$, respectively. Finally, the Euclidean norm for column and row vectors is defined as $\|\mathbf{a}\|_2 \triangleq \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$ and $\|\tilde{\mathbf{a}}\|_2 \triangleq \sqrt{\langle \tilde{\mathbf{a}}, \tilde{\mathbf{a}} \rangle}$, respectively.

4.2 Model and assumptions

Before the design strategy is introduced we discuss the mixing model and assumptions on the source signals.

4.2.1 Introduction of the mixing model

The observed signals are assumed to be complex instantaneous mixtures of multiple source signals. These mixtures can either be natural instantaneous mixtures, such as in biomedical or narrowband telecommunication applications, or they can be obtained from a DFT filterbank or short time Fourier transform (STFT) such as in speech enhancement applications.

The complex instantaneous mixing system is modeled by a complex matrix $\mathbf{A}$ with D rows and S columns, corresponding to the number of sensors and sources,

respectively, i.e.,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}^1 & \cdots & \mathbf{a}^S \end{bmatrix} = \begin{bmatrix} a_1^1 & \cdots & a_1^S \\ \vdots & \ddots & \vdots \\ a_D^1 & \cdots & a_D^S \end{bmatrix}. \tag{4.1}$$

The structure in the sensor signals for complex instantaneous mixtures is assumed as follows:

$$\mathbf{x}[n] \equiv \mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n] = \sum_{j=1}^{S} \mathbf{a}^j s_j[n] + \boldsymbol{\nu}[n] \tag{4.2}$$

where the complex-valued column vectors $\mathbf{x}[n]$, $\mathbf{s}[n]$, and $\boldsymbol{\nu}[n]$ represent the complex sensor, source, and noise signals at time instance $n \in \mathbb{Z}$, respectively, i.e.,

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x_1[n] \\ \vdots \\ x_D[n] \end{bmatrix}, \quad \mathbf{s}[n] \triangleq \begin{bmatrix} s_1[n] \\ \vdots \\ s_S[n] \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\nu}[n] \triangleq \begin{bmatrix} \nu_1[n] \\ \vdots \\ \nu_D[n] \end{bmatrix}.$$

The extraction filter is represented by a complex-valued row vector $\tilde{\mathbf{w}}$ of length D and produces a complex output signal $y[n]$, i.e.,

$$y[n] = \sum_{i=1}^{D} w^i x_i[n] \equiv \tilde{\mathbf{w}}\mathbf{A}\mathbf{s}[n] + \tilde{\mathbf{w}}\boldsymbol{\nu}[n]. \tag{4.3}$$

Extraction of the desired source is performed if the extraction filter $\tilde{\mathbf{w}}$ enhances the desired source. We show that different types of extraction filters can be obtained with the presented signal extraction algorithms if the desired source is *extractable* from the mixture and certain conditions on the second order statistics of the source and noise signals hold.

Definition 4.2.1 (Extractability)**:** The j'th source signal is called extractable if there exists an extraction filter $\tilde{\mathbf{w}}$ such that $\tilde{\mathbf{w}}\mathbf{A} = \tilde{\mathbf{e}}^j$, where $\tilde{\mathbf{e}}^j$ is a row vector which is only non-zero at the j'th column.

Theorem 4.2.1: If the j'th source is extractable, then its corresponding mixing column vector is linearly independent with respect to the remaining mixing column vectors. However, these remaining mixing column vectors may be linearly dependent among each other.

Proof. Suppose that the d'th mixing column is depending linearly on the remaining columns, i.e., $\mathbf{a}^d = \sum_j \xi_j \mathbf{a}^j$ for $j \in \mathcal{S}\backslash d$, with the index set $\mathcal{S} = [1, S]$. If the d'th source is extractable, then there exists a filter $\tilde{\mathbf{w}}$ such that $\tilde{\mathbf{w}}\mathbf{a}^j = 0 \; \forall \, j \in \mathcal{S}\backslash d$ and $\tilde{\mathbf{w}}\mathbf{a}^d = 1$. Combining these two results we have the following contradiction:

$$\tilde{\mathbf{w}}\mathbf{a}^d = \tilde{\mathbf{w}} \sum_{j \in \mathcal{S}\backslash d} \xi_j \mathbf{a}^j = \sum_{j \in \mathcal{S}\backslash d} \xi_j \tilde{\mathbf{w}}\mathbf{a}^j = \sum_{j \in \mathcal{S}\backslash d} \xi_j \cdot 0 = 0 \neq 1 \tag{4.4}$$

which shows that mixing column $\mathbf{a}^d$ must be linearly independent if the corresponding source is extractable. □

In order to ensure immediate extraction of the desired source in Section 4.3, we require that all sources are extractable, i.e., the mixing system must contain only linearly independent columns.

4.2.2 Assumptions on the second order statistics of the source and noise signals

The source and noise signals are characterized by their second order statistics (SOS). We first introduce the definition of the auto- and crosscorrelation function. Subsequently we give the definition of a noise-free region of support (NF-ROS) in order to deal with noise. Finally, we discuss the required *identifiability* conditions.

Definition 4.2.2 (Correlation functions)**:** The auto- and crosscorrelation functions for two signals $p_i[n]$ and $q_j[n]$ at time n and lag k are defined as follows:

$$r_{ij}^{pq}[n,k] \triangleq \mathbb{E}\{p_i[n]q_j^*[n-k]\} \tag{4.5}$$

where $\mathbb{E}$ is the mathematical expectation operator. If the two signals are the same then we call the function an autocorrelation function and use the following short hand notation: $r_{ii}^{p}[n,k] \triangleq r_{ii}^{pp}[n,k]$.

Both the time index n and lag index k are incorporated in the definition of the correlation functions. In this way we are able to exploit the non-stationarity of signals. Furthermore, we assume that the mathematical expectation operator can be approximated by an ensemble average of the product $p_i[n]q_j[n-k]$ near the considered time instance n, which is a widely used assumption.

Using Definition 4.2.2 for the sensor, source, and noise signals, leads to the following sets of auto- and crosscorrelation functions, respectively:

$$r_{i_1 i_2}^{x}[n,k] \triangleq \mathbb{E}\{x_{i_1}[n]x_{i_2}^*[n-k]\} \quad \forall\,(i_1,i_2) \in \mathcal{D}\times\mathcal{D} \tag{4.6}$$

$$r_{i_1 i_2}^{s}[n,k] \triangleq \mathbb{E}\{s_{j_1}[n]s_{j_2}^*[n-k]\} \quad \forall\,(j_1,j_2) \in \mathcal{S}\times\mathcal{S} \tag{4.7}$$

$$r_{i_1 i_2}^{\nu}[n,k] \triangleq \mathbb{E}\{\nu_{i_1}[n]\nu_{i_2}^*[n-k]\} \quad \forall\,(i_1,i_2) \in \mathcal{D}\times\mathcal{D} \tag{4.8}$$

where $\mathcal{D} = [1, D]$ and $\mathcal{S} = [1, S]$ are index sets for the sensor and source elements, respectively.

We also define the following crosscorrelation functions between source and noise signals:

$$r_{ij}^{\nu s}[n,k] \triangleq \mathbb{E}\{\nu_i[n]s_j^*[n-k]\} \quad \forall\,(i,j) \in \mathcal{D}\times\mathcal{S} \tag{4.9}$$

$$r_{ji}^{s\nu}[n,k] \triangleq \mathbb{E}\{s_j[n]\nu_i^*[n-k]\} \quad \forall\,(i,j) \in \mathcal{D}\times\mathcal{S}. \tag{4.10}$$

Using these auto- and crosscorrelation functions, a noise-free region of support (NF-ROS) is defined such that we can deal with noise in a systematic way.

Definition 4.2.3 (Noise-free region of support)**:** A noise-free region of support (NF-ROS), also denoted by Ω, consists of a set of time-lag pairs (n, k) for which the noise signals are not correlated, i.e.,

$$\forall\, \Omega_\kappa \in \Omega : \begin{cases} r^{\nu}_{i_1 i_2}[\Omega_\kappa] = 0 & \forall\, (i_1, i_2) \in \mathcal{D} \times \mathcal{D} \\ r^{\nu s}_{ij}[\Omega_\kappa] = 0 & \forall\, (i, j) \in \mathcal{D} \times \mathcal{S} \\ r^{s\nu}_{ji}[\Omega_\kappa] = 0 & \forall\, (i, j) \in \mathcal{D} \times \mathcal{S} \end{cases} \tag{4.11}$$

where $\Omega \triangleq \{(n, k)_\kappa \,|\, \kappa \in \mathcal{K}\}$ and $\mathcal{K} = [1, K]$ is the index set for time-lag pairs in the NF-ROS.

Within the NF-ROS, we require two additional conditions on the SOS of the source signals in order to be able to identify extraction filters. First, we require that the source signals are mutually uncorrelated, i.e.,

$$r^{s}_{j_1 j_2}[\Omega_\kappa] = 0 \quad \forall\, \{(j_1, j_2, \kappa) \in \mathcal{S} \times \mathcal{S} \times \mathcal{K} \,|\, j_1 \neq j_2\}. \tag{4.12}$$

Second, we require that the source autocorrelation functions are linearly independent in the NF-ROS [36], i.e.,

$$\sum_{j=1}^{S} \xi^j r^{s}_{jj}[\Omega_\kappa] = 0 \quad \forall\, \kappa \in \mathcal{K} \quad \Longleftrightarrow \quad \xi^j = 0 \quad \forall\, j \in \mathcal{S}. \tag{4.13}$$

Notice that the definition of a NF-ROS and the additional assumptions within the NF-ROS are sufficient, but not necessary conditions for blind identification of the extraction filter. For example, alternative techniques exist that exploit assumptions on higher order statistics.

If the presented assumptions hold, then it is possible to extract sources from observed mixtures. However, additional prior information is required such that immediate extraction of the desired source can be guaranteed. Types of and conditions on prior information are discussed in the following section, where we introduce extraction algorithms that exploit this prior information in the form of a mold.

4.3 Signal extraction based on a mold

We present two signal extraction algorithms that require a small amount of a priori information in the form of a mold. An overview of the algorithms is depicted in Figure 4.1. First we discuss the structure in the sensor correlation data. Subsequently, we identify the desired extraction filter from noise-free sensor correlation matrices based on the mold. Finally, we present noise reduction techniques.

4.3.1 Structure in the second order statistics of sensor signals

Identification of the desired extraction filter is accomplished by exploiting the structure in the sensor correlation functions. Since the source signals are mutually uncorrelated, see (4.12), the sensor correlation functions depend only on the mixing

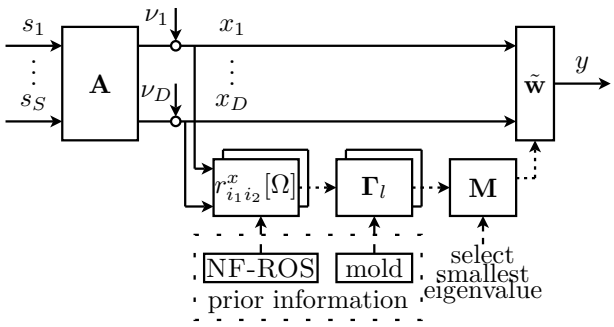


Figure 4.1: An overview of signal extraction algorithms that exploits a small amount of prior information in the form of a mold. Based on the noise-free sensor correlation data $r^x_{i_1 i_2}[\Omega]$ and the mold we take linear combinations of correlation matrices. From these linear combinations $\mathbf{\Gamma}_l$ we construct a matrix $\mathbf{M}$ whose eigenvector that corresponds to the smallest eigenvalue is the desired extraction filter $\tilde{\mathbf{w}}$.

parameters and the source autocorrelation functions, i.e.,

$$r^x_{i_1 i_2}[\Omega_\kappa] = \sum_{j=1}^{S} a^j_{i_1} a^{j,*}_{i_2} r^s_{jj}[\Omega_\kappa] \quad \forall\,(i_1, i_2, \kappa) \in \mathcal{D} \times \mathcal{D} \times \mathcal{K}. \tag{4.14}$$

Evaluating (4.14) over all indices i_1, i_2, and κ, leads to a cube of data of size $D \times D \times K$.

If we consider the slices of the cube of sensor correlation data in matrix form, i.e., sets of sensor correlation matrices, then we distinguish the following three unique structures of size $D \times D$ and $D \times K$, and $D \times K$, respectively:

$$\mathbf{R}^x_\kappa \triangleq \begin{bmatrix} r^x_{11}[\Omega_\kappa] & r^x_{12}[\Omega_\kappa] & \cdots & r^x_{1D}[\Omega_\kappa] \\ r^x_{21}[\Omega_\kappa] & r^x_{22}[\Omega_\kappa] & \cdots & r^x_{2D}[\Omega_\kappa] \\ \vdots & \ddots & \ddots & \vdots \\ r^x_{D1}[\Omega_\kappa] & r^x_{D2}[\Omega_\kappa] & \cdots & r^x_{DD}[\Omega_\kappa] \end{bmatrix} \quad \forall\,\kappa \in \mathcal{K} \tag{4.15}$$

$$\mathbf{C}^x_{\bullet i} \triangleq \begin{bmatrix} r^x_{1i}[\Omega_1] & r^x_{1i}[\Omega_2] & \cdots & r^x_{1i}[\Omega_K] \\ r^x_{2i}[\Omega_1] & r^x_{2i}[\Omega_2] & \cdots & r^x_{2i}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r^x_{Di}[\Omega_1] & r^x_{Di}[\Omega_2] & \cdots & r^x_{Di}[\Omega_K] \end{bmatrix} \quad \forall\, i \in \mathcal{D} \tag{4.16}$$

$$\mathbf{C}^x_{i\bullet} \triangleq \begin{bmatrix} r^x_{i1}[\Omega_1] & r^x_{i1}[\Omega_2] & \cdots & r^x_{i1}[\Omega_K] \\ r^x_{i2}[\Omega_1] & r^x_{i2}[\Omega_2] & \cdots & r^x_{i2}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r^x_{iD}[\Omega_1] & r^x_{iD}[\Omega_2] & \cdots & r^x_{iD}[\Omega_K] \end{bmatrix} \quad \forall\, i \in \mathcal{D} \tag{4.17}$$

where $\mathbf{R}^x_\kappa$ is the collection of sensor correlation function values for time-lag pair κ and $\mathbf{C}^x_{\bullet i}$ and $\mathbf{C}^x_{i\bullet}$ are collections of sensor correlation function values for the i'th sensor correlated with all sensors, for all time-lag pairs in the NF-ROS.

The structure in these sensor correlation matrices in terms of the mixing system and the source autocorrelation functions is derived from (4.14) and is as follows (see Appendix 4.B for a derivation):

$$\mathbf{R}^x_\kappa \equiv \mathbf{A}\,\mathrm{diag}\left(\mathbf{r}^s[\Omega_\kappa]\right)(\mathbf{A})^H \quad \forall\,\kappa \in \mathcal{K} \tag{4.18}$$

$$\mathbf{C}^x_{\bullet i} \equiv \mathbf{A}\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i^*\right)\mathbf{C}^s \quad \forall\, i \in \mathcal{D} \tag{4.19}$$

$$\mathbf{C}^x_{i\bullet} \equiv \mathbf{A}^*\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i\right)\mathbf{C}^s \quad \forall\, i \in \mathcal{D} \tag{4.20}$$

where the row vector $\tilde{\mathbf{a}}_i$ is the i'th row of the mixing matrix $\mathbf{A}$, $\mathbf{C}^s$ is an $S \times K$ source autocorrelation matrix with the following structure:

$$\mathbf{C}^s \triangleq \begin{bmatrix} r^s_{11}[\Omega_1] & r^s_{11}[\Omega_2] & \cdots & r^s_{11}[\Omega_K] \\ r^s_{22}[\Omega_1] & r^s_{22}[\Omega_2] & \cdots & r^s_{22}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r^s_{SS}[\Omega_1] & r^s_{SS}[\Omega_2] & \cdots & r^s_{SS}[\Omega_K] \end{bmatrix} \tag{4.21}$$

and the column vector $\mathbf{r}^s[\Omega_\kappa]$ contains the autocorrelation function values for all sources at time-lag pair Ω_κ, i.e., it is the κ'th column of the source autocorrelation matrix $\mathbf{C}^s$. Notice that the matrix $\mathbf{C}^x_{\bullet i}$ is essentially different from $\mathbf{C}^x_{i\bullet}$ for complex data due to the conjugation pattern in (4.19) and (4.20).

The sensor correlation matrices have a size of either $D \times D$ or $D \times K$; however, their individual ranks are at most equal to the number of sources S. In case the number of sensors is larger than the number of sources, i.e., $D > S$, then there exists a subspace matrix $\mathbf{P}_s$ that compresses the mixing matrix $\mathbf{A}$ from size $D \times S$ to size $S \times S$, without losing information. Similarly, if the number of time-lag pairs in the NF-ROS is larger than the number of sources, i.e., $K > S$, then there exists a subspace matrix $\mathbf{Q}^s$ that reduces the source autocorrelation matrix $\mathbf{C}^s$ from size $S \times K$ to size $S \times S$, without losing information. From the structure in the sensor correlation matrices, see (4.18), (4.19), and (4.20), it follows that all sensor correlation matrices can be compressed by applying only two subspace matrices, i.e.,

$$\bar{\mathbf{R}}^x_\kappa = \mathbf{P}_s\mathbf{R}^x_\kappa\left(\mathbf{P}_s\right)^H,\ \bar{\mathbf{C}}^x_{\bullet i} = \mathbf{P}_s\mathbf{C}^x_{\bullet i}\mathbf{Q}^s,\ \text{and}\ \bar{\mathbf{C}}^x_{i\bullet} = \mathbf{P}^*_s\mathbf{C}^x_{i\bullet}\mathbf{Q}^s. \tag{4.22}$$

The method to calculate these subspace matrices from the sensor correlation data is discussed in Appendix 4.A. These reduced sensor correlation matrices and the reduction matrices can be used for identification of the desired extraction filter.

4.3.2 Identification of the desired extraction filter

We identify the desired extraction filter by exploiting the structure in the noise-free sensor correlation matrices. An overview of this procedure is depicted in Figure 4.1. First we show that the eigenvectors of a matrix $\mathbf{M}$, which is built from a specific combination of the sensor correlation matrices, are extraction filters. Subsequently, we exploit the structure in the corresponding eigenvalues to ensure that the desired extraction filter is the eigenvector that corresponds to the smallest eigenvalue. This

is accomplished by incorporating some amount of a priori information in the form of a *mold* into the algorithm.

The matrix $\mathbf{M}$ is built from a specific combination of linear combinations of sensor correlation matrices. Depending on the type of a priori information, either the matrices $\bar{\mathbf{R}}^x_\kappa$ or $\{\bar{\mathbf{C}}^x_{\bullet i}, \bar{\mathbf{C}}^x_{i\bullet}\}$ are used. The linear combinations of these sensor correlation matrices are defined as follows [36]:

$$\mathbf{\Gamma}_l \triangleq \sum_{\kappa=1}^{K} \xi_l^{\kappa,*} \bar{\mathbf{R}}^x_\kappa \quad \text{and} \quad \breve{\mathbf{\Gamma}}_l \triangleq \sum_{\kappa=1}^{K} \xi_l^{\kappa} \left(\bar{\mathbf{R}}^x_\kappa\right)^H \tag{4.23}$$

or

$$\mathbf{\Gamma}_l \triangleq \sum_{i=1}^{D} \xi_i^{l,*} \bar{\mathbf{C}}^x_{i\bullet} \quad \text{and} \quad \breve{\mathbf{\Gamma}}_l \triangleq \sum_{i=1}^{D} \xi_i^{l} \bar{\mathbf{C}}^x_{\bullet i} \tag{4.24}$$

where the complex weights ξ_l^κ and ξ_i^l are elements of the weight vectors $\tilde{\boldsymbol{\xi}}_l$ and $\boldsymbol{\xi}^l$, respectively, for $(\kappa, l) \in \mathcal{K} \times \mathcal{L}$ and $(i, l) \in \mathcal{D} \times \mathcal{L}$ with $\mathcal{L} = [1, L]$. These L weight vectors $\tilde{\boldsymbol{\xi}}_l$ or $\boldsymbol{\xi}^l$ are free parameters that can be chosen or designed, based on available a priori information. In this section we discuss how to choose these free parameters based on a mold, such that immediate extraction of the desired source is guaranteed. In Section 4.4.3 we discuss in detail the design of these vectors based on a parametrization of the mixing system.

The matrix $\mathbf{M}$ is formed from the linear combinations $\mathbf{\Gamma}_l$ and $\breve{\mathbf{\Gamma}}_l$ of sensor correlation matrices and has the following structure:

$$\mathbf{M} = \sum_{l=2}^{L} \breve{\mathbf{\Gamma}}_l \left(\mathbf{\Gamma}_1\right)^{-1} \mathbf{\Gamma}_l \left(\breve{\mathbf{\Gamma}}_1\right)^{-1} \equiv \bar{\mathbf{A}} \mathbf{\Lambda} \left(\bar{\mathbf{A}}\right)^{-1} \tag{4.25}$$

where $\bar{\mathbf{A}} = \mathbf{P}_s \mathbf{A}$ is the size-reduced mixing matrix and the diagonal matrix $\mathbf{\Lambda}$ contains the following elements (see Appendix 4.B for a derivation):

$$\lambda^j = \sum_{l=2}^{K} \frac{\left|\left\langle \tilde{\boldsymbol{\xi}}_l, \tilde{\mathbf{r}}^s_{jj} \right\rangle\right|^2}{\left|\left\langle \tilde{\boldsymbol{\xi}}_1, \tilde{\mathbf{r}}^s_{jj} \right\rangle\right|^2} \quad \text{or} \quad \lambda^j = \sum_{l=2}^{D} \frac{\left|\left\langle \boldsymbol{\xi}^l, \mathbf{a}^j \right\rangle\right|^2}{\left|\left\langle \boldsymbol{\xi}^1, \mathbf{a}^j \right\rangle\right|^2} \tag{4.26}$$

for linear combinations of the matrices $\bar{\mathbf{R}}^x_\kappa$ or $\{\bar{\mathbf{C}}^x_{\bullet i}, \bar{\mathbf{C}}^x_{i\bullet}\}$, respectively, with the row vector $\tilde{\mathbf{r}}^s_{jj}$ containing the autocorrelation function values of source j for all lags in the NF-ROS, i.e., the j'th row of the matrix in (4.21).

The decomposition of the matrix $\mathbf{M}$ in (4.25) is an eigenvalue decomposition, where the eigenvectors are proportional to the size-reduced mixing columns $\bar{\mathbf{a}}^j = \mathbf{P}_s \mathbf{A}$ and the eigenvalues are given by λ^j from (4.26), i.e., the following equation holds:

$$\mathbf{M} \bar{\mathbf{a}}^j = \lambda^j \bar{\mathbf{a}}^j \quad \forall\, j \in \mathcal{S}. \tag{4.27}$$

The eigenvalue decomposition in (4.27) can be used to calculate the mixing column vectors $\mathbf{a}^j = \left(\mathbf{P}_s\right)^\dagger \bar{\mathbf{a}}^j$, up to an arbitrary scaling. These mixing column vectors can

be used to form extraction filters, as we will show in Section 4.3.3. However, they do not fully exploit the extractability of the sources as defined in Definition 4.2.1, i.e., the interfering sources are not fully suppressed. In fact, we are interested in an extraction filter that is orthogonal to the mixing columns of the interfering sources. This filter is obtained by exploiting the inverse of the mixing system, i.e., for the j'th row vector $\tilde{\boldsymbol{\mu}}_j$ from the inverse of the mixing matrix it holds that

$$\tilde{\boldsymbol{\mu}}_{j_1}\bar{\mathbf{a}}^{j_2} = \begin{cases} 1 & \text{if} \quad j_1 = j_2 \\ 0 & \text{if} \quad j_1 \neq j_2 \end{cases} \qquad \forall\,(j_1, j_2) \in \mathcal{S} \times \mathcal{S}. \tag{4.28}$$

One way to obtain these extraction filters $\tilde{\boldsymbol{\mu}}_j$ is by taking the inverse of the matrix that contains all the eigenvectors calculated with (4.27). However, a more efficient way to obtain these extraction filters is by calculating the left eigenvectors of the matrix $\mathbf{M}$, i.e.,

$$\tilde{\boldsymbol{\mu}}_j \mathbf{M} = \tilde{\boldsymbol{\mu}}_j \lambda^j \quad \forall\, j \in \mathcal{S} \tag{4.29}$$

where $\tilde{\boldsymbol{\mu}}_j$ are proportional to rows from the inverse of the size-reduced mixing matrix $\bar{\mathbf{A}}$ and λ^j are again the eigenvalues as in (4.26).

Now that we are able to calculate extraction filters, we still have to identify the filter that extracts the desired source. A remarkable property of the eigenvalues is that they depend on information of only the source that is extracted by its corresponding eigenvectors, as follows from (4.26). Therefore, by choosing or designing the weights ξ_l^κ or ξ_i^l in (4.23) and (4.24) based on available a priori information about the desired source, we are able to characterize the eigenvalue that corresponds to the desired source.

The a priori information exploited in this section is represented by a mold. The mold consists of either a guess of the autocorrelation function of the desired source, represented by the row vector $\tilde{\mathbf{r}}_0$; or, a guess of the mixing column of the desired source, represented by the column vector $\mathbf{a}^0$.

If the mold $\tilde{\mathbf{r}}_0$ is available, then linear combinations of the sensor correlation matrices $\bar{\mathbf{R}}_\kappa^x$ are used, while for the mold $\mathbf{a}^0$ the sensor correlation matrices $\{\bar{\mathbf{C}}_{\bullet i}^x, \bar{\mathbf{C}}_{i\bullet}^x\}$ are taken.

Theorem 4.3.1: If all sources are extractable, i.e., the mixing system has full rank, and the following condition holds for the mold that is available:

$$\frac{|\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}_{dd}^s \rangle|}{\|\tilde{\mathbf{r}}_{dd}^s\|_2} > \frac{|\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}_{jj}^s \rangle|}{\|\tilde{\mathbf{r}}_{jj}^s\|_2} \quad \forall\, j \in \mathcal{S}\backslash d \tag{4.30}$$

respectively

$$\frac{|\langle \mathbf{a}^0, \mathbf{a}^d \rangle|}{\|\mathbf{a}^d\|_2} > \frac{|\langle \mathbf{a}^0, \mathbf{a}^j \rangle|}{\|\mathbf{a}^j\|_2} \quad \forall\, j \in \mathcal{S}\backslash d \tag{4.31}$$

where d is the index for the desired source, then the smallest eigenvalue of the matrix $\mathbf{M}$ is guaranteed to correspond to the desired source if the vectors $\tilde{\boldsymbol{\xi}}_l$ or $\boldsymbol{\xi}^l$ form an

orthonormal basis, i.e.,

$$\left\langle \tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\boldsymbol{\xi}}_{l_2} \right\rangle = \begin{cases} 1 & \text{if} \quad l_1 = l_2 \\ 0 & \text{if} \quad l_1 \neq l_2 \end{cases} \qquad \forall\, (l_1, l_2) \in \mathcal{K} \times \mathcal{K} \tag{4.32}$$

or

$$\left\langle \boldsymbol{\xi}^{l_1}, \boldsymbol{\xi}^{l_2} \right\rangle = \begin{cases} 1 & \text{if} \quad l_1 = l_2 \\ 0 & \text{if} \quad l_1 \neq l_2 \end{cases} \qquad \forall\, (l_1, l_2) \in \mathcal{D} \times \mathcal{D} \tag{4.33}$$

with the first vector being the normalized mold, i.e., $\tilde{\boldsymbol{\xi}}_1 = \tilde{\mathbf{r}}_0 / \|\tilde{\mathbf{r}}_0\|_2$ or $\boldsymbol{\xi}^1 = \mathbf{a}^0 / \|\mathbf{a}^0\|_2$.

Proof. We use the following short hand notation: $b_l^j \triangleq \langle \tilde{\boldsymbol{\xi}}_l, \tilde{\mathbf{r}}_{jj}^s \rangle$ and $b_l^j \triangleq \langle \boldsymbol{\xi}^l, \mathbf{a}^j \rangle$. Then the eigenvalues in (4.26) have the following structure:

$$\lambda^j = \sum_{l=2}^{L} \frac{|b_l^j|^2}{|b_1^j|^2} = \frac{1}{|b_1^j|^2} \sum_{l=2}^{L} |b_l^j|^2 \quad \forall\, j \in \mathcal{S} \tag{4.34}$$

where $L = K$ or $L = D$ for the mold $\tilde{\mathbf{r}}_0$ or $\mathbf{a}^0$, respectively.

Since the vectors $\boldsymbol{\xi}_l$ or $\boldsymbol{\xi}^l$ form an orthonormal basis, we have

$$\sum_{l=2}^{K} |b_l^j|^2 = \left|\left|\tilde{\mathbf{r}}_{jj}^s\right|\right|_2^2 - |b_1^j|^2 \quad \text{or} \quad \sum_{l=2}^{D} |b_l^j|^2 = \left|\left|\mathbf{a}^j\right|\right|_2^2 - |b_1^j|^2 \tag{4.35}$$

such that (4.34) can be written as

$$\lambda^j = \frac{\|\tilde{\mathbf{r}}_{jj}^s\|_2^2 - |b_1^j|^2}{|b_1^j|^2} \quad \text{or} \quad \lambda^j = \frac{\|\mathbf{a}^j\|_2^2 - |b_1^j|^2}{|b_1^j|^2}. \tag{4.36}$$

The condition on the mold in (4.30) or (4.31) translates to respectively the following condition:

$$\frac{|b_l^d|}{\|\tilde{\mathbf{r}}_{dd}^s\|_2} > \frac{|b_l^j|}{\|\tilde{\mathbf{r}}_{jj}^s\|_2} \quad \text{or} \quad \frac{|b_l^d|}{\|\mathbf{a}^j\|_2} > \frac{|b_l^j|}{\|\mathbf{a}^j\|_2} \quad \forall\, j \in \mathcal{S} \backslash d. \tag{4.37}$$

From (4.37) it follows that the nominator of the eigenvalue that corresponds to the desired source in (4.36) has the smallest value, while simultaneously the corresponding denominator has the largest value. Consequently, the smallest eigenvalue of the matrix $\mathbf{M}$ corresponds to the desired source. □

4.3.3 Applying extraction filters and noise reduction

Once the desired eigenvector $\tilde{\boldsymbol{\mu}}_d$ or $\bar{\mathbf{a}}^d$ has been identified, up to the unknown scaling, noise reduction can be applied. The eigenvector $\bar{\mathbf{a}}^d$ from (4.27) can be used

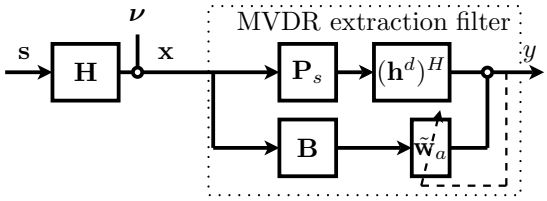


Figure 4.2: Scheme of the MVDR extraction filter, with $\mathbf{B} = \left((\bar{\mathbf{a}}^d)^H \mathbf{P}_s\right)^{\perp}$.

.

together with the subspace matrix $\mathbf{P}_s$ to calculate the following minimum variance distortionless response (MVDR) filter:

$$\begin{aligned} \min_{\tilde{\mathbf{w}}_a} \quad & \mathbb{E}\{y[n]y^*[n]\} \\ \text{s.t.} \quad & y[n] = \left(\left(\bar{\mathbf{a}}^d\right)^H \mathbf{P}_s - \tilde{\mathbf{w}}_a \left(\left(\bar{\mathbf{a}}^d\right)^H \mathbf{P}_s \right)^{\perp} \right) \mathbf{x}[n]. \end{aligned} \tag{4.38}$$

An implementation based on the generalized sidelobe canceller (GSC) structure that uses adaptive filtering to calculate the noise reduction filter $\tilde{\mathbf{w}}_a$ is given in Figure 4.2.

Although the MVDR filter does not take full potential of the extractability of the sources it is capable of balancing between noise reduction and interference suppression, which might be a powerful feature in noisy scenarios.

Using the eigenvector $\tilde{\boldsymbol{\mu}}_d$ and the subspace matrix $\mathbf{P}_s$ leads to the extraction of the desired source, i.e.,

$$y[n] = \tilde{\boldsymbol{\mu}}_d \mathbf{P}_s \mathbf{x}[n]. \tag{4.39}$$

However, noise reduction is only possible if the system is overdetermined, i.e., $D \geq S$. In that case, the $D - S$ dimensional space orthogonal to $\mathbf{P}^s$ can be used for noise reduction by solving the following optimization problem:

$$\begin{aligned} \min_{\tilde{\mathbf{w}}_a} \quad & \mathbb{E}\{y[n]y^*[n]\} \\ \text{s.t.} \quad & y[n] = \left(\tilde{\boldsymbol{\mu}}_d \mathbf{P}_s - \tilde{\mathbf{w}}_a \left(\mathbf{P}_s\right)^{\perp} \right) \mathbf{x}[n]. \end{aligned} \tag{4.40}$$

This filter is called the linearly constrained minimum variance (LCMV) filter from now on. The linear constraints are the zero constraints on the interfering sources and the minimum variance property is obtained by the noise reduction. An implementation based on the GSC structure that uses adaptive filtering to calculate the noise reduction filter $\tilde{\mathbf{w}}_a$ is given in Figure 4.3.

In the following section we present a measure that can be useful for choosing between the LCMV and MVDR objective function.

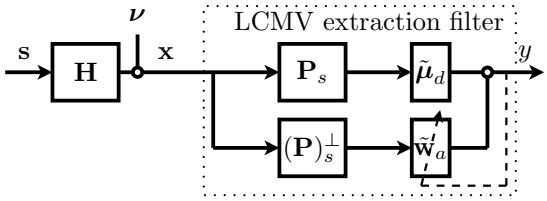


Figure 4.3: Scheme of the LCMV extraction filter.

4.3.4 Objective selection

The relatively large set of constraints for the LCMV filter forms the basis of a noise gain problem with these filters. In order to clarify this issue we have depicted the mixing column and the LCMV filter as vectors for two example mixing scenarios in Figure 4.4. The mixtures consist of two sources observed by two sensors. The LCMV extraction filter, represented by the vector $\tilde{\boldsymbol{\mu}}_d$, is orthogonal to the mixing column vector $\mathbf{a}^i$ that corresponds to the interfering source. For these examples we assume that all the vectors have unit length and that the sensor noise is spatially white. Consequently, the noise power in an extracted signal is independent of the direction of the extraction filter. On the left hand side of Figure 4.4 a relatively ill conditioned mixing system is presented in which the angle between the mixing column vectors is relatively small. For this ill conditioned mixture, the projection of the mixing column vector of the desired source $\mathbf{a}^d$ onto the LCMV extraction filter $\boldsymbol{\mu}_d$ results into a relatively short vector. Consequently, the power of the desired source in the extracted signal is relatively low with respect to the power of the noise. On the right hand side a relatively well conditioned mixing system is presented. For this mixing system the projection of the mixing column of the desired source onto the LCMV extraction filter is relatively large. Consequently, for the well conditioned mixing system the power of the desired source in the extracted signal is relatively high with respect to the noise. If the mixing system is ill conditioned, then the SNR of the signal extracted by the LCMV filter reaches impermissible levels. In such a case the MVDR filter is typically preferred. If we are able to measure the conditioning of the mixing system, then we can use this measure to support the (automated) decision for one of the objective functions. One way to determine the conditioning of the mixing system is to identify all mixing columns with a blind system identification algorithm and calculate the condition number of the identified mixing system, which is typically defined as the ratio of the largest and smallest singular value of the matrix under consideration. A disadvantage of this method is that the entire mixing system has to be identified. Furthermore, the fact that one is interested in the extraction of only one of the sources is not taken into account. Therefore, a measure for the conditioning of the mixing system is to be found that takes advantage of this extraction property. In the remainder of this section we present a measure that is useful for determining the conditioning of the mixing system, related to the fact that extraction of only the desired source is desired.

With the previously presented source extraction algorithm we are able to identify

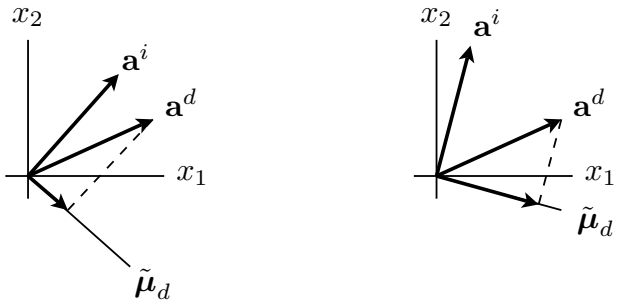


Figure 4.4: Geometrical interpretation of the LCMV filter for two mixtures consisting of two sources observed by two sensors.

the basic elements of the MVDR and the LCMV filters by calculating respectively the right and left eigenvector of a matrix $\mathbf{M}$. These eigenvectors, respectively $\bar{\mathbf{a}}^d$ and $\tilde{\boldsymbol{\mu}}_d$, have an arbitrary length and an arbitrary phase, which is a reflection of the scaling indeterminacy in blind signal processing. A way to measure the conditioning of the mixing system while considering the fact that one is interested in the extraction of only the desired source is to calculate the Hermitian angle [67] between these two eigenvectors. This Hermitian angle θ_H is defined in Definition 4.3.1 and is the angle between two vectors, independent of their potential phase shift. Since γ_H in Definition 4.3.1 is always positive and real-valued, we have for the Hermitian angle that $0 \leq \theta_H \leq 90$ degrees where the Hermitian angle is zero degrees if the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ point in the same direction and is 90 degrees if the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ are orthogonal to each other.

Definition 4.3.1 (Hermitian angle)**:** The Hermitian angle θ_H between the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ is defined as follows:

$$\gamma_H = \cos\theta_H \triangleq \frac{|\langle \tilde{\boldsymbol{\mu}}_d, \bar{\mathbf{a}}^d \rangle|}{\|\tilde{\boldsymbol{\mu}}_d\|_2 \cdot \|\bar{\mathbf{a}}^d\|_2} \quad \Longleftrightarrow \quad \theta_H = \arccos\gamma_H \tag{4.41}$$

where $0 \leq \gamma_H \leq 1$ is a real-valued number, even if the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ are complex-valued.

In order to show how the Hermitian angle can be used in practice we assume to have observations that are corrupted by spatially white sensor noise. This means that the lag zero sensor correlation matrix $\mathbf{R}^x[n,0]$ has the following structure:

$$\mathbf{R}^x[n,0] \triangleq \begin{bmatrix} r^x_{11}[n,0] & \cdots & r^x_{1D}[n,0] \\ \vdots & \ddots & \vdots \\ r^x_{D1}[n,0] & \cdots & r^x_{DD}[n,0] \end{bmatrix} = \mathbf{A}\mathbf{R}^s[n,0](\mathbf{A})^H + \sigma^2_\nu \mathbf{I} \tag{4.42}$$

where σ^2_ν is the power of the sensor noise and $\mathbf{R}^s[n,0] \triangleq \text{diag}\left(\begin{bmatrix}\sigma^2_{s_1} & \cdots & \sigma^2_{s_S}\end{bmatrix}\right)$ is a diagonal matrix containing the powers of the sources at time n, i.e., $\sigma^2_{s_j} = r^s_{jj}[n,0]$ for $j \in \mathcal{S}$. We have learned in the beginning of this section that the noise gain issue

is observed in the desired signal to noise ratio ($\mathrm{SNR}_{\mathrm{LCMV}}$) of the signal extracted by the LCMV filter $\tilde{\boldsymbol{\mu}}_d$. For the model in (4.42) the $\mathrm{SNR}_{\mathrm{LCMV}}$ is given as follows:

$$\mathrm{SNR}_{\mathrm{LCMV}} = \frac{|\langle \tilde{\boldsymbol{\mu}}_d, \bar{\mathbf{a}}^d \rangle|^2}{\|\tilde{\boldsymbol{\mu}}_d\|_2^2} \cdot \frac{\sigma_{s_d}^2}{\sigma_\nu^2} = \|\bar{\mathbf{a}}^d\|_2^2 \cdot (\cos\theta_H)^2 \cdot \frac{\sigma_{s_d}^2}{\sigma_\nu^2} \tag{4.43}$$

where we used that $|\langle \tilde{\boldsymbol{\mu}}_d, \bar{\mathbf{a}}^d \rangle| = \|\tilde{\boldsymbol{\mu}}_d\|_2 \|\bar{\mathbf{a}}^d\|_2 \cos\theta_H$. Due to lack of information about the length of the mixing column corresponding to the desired source $\|\bar{\mathbf{a}}^d\|_2^2$, the power of the desired source $\sigma_{s_d}^2$, and the power of the noise σ_ν^2, the exact value of $\mathrm{SNR}_{\mathrm{LCMV}}$ cannot be calculated. However, the Hermitian angle can be used to determine the relative amount of energy of the desired source that is obtained in the extracted signal. If $\gamma_H \approx 1$, i.e., $\theta_H \approx 0°$, then most of the observed desired source energy is available in the extracted signal, i.e., the mixing system is well conditioned for extraction of the desired source. On the contrary, if $\gamma_H \approx 0$, i.e., $\theta_H \approx 90°$, then only a little amount of desired source energy is available in the extracted signal. Consequently, the mixing system is ill conditioned for extraction of the desired source. Notice that this measure is independent of the number of sources, as long as the number of sensors is equal to or larger than the number of sources.

In order to obtain more insight into the conditioning of the mixing system we turn back to a geometrical interpretation. As stated before, the LCMV extraction filter $\tilde{\boldsymbol{\mu}}_d$ is orthogonal to the mixing columns that correspond to the interfering sources. Therefore, a large angle between the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ is observed if the angle between the mixing columns of the desired source and (at least) one of the interfering sources is small. On the contrary, a small angle between the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ is observed if the angles between the mixing column corresponding to the desired source and the mixing column vectors corresponding to all the interfering sources is large. An example of such a mixing system is an orthogonal mixing system. In that case, the angle between two mixing column vectors is 90°. However, an orthogonal mixing system is not required in order to obtain an Hermitian angle of $\theta_H = 90°$ between the vectors $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$. An Hermitian angle of $\theta_H = 90°$ is obtained as long as the mixing columns that correspond to the interfering sources are orthogonal to the mixing column that corresponds to the desired source. This property holds even if the angle between the mixing column vectors of two interferers becomes very small. Consequently, the Hermitian angle is a useful measure to indicate the conditioning of a mixing system that takes into account that extraction of only the desired source is desired. For other, application specific noise models, the usage of this Hermitian angle has to be reconsidered.

We have shown that extraction of the desired source is possible, up to an undetermined scale, based on an a priori guess of the either the autocorrelation function or the mixing column vector of the desired source, represented as a mold. Based on this mold, a specific design of the weights $\tilde{\boldsymbol{\xi}}_l$ or $\boldsymbol{\xi}^l$ in the form of an orthonormal basis was used. Finally, we introduced a measure that can be useful in (automated) selection of the LCMV or the MVDR objective function. In the following section we introduce new design tools such that extraction algorithms can be designed based on a priori information about physical parameters of the sources.

4.4 Algorithm design exploiting a parameterized mixing model

In the previous section, direct extraction of the desired source is guaranteed by means of a mold. By means of a parameterized modeling of the environment, new signal extraction algorithms can be designed that guarantee extraction of the desired source based on a priori information about physical parameters of the sources, e.g., in the form of direction of arrival (DOA) information. We present tools that can be used for designing signal extraction algorithms and we give example designs.

4.4.1 Parameterized modeling of the mixing system

Parameterized modeling is a widely used tool for the design of beamformers or spatial filters in the fields of acoustics and telecommunication. Here we repeat certain basic steps of beamformer design and evaluation since we use them to introduce new design tools for developing signal extraction algorithms.

Beamformer design is typically based on a parameterized model of the environment in terms of a DOA parameter [5]. In case the sources are located at approximately the same height, at a relatively large distance from the sensor array, then the following parametrization of the mixing model is often used:

$$\mathbf{x}[n] = \sum_{j=1}^{S} \mathbf{a}[\theta_j, \mathcal{P}] s_j[n] + \boldsymbol{\nu}[n] \tag{4.44}$$

where $\mathbf{a}[\theta_j, \mathcal{P}]$ is a frequency dependent array response vector, θ_j represents the DOA of the j'th source, and $\mathcal{P}$ contains the sensor positions. The elements in the array response vector are modeled as $a_i(\theta_j) = \exp(-\jmath 2\pi\tau_i(\theta_j))$ for $(i, j) \in \mathcal{D} \times \mathcal{S}$, where $\tau_i(\theta_j)$ represents the time that the signal of the j'th source takes to go from a reference point to the i'th sensor.

This model is developed for environments with no, or at least very low, reverberation. Since the goal of this work is to provide insight we keep using this model, even though it is too simple for many practical scenarios. Furthermore, the presented signal extraction algorithms have the capability to correct modeling errors up to a certain extent, which we evaluate in Section 4.5.

Beamformers are designed by constraining the gain of the beamformer for certain DOA values. The constraints, possibly combined with an optimization objective [5], determine the beampattern or spatial response of the beamformer; hence the name spatial filter.

Definition 4.4.1 (Beampattern)**:** A beampattern displays the response of a beamformer as function of the model parameters, i.e., it shows the amplification of the beamformer for all possible model parameter values. The beampattern for beam-

former $\boldsymbol{\xi}$ and model (4.44) is defined as follows:

$$B(\theta) \triangleq |\langle \boldsymbol{\xi}, \mathbf{a}(\theta) \rangle|^2 = \left| \sum_{i=1}^{D} \xi_i^* a_i(\theta) \right|^2 \quad \forall\, \theta \in \Theta \tag{4.45}$$

where $\boldsymbol{\xi}$ contains the weights of the beamformer, $\mathbf{a}(\theta)$ is the parameterized array response vector for DOA parameter θ, and Θ contains all possible angles of arrival, i.e., $\Theta = [-180, 180)$ degrees or $\Theta = [-\pi, \pi)$ radians.

We exploit this parameterized mixing model to define new tools. The tools are used to design signal extraction algorithms that select and extract the desired source signal based on physical parameters.

4.4.2 New design tools: selection pattern and selection beamformers

Parametrization of the mixing system is useful for signal extraction in two ways. The parametrization can be used to design new signal extraction algorithms and it helps to evaluate signal extraction algorithms.

Suppose we have a parameterized model such as (4.44) that describes the mixing model for our mixing scenario in (4.2), i.e., a parameterized array response vector $\mathbf{a}(\theta_j)$ instead of the mixing column vector $\mathbf{a}^j$. In that case, the eigenvalues of the matrix $\mathbf{M}$ in (4.26) are parameterized by the DOA parameter θ and obtain the following structure:

$$\lambda^j = \sum_{l=2}^{L} \frac{|\langle \boldsymbol{\xi}^l, \mathbf{a}(\theta_j) \rangle|^2}{|\langle \boldsymbol{\xi}^1, \mathbf{a}(\theta_j) \rangle|^2} \quad \forall\, \theta \in \Theta. \tag{4.46}$$

As follows from (4.46), it is possible to predict eigenvalues for sources arriving from any direction. This property can be used as a tool for signal extraction algorithm design and evaluation in the form of a selection pattern.

Definition 4.4.2 (Selection pattern)**:** A selection pattern displays the expected eigenvalues for all possible parameter values. For the parameterized model in (4.44) the selection pattern gives the expected eigenvalues for all DOAs and is given by

$$\lambda(\theta) = \sum_{l=2}^{L} \frac{|\langle \boldsymbol{\xi}^l, \mathbf{a}(\theta) \rangle|^2}{|\langle \boldsymbol{\xi}^1, \mathbf{a}(\theta) \rangle|^2} \quad \forall\, \theta \in \Theta. \tag{4.47}$$

Visualization of a selection pattern is performed by plotting the reciprocal or the inverse tangent of the reciprocal of the expected eigenvalue $\lambda(\theta)$ in (4.47). This visualization is chosen since it leads to more intuitive and better interpretable figures. Consequently, the source with a DOA that corresponds to the highest value in a visualized selection pattern will be extracted if the smallest eigenvalue is selected.

If we compare the structure of beampatterns from Definition 4.4.1 with the structure of selection patterns in Definition 4.4.2, then we observe that both are depending

on the inner product of a set of weights $\boldsymbol{\xi}$ and the parameterized mixing column vector $\mathbf{a}(\theta)$. In Definition 4.4.1 the vector $\boldsymbol{\xi}$ is called a beamformer, which leads to the definition of selection beamformers.

Definition 4.4.3 (Selection beamformer)**:** The vectors $\boldsymbol{\xi}^l$ for $l \in \mathcal{L}$, which are used to take linear combinations of sensor correlation matrices in a signal extraction algorithm, are responsible for the selection of the desired source. Therefore, these vectors $\boldsymbol{\xi}^l$ are called selection beamformers.

From (4.47) it follows that the selection pattern of a signal extraction algorithm consists of a weighting of beampatterns from several selection beamformers. This insight can be used to incorporate different kinds of information about the DOA of the desired and undesired sources into the signal extraction algorithm, i.e., it can be used to shape the selection pattern. The design of these signal extraction algorithms is more complex than the design of individual beamformers since multiple beampatterns are combined. In the following section we discuss example designs that make use of the presented tools.

4.4.3 Design strategies using new design tools

The new design tools help to incorporate a priori information in signal extraction algorithms. In this section we present design strategies based on several types of a priori information.

First, if we take only two selection beamformers into account and choose selection beamformer $\boldsymbol{\xi}^2$ as an identity vector $\mathbf{e}_i$ for some $i \in \mathcal{D}$, i.e., a one at the i'th row and zeros elsewhere, then it follows that the selection pattern is inversely proportional to the beampattern of selection beamformer $\boldsymbol{\xi}^1$ because the beampattern of an omnidirectional sensor is constant, i.e.,

$$\lambda(\theta) = \frac{|a_i(\theta)|^2}{\left|\langle \boldsymbol{\xi}^1, \mathbf{a}(\theta)\rangle\right|^2} = \frac{1}{\left|\langle \boldsymbol{\xi}^1, \mathbf{a}(\theta)\rangle\right|^2}. \tag{4.48}$$

This observation makes it easier to discuss the different types of selection pattern designs. However, a problem of this approach may be a lack of robustness of the extraction algorithm due to a low distribution of the eigenvalues. If we also design the selection beamformer $\boldsymbol{\xi}^2$, or even more selection beamformers, then we may increase the robustness of the method, which is addressed in Section 4.5.

All the presented design strategies are complemented with an example design that is based on a single scenario. In this scenario, eight sensors are distributed evenly in a circular array with a mutual spacing of $d/\lambda = 1/4$ for consecutive sensors. The first sensor is placed in the 0 degrees direction and we use 30 degrees as a guess for the DOA of the desired source.

Design strategy I Suppose we have a guess of the DOA of the desired source, which is given by θ_e. If we design a beamformer $\boldsymbol{\xi}^1$ that has the highest gain for

$\theta = \theta_e$, then a signal coming from that direction will be extracted if the smallest eigenvalue of $\mathbf{M}$ is selected. If the sensor array is properly designed, then the design of such a selection beamformer can be accomplished with the following optimization problem:

$$\min_{\boldsymbol{\xi}} \quad ||\boldsymbol{\xi}||_2 \quad \text{s.t.} \quad \langle \boldsymbol{\xi}, \mathbf{a}(\theta_e) \rangle = 1 \quad \Rightarrow \quad \boldsymbol{\xi} = \frac{\mathbf{a}(\theta_e)}{||\mathbf{a}(\theta_e)||_2^2}. \tag{4.49}$$

An example design with $\boldsymbol{\xi}^1 = \boldsymbol{\xi}$ from (4.49) and $\boldsymbol{\xi}^2 = \mathbf{e}_1$, for the eight sensor circular array configuration, is depicted in Figure 4.5. Again, the source with a DOA that corresponds to the highest value in Figure 4.5, thus the smallest eigenvalue, will be extracted.

Design strategy II If there is no source present in exactly the estimated direction θ_e, then the source that is responsible for the smallest eigenvalue will be extracted. If we want this source to be the one with its DOA closest to the estimated DOA θ_e, then we must design the selection pattern such that it is a monotonically decreasing and symmetric function with respect to DOA estimate θ_e. Exact design of such a beamformer is difficult in general, except for certain special cases where symmetry is natural. However, we can use convex optimization techniques to obtain an as symmetric as possible beampattern. The following optimization problem can be used to obtain symmetric beampatterns:

$$\begin{aligned} \min_{\boldsymbol{\xi}} \quad & \|\langle \boldsymbol{\xi}, \mathbf{a}(\theta_e + \theta) - \mathbf{a}(\theta_e - \theta) \rangle\|_\infty \\ \text{s.t.} \quad & \langle \boldsymbol{\xi}, \mathbf{a}(\theta_e) \rangle = 1 \quad \wedge \quad \langle \boldsymbol{\xi}, \mathbf{a}(\theta_e - 180) \rangle = 0. \end{aligned} \tag{4.50}$$

The unit constraint in the estimated direction θ_e and the zero constraint in the opposite direction $\theta_e - 180$ degrees are used to push the beampattern into a monotonically decreasing function w.r.t. θ_e; however, it depends on the array configuration and the value for θ_e whether the function is monotonically decreasing.

An example design with $\boldsymbol{\xi}^1 = \boldsymbol{\xi}$ from (4.50) and $\boldsymbol{\xi}^2 = \mathbf{e}_1$ is depicted in Figure 4.6. The selection pattern is approximately symmetric and a monotonically decreasing function w.r.t. the guessed DOA $\theta_e = 30$ degrees.

Design strategy III Another signal extraction algorithm design is an algorithm that only extracts a source signal if the source is active in a certain DOA area, which we call a region of interest (ROI). The left and right border of this ROI is denoted by θ_l and θ_r, respectively. Such a signal extraction algorithm should have eigenvalues below a threshold λ_T only for DOAs inside the ROI. The following optimization problem will produce selection patterns for such an algorithm in general:

$$\begin{aligned} \min_{\boldsymbol{\xi}} \quad & |\langle \boldsymbol{\xi}, \mathbf{a}(\theta) \rangle| \quad \forall\, \theta \in \Theta \setminus \text{ROI} \\ \text{s.t.} \quad & \langle \boldsymbol{\xi}, \mathbf{a}(\theta_e) \rangle = 1 \quad \text{for } \theta_e = \frac{\theta_l + \theta_r}{2} \in \text{ROI}\,. \end{aligned} \tag{4.51}$$

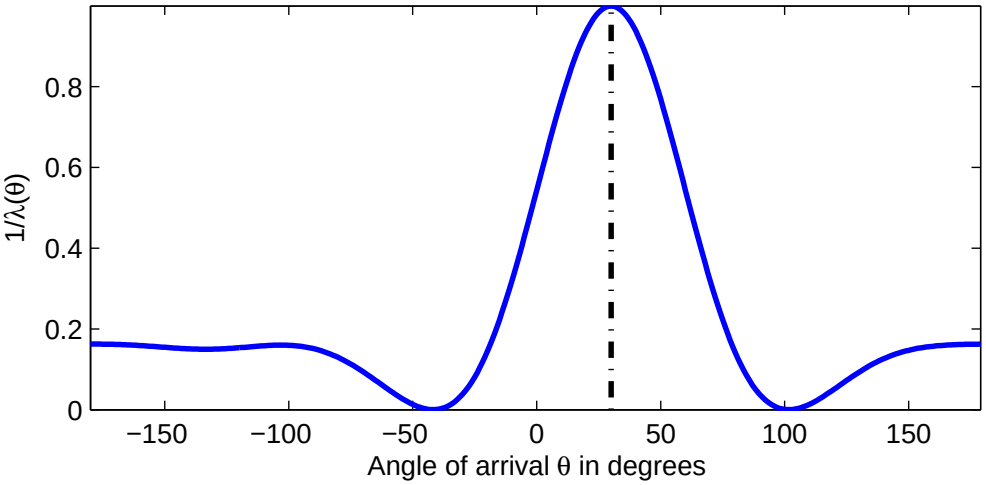


Figure 4.5: Selection pattern with the highest value, i.e., smallest eigenvalue, for guessed DOA $\theta_e = 30$ degrees.

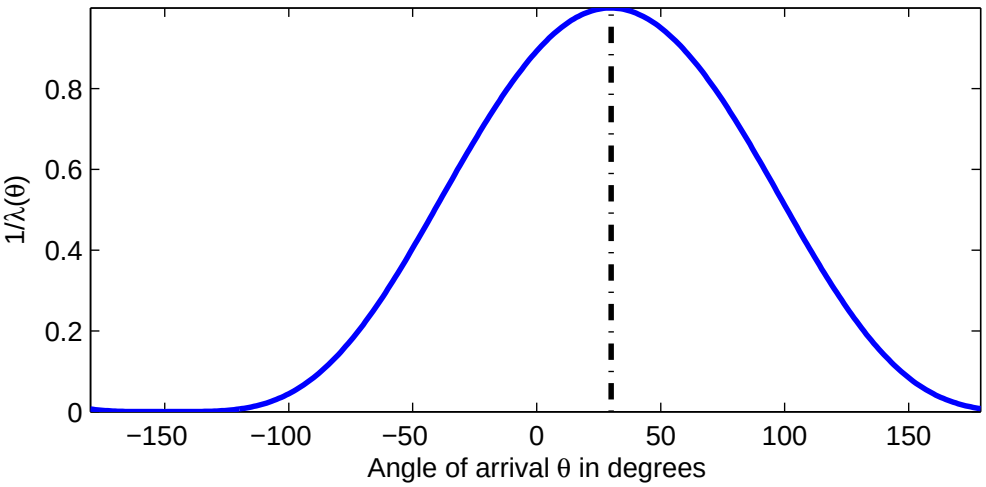


Figure 4.6: A symmetric selection pattern with respect to guessed DOA $\theta_e = 30$ degrees.

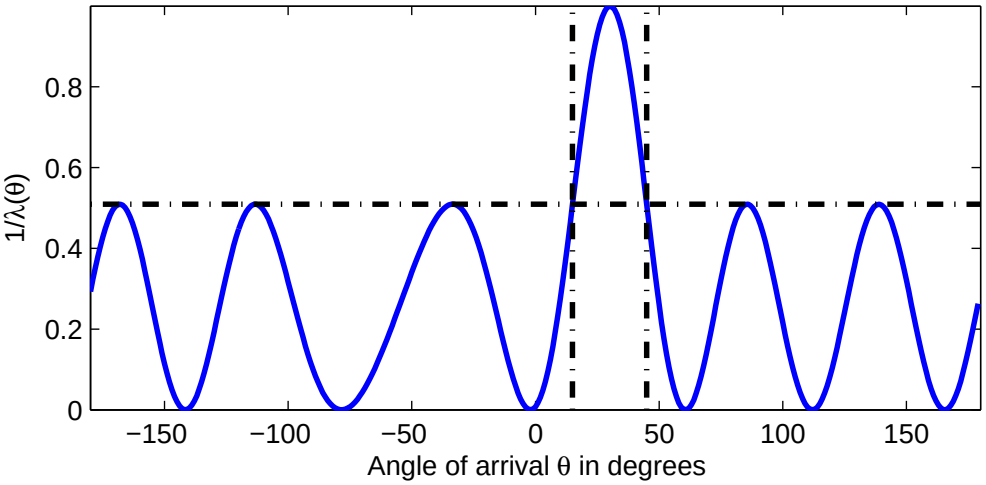


Figure 4.7: Selection pattern for a region of interest, which is given by the angles $\theta_l \leq \theta \leq \theta_r$ with $\theta_l = 15$ degrees and $\theta_r = 45$ degrees, indicated by the vertical black lines. The horizontal black line depicts the inverse of the threshold $\lambda_T = 1.96$.

Originally, this optimization problem is used to minimize sidelobes of a beamformer with a main lobe in the ROI. Here we use it because it is likely to result in a beamformer that has a gain higher than a threshold $t = 1/\lambda_T$ for DOA angles inside the ROI and a gain lower than $t = 1/\lambda_T$ for DOA angles outside the ROI.

An example selection pattern design with $\boldsymbol{\xi}^1 = \boldsymbol{\xi}$ from (4.51) and $\boldsymbol{\xi}^2 = \mathbf{e}_1$, for the eight sensor circular array, is depicted in Figure 4.7. The threshold $t = 0.51$ leads to the eigenvalue threshold $\lambda_T = 1.96$. Thus, sources with an eigenvalue lower than 1.96 have to be active within the ROI.

Alternative extraction algorithm designs can exploit other types of information. One can think of the following scenarios:

- A preferred DOA guess θ_e is available and also information about a possible error w.r.t. this guess, e.g., if no source is present at θ_e, then a source on the left of θ_e is preferred.
- Multiple regions of interest may be available, which should be evaluated simultaneously.
- Information about the DOA of interfering sources is available and these sources should not be extracted.
- Information w.r.t. a subset of the sensors is available, e.g., a wireless sensor array configuration with multiple sensors per node, see [65].

As follows from these design examples and sketches of alternative scenarios, the proposed framework to incorporate a priori information about the mixing system into signal extraction algorithms is very flexible. Furthermore, this approach has several advantages over traditional beamforming. First, the parametrization of the mixing

system is not required to be exact. If the desired source is selected, then the correct extraction filter will be identified. Additionally, it is not required to know all the sensor locations. If information about the location of a subset of sensors is available and we have an estimate of the DOA of the desired source w.r.t. this small sensor array, then the extraction filter exploiting all available sensors will be calculated, as long as the filter that corresponds to the desired source is selected.

4.5 Experiments

In the following series of experiments we validate the results obtained in the previous sections. We show that the presented signal extraction algorithm can be applied for different applications and that the algorithms are robust to modeling errors.

4.5.1 Objective selection

In this experiment we calculate the Hermitian angle between $\tilde{\boldsymbol{\mu}}_d$ and $\bar{\mathbf{a}}^d$ for a variety of mixing systems. In this experiment we assume to observe a complex-valued mixture of two narrowband sources measured by two sensors with a relative mutual spacing of ${}^{d}/\lambda = 1/2$. The sensors are corrupted by spatially white sensor noise with power $\sigma_\nu^2 = 0.1$. Both the desired source and the interfering source have unit power, i.e., $\sigma_{s_d}^2 = 1$ and $\sigma_{s_i}^2 = 1$. The direction of arrival (DOA) of the desired source is $\theta_d = 30°$ and the DOA of the interfering source θ_i varies during the simulation, which makes the variety in mixing systems. For an anechoic environment with sources in far field we are able to model the mixing column vectors for the desired and interfering source by the following array response vectors, respectively:

$$\mathbf{a}^d = {}^{1}/\sqrt{2} \cdot \begin{bmatrix} \mathrm{e}^{\jmath \frac{\pi}{4} \sin(\theta_d)} \\ \mathrm{e}^{-\jmath \frac{\pi}{4} \sin(\theta_d)} \end{bmatrix} \quad \text{and} \quad \mathbf{a}^i = {}^{1}/\sqrt{2} \cdot \begin{bmatrix} \mathrm{e}^{\jmath \frac{\pi}{4} \sin(\theta_i)} \\ \mathrm{e}^{-\jmath \frac{\pi}{4} \sin(\theta_i)} \end{bmatrix}. \tag{4.52}$$

Based on this information we calculated the MVDR and the LCMV extraction filters. For each of the filters we have calculated the contribution of the desired source and the interfering source in the extracted signal, which we have depicted in respectively the top and bottom graph on the left of Figure 4.8. We calculated the desired source to noise ratio for each DOA θ_i for both the MVDR and the LCMV filter and depicted it in the top right of Figure 4.8. The graph in the bottom right of Figure 4.8 depicts the Hermitian angle between the LCMV extraction filter and the mixing column vector corresponding to the desired source.

From Figure 4.8 it follows that the LCMV filter is able to suppress the interfering source for all DOA's of that source. For most of these DOA's the LCMV filter has a relatively good contribution of the desired source such that the SNR of the LCMV filter is similar to the SNR of the MVDR filter. However, for a DOA of the interfering source close to the DOA of the desired source at 30° the contribution of the desired source in the signal extracted by the LCMV filter drops, as expected. Consequently, the SNR for those directions become impermissibly low for many applications. The

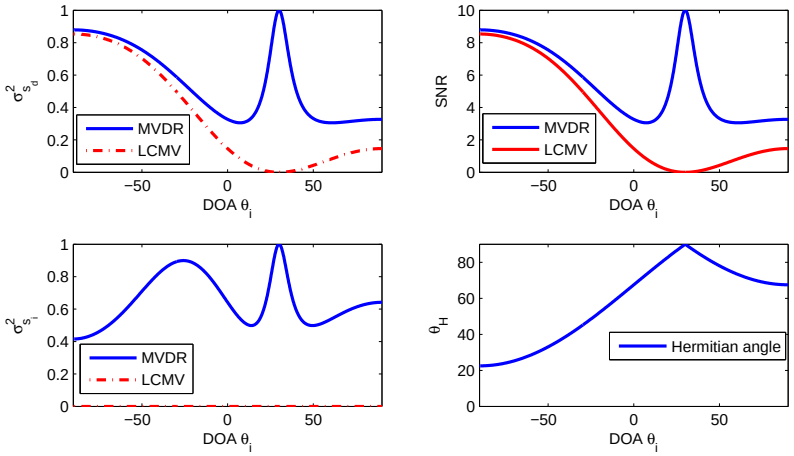


Figure 4.8: The powers of the desired source and interfering source components are depicted in respectively the top and bottom graph on the left hand side for both the MVDR and the LCMV extraction filters for varying DOA of the interfering source. The top graph on the right shows the SNR for both extraction filters and in the graph on the bottom right the Hermitian angle between the LCMV filter and the mixing column corresponding to the desired source is depicted.

Hermitian angle θ_H in the graph on the bottom right of Figure 4.8 shows values near 90° for those DOA's near 30°, again this corresponds to our expectations.

4.5.2 Extraction of a fetal electrocardiogram

The signal extraction algorithm from Section 4.3, based on the mold that contains a guess of the autocorrelation function of the desired source, is used in this experiment to extract the fetal ECG signal from a ten seconds ECG measurement taken from the Daisy database [68]. The experiment is mainly conducted to verify the flexibility of the algorithm design framework and no further pre- or postprocessing has been performed to increase the quality of the signals.

We use the five measurements from the abdomen since they are more likely to contain the fetal ECG signal than measurements from the thorax. These signals, which are sampled at a frequency of 250 Hz, are depicted in Figure 4.9. The peak in the sum of absolute values of the sensor autocorrelation functions, in the range from 40 to 150 BPM, is used to estimate the heartbeat of the mother. The estimated heartbeat of the mother is 81 beats per minute (BPM), which corresponds to a period of 0.7 seconds. We assume that the heartbeat of a fetus is higher than that of the mother and guess it between 100 BPM and 150 BPM, which is expected to result in a peak in the autocorrelation function between lags 100 and 150. Consequently, we use the autocorrelation function values for lags 1 until 150 and create a mold vector $\tilde{\mathbf{r}}_0$ that has ones for lags 100 to 150 and zeros elsewhere.

The extracted signal is depicted in Figure 4.10. From a comparison with Figure 4.9 it follows that a signal with a shorter period is extracted, which was the goal of this experiment. Further research could focus on the design of a good mold vector $\tilde{\mathbf{r}}_0$.

4.5.3 Mismatch in prior information

We know from the previous sections that the optimal extraction filter is obtained as long as the desired source is selected. This makes the algorithm robust against mismatches in the available a priori information. In this section we evaluate the sensitivity of the eigenvalues for mismatches in prior information.

In the experiment we mixed 10000 samples of three complex sources, each having an autoregressive structure with a single complex pole. The mixing system is generated by means of the far field parameterized model of (4.44) for a linear array of consisting of four elements. The relative mutual spacing is chosen as $d/\lambda = 1/4$ such that we obtain a proper array in terms of spatial aliasing. The power of the desired signal and the interfering signals is set to one and the sensor noise is complex white Gaussian with a signal to noise ratio of 20 dB per sensor. The zero degree direction is defined perpendicular to the array and the two interfering sources have a DOA of -17 and +28 degrees and their poles are $0.95\,\mathrm{e}^{\jmath 2.31}$ and $0.56\,\mathrm{e}^{-\jmath 0.61}$, respectively. In the experiment we move the desired source, which has a pole at $0.91\,\mathrm{e}^{\jmath 0.54}$, from a DOA of -20 degrees to +20 degrees.

Four extraction algorithms are evaluated in this experiment. Each algorithm is based on the guess $\theta_e = 0$ degrees for the DOA of the desired source; and, the corresponding array response vector is denoted as $\mathbf{a}(\theta_e)$. The four extraction algorithms differ only in the vectors $\boldsymbol{\xi}^l$.

- Alg. 1 has four selection beamformers that form an orthonormal basis, with $\boldsymbol{\xi}^1 = \mathbf{a}(\theta_e)$.
- Alg. 2 uses only the second and third sensor. The vector $\boldsymbol{\xi}^1$ consists of the second and third element of $\mathbf{a}(\theta_e)$ and $\boldsymbol{\xi}^2$ is orthonormal w.r.t. $\boldsymbol{\xi}^1$.
- Alg. 3 uses $\mathbf{a}(\theta_e)$ as selection beamformer $\boldsymbol{\xi}^1$ and the first sensor, i.e., $\boldsymbol{\xi}^2 = \mathbf{e}_1$.
- Alg. 4 uses $\mathbf{a}(\theta_e)$ as $\boldsymbol{\xi}^1$ and a delay and subtract beamformer as second selection beamformer, i.e., $\boldsymbol{\xi}^2 = \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix}^T$, which has a symmetric beampattern and a zero in the guessed direction $\theta_e = 0$ degrees.

The selection patterns for these algorithms are depicted in Figure 4.11.

First we depict in Figure 4.12 the expected and actual eigenvalues of the matrix $\mathbf{M}$ for every algorithm for a range of actual DOA's of the desired source. Each algorithm produces the eigenvalues that we expected. Furthermore, only one set of eigenvalues changes with a changing DOA of the desired source. Consequently, these eigenvalues must correspond to the desired source. For an absolute actual DOA below 17 degrees we observe that the eigenvalues that correspond to the desired source are the smallest. Outside this region, the smallest eigenvalue corresponds to the interferer at -17 degrees. Furthermore, we observe a different arrangement of the eigenvalues

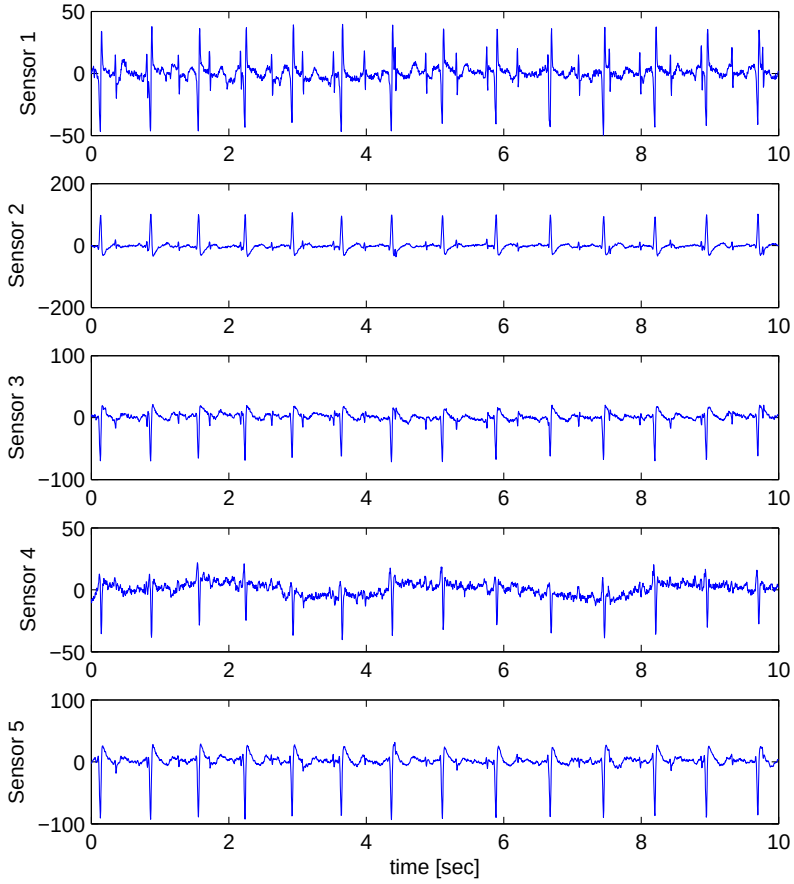


Figure 4.9: Original ECG measurements taken from the abdomen.

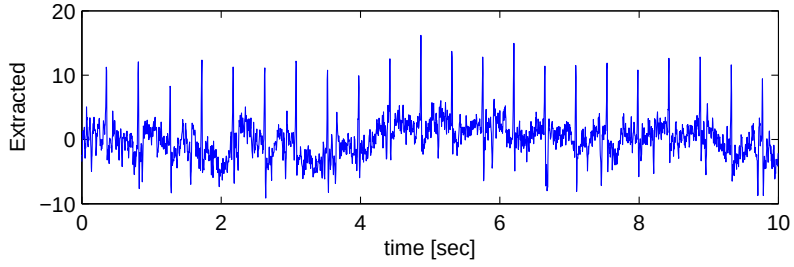


Figure 4.10: Signal extracted from the ECG measurements.

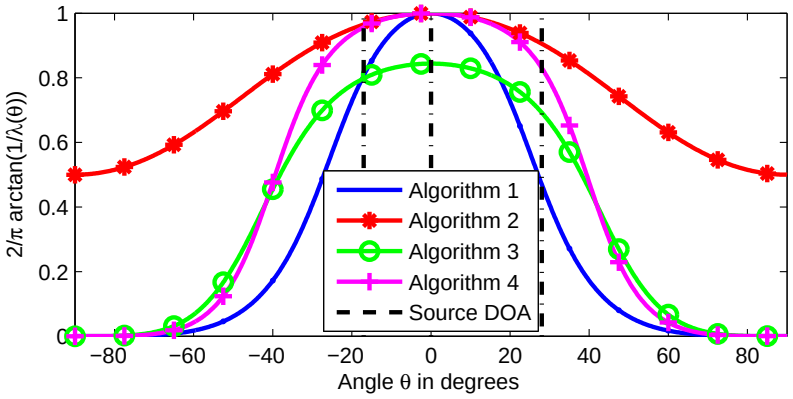


Figure 4.11: Selection patterns for the four algorithms and nominal DOA's for the interferers and the desired source at -17, 28, and 0 degrees, respectively. Due to the linear array configuration these patterns repeat after 180 degrees.

among the algorithms. For Algorithm 3 the smallest eigenvalue does not tend to become a zero eigenvalue, while this is the case for the other algorithms; and, the use of two additional sensors in Algorithm 4 leads to an increase in spread of the eigenvalues w.r.t. Algorithm 2. Finally, the eigenvalues of Algorithm 1 have the highest spread; however, it also uses more than two selection beamformers. The mismatch in, especially the largest, eigenvalues is obtained due to a limited amount of data. If we increase the number of samples the actual eigenvalues better match the expected eigenvalues.

The most interesting performance measure to evaluate extraction algorithms is the signal to interference ratio (SIR), which we define as follows:

$$\text{SIR} \triangleq 10 \log_{10} \frac{\sigma_{s_d}^2}{\sum_{j \in S \setminus d} \sigma_{s_j}^2} \tag{4.53}$$

where $\sigma_{s_j}^2$ is the power of the j'th source in the extracted signal.

We depicted the SIR results for the four algorithms in Figure 4.13. In the upper graph we depict the SIR for MVDR filters that are based on the identified mixing column, followed by noise reduction in batch modus. In the lower graph we depict the SIR for LCMV filters based on the identified extraction filters and noise reduction. Both graphs also display the performance of the ideal MVDR filter based on knowledge of the actual DOA of the desired source and the mismatched MVDR filter that is based on the DOA guess θ_e.

We observe that the four algorithms identify the ideal MVDR filter for all DOA values, as long as the smallest eigenvalue corresponds to the desired source, i.e., the actual DOA of the desired source has to be between -17 and 17 degrees. Furthermore, we observe that the LCMV filters have a similar, good performance if the actual DOA of the desired source is between -17 and 17 degrees. If we compare these results with the mismatched MVDR filter, then we observe that for both the MVDR and LCMV

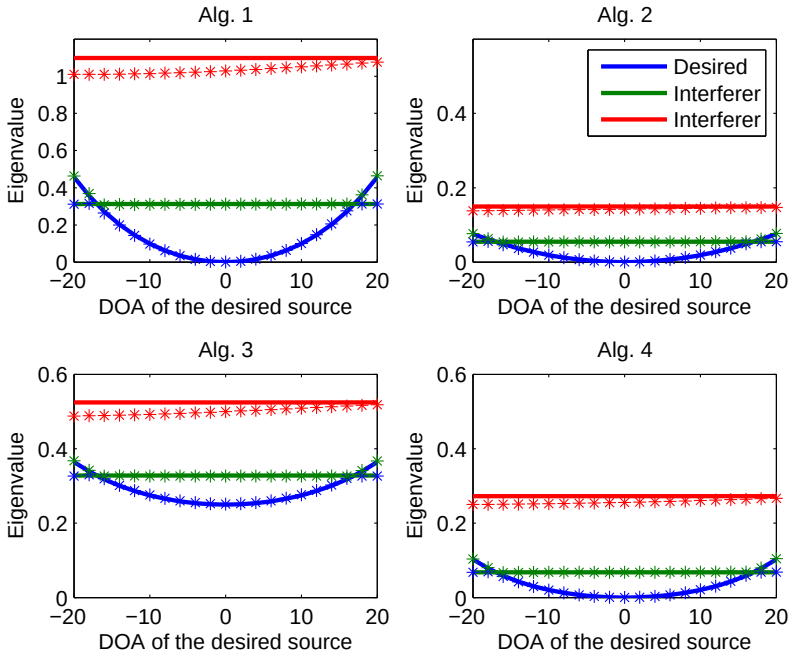


Figure 4.12: Expected (line) and actual (asterisk) eigenvalues for the four extraction algorithms. The smallest eigenvalues are supposed to correspond to the desired source.

filter, the extraction algorithms are much more robust w.r.t. errors in the a priori information. The only condition is that the smallest eigenvalue has to correspond to the desired source. Finally, notice that extraction algorithm 2 uses information about only two sensors, while the identified extraction filter uses four sensors.

4.5.4 Signal extraction in a reverberant environment

In Section 4.5.3 we investigated the influence of a mismatch in a priori information. Here we make the problem harder by taking reverberation into account, i.e., a mismatch in the parameterized model of the environment.

The scenario is the same as in Section 4.5.3, except that the sources are now placed in a room of size $6 \times 4 \times 3$ meters. The center of the array is located at position $(1, 1.5, 1.2)$ meters, with the sensor positions varying along the 4 meter wall. The sources are positioned at 1.5 meters from the array, with the angles -17 and 28 degrees for the interferers and a range of angles from -20 to 20 degrees for the desired source. We used the room impulse response generator [69] with a reverberation time of 300 ms to obtain realistic echoes and reverberation. We transformed the impulse responses to the frequency domain and picked the mixing matrix at 1 kHz, which lead again to a relative sensor spacing of $d/\lambda = 1/4$.

We depicted the eigenvalues for the four algorithms from Section 4.5.3 in Fig-

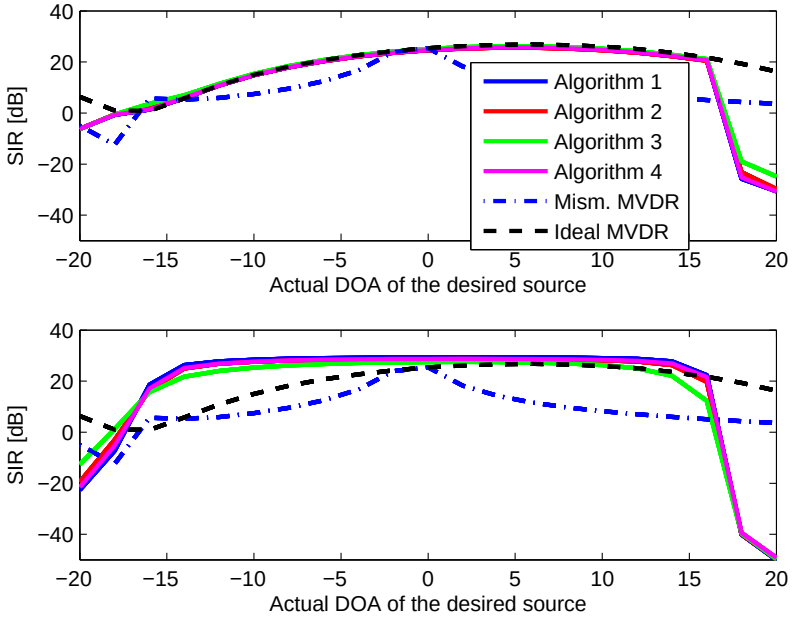


Figure 4.13: SIR for MVDR and LCMV filters in respectively the upper and lower graph, for four extraction algorithms. The black dashed line represents the SIR for an MVDR filter with knowledge about the actual DOA of the desired source and the blue dash-dotted line represents the SIR for an MVDR filter with $\theta_e = 0$ degrees.

ure 4.14. We observe that the eigenvalues match the expected eigenvalues for the mixture at hand. Furthermore, the changing eigenvalues over varying DOA of the desired source are the smallest for quite a large region, which shows a good robustness w.r.t. the given a priori information.

In Figure 4.15 we have depicted the SIR for the four algorithms for a varying DOA of the desired source. Again, the upper graph displays the MVDR filters, while the lower graph depicts the LCMV filters. Additionally, we have plotted in both graphs the SIR for a mismatched MVDR filter based on the a priori guess $\theta_e = 0$ and we have plotted the SIR result for the ideal MVDR filter, which has full knowledge about the mixing column that corresponds to the desired source.

We observe that all four extraction algorithms are able to extract the desired source for DOA's of the desired source in the range from -17 up to 10 degrees. This reduction in range of actual DOA values is a consequence of the modeling errors obtained by the reverberation. For the range of DOA's where the smallest eigenvalues correspond to the desired source, the MVDR filters obtain a similar performance as the ideal MVDR filter, while the mismatched MVDR filter is not able to improve the SIR. Furthermore, the LCMV filters show a very robust result in this region since the SIR is almost flat.

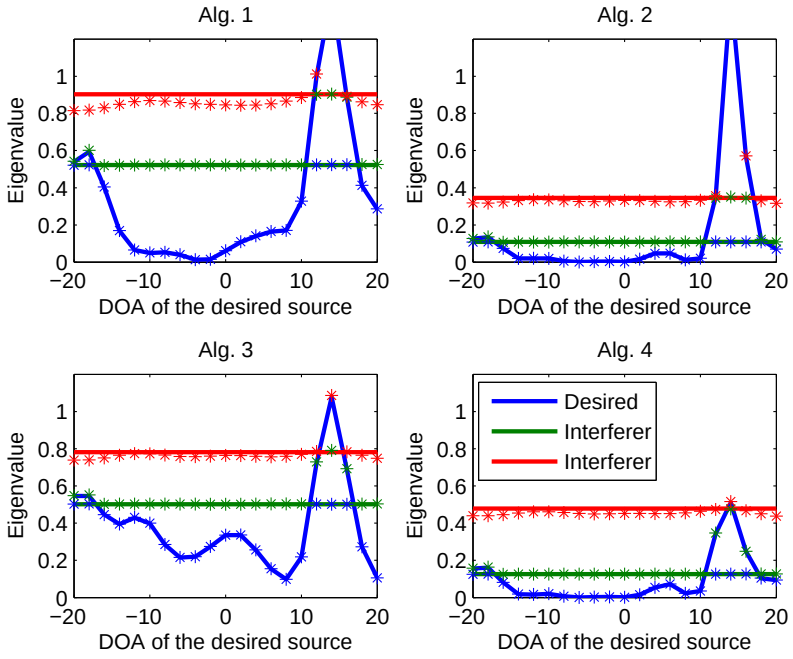


Figure 4.14: Expected (line) and actual (asterisk) eigenvalues for the multipath simulation.

4.6 Discussion and conclusions

The objectives of this work are to provide insight into signal extraction for real and complex instantaneous mixtures of multiple sources and to develop a design procedure for extraction algorithms that immediately extract the desired signal. The developed signal extraction algorithms and filters are required to be flexible, efficient, and robust.

The presented design procedure results in extraction algorithms for both real and complex instantaneous mixtures and identify extraction filters with different objectives, i.e., the MVDR filter, which balances between noise and interference reduction, and the LCMV filter, which cancels interfering source signals prior to noise reduction. Simulation results have shown that these filters can be identified in the presence of modeling errors. Furthermore, the extraction algorithms are robust to noise through the use of a noise-free region of support.

By incorporating a priori information into signal extraction algorithms immediately the desired extraction filter is identified and no deflation method is required. In this way, we deal with the permutation problem, which is a characteristic of blind signal processing (BSP). We have shown that the signal extraction algorithms can be designed by using a mold that contains either a guess of the autocorrelation function or a guess of the mixing column of the desired source. We derived conditions for this mold such that immediate extraction of the desired signal is guaranteed. Furthermore,

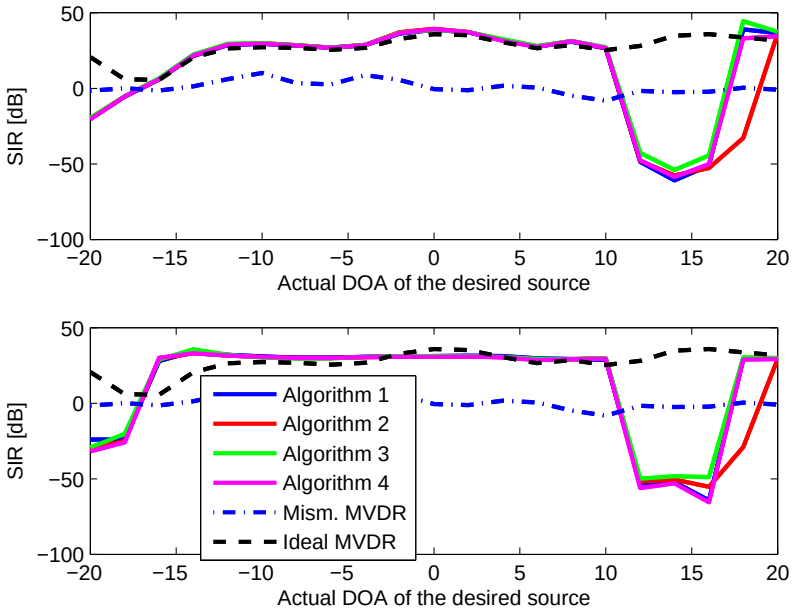


Figure 4.15: SIR for MVDR and LCMV filters in respectively the upper and lower graph, for the multipath simulation.

for a parameterized mixing system we developed new design and evaluation tools in the form of selection beamformers and selection patterns. These tools can be used to design signal extraction algorithms that guarantee immediate extraction of the desired signal based on physical parameters such as direction of arrival information. A similar approach can be followed for a priori information about the autocorrelation function, e.g., by using linear prediction techniques or information from other sensors [65].

We have not addressed the scaling indeterminacy, which is another characteristic of BSP. Ways to deal with this problem are for example to estimate the contribution of the desired source in a reference signal or to normalize the extraction filter or the output of the extraction filter; however, the scaling problem is of greater concern in wideband applications.

We conclude that the presented work provides insight into the design of signal extraction algorithms for real and complex mixtures of multiple sources and that the presented design procedure and design and evaluation tools are flexible and lead to robust and efficient signal extraction algorithms.

Topics for future research are the application of these types of algorithms in real life applications, design of extraction algorithms based on other types of a priori information such as different physical parameterizations; and, development of signal extraction algorithms for convolutive mixtures.

4.A Subspaces of the sensor correlation matrices

Due to their rank structure, all the sensor correlation matrices $\mathbf{R}^x_\kappa$ or $\{\mathbf{C}^x_{\bullet i}, \mathbf{C}^x_{i\bullet}\}$ can be reduced to size $S \times S$ by means of only two reduction matrices $\mathbf{P}_s$ and $\mathbf{Q}^s$. These reduction matrices are identified by applying subspace techniques to concatenated sensor correlation matrices. The concatenated matrices $\mathbf{C}^x$ and $\bar{\mathbf{C}}^x$ and their structure in terms of the mixing matrix $\mathbf{A}$ and the source autocorrelation matrix $\mathbf{C}^s$ are as follows:

$$\mathbf{C}^x \triangleq \begin{bmatrix} \mathbf{C}^x_{1\bullet} \\ \vdots \\ \mathbf{C}^x_{D\bullet} \end{bmatrix} \equiv (\mathbf{A} \diamond \mathbf{A}^*) \cdot \mathbf{C}^s \tag{4.54}$$

$$\bar{\mathbf{C}}^x \triangleq \begin{bmatrix} \mathbf{C}^x_{\bullet 1} & \cdots & \mathbf{C}^x_{\bullet D} \end{bmatrix} \equiv \mathbf{A} \cdot \left(\mathbf{A}^* \diamond (\mathbf{C}^s)^T \right)^T \tag{4.55}$$

where $\diamond$ is the Khatri-Rao product, i.e., the column-wise Kronecker product [55].

The ranks of these matrices depend on the structure of the mixing matrix and the source autocorrelation matrix. If all sources are extractable, then the mixing column vectors are linearly independent and the mixing matrix $\mathbf{A}$ has rank S. Furthermore, the assumption of linearly independent source autocorrelation functions in (4.13) makes that the source autocorrelation matrix $\mathbf{C}^s$ has rank S as well.

Since the Khatri-Rao product of two full rank matrices is full rank, in general [36], it turns out that the matrices $\mathbf{C}^x$ and $\bar{\mathbf{C}}^x$ have rank S. Therefore, a singular value decomposition (SVD) can be used to identify signal subspaces. The SVD of $\mathbf{C}^x$ is used to identify reduction matrix $\mathbf{Q}^s$ and the SVD of $\bar{\mathbf{C}}^x$ is used to identify reduction matrix $\mathbf{P}_s$. Reduction matrix $\mathbf{Q}^s$ is used to compress the source autocorrelation matrix $\mathbf{C}^s$, while reduction matrix $\mathbf{P}_s$ is used to compress the mixing matrix $\mathbf{A}$.

The singular value decomposition of $\mathbf{C}^x$ has the following structure:

$$\text{SVD}\left(\mathbf{C}^x\right) = \begin{bmatrix} \mathbf{U}^s & \mathbf{U}^\nu \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s \\ \mathbf{V}_\nu \end{bmatrix} \tag{4.56}$$

where $\mathbf{U}^s$ has S columns of length D^2, $\mathbf{\Sigma}^s$ is an $S \times S$ matrix, and $\mathbf{V}_s$ has S rows of length K. The reduction matrix $\mathbf{Q}^s$, that compresses the source autocorrelation matrix $\mathbf{C}^s$ is given by $\mathbf{Q}^s = (\mathbf{V}_s)^H$. The number of columns of $\mathbf{C}^s$ is compressed by means of the product $\mathbf{C}^s\mathbf{Q}^s$, which results in an $S \times S$ matrix.

Similarly, the mixing matrix $\mathbf{A}$ is reduced by means of the reduction matrix $\mathbf{P}_s$. The SVD of the matrix $\bar{\mathbf{C}}^x$ is as follows:

$$\text{SVD}\left(\bar{\mathbf{C}}^x\right) = \begin{bmatrix} \mathbf{U}^s & \mathbf{U}^\nu \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}^s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s \\ \mathbf{V}_\nu \end{bmatrix} \tag{4.57}$$

where $\mathbf{U}^s$ has S columns of length D, $\mathbf{\Sigma}^s$ is an $S \times S$ matrix, and $\mathbf{V}_s$ has S rows of length $D \cdot K$. The columns of the matrix $\mathbf{U}^s$ span the same space as the columns of the mixing matrix $\mathbf{A}$. Therefore, the reduction matrix $\mathbf{P}_s$ is given as $\mathbf{P}_s = (\mathbf{U}^s)^H$.

4.B Derivations of selected equations

The derivations of (4.18), (4.19), and (4.20) are very similar to the derivations of (3.6) and (3.7) in Appendix 3.A. Consequently, here we follow the same approaches for the derivations. The only difference is that we have to take the conjugation into account in a proper way.

Derivation of (4.18) The sensor correlation matrix $\mathbf{R}_{\kappa}^{x}$ is obtained in the following way:

$$\mathbf{R}_{\kappa}^{x} = \mathbb{E}\{\mathbf{x}[n](\mathbf{x}[n-k])^{H}\} = \begin{bmatrix} r_{11}^{x}[\Omega_{\kappa}] & \cdots & r_{1D}^{x}[\Omega_{\kappa}] \\ r_{21}^{x}[\Omega_{\kappa}] & \cdots & r_{2D}^{x}[\Omega_{\kappa}] \\ \vdots & \ddots & \ddots \\ r_{D1}^{x}[\Omega_{\kappa}] & \ddots & r_{DD}^{x}[\Omega_{\kappa}] \end{bmatrix} \tag{4.58}$$

where $\Omega_{\kappa} = (n,k)_{\kappa}$ is a time-lag pair in the NF-ROS. By substituting the vector with observations $\mathbf{x}[n]$ by the model in (4.2) we obtain the following structure:

$$\begin{aligned} \mathbf{R}_{\kappa}^{x} &= \mathbb{E}\{(\mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n])(\mathbf{A}\mathbf{s}[n-k] + \boldsymbol{\nu}[n-k])^{H}\} \\ &= \mathbf{A}\mathbb{E}\{\mathbf{s}[n](\mathbf{s}[n-k])^{H}\}(\mathbf{A})^{H} + \mathbb{E}\{\boldsymbol{\nu}[n](\boldsymbol{\nu}[n-k])^{H}\} \\ &\quad + \mathbf{A}\mathbb{E}\{\mathbf{s}[n](\boldsymbol{\nu}[n-k])^{H}\} + \mathbb{E}\{\boldsymbol{\nu}[n](\mathbf{s}[n-k])^{H}\}(\mathbf{A})^{H}. \end{aligned}$$

Applying the assumptions on the source and noise signals that hold for time-lag pairs in the NF-ROS, as given in Definition 4.2.3, leads to the following structure in the sensor correlation matrix:

$$\mathbf{R}_{\kappa}^{x} = \mathbf{A}\mathbb{E}\{\mathbf{s}[n](\mathbf{s}[n-k])^{T}\}(\mathbf{A})^{H} = \mathbf{A}\mathbf{R}_{\kappa}^{s}(\mathbf{A})^{H} \tag{4.59}$$

with the source autocorrelation matrix a diagonal matrix, i.e., $\mathbf{R}_{\kappa}^{s} = \operatorname{diag}(\mathbf{r}^{s}[\Omega_{\kappa}])$.

Derivation of (4.19) The sensor correlation matrix $\mathbf{C}_{\bullet i}^{x}$ is build from the concatenation of sensor correlation vectors with the following structure:

$$\mathbb{E}\{\mathbf{x}[n]x_{i}^{*}[n-k]\} = \begin{bmatrix} r_{1i}^{x}[\Omega_{\kappa}] \\ \vdots \\ r_{Di}^{x}[\Omega_{\kappa}] \end{bmatrix}. \tag{4.60}$$

By substituting the observations $x_i[n]$ with the model in (4.2) and immediately taking into account that the noise terms do not contribute to the sensor correlation data for time-lag pairs taken from the NF-ROS, the following structure in the sensor correlation vectors is obtained:

$$\mathbb{E}\{\mathbf{x}[n]x_{i}^{*}[n-k]\} = \mathbb{E}\{\mathbf{A}\mathbf{s}[n]\sum_{j=1}^{S} a_{i}^{j,*}s_{j}^{*}[n-k]\} = \mathbf{A}\sum_{j=1}^{S} a_{i}^{j,*}\mathbb{E}\{\mathbf{s}[n]s_{j}^{*}[n-k]\} \tag{4.61}$$

where $(n,k)_\kappa \in \Omega$.

Due to the assumption that the source signals are mutually uncorrelated, the vector $\mathbb{E}\{\mathbf{s}[n]s_j^*[n-k]\}$ is only non-zero for the j'th row. Consequently, the sensor correlation vectors can be represented in the following way:

$$\mathbb{E}\{\mathbf{x}[n]x_i^*[n-k]\} = \mathbf{A}\begin{bmatrix} a_i^{1,*} r_{11}^s[\Omega_\kappa] \\ \vdots \\ a_i^{S,*} r_{SS}^s[\Omega_\kappa] \end{bmatrix} = \mathbf{A}\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i^*\right)\mathbf{r}^s[\Omega_\kappa] \tag{4.62}$$

where $\tilde{\mathbf{a}}_i^*$ is the i'th row of the conjugated mixing matrix $\mathbf{A}^*$. Concatenating such vectors for all time-lag pairs in the NF-ROS leads to the matrix $\mathbf{C}_{\bullet i}^x$ with the following structure:

$$\mathbf{C}_{\bullet i}^x = \mathbf{A}\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i^*\right)\mathbf{C}^s \tag{4.63}$$

where $\mathbf{C}^s = \begin{bmatrix}\mathbf{r}^s[\Omega_1] & \cdots & \mathbf{r}^s[\Omega_K]\end{bmatrix}$.

Derivation of (4.20) By exploiting the following property:

$$r_{i_1 i_2}^x[\Omega_\kappa] = \mathbb{E}\{x_{i_1}[n]x_{i_2}^*[n-k]\} = \mathbb{E}\{x_{i_2}^*[n-k]x_{i_1}[n]\} \quad \forall\,\Omega_\kappa = (n,k)_\kappa \in \Omega$$

it follows that the sensor correlation matrix $\mathbf{C}_{i\bullet}^x$ is build from the concatenation of sensor correlation vectors with the following structure:

$$\mathbb{E}\{x_i[n]\mathbf{x}^*[n-k]\} = \mathbb{E}\{\mathbf{x}^*[n-k]x_i[n]\} = \begin{bmatrix} r_{i1}^x[\Omega_\kappa] \\ \vdots \\ r_{iD}^x[\Omega_\kappa] \end{bmatrix}. \tag{4.64}$$

By substituting the observations $x_i[n]$ with the model in (4.2) and immediately taking into account that the noise terms do not contribute to the sensor correlation data for time-lag pairs taken from the NF-ROS, the following structure in the sensor correlation vectors is obtained:

$$\mathbb{E}\{\mathbf{x}^*[n-k]x_i[n]\} = \mathbb{E}\{\mathbf{A}^*\mathbf{s}^*[n-k]\sum_{j=1}^{S} a_i^j s_j[n]\} = \mathbf{A}^*\sum_{j=1}^{S} a_i^j \mathbb{E}\{s_j[n]\mathbf{s}^*[n-k]\} \tag{4.65}$$

where $(n,k)_\kappa \in \Omega$.

Due to the assumption that the source signals are mutually uncorrelated, the vector $\mathbb{E}\{s_j[n]\mathbf{s}^*[n-k]\}$ is only non-zero for the j'th row. Consequently, the sensor correlation vectors can be represented in the following way:

$$\mathbb{E}\{\mathbf{x}^*[n-k]x_i[n]\} = \mathbf{A}^*\begin{bmatrix} a_i^1 r_{11}^s[\Omega_\kappa] \\ \vdots \\ a_i^S r_{SS}^s[\Omega_\kappa] \end{bmatrix} = \mathbf{A}^*\,\mathrm{diag}\left(\tilde{\mathbf{a}}_i\right)\mathbf{r}^s[\Omega_\kappa] \tag{4.66}$$

where $\tilde{\mathbf{a}}_i$ is the i'th row of the mixing matrix $\mathbf{A}$. Concatenating such vectors for all time-lag pairs in the NF-ROS leads to the matrix $\mathbf{C}^x_{i\bullet}$ with the following structure:

$$\mathbf{C}^x_{i\bullet} = \mathbf{A}^* \operatorname{diag}(\tilde{\mathbf{a}}_i)\, \mathbf{C}^s \tag{4.67}$$

where $\mathbf{C}^s = \begin{bmatrix} \mathbf{r}^s[\Omega_1] & \cdots & \mathbf{r}^s[\Omega_K] \end{bmatrix}$.

Derivation of (4.26) In order to derive (4.26) we first derive the structure in the linear combinations of sensor correlation matrices that were defined in (4.23) and (4.24). Using the structure in the sensor correlation matrices as defined in (4.18), (4.19), and (4.20), the structure for these linear combinations of sensor correlation matrices is given as follows:

$$\begin{aligned}
\mathbf{\Gamma}_l &= \sum_{\kappa=1}^{K} \xi_l^{\kappa,*} \bar{\mathbf{A}} \mathbf{R}^s_\kappa (\bar{\mathbf{A}})^H = \bar{\mathbf{A}} \left(\sum_{\kappa=1}^{K} \xi_l^{\kappa,*} \mathbf{R}^s_\kappa \right) (\bar{\mathbf{A}})^H = \bar{\mathbf{A}} \mathbf{\Lambda}_l (\bar{\mathbf{A}})^H \\
\breve{\mathbf{\Gamma}}_l &= \sum_{\kappa=1}^{K} \xi_l^{\kappa} \bar{\mathbf{A}} (\mathbf{R}^s_\kappa)^H (\bar{\mathbf{A}})^H = \bar{\mathbf{A}} \left(\sum_{\kappa=1}^{K} \xi_l^{\kappa} (\mathbf{R}^s_\kappa)^H \right) (\bar{\mathbf{A}})^H = \bar{\mathbf{A}} (\mathbf{\Lambda}_l)^H (\bar{\mathbf{A}})^H
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{\Gamma}_l &= \sum_{i=1}^{D} \xi_i^{l,*} \bar{\mathbf{A}}^* \operatorname{diag}(\tilde{\mathbf{a}}_i)\, \mathbf{C}^s = \bar{\mathbf{A}}^* \left(\sum_{i=1}^{D} \xi_i^{l,*} \operatorname{diag}(\tilde{\mathbf{a}}_i) \right) \mathbf{C}^s = \bar{\mathbf{A}}^* \mathbf{\Lambda}_l \mathbf{C}^s \\
\breve{\mathbf{\Gamma}}_l &= \sum_{i=1}^{D} \xi_i^{l} \bar{\mathbf{A}} \operatorname{diag}(\tilde{\mathbf{a}}_i^*)\, \mathbf{C}^s = \bar{\mathbf{A}} \left(\sum_{i=1}^{D} \xi_i^{l} \operatorname{diag}(\tilde{\mathbf{a}}_i^*) \right) \mathbf{C}^s = \bar{\mathbf{A}} (\mathbf{\Lambda}_l)^H \mathbf{C}^s
\end{aligned}$$

where we used $\mathbf{R}^s_\kappa = \operatorname{diag}(\mathbf{r}^s[\Omega_\kappa])$.

The matrices $\mathbf{\Lambda}_l$ are diagonal matrices, with the j'th element having the following structure for the two different sensor correlation matrix structures:

$$\sum_{\kappa=1}^{K} \xi_l^{\kappa,*} r^s_{jj}[\Omega_\kappa] = \langle \tilde{\boldsymbol{\xi}}_l, \tilde{\mathbf{r}}^s_{jj} \rangle \quad \text{and} \quad \sum_{i=1}^{D} \xi_i^{l,*} a_i^j = \langle \tilde{\boldsymbol{\xi}}^l, \mathbf{a}^j \rangle. \tag{4.68}$$

In (4.25) the linear combinations of sensor correlation matrices are combined into a single matrix $\mathbf{M}$ that has a specific eigenstructure. Computing for two arbitrary linear combinations of sensor correlation matrices $\mathbf{\Gamma}_{l_1}$ and $\breve{\mathbf{\Gamma}}_{l_2}$ the product $\breve{\mathbf{\Gamma}}_{l_2} (\mathbf{\Gamma}_{l_1})^{-1}$ leads to the following:

$$\breve{\mathbf{\Gamma}}_{l_2} (\mathbf{\Gamma}_{l_1})^{-1} = \bar{\mathbf{A}} (\mathbf{\Lambda}_{l_2})^H (\bar{\mathbf{A}})^H \left(\bar{\mathbf{A}} \mathbf{\Lambda}_{l_1} (\bar{\mathbf{A}})^H \right)^{-1} = \bar{\mathbf{A}} (\mathbf{\Lambda}_{l_2})^H (\mathbf{\Lambda}_{l_1})^{-1} \left(\bar{\mathbf{A}} \right)^{-1} \tag{4.69}$$

and

$$\breve{\mathbf{\Gamma}}_{l_2} (\mathbf{\Gamma}_{l_1})^{-1} = \bar{\mathbf{A}} (\mathbf{\Lambda}_{l_2})^H \bar{\mathbf{C}}^s \left(\bar{\mathbf{A}} \mathbf{\Lambda}_{l_1} \bar{\mathbf{C}}^s \right)^{-1} = \bar{\mathbf{A}} (\mathbf{\Lambda}_{l_2})^H (\mathbf{\Lambda}_{l_1})^{-1} \left(\bar{\mathbf{A}} \right)^{-1}. \tag{4.70}$$

The products of diagonal matrices $(\breve{\mathbf{\Lambda}}_{l_2})^H (\mathbf{\Lambda}_{l_1})^{-1}$ form again diagonal matrices, such that (4.69) and (4.70) form eigenvalue decompositions. The eigenvalues, i.e., the elements on the diagonal of the matrix products $\breve{\mathbf{\Lambda}}_{l_2} (\mathbf{\Lambda}_{l_1})^{-1}$, have the following structure:

$$\frac{\langle \tilde{\boldsymbol{\xi}}_{l_2}, \tilde{\mathbf{r}}^s_{jj} \rangle^*}{\langle \tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\mathbf{r}}^s_{jj} \rangle} \quad \text{and} \quad \frac{\langle \boldsymbol{\xi}^{l_2}, \mathbf{a}^j \rangle^*}{\langle \boldsymbol{\xi}^{l_1}, \mathbf{a}^j \rangle}. \tag{4.71}$$

The matrix products $(\mathbf{\Lambda}_{l_2})^H \left(\breve{\mathbf{\Lambda}}_{l_1}\right)^{-1}$ have the same structure as (4.69) and (4.70), except for conjugated eigenvalues, i.e.,:

$$\frac{\langle \tilde{\boldsymbol{\xi}}_{l_2}, \tilde{\mathbf{r}}^s_{jj} \rangle}{\langle \tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\mathbf{r}}^s_{jj} \rangle^*} \quad \text{and} \quad \frac{\langle \boldsymbol{\xi}^{l_2}, \mathbf{a}^j \rangle}{\langle \boldsymbol{\xi}^{l_1}, \mathbf{a}^j \rangle^*}. \tag{4.72}$$

By multiplying the matrix products $\breve{\mathbf{\Gamma}}_{l_2}(\mathbf{\Gamma}_{l_1})^{-1}$ and $\mathbf{\Gamma}_{l_2}(\breve{\mathbf{\Gamma}}_{l_1})^{-1}$ the eigenvectors remain the same, but the eigenvalues of the two matrix products are multiplied, i.e.,

$$\frac{\left|\langle \tilde{\boldsymbol{\xi}}_{l_2}, \tilde{\mathbf{r}}^s_{jj} \rangle\right|^2}{\left|\langle \tilde{\boldsymbol{\xi}}_{l_1}, \tilde{\mathbf{r}}^s_{jj} \rangle\right|^2} \quad \text{and} \quad \frac{\left|\langle \boldsymbol{\xi}^{l_2}, \mathbf{a}^j \rangle\right|^2}{\left|\langle \boldsymbol{\xi}^{l_1}, \mathbf{a}^j \rangle\right|^2}. \tag{4.73}$$

Finally, summing the products $\breve{\mathbf{\Gamma}}_{l_2}(\mathbf{\Gamma}_{l_1})^{-1}\mathbf{\Gamma}_{l_2}(\breve{\mathbf{\Gamma}}_{l_1})^{-1}$ for $l_1 = 1$ and $2 \leq l_2 \leq L$ leads to the matrix $\mathbf{M}$ in (4.25) with the summed eigenvalues as given in (4.26).

5

Reference based source extraction exploiting second order temporal structure

Several algorithms can be found in the literature that exploit reference signals in order to extract a desired source from a mixture of multiple sources; however, immediate extraction of the desired source is not always guaranteed for these algorithms. Furthermore, these algorithms utilize higher order statistics, which are difficult to estimate in practice. In this chapter we show that the informed source extraction algorithms from Chapter 3 and Chapter 4 can be reformulated as reference based source extraction algorithms and that a proper design of the reference signals ensures extraction of the desired source. We consider the reference based source extraction problem in the context of a wireless acoustic sensor network.

First we show that both information about the mixing parameters and the autocorrelation function of the desired source can be used for creating or deriving proper reference signals that lead to extraction of the desired source. Furthermore, for applications where the number of sensors is much larger than the number of sources we show that using a subset of the sensors as a reference for source selection and another, non-overlapping, subset of sensors for source extraction is advantageous with respect to spatially uncorrelated noise. Finally, we validate our algorithms via computer simulations.

5.1 Introduction

Traditionally, in speech enhancement applications desired speech is captured with a microphone close to or worn by the speaker. In modern applications microphone arrays capture desired speech from a distance, allowing for hands-free communication. In these applications, knowledge about the array geometry and the location of the desired speaker is combined with beamforming techniques in order to extract the desired speech from the microphone signals [5]. Nowadays, many researchers focus on the development of wireless (acoustic) sensor networks. An advantage of using a wireless acoustic sensor network (WASN) instead of a microphone array for speech capture is the wider spatial coverage, allowing for both hands-free and legs-free communication. However, the array geometry or spatial distribution of the wireless sensor nodes is typically not known and cannot be determined accurately.

In the literature, many blind source extraction (BSE) algorithms are presented in different fields of application that do not require knowledge about the array geometry [7, 8, 11]. The goal of a BSE algorithm is to extract one source from a mixture of mutually independent sources while observing only the mixed signals. In order to guarantee extraction of the desired source with a BSE algorithm, a classifier is required in a second stage to determine if the desired source has been extracted. If not the desired source has been extracted, then a new source has to be extracted and classified. This two-stage procedure is naturally inefficient and is only applicable in batch mode.

In order to extract immediately the desired source from a mixture of multiple sources, researchers have started to incorporate a priori information in the form of reference signals in their source extraction algorithms [34, 41, 42]. These reference based algorithms rely on a threshold parameter that is related to a certain distance measure between the reference signal and the extracted signal; however, no technique is presented to choose the parameter value such that immediate extraction of the desired source is ensured. Furthermore, these algorithms only work for real-valued instantaneous mixtures and exploit higher order statistics of the sources, which are difficult to estimate in practice. Another, well known reference based method for source extraction based on second order statistics is the minimum mean squared error (MMSE) approach, which leads to a (multichannel) Wiener filter [12]. This algorithm extracts the desired source from the mixture if the reference is correlated only with the desired source. Since the acquisition of such a reference is quite a challenge in practice, alternative algorithms are desired.

In this chapter we focus on design techniques for source extraction algorithms that exploit reference signals. In Chapters 3 and 4 we have shown that informed source extraction algorithms can be designed based on different types of a priori information about physical parameters of the sources. For these algorithms, immediate extraction of the desired source is guaranteed if certain conditions on the a priori information hold. In the current chapter we show that the informed source extraction algorithms from Chapters 3 and 4 can be formulated as reference based source extraction algorithms. Subsequently we derive new reference based source extraction algorithms for

applications where the number of sensors is larger than the number of sources. By using a subset of the sensors to generate reference signals and the remaining sensors for source extraction the source extraction filter identification algorithm is shown to be insensitive to spatially uncorrelated noise such as sensor noise. The step towards using reference signals opens up new opportunities for the design of sensing devices and algorithms that generate reference signals that can be used to guarantee immediate extraction of the desired source.

The outline of this work is as follows. In Section 5.2 we discuss the model and assumptions for the mixing scenario. Subsequently, we present two source extraction algorithms based on temporal reference signals in Section 5.3 and two source extraction algorithms based on spatial reference signals in Section 5.4. In Section 5.5 we discuss results from computer simulations and in Section 5.6 we discuss our work and present conclusions.

Notation of vectors and matrices is in bold face lower and upper case letters respectively. A row vector carries an additional tilde symbol, e.g., $\tilde{\mathbf{w}}$. Superscript and subscript indices denote column and row indices, respectively, e.g. a_i^j is the element on the i'th row and j'th column of the matrix $\mathbf{A}$. An index set is denoted with a calligraphic letter. Unless stated otherwise, the set has indices from 1 until the letter represented by the calligraphic letter, e.g., $\mathcal{S} = [1, S]$.

5.2 Model and assumptions

First we introduce the structure of and assumptions on the mixing model, followed by a discussion on the structure and the goal of extraction filters. Subsequently, we present two strategies for obtaining reference signals that can be used to identify source extraction filters. Finally, we discuss statistical properties of the signals.

5.2.1 Mixing model

An overview of the model for the considered scenario is depicted in Figure 5.1. We assume that D sensors observe an instantaneous mixture of S mutually uncorrelated source signals, and additive noise. These instantaneous mixtures have the following structure:

$$x_i[n] = \sum_{j \in \mathcal{S}} a_i^j s_j[n] + \nu_i[n] \qquad \forall\, i \in \mathcal{D} \tag{5.1}$$

where $s_j[n]$, $x_i[n]$, and $\nu_i[n]$ represent respectively the signal of the j'th source, the i'th sensor, and the noise at the i'th sensor, at time-frame n. The complex-valued scalar a_i^j represents the mixing coefficient from the j'th source to the i'th.

By considering all sensor signals simultaneously we can use the following vector

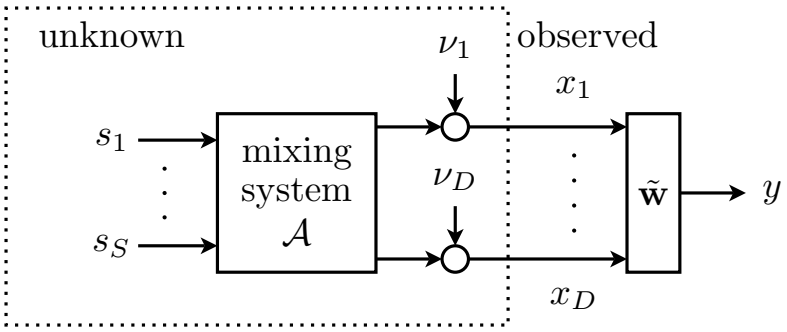


Figure 5.1: Overview of the mixing system where a mixture of S source signals is observed by D sensors. The D sensor signals are combined by the extraction filter $\tilde{\mathbf{w}}$. The mixing system $\mathcal{A}$ is assumed to be a linear, time invariant mixing system.

matrix notation:

$$\mathbf{x}[n] = \sum_{j \in \mathcal{S}} \mathbf{a}^j s_j[n] + \boldsymbol{\nu}[n] = \mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n] \tag{5.2}$$

where the complex-valued vectors containing respectively the sensor, source, and noise signals are defined as follows:

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x_1[n] \\ \vdots \\ x_D[n] \end{bmatrix}, \quad \mathbf{s}[n] \triangleq \begin{bmatrix} s_1[n] \\ \vdots \\ s_S[n] \end{bmatrix}, \quad \boldsymbol{\nu}[n] \triangleq \begin{bmatrix} \nu_1[n] \\ \vdots \\ \nu_D[n] \end{bmatrix}$$

and the mixing matrix of size $D \times S$ is defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}^1 & \cdots & \mathbf{a}^S \end{bmatrix} = \begin{bmatrix} a_1^1 & \cdots & a_1^S \\ \vdots & \ddots & \vdots \\ a_D^1 & \cdots & a_D^S \end{bmatrix}. \tag{5.3}$$

We assume that we have at least the same amount of sensors as sources, i.e., $D \geq S$. Finally, we assume that the mixing columns are linearly independent such that the mixing matrix has rank S.

5.2.2 Source extraction filters

The objective of a source extraction algorithm is to identify a filter $\tilde{\mathbf{w}}$ that extracts the desired source from the observations, i.e.,

$$y[n] = \tilde{\mathbf{w}}\mathbf{x}[n] = \tilde{\mathbf{w}}\mathbf{A}\mathbf{s}[n] + \tilde{\mathbf{w}}\boldsymbol{\nu}[n] \tag{5.4}$$

Ideally, the output $y[n]$ corresponds exactly to the desired source $s_d[n]$, with $d \in \mathcal{S}$ corresponding to the index of the desired source. However, from linear algebra it

follows immediately that exact recovery of the desired source is not possible in the presence of noise. Therefore an alternative goal has to be chosen for the extraction filter. Different objectives for the extraction filter have been considered in the literature. Natural objectives for source extraction are to maximize the desired source to interference plus noise ratio (SINR) and to maximize the desired source to noise ratio (SNR) under the constraint that the interfering sources are canceled. These objectives lead to respectively the minimum variance distortionless response (MVDR) filter and the linearly constrained minimum variance (LCMV) filter, which are widely used in the field of speech enhancement. These filters are the solutions to respectively the following optimization problems:

$$\tilde{\mathbf{w}}_{\text{MVDR}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{a}^d = 1 \tag{5.5}$$

$$\tilde{\mathbf{w}}_{\text{LCMV}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{A} = \tilde{\mathbf{e}}_d \tag{5.6}$$

where $\mathbb{E}\{\cdot\}$ is the mathematical expectation operator, $d \in \mathcal{S}$ is the index corresponding to the desired source, and $\tilde{\mathbf{e}}_d$ is a row vector of length S with $S-1$ zeros and a one at the d'th position. Notice that the LCMV filter only exists if the number of sensors is at least equal to the number of sources. Since the mixing system is unknown, both these MVDR and LCMV filters cannot be calculated directly.

In the literature mainly source extraction algorithms are found that extract the desired source according to the LCMV filtering principle. These filters cancel or block the undesired or interfering sources such that a single source is obtained at the output. A drawback of these filters is that noise is amplified to impermissible levels if the mixing system is ill conditioned. Consequently, identification of the MVDR extraction filter, which balances between interference and noise reduction, is also desired. Later in this chapter we derive an algorithm that can be used to identify both these source extraction filters.

5.2.3 Reference signals and systems

In many papers on reference-based source extraction the reference signals are given without a discussion on the way how to obtain or design them. In this work we consider two filter types to obtain reference signals, i.e., temporal and spatial filters. We show that immediate extraction of the desired source can be guaranteed by incorporating a priori information into the design of these reference filters. For each reference filter type we consider two approaches to generate reference signals, leading to their own algorithms and conditions.

In the first approach we generate the reference signals $\boldsymbol{\upsilon}[n]$ by filtering all the observations with a reference system $\mathbf{T}$, as is depicted in Figure 5.2. The structure of reference signals for a reference system consisting of temporal filters is as follows:

$$\upsilon_{l,i}[n] = (t_l * x_i)[n] \triangleq \sum_{\tau \in \mathcal{T}} t_l[\tau] x_i[n-\tau] \quad \forall\,(i,l) \in \mathcal{D} \times \mathcal{L} \tag{5.7}$$

where $t_1, \cdots, t_L$ represent the impulse responses of the L temporal filters and $\mathcal{T}$ represents the support of the impulse responses. If the reference system consists of

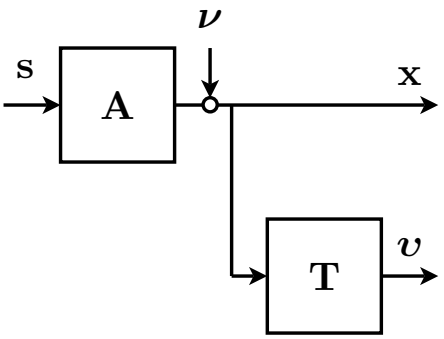


Figure 5.2: In the first approach the reference signals $\boldsymbol{\upsilon}$ are derived by filtering all observations $\mathbf{x}$ and the desired source is obtained by filtering all available sensor signals with an extraction filter.

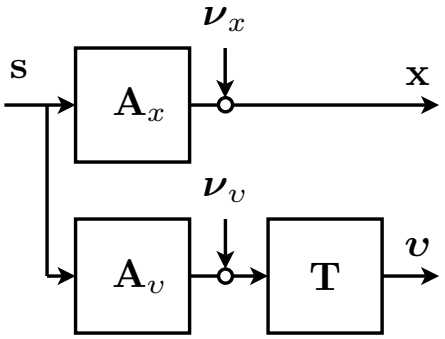


Figure 5.3: In the second approach the reference signals $\boldsymbol{\upsilon}$ are derived by filtering a subset of D_υ sensor signals. The desired source is extracted by filtering the signals from the remaining signals $\mathbf{x}$.

spatial filters, then the reference signals have the following structure:

$$\upsilon_l[n] = \sum_{i=1}^{D} t_l^i x_i[n] = \tilde{\mathbf{t}}_l \mathbf{x}[n] \quad \forall\, l \in \mathcal{L} \tag{5.8}$$

where the length of the L spatial filters $\tilde{\mathbf{t}}_l = \begin{bmatrix} t_l^1 & \cdots & t_l^D \end{bmatrix}$ is D. Notice that the number of reference signals derived from temporal filters equals DL, while the number of reference signals derived from spatial filters equals L.

In Sections 5.3 and 5.4 we show for respectively temporal and spatial reference filters that this approach for generating reference signals can be used to derive source extraction algorithms that are closely linked to the informed source extraction algorithms from Chapter 4. Consequently, we can reuse the design techniques from that chapter in order to design the reference filters such that immediate extraction of the desired source is guaranteed.

In the second approach, depicted in Figure 5.3, we split the set of sensor signals into two non-overlapping subsets $\mathcal{D}_x$ and $\mathcal{D}_\upsilon$ of size D_x and D_υ, respectively, such that $D_x + D_\upsilon = D$. The observation subset $\mathcal{D}_x$ consists of D_x sensor signals and the desired source is extracted by filtering these signals with an extraction filter; therefore, we require that $D_x \geq S$. The remaining D_υ sensor signals form the reference subset $\mathcal{D}_\upsilon$. These signals are filtered with the reference system $\mathbf{T}$ such that identification of the filter parameters that extract the desired source can be ensured. The structure of the reference signals obtained from the temporal reference filters is as follows:

$$\upsilon_{l,i}[n] = (t_l * x_i)[n] \quad \forall\, (i,l) \in \mathcal{D}_\upsilon \times \mathcal{L}. \tag{5.9}$$

For the system consisting of spatial reference filters the reference signals have the

following structure:

$$v_l[n] = \sum_{i \in \mathcal{D}_v} t_l^i x_i[n] = \tilde{\mathbf{t}}_l(\mathbf{A}_v \mathbf{s}[n] + \boldsymbol{\nu}_v[n]) \quad \forall\, l \in \mathcal{L} \tag{5.10}$$

where the length of the spatial reference filters is D_v and $\boldsymbol{\nu}_v[n]$ represents the reference signal subset of $\boldsymbol{\nu}[n]$. For this approach the number of reference signals is $D_v L$ and L for temporal and spatial reference filters, respectively.

In Sections 5.3 and 5.4 we show for respectively temporal and spatial reference filters that this second approach for generating reference signals is advantageous in case little a priori information about the noise is available. The cost for this reduction in required a priori information is in the sense that a sub-optimal filter is obtained with respect to the number of sensors, i.e., a filter of length D_x is identified while the number of sensors equals $D = D_x + D_v$.

For both approaches the reference filters can be implemented in the digital domain. The temporal reference filter can be implemented for example in the form of a finite impulse response (FIR) filter, while the spatial reference filter can be implemented in the form of a beamformer. Additionally, the second approach allows for the use of alternative sensing devices, specifically designed for generating reference signals. Before we discuss the reference based source extraction algorithms and the reference filter design we present our assumptions on the second order temporal structure (SOTS) of the signals.

5.2.4 Second order statistics of signals

In order to be able to identify extraction filters we have to make some assumptions on statistical features of the signals. A major advantage of second order statistics (SOS) compared to higher order statistics is the robustness and efficiency of the estimation procedure. In order to obtain a good estimate of higher order statistical coefficients typically a large amount of stationary data is required. Consequently, we consider only the SOTS of the signals. We first introduce definitions for the SOTS of signals in terms of auto- and crosscorrelation functions. Subsequently, we discuss the SOTS of the sensor signals and present assumptions on the source and noise signals. These assumptions on the source and noise signals are valid in many practical situations and make source extraction filter identification feasible.

Definition 5.2.1 (Auto- and crosscorrelation functions)**:** The auto- and crosscorrelation functions for complex-valued signals $p_i[n]$ and $q_j[n]$ at time-frame n and lag k are defined as follows:

$$r_{ij}^{pq}[n,k] \triangleq \mathbb{E}\{p_i[n] q_j^*[n-k]\} \tag{5.11}$$

where $\mathbb{E}$ is the mathematical expectation operator. The function is called an autocorrelation function if a signal is correlated with itself. Furthermore, if the two signals carry the same symbol, say p, then we use the following short-hand notation: $r_{ij}^{p}[n,k] = r_{ij}^{pp}[n,k]$.

We assume that the mathematical expectation operator $\mathbb{E}$ can be approximated by taking local averages over the time-frame index n, which is a widely used assumption. By using both the time-frame n and lag index k we are able to apply our techniques for both stationary and non-stationary signals. If the signals are non-stationary, then the autocorrelation functions are likely to vary over different time-frame instances n.

Based on Definition 5.2.1, we define the following sensor, reference, source, and noise correlation functions, respectively:

$$\begin{aligned}
r^{x}_{i_1 i_2}[n,k] &\triangleq \mathbb{E}\{x_{i_1}[n]x^*_{i_2}[n-k]\} \;\; \forall\,(i_1,i_2)\in\mathcal{D}\times\mathcal{D}\\
r^{\upsilon}_{l_1 l_2}[n,k] &\triangleq \mathbb{E}\{\upsilon_{l_1}[n]\upsilon^*_{l_2}[n-k]\} \;\; \forall\,(l_1,l_2)\in\mathcal{L}\times\mathcal{L}\\
r^{s}_{j_1 j_2}[n,k] &\triangleq \mathbb{E}\{s_{j_1}[n]s^*_{j_2}[n-k]\} \;\; \forall\,(j_1,j_2)\in\mathcal{S}\times\mathcal{S}\\
r^{\nu}_{i_1 i_2}[n,k] &\triangleq \mathbb{E}\{\nu_{i_1}[n]\nu^*_{i_2}[n-k]\} \;\; \forall\,(i_1,i_2)\in\mathcal{D}\times\mathcal{D}.
\end{aligned}$$

Notice that these correlation functions are only called autocorrelation functions if their indices are the same, e.g., $i_1 = i_2$. Next to these auto- and crosscorrelation functions between signals of the same class, we define the crosscorrelation functions between the noise and source signals as follows:

$$\begin{aligned}
r^{\nu s}_{ij}[n,k] &\triangleq \mathbb{E}\{\nu_i[n]s^*_j[n-k]\} \;\; \forall\,(i,j)\in\mathcal{D}\times\mathcal{S}\\
r^{s\nu}_{ji}[n,k] &\triangleq \mathbb{E}\{s_j[n]\nu^*_i[n-k]\} \;\; \forall\,(i,j)\in\mathcal{D}\times\mathcal{S}.
\end{aligned}$$

For the second type of reference system the total number of sensors equals D_x instead of D; however, for ease of notation we keep using here the symbol $\mathcal{D}$. By denoting the crosscorrelation functions between the observations and the reference signals obtained from the temporal reference filters by $r^{x\upsilon}_{i_1(l,i_2)}[n,k]$ and $r^{\upsilon x}_{(l,i_1)i_2}[n,k]$, we obtain the following definitions:

$$\begin{aligned}
r^{x\upsilon}_{i_1(l,i_2)}[n,k] &\triangleq \mathbb{E}\{x_{i_1}[n]\upsilon^*_{l,i_2}[n-k]\} \;\; \forall\,(i_1,i_2,l)\in\mathcal{D}\times\mathcal{D}\times\mathcal{L}\\
r^{\upsilon x}_{(l,i_1)i_2}[n,k] &\triangleq \mathbb{E}\{\upsilon_{l,i_1}[n]x^*_{i_2}[n-k]\} \;\; \forall\,(i_1,i_2,l)\in\mathcal{D}\times\mathcal{D}\times\mathcal{L}.
\end{aligned}$$

Finally, we obtain the following definitions of the crosscorrelation functions between the observations and the reference signals obtained from the spatial reference filters:

$$\begin{aligned}
r^{x\upsilon}_{il}[n,k] &\triangleq \mathbb{E}\{x_i[n]\upsilon^*_l[n-k]\} \;\; \forall\,(i,l)\in\mathcal{D}\times\mathcal{L}\\
r^{\upsilon x}_{li}[n,k] &\triangleq \mathbb{E}\{\upsilon_l[n]x^*_i[n-k]\} \;\; \forall\,(i,l)\in\mathcal{D}\times\mathcal{L}.
\end{aligned}$$

In the following section we use the auto- and crosscorrelation functions defined in this section in order to present and discuss assumptions on the SOTS of the source and noise signals. These assumptions are widely used and hold in many applications [8, 11, 36]. Furthermore, they are sufficient to make extraction filter identification based on SOTS feasible [61].

5.2.5 Assumptions on the sensor and reference signals crosscorrelation functions

The first assumption is that the source signals are mutually uncorrelated, i.e.,

$$r^{s}_{j_1 j_2}[n,k] = 0 \qquad \forall\,\{(j_1,j_2) \in \mathcal{S}\times\mathcal{S} |\, j_1 \neq j_2\}. \tag{5.12}$$

This assumption is widely used in the field of blind signal processing. The assumption is less restrictive than the assumption of independent sources, which means that the sources are not related at any order statistics. Besides assumptions on the source signals we also require assumptions on the SOS of noise signals. The assumptions we make on the SOS of the noise depends on the method used to obtain reference signals. For the first method, see Figure 5.2, we present the assumptions in the form of a noise-free region of support (NF-ROS).

Definition 5.2.2 (Noise-free region of support)**:** The NF-ROS, also denoted with Ω, consists of a set of time-lag pairs $\Omega_\kappa = (n,k)_\kappa$ for which the following conditions hold [36, 70]:

$$\forall\,\Omega_\kappa \in \Omega : \begin{cases} r^{\nu}_{i_1 i_2}[\Omega_\kappa] = 0 & \forall\,(i_1,i_2) \in \mathcal{D}\times\mathcal{D} \\ r^{\nu s}_{ij}[\Omega_\kappa] = 0 & \forall\,(i,j) \in \mathcal{D}\times\mathcal{S} \\ r^{s\nu}_{ji}[\Omega_\kappa] = 0 & \forall\,(i,j) \in \mathcal{D}\times\mathcal{S} \end{cases}$$

where $\Omega \triangleq \{(n,k)_\kappa \,|\, \kappa \in \mathcal{K}\}$ and $\mathcal{K} = [1,K]$ is the index set for time-lag pairs in the NF-ROS. The number of elements in the NF-ROS is indicated by the letter K, i.e., $K = |\Omega|$.

In short, the NF-ROS consists of a set of time-lag pairs for which only the source signals are correlated.

For the second method, see Figure 5.3, we change the first requirement in the definition of the NF-ROS. Instead of having no noise correlation we put the following requirements:

$$\forall\,\Omega_\kappa \in \Omega : \begin{cases} r^{\nu}_{i_1 i_2}[\Omega_\kappa] = 0 & \forall\,\{(i_1,i_2) \in \mathcal{D}_x\times\mathcal{D}_v |\, i_1 \neq i_2\} \\ r^{\nu}_{i_1 i_2}[\Omega_\kappa] = 0 & \forall\,\{(i_1,i_2) \in \mathcal{D}_v\times\mathcal{D}_x |\, i_1 \neq i_2\} \\ r^{\nu s}_{ij}[\Omega_\kappa] = 0 & \forall\,(i,j) \in \mathcal{D}\times\mathcal{S} \\ r^{s\nu}_{ji}[\Omega_\kappa] = 0 & \forall\,(i,j) \in \mathcal{D}\times\mathcal{S} \end{cases}.$$

Note that we require the noise to be spatially uncorrelated between a sensor from the observation set $\mathcal{D}_x$ and a sensor from the reference set $\mathcal{D}_v$. An example of such a noise is spatially uncorrelated sensor noise. Even though this set of time-lag pairs is not free of noise, we keep referring to the NF-ROS and mention explicitly if spatially uncorrelated noise is allowed.

In the next sections we present algorithms that can be used to extract the desired source based on the models and assumptions presented in this section.

5.3 Temporal reference filter based source extraction algorithms

In this section we present two reference based source extraction algorithms that exploit reference signals obtained by filtering observations with temporal filters. First, we discuss the structure in sensor-reference crosscorrelation functions, where we immediately consider the NF-ROS and the application of reference filters. Subsequently, we show for the approach from Figure 5.2 that the structure in the sensor-reference crosscorrelation functions is very similar to the structure of the linear combinations of sensor correlation matrices in (4.23) of Chapter 4. Based on this observation we show that the reference system can be designed such that immediate extraction of the desired source is guaranteed. Finally, we present a temporal reference based source extraction algorithm based on the approach from Figure 5.3. This algorithm deals with spatially uncorrelated noise in an alternative way at the cost of using additional sensors.

5.3.1 Structure in the SOS of the sensor and reference signals

By following the definitions and the assumptions from Section 5.2.4 and considering the structure of the observations and the reference signals in (5.1) and (5.7), respectively, we observe the following linear relation between the sensor-reference crosscorrelation functions and the sensor correlation functions from (4.14) in Chapter 4:

$$r^{xv}_{i_1(l,i_2)}[n,k] = \sum_{\tau\in\mathcal{T}} t^*_l[\tau] r^x_{i_1 i_2}[n,k+\tau] \quad \forall\,(i_1,i_2,l)\in\mathcal{D}\times\mathcal{D}\times\mathcal{L} \tag{5.13}$$

where $\mathcal{T}$ represents the support of the temporal filters and $t_l[\tau]$ represents the impulse response of the l'th reference filter.

By using the definition of the NF-ROS in Definition 5.2.2 we know that the sensor-reference crosscorrelation functions are free of noise at time-lag pair (n,k) if the time-lag pairs $(n,k+\tau)$ fall in the NF-ROS for all $\tau\in\mathcal{T}$. Consequently, a proper design of the reference filters and selection of time-lag pairs based on the NF-ROS is important. In Section 5.3.2.1 we discuss this design of the temporal reference filters and the choice of time-lag pairs such that the sensor-reference crosscorrelation functions are free of noise. For now we assume that the time-lag pairs and the support of the temporal reference filters are chosen in such a way that the sensor-reference crosscorrelation data is free of noise. Using this assumption, we obtain the following, noise-free, structure in these crosscorrelation functions:

$$r^{xv}_{i_1(l,i_2)}[n,k] = \sum_{j\in\mathcal{S}} a^j_{i_1} a^{j,*}_{i_2} \sum_{\tau\in\mathcal{T}} t^*_l[\tau] r^s_{jj}[n,k+\tau] \quad \forall\,(n,k+\tau)\in\Omega \tag{5.14}$$

where we used the structure of the noise-free sensor correlation functions presented in (4.14) of Chapter 4, which is as follows:

$$r^x_{i_1 i_2}[n,k] = \sum_{j\in\mathcal{D}} a^j_{i_1} a^{j,*}_{i_2} r^s_{jj}[n,k]. \tag{5.15}$$

By evaluating (5.13) over all sensor indices i_1 and i_2 for each temporal reference filter we obtain the following sets of sensor-reference crosscorrelation matrices:

$$\mathbf{R}_l^{xv}[n,k] \triangleq \begin{bmatrix} r^{xv}_{1(l,1)}[n,k] & r^{xv}_{1(l,2)}[n,k] & \cdots & r^{xv}_{1(l,D)}[n,k] \\ r^{xv}_{2(l,1)}[n,k] & r^{xv}_{2(l,2)}[n,k] & \cdots & r^{xv}_{2(l,D)}[n,k] \\ \vdots & \ddots & \ddots & \vdots \\ r^{xv}_{D(l,1)}[n,k] & r^{xv}_{D(l,2)}[n,k] & \cdots & r^{xv}_{D(l,D)}[n,k] \end{bmatrix} \quad \forall\, l \in \mathcal{L}. \tag{5.16}$$

By exploiting the structure in the sensor-reference crosscorrelation data presented in (5.14), we obtain the following structure in the sensor-reference crosscorrelation matrices in terms of the mixing system, temporal reference filters, and the source autocorrelation functions:

$$\mathbf{R}_l^{xv}[n,k] \equiv \mathbf{A} \sum_{\tau\in\mathcal{T}} t_l^*[\tau]\, \mathrm{diag}\left(\mathbf{r}^s[n,k+\tau]\right) (\mathbf{A})^H = \sum_{\tau\in\mathcal{T}} t_l^*[\tau]\mathbf{R}^x[n,k+\tau]$$

where $\mathbf{r}^s[n,k] = \left(\begin{bmatrix} r^s_{11}[n,k] & \cdots, r^s_{SS}[n,k] \end{bmatrix}\right)^T$.

The sensor-reference crosscorrelation matrices are linear combinations of a set of sensor correlation matrices $\mathbf{R}^x[n,k+\tau]$, where the weights are determined by the impulse response of the considered temporal filter. The sizes of these sensor-reference crosscorrelation matrices are $D \times D$; however, from their structure it follows immediately that their rank is at most S. Consequently, by applying the subspace matrix that is related to the mixing matrix, i.e., $\mathbf{P}_s$, to this sensor-reference crosscorrelation matrix we obtain, without loosing information, the following size $S \times S$ crosscorrelation matrices:

$$\bar{\mathbf{R}}_l^{xv}[n,k] \triangleq \mathbf{P}_s \mathbf{R}_l^{xv}[n,k](\mathbf{P}_s)^H = \sum_{t\in\mathcal{T}} t_l^*[\tau]\bar{\mathbf{R}}^x[n,k+\tau] \quad \forall\, l \in \mathcal{L} \tag{5.17}$$

where $\bar{\mathbf{R}}^x[n,k+\tau] = \bar{\mathbf{A}}\mathbf{R}^s[n,k+\tau](\bar{\mathbf{A}})^H$ are reduced sensor correlation matrices with the same structure as the matrix at the left hand side in (4.22). The subspace matrix $\mathbf{P}_s$ was already presented in Section 4.3.1 and can be estimated for example from the singular value decomposition of one of the sensor reference crosscorrelation matrices $\mathbf{R}_l^{xv}[n,k]$, i.e.,

$$\mathrm{SVD}\left(\mathbf{R}_l^{xv}[n,k]\right) = \mathbf{U}^s \mathbf{\Sigma}_S \mathbf{V}_s \tag{5.18}$$

where $\mathbf{\Sigma}_S$ is a matrix of size $S \times S$ containing singular values on its diagonal and the matrices $\mathbf{U}^s$ and $\mathbf{V}_s$ with sizes $D \times S$ and $S \times D$, respectively, contain their corresponding left and right singular vectors. These left and right singular vectors span the same space as the mixing column vectors in the mixing matrix $\mathbf{A}$. Consequently, the reduction matrix $\mathbf{P}_s$ can be found as $\mathbf{P}_s = (\mathbf{U}^s)^H$ or $\mathbf{P}_s = \mathbf{V}_s$. Once the reduction matrix $\mathbf{P}_s$ is identified, we can apply it either to the correlation matrices or directly to the observations $\mathbf{x}[n]$ leading to $\bar{\mathbf{x}}[n] = \mathbf{P}_s\mathbf{x}[n]$. This latter reduction method will be used in the remainder of this chapter.

In the following section we use these sensor-reference crosscorrelation matrices and show that with a proper design of the temporal reference filters, these matrices can be used to identify source extraction filters.

5.3.2 Temporal reference based source extraction as informed source extraction

From (5.17) it follows that the sensor-reference crosscorrelation matrices are very similar to the linear combinations of sensor correlation matrices $\mathbf{\Gamma}_l$ and $\check{\mathbf{\Gamma}}_l$ in (4.23) of Chapter 4; however, some differences exist. In order to link temporal reference based source extraction and the informed source extraction from Chapter 4, we assume in the first part that the observed signals are stationary such that the SOS are independent of the time-index n. Later in this section we consider non-stationary signals.

For stationary signals it holds that the lag zero sensor-reference crosscorrelation matrix pairs $(\bar{\mathbf{R}}_l^{xv}[n,0], (\bar{\mathbf{R}}_l^{xv}[n,0])^H)$ are equal to the linear combinations of sensor correlation matrix pairs $(\mathbf{\Gamma}_l, \check{\mathbf{\Gamma}}_l)$ in case the following condition holds:

$$\{t_l[k] = \xi_l^\kappa \,|\, \Omega_\kappa = k, \forall\, \Omega_\kappa \in \Omega\} \quad \Longrightarrow \quad \begin{cases} \bar{\mathbf{R}}_l^{xv}[n,0] = \mathbf{\Gamma}_l \\ (\bar{\mathbf{R}}_l^{xv}[n,0])^H = \check{\mathbf{\Gamma}}_l \end{cases} \tag{5.19}$$

where we used the property that $\mathbf{\Gamma}_l$ and $\check{\mathbf{\Gamma}}_l$ in (4.23) of Chapter 4 differ only by a Hermitian transposition, i.e., $\check{\mathbf{\Gamma}}_l = (\mathbf{\Gamma}_l)^H$. The equality in (5.19) follows immediately when considering (4.23) and (5.17).

For non-stationary signals the link between informed and reference based source extraction is not as trivial as for stationary signals. In order to show the connection between reference and informed based source extraction we decompose the NF-ROS into multiple time-lag pair groups where all time-lag pairs in a single group share their time index, i.e., $\Omega = \bigcup \Omega_\eta$, with η the index for time instance n_η. If we apply this decomposition to the linear combinations of sensor correlation matrices $\mathbf{\Gamma}_l$ and $\check{\mathbf{\Gamma}}_l$ for non-stationary signals, then we obtain the following relations:

$$\mathbf{\Gamma}_l = \sum_\eta \bar{\mathbf{R}}_l^{xv}[n_\eta, 0] \quad \text{and} \quad \check{\mathbf{\Gamma}}_l = \sum_\eta (\bar{\mathbf{R}}_l^{xv}[n_\eta, 0])^H. \tag{5.20}$$

This decomposition means that a set of reference filters is required in order to build the matrices $\mathbf{\Gamma}_l$ and $\check{\mathbf{\Gamma}}_l$, i.e., for each time instance considered in the NF-ROS we require a filter. Subsequently, statistical parameters have to be estimated and combined over the time-indices. In practice, this might be a cumbersome procedure; however, we discuss this approach here for the insight it delivers. Furthermore, this approach allows for new ways to develop and design source extraction algorithms. In the following section we discuss the design of temporal reference based source extraction algorithms.

5.3.2.1 Design of temporal reference based source extraction algorithms

The equality between the lag-zero sensor-reference crosscorrelation matrices and the linear combinations of sensor correlation matrices, discussed in the previous section, makes that the design techniques presented in Chapters 3 and 4 can be re-used for a proper design of temporal reference filters such that immediate extraction of the desired source can be guaranteed.

In Chapters 3 and 4 we have seen that extraction of the desired source can be guaranteed based on an estimate $\tilde{\mathbf{r}}_0$ of the autocorrelation function of the desired source for time-lag pairs in the NF-ROS, which we call a mold. For the mold we assume that the condition in (4.30) holds, i.e.,

$$\frac{|\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}^s_{dd} \rangle|}{\|\tilde{\mathbf{r}}^s_{dd}\|_2} > \frac{|\langle \tilde{\mathbf{r}}_0, \tilde{\mathbf{r}}^s_{jj} \rangle|}{\|\tilde{\mathbf{r}}^s_{jj}\|_2} \quad \forall\, j \in \mathcal{S} \backslash d \tag{5.21}$$

which means that $\tilde{\mathbf{r}}^s_0$ is a 'good' estimate of the autocorrelation function corresponding to the desired source. Based on the theory developed in Chapter 4, immediate extraction of the desired source is guaranteed if the impulse response of one reference filter matches the mold and if the other reference filters are chosen as an orthonormal basis. In summary the following condition holds for the design of the reference filters:

$$\tilde{\mathbf{t}}_{l_1} (\tilde{\mathbf{t}}_{l_2})^H = \begin{cases} ||\tilde{\mathbf{r}}_0||_2^2 & \text{for} \quad \{(l_1, l_2) \in \mathcal{L} \times \mathcal{L} |\, l_1 = l_2\} \\ 0 & \text{for} \quad \{(l_1, l_2) \in \mathcal{L} \times \mathcal{L} |\, l_1 \neq l_2\} \end{cases} \tag{5.22}$$

where $\tilde{\mathbf{t}}_1$ equals the mold $\tilde{\mathbf{r}}_0$ and where we did not take the decomposition for non-stationary signals into consideration in the notation.

A disadvantage of the temporal reference based source extraction algorithm is that all D observations have to be filtered by the L reference filters, leading to DL filter operations. Although this approach might seem cumbersome, it strongly depends on the application and the implementation if the use of reference based or informed based source extraction is more appropriate. Furthermore, we have shown only the connection between reference based and informed source extraction, the implementation of the reference filters in the form of digital FIR filters, and, we considered the sensor-reference crosscorrelation data for only lag zero. Consequently, we have restricted ourselves a lot in the design of temporal reference based source extraction algorithms. We only introduced these limitations to discuss the connection with informed source extraction and proof the validity of reference based source extraction. Dropping these limitations leads to opportunities for new design techniques of reference based source extraction algorithms. An example algorithm that is closely related to temporal reference based source extraction is known in the literature as linear prediction based source extraction [14, 51]. We considered this algorithm already in [52] and it can be seen as a special case of temporal reference based source extraction for identification of the LCMV filter. Alternative strategies are to implement reference filters in the analog domain and to consider the sensor-reference crosscorrelation function for not only lag zero.

5.3.2.2 Summary of temporal reference based source extraction algorithm

An overview of the temporal reference based source extraction algorithm is depicted in Figure 5.4. This algorithm can be used to identify the main components $\tilde{\boldsymbol{\mu}}$ and $\mathbf{h}$ of the LCMV and MVDR extraction filter $\tilde{\mathbf{w}}$, respectively. A summary of the steps in the temporal reference based source extraction algorithm is as follows:

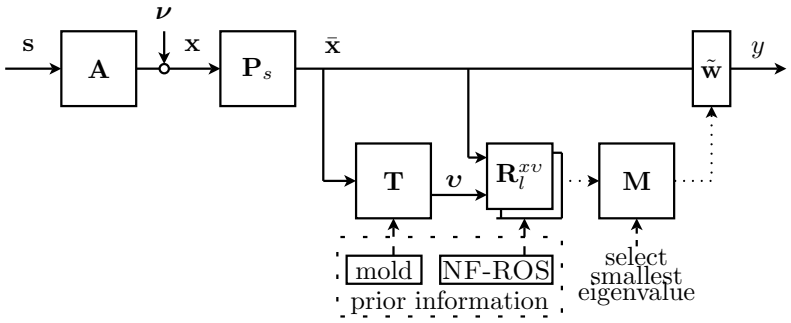


Figure 5.4: Overview of the temporal reference based source extraction algorithm. The *mold* represents the a priori information about the mixing system that is used to design the temporal reference system $\mathbf{T}$.

1. Design a temporal reference system $\mathbf{T}$ and apply it to the observations, i.e., $\upsilon_{l,i}[n] = (t_l * x_i)[n] \quad \forall\,(i,l) \in \mathcal{D} \times \mathcal{L}$.
2. Identify subspace matrix $\mathbf{P}_s$ and calculate the reduced size correlation matrices $\bar{\mathbf{R}}_l^{x\upsilon}$ as in (5.17) or reduce the observations directly.
3. Calculate sensor-reference crosscorrelation matrices $\mathbf{\Gamma}_l$ and $\check{\mathbf{\Gamma}}_l$ as in (5.20).
4. Combine the reduced crosscorrelation matrices in the following way:

$$\mathbf{M} = \sum_{l=2}^{L} \check{\mathbf{\Gamma}}_l \left(\mathbf{\Gamma}_1\right)^{-1} \mathbf{\Gamma}_l \left(\check{\mathbf{\Gamma}}_1\right)^{-1}. \tag{5.23}$$

5. Calculate the left or right eigenvector, $\tilde{\boldsymbol{\mu}}$ or $\mathbf{h}$, that corresponds to the smallest eigenvalue of $\mathbf{M}$.
6. Extract the desired source and apply noise reduction if desired.

The underlying structure of $\mathbf{M} \equiv \bar{\mathbf{A}}\mathbf{\Lambda}\left(\bar{\mathbf{A}}\right)^{-1}$ makes that the left eigenvectors of $\mathbf{M}$ correspond to rows from the inverse of the reduced mixing system $\bar{\mathbf{A}}$, while the right eigenvectors correspond to columns from the reduced mixing system. The eigenvalues, i.e., the elements on the diagonal of $\mathbf{\Lambda}$, are positive and have the following structure:

$$\lambda^j = \sum_{l=2}^{L} \frac{|\langle \tilde{\mathbf{t}}_l, \tilde{\mathbf{r}}_{jj}^2 \rangle|^2}{|\langle \tilde{\mathbf{t}}_1, \tilde{\mathbf{r}}_{jj}^s \rangle|^2} \quad \forall\, j \in \mathcal{S}. \tag{5.24}$$

Each eigenvalue depends on the autocorrelation function of one of the sources and the parameters of the reference system. Consequently, a proper design of the reference system as in Section 5.3.2.1 makes that the left and right eigenvectors that correspond to the smallest eigenvalue can be used to extract the desired source.

The MVDR extraction filter is found by solving the following optimization problem:

$$\begin{aligned} w_{a_{opt}} = \arg\min_{\tilde{\mathbf{w}}_a} \quad & \mathbb{E}\{y[n]y^*[n]\} \\ \text{s.t.} \quad & y[n] = \left((\mathbf{h})^H \mathbf{P}_s - w_a \left((\mathbf{h})^H \mathbf{P}_s \right)^{\perp} \right) \mathbf{x}[n] \end{aligned} \tag{5.25}$$

which leads to $\tilde{\mathbf{w}}_{\mathrm{MVDR}} = (\mathbf{h})^H \mathbf{P}_s - w_{a_{opt}} \left((\mathbf{h})^H \mathbf{P}_s \right)^{\perp}$.

The LCMV extraction filter, including noise reduction, is found by solving the following optimization problem:

$$\begin{aligned} w_{a_{opt}} = \arg\min_{\tilde{\mathbf{w}}_a} \quad & \mathbb{E}\{y[n]y^*[n]\} \\ \text{s.t.} \quad & y[n] = \left(\tilde{\boldsymbol{\mu}} \mathbf{P}_s - w_a \left(\mathbf{P}_s \right)^{\perp} \right) \mathbf{x}[n] \end{aligned} \tag{5.26}$$

which leads to $\tilde{\mathbf{w}}_{\mathrm{LCMV}} = \tilde{\boldsymbol{\mu}} \mathbf{P}_s - w_{a_{opt}} \left(\mathbf{P}_s \right)^{\perp}$.

These extraction filters maximize the SINR or maximize the SNR under the constraint that the interfering sources are canceled, respectively, and exploit all available sensors for this task. In the following section we present an algorithm that uses a subset of the sensors to ensure identification of the desired filter. This filter is applied to the remaining sensors in the system, i.e., not the ones used for desired source selection. An advantage of this new approach is that spatially uncorrelated noise is allowed to have non-zero autocorrelation function values for time-lag pairs in the NF-ROS.

5.3.3 Source extraction in the presence of spatially uncorrelated noise

So far, we have assumed that a NF-ROS exists and is known. In practice, this information is often difficult to obtain and a NF-ROS does not have to exist. In this section we assume that spatially uncorrelated noise has an unknown contribution in the SOTS of the considered sensor signals.

For the model in Figure 5.2 the presence of spatially uncorrelated signals leads to the following structure in the crosscorrelation functions of sensor and reference signals:

$$r^{x\upsilon}_{i_1(l,i_2)}[n,k] = \sum_{j=1}^{S} a^j_{i_1} a^{j,*}_{i_2} \sum_{\tau \in \mathcal{T}} t^*_l[\tau] r^s_{jj}[n, k+\tau] + \sum_{\tau \in \mathcal{T}} t^*_l[\tau] r^{\nu}_{i_1 i_2}[n, k-\tau]. \tag{5.27}$$

It follows that the crosscorrelation functions $r^{x\upsilon}_{i_1(l,i_2)}[n,k]$ are only contaminated by spatially uncorrelated noise if $i_1 = i_2$. Therefore, we split the sensors into two subsets. One set is used for generating reference signals and the other set is used for source extraction, as is depicted in Figure 5.3. This leads to the following sensor reference crosscorrelation matrices that are free of spatially uncorrelated noise:

$$\mathbf{R}^{x\upsilon}_l[n,k] = \mathbf{A}_x \sum_{\tau \in \mathcal{T}} t^*_l[\tau] \operatorname{diag}\left(\mathbf{r}^s[n, k+\tau] \right) \left(\mathbf{A}_\upsilon \right)^H \tag{5.28}$$

where $\mathbf{A}_x$ and $\mathbf{A}_\upsilon$ are the mixing matrices for two non-overlapping subsets of sensor signals, as is discuss in Section 5.2.3.

In order to identify extraction filters from the crosscorrelation matrices in (5.28) we require that the rank of both $\mathbf{A}_x$ and $\mathbf{A}_\upsilon$ is at least S. Consequently, we have the

following condition for the total number of sensor: $D \geq 2S$, i.e., each subset should consist of at least S sensors. Furthermore, we have to reduce the sensor correlation matrices from size $D_x \times D_\upsilon$ to size $S \times S$. This reduction can be accomplished by two reduction matrices $\mathbf{P}_s$ and $\check{\mathbf{P}}_s$ from which the rows span the same space as the space spanned by the columns of the mixing matrices $\mathbf{A}_x$ and $\mathbf{A}_\upsilon$, respectively. These matrices can also be identified from the singular value decomposition of a sensor reference crosscorrelation matrices, i.e.,

$$\text{SVD}\left(\mathbf{R}_l^{x\upsilon}[n,k]\right) = \mathbf{U}^s \mathbf{\Sigma}_S \mathbf{V}_s \tag{5.29}$$

where $\mathbf{P}_s = (\mathbf{U}^s)^H$ and $\check{\mathbf{P}}_s = \mathbf{V}_s$ leads to the desired reduction matrices. Again, these matrices can be applied either to the calculated sensor reference crosscorrelation matrices or directly to the data. Applying these reduction matrices leads to the following reduced sensor reference crosscorrelation matrices:

$$\bar{\mathbf{R}}_l^{x\upsilon}[n,k] = \mathbf{P}_s \mathbf{R}_l^{x\upsilon}[n,k](\check{\mathbf{P}}_s)^H = \bar{\mathbf{A}}_x \sum_{\tau\in\mathcal{T}} t_l^*[\tau]\,\text{diag}\left(\mathbf{r}^s[n,k+\tau]\right)\left(\bar{\mathbf{A}}_\upsilon\right)^H \tag{5.30}$$

where $\bar{\mathbf{A}}_x = \mathbf{P}_s\mathbf{A}_x$ and $\bar{\mathbf{A}}_\upsilon = \check{\mathbf{P}}_s\mathbf{A}_\upsilon$ are reduced mixing matrices.

Since the symmetry of the matrices $\bar{\mathbf{R}}_l^{x\upsilon}[n,k]$ is lost we are not able to calculate the matrix $\check{\mathbf{\Gamma}}_l$ by taking the transpose of $\mathbf{\Gamma}_l$. In order to be able to build a matrix with the structure of $\check{\mathbf{\Gamma}}_l$ we require sensor correlation matrices with the structure as in (5.30), but now with conjugated temporal reference filters and conjugated source autocorrelation function values, without changing the order or conjugation of the mixing systems $\bar{\mathbf{A}}_x$ and $\bar{\mathbf{A}}_\upsilon$. This means that we have to calculate sensor reference crosscorrelation matrices that have the following structure:

$$\check{\mathbf{R}}_l^{x\upsilon}[n,k] = \bar{\mathbf{A}}_x \sum_{\tau\in\mathcal{T}} t_l[\tau]\,\text{diag}\left(\mathbf{r}^s[n,k+\tau]\right)^*\left(\bar{\mathbf{A}}_\upsilon\right)^H. \tag{5.31}$$

These new sensor reference crosscorrelation matrices can be calculated by using the following property of complex-valued non-stationary signals:

$$(r_{jj}^s[n,k])^* = (\mathbb{E}\{s_j[n]s_j^*[n-k]\})^* = \mathbb{E}\{s_j^*[n]s_j[n-k]\} = r_{jj}^s[n-k,-k] \tag{5.32}$$

which reduces to the well known property $(r_{jj}^s[k])^* = r_{jj}^s[-k]$ for stationary signals. Consequently, the sensor reference crosscorrelation matrices with the structure in (5.31) can be calculated as follows:

$$\check{\mathbf{R}}_l^{x\upsilon}[n,k] = \bar{\mathbf{R}}_{l,*}^{x\upsilon}[n-k,-k] \tag{5.33}$$

where $l,*$ means that the conjugated temporal reference filter should be used.

Now the matrices $\mathbf{\Gamma}_l$ and $\check{\mathbf{\Gamma}}_l$ can be calculated as linear combinations of the reduced sensor reference crosscorrelation matrices $\bar{\mathbf{R}}_l^{x\upsilon}[n,0]$ and $\check{\mathbf{R}}_l^{x\upsilon}[n,0]$, as was shown on the left hand side in (5.20). Again, these matrices can be used to construct the following matrix $\mathbf{M}$ from which source extraction filters can be identified as eigenvectors:

$$\mathbf{M} = \sum_{l=2}^{L} \check{\mathbf{\Gamma}}_l \left(\mathbf{\Gamma}_1\right)^{-1} \mathbf{\Gamma}_l \left(\check{\mathbf{\Gamma}}_1\right)^{-1} \equiv \bar{\mathbf{A}}_x \mathbf{\Lambda} \left(\bar{\mathbf{A}}_x\right)^{-1}. \tag{5.34}$$

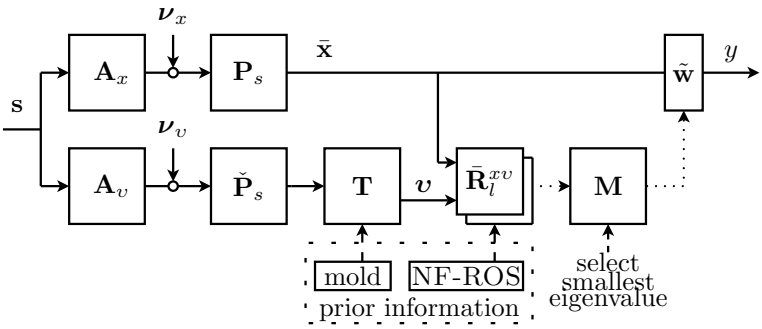


Figure 5.5: Overview of the temporal reference based source extraction algorithm that deals with spatially uncorrelated noise.

Each eigenvalue of the matrix $\mathbf{M}$ has a structure as in (5.24), i.e., it corresponds only to the temporal reference filters and the autocorrelation function of one of the sources; and, its corresponding left and right eigenvectors are extraction filters for that source with respect to the reduced mixing matrix $\bar{\mathbf{A}}_x$.

An overview of this algorithm for spatially uncorrelated noise is depicted in Figure 5.5. Extraction of the desired source with this new algorithm can be guaranteed by meeting the following two conditions. First, the reference system has to be designed in a similar way as is discussed in Section 5.3.2.1. Second, both mixing matrices $\mathbf{A}_x$ and $\mathbf{A}_\upsilon$ have to be well conditioned, i.e., they have to have rank S.

5.4 Spatial reference filter based source extraction algorithms

In this section we present two reference based source extraction algorithms for reference signals obtained from spatial filters. These algorithms are derived as follows. First, we discuss the structure in sensor-reference crosscorrelation functions. We show for the approach from Figure 5.2 that this structure is very similar to the structure of the linear combinations of sensor correlation matrices in (4.24) of Chapter 4. Based on this observation we show that the reference system can be designed such that immediate extraction of the desired source is guaranteed. Finally, we present a spatial reference based source extraction algorithm based on the approach from Figure 5.3. This algorithm deals with spatially uncorrelated noise in an alternative way at the cost of using additional sensors.

5.4.1 Structure in the SOS of the sensor and reference signals

The structure in the sensor-reference crosscorrelation functions is strongly depending on the type of reference filter that is used. In this section we discuss the structure of these crosscorrelation functions for reference signals obtained with spatial filters. For time-lag pairs in the NF-ROS, the structure of the sensor-reference crosscorrelation

functions is obtained by considering the structure of the observations (5.1) and the structure of the reference signals in (5.8); and, is as follows:

$$r_{il}^{xv}[\Omega_\kappa] = \sum_{j=1}^{S} a_i^j b_l^{j,*} r_{jj}^s[\Omega_\kappa] \quad \forall\,(i,l,\kappa) \in \mathcal{D} \times \mathcal{L} \times \mathcal{K} \tag{5.35}$$

$$r_{li}^{vx}[\Omega_\kappa] = \sum_{j=1}^{S} b_l^j a_i^{j,*} r_{jj}^s[\Omega_\kappa] \quad \forall\,(i,l,\kappa) \in \mathcal{D} \times \mathcal{L} \times \mathcal{K} \tag{5.36}$$

where $b_l^j = \tilde{\mathbf{t}}_l \mathbf{a}^j$ and we define the reference mixing system $\mathbf{B} \triangleq \mathbf{TA}$.

The structure of these crosscorrelation functions is very similar to the structure in the sensor correlation functions in (4.14) of Chapter 4, which we repeat here for completeness:

$$r_{i_1 i_2}^x[\Omega_\kappa] = \sum_{j=1}^{S} a_{i_1}^j a_{i_2}^{j,*} r_{jj}^s[\Omega_\kappa] \quad \forall\,(i_1,i_2,k) \in \mathcal{D} \times \mathcal{D} \times \mathcal{K}.$$

By exploiting the fact that the reference mixing column vectors $\mathbf{b}^j$ are linear combinations of mixing column vectors $\mathbf{a}^j$, it follows that the sensor-reference crosscorrelation functions $r_{il}^{xv}[\Omega_\kappa]$ and $r_{li}^{vx}[\Omega_\kappa]$ are linear combinations of the sensor correlation functions $r_{i_i i_2}^x[\Omega_\kappa]$, i.e.,

$$r_{i_1 l}^{xv}[\Omega_\kappa] = \sum_{i_2=1}^{D} t_l^{i_2,*} r_{i_1 i_2}^x[\Omega_\kappa], \quad r_{li_2}^{vx}[\Omega_\kappa] = \sum_{i_1=1}^{D} t_l^{i_1} r_{i_1 i_2}^x[\Omega_\kappa].$$

By evaluating the sets of crosscorrelation functions over all sensor indices i, reference indices l, and time-lag pair indices κ we are able to generate a cube of crosscorrelation data for each set of crosscorrelation functions. For the models in (5.35) and (5.36) three essentially unique slices of correlation data exist that can be represented by three differently structured correlation matrices. These matrices are constructed by fixing one of the three indices, i, l, or κ, and evaluating the functions over the other two indices. In this section we consider the sensor-reference crosscorrelation matrices obtained by fixing the reference signal index l and evaluating over the index i and κ in the following way:

$$\mathbf{C}_l^{xv} \triangleq \begin{bmatrix} r_{1l}^{xv}[\Omega_1] & r_{1l}^{xv}[\Omega_2] & \cdots & r_{1l}^{xv}[\Omega_K] \\ r_{2l}^{xv}[\Omega_1] & r_{2l}^{xv}[\Omega_2] & \cdots & r_{2l}^{xv}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r_{Dl}^{xv}[\Omega_1] & r_{Dl}^{xv}[\Omega_2] & \cdots & r_{Dl}^{xv}[\Omega_K] \end{bmatrix} \quad \forall\, l \in \mathcal{L} \tag{5.37}$$

$$\mathbf{C}_l^{vx} \triangleq \begin{bmatrix} r_{l1}^{vx}[\Omega_1] & r_{l1}^{vx}[\Omega_2] & \cdots & r_{l1}^{vx}[\Omega_K] \\ r_{l2}^{vx}[\Omega_1] & r_{l2}^{vx}[\Omega_2] & \cdots & r_{l2}^{vx}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r_{lD}^{vx}[\Omega_1] & r_{lD}^{vx}[\Omega_2] & \cdots & r_{lD}^{vx}[\Omega_K] \end{bmatrix} \quad \forall\, l \in \mathcal{L}. \tag{5.38}$$

The structure of these sensor-reference crosscorrelation matrices in terms of the mixing matrix, the reference mixing matrix, and the source autocorrelation functions follows from (5.35) and (5.36), i.e.,

$$\mathbf{C}_l^{xv} \equiv \mathbf{A}\ \mathrm{diag}(\tilde{\mathbf{b}}_l^*)\mathbf{C}^s \qquad \forall\, l \in \mathcal{L} \tag{5.39}$$

$$\mathbf{C}_l^{vx} \equiv \mathbf{A}^*\,\mathrm{diag}(\tilde{\mathbf{b}}_l)\mathbf{C}^s \qquad \forall\, l \in \mathcal{L} \tag{5.40}$$

where the row vector $\tilde{\mathbf{b}}_l$ is the l'th row of the reference mixing matrix $\mathbf{B}$ and $\mathbf{r}^s[\Omega_\kappa]$ is the κ'th column of the following source autocorrelation matrix $\mathbf{C}^s$ of size $S \times K$:

$$\mathbf{C}^s \triangleq \begin{bmatrix} r_{11}^s[\Omega_1] & r_{11}^s[\Omega_2] & \cdots & r_{11}^s[\Omega_K] \\ r_{22}^s[\Omega_1] & r_{22}^s[\Omega_2] & \cdots & r_{22}^s[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r_{SS}^s[\Omega_1] & r_{SS}^s[\Omega_2] & \cdots & r_{SS}^s[\Omega_K] \end{bmatrix}. \tag{5.41}$$

Each of these correlation matrices $\mathbf{C}_l^{xv}$ or $\mathbf{C}_l^{vx}$ has size $D \times K$; however, their rank is at most S due to their structure. Therefore, we can apply two subspace matrices $\mathbf{P}_s$ and $\mathbf{Q}^s$ that reduce the matrices to size $S \times S$ without losing information in the following way:

$$\bar{\mathbf{C}}_l^{xv} \triangleq \mathbf{P}_s\mathbf{C}_l^{xv}\mathbf{Q}^s \equiv \bar{\mathbf{A}}\,\mathrm{diag}\left(\tilde{\mathbf{b}}_l^*\right)\bar{\mathbf{C}}^s \tag{5.42}$$

$$\bar{\mathbf{C}}_l^{vx} \triangleq \mathbf{P}_s^*\mathbf{C}_l^{vx}\mathbf{Q}^s \equiv \bar{\mathbf{A}}^*\,\mathrm{diag}\left(\tilde{\mathbf{b}}_l\right)\bar{\mathbf{C}}^s \tag{5.43}$$

where $\bar{\mathbf{C}}_l^{xv}$ and $\bar{\mathbf{C}}_l^{vx}$ represent the reduced crosscorrelation matrices and $\bar{\mathbf{A}} \triangleq \mathbf{P}_s\mathbf{A}$ is defined as the reduced mixing system and $\bar{\mathbf{C}}^s \triangleq \mathbf{C}^s\mathbf{Q}^s$ is defined as the reduced source autocorrelation matrix. Due to the structure of the correlation matrices, the same two subspace matrices $\mathbf{P}_s$ and $\mathbf{Q}^s$ can be applied to the entire set of sensor-reference crosscorrelation matrices $\mathbf{C}_l^{xv}$ and $\mathbf{C}_l^{vx}$.

The reduction matrices $\mathbf{P}_s$ and $\mathbf{Q}^s$ can be identified for example by applying a singular value decomposition to one of the sensor-reference correlation matrices, i.e.,

$$\mathrm{SVD}\left(\mathbf{C}_l^{xv}\right) = \mathbf{U}^s\mathbf{\Sigma}_S\mathbf{V}_s \quad \text{for} \quad l \in \mathcal{L} \tag{5.44}$$

where $\mathbf{\Sigma}_S$ is a matrix of size $S \times S$ containing singular values on its diagonal and the matrices $\mathbf{U}^s$ and $\mathbf{V}_s$ with sizes $D \times S$ and $S \times K$, respectively, contain their corresponding left and right singular vectors. These left and right singular vectors span the same space as the mixing column vectors in $\mathbf{A}$ and the autocorrelation functions in $\mathbf{C}^s$, respectively. Consequently, the reduction matrices are found as $\mathbf{P}_s = (\mathbf{U}^s)^H$ and $\mathbf{Q}^s = (\mathbf{V}_s)^H$. The only condition is that the considered sensor-reference crosscorrelation matrix in (5.44) has rank S. Other approaches to identify these subspace matrices are described in Chapter 3 and [70]. Once the reduction matrix $\mathbf{P}_s$ is identified, we can apply it either to the correlation matrices or directly to the observations $\mathbf{x}[n]$ leading to $\bar{\mathbf{x}}[n] = \mathbf{P}_s\mathbf{x}[n]$. This latter reduction method will be used in the remainder of this chapter.

5.4.2 Spatial reference based source extraction as informed source extraction

Up to now we have derived the structure of sensor-reference crosscorrelation matrices that are reduced to a size of $S \times S$. In this section we show how these correlation matrices can be used to identify extraction filters that extract the desired source.

By exploiting the linear relation between the reference mixing elements b_l^j and the mixing elements, we are able to decompose (5.39) and (5.40) respectively as follows:

$$\bar{\mathbf{C}}_l^{xv} = \sum_{i=1}^{D} t_l^{i,*} \bar{\mathbf{C}}_{\bullet i}^{x} \quad \text{and} \quad \bar{\mathbf{C}}_l^{vx} = \sum_{i=1}^{D} t_l^{i} \bar{\mathbf{C}}_{i\bullet}^{x} \tag{5.45}$$

where $\bar{\mathbf{C}}_{\bullet i}^{x}$ and $\bar{\mathbf{C}}_{i\bullet}^{x}$ are the reduced sensor correlation matrices in (4.22) of Chapter 4.

From these decompositions it follows that the sensor-reference crosscorrelation matrices have exactly the same structure as the linear combinations of sensor correlation matrices $\check{\mathbf{\Gamma}}_l$ and $\mathbf{\Gamma}_l$ as defined in (4.24) of Chapter 4. Moreover, if we choose the spatial filters $\tilde{\mathbf{t}}_l$ equal to the Hermitian transpose of the selection beamformers $\boldsymbol{\xi}^l$ from Chapter 4, i.e., $\tilde{\mathbf{t}}_l = (\boldsymbol{\xi}^l)^H$, then the matrices are equal to each other. Consequently, we can reuse the design procedures for the selection beamformers to design reference systems and we can reuse the informed source extraction algorithm to identify the filter that extracts the desired source.

5.4.2.1 Design of selection beamformers or spatial reference filters

The design method for the reference system discussed here requires an estimate of the mixing coefficients that model the transfer from the desired source to all the sensors that are used to generate reference signals. This estimate is denoted by the symbol $\mathbf{a}^0$ and is called a mold. From Chapter 4 it follows that extraction of the desired source can be guaranteed if the following condition holds:

$$\frac{|\langle \mathbf{a}^0, \mathbf{a}^d \rangle|}{\|\mathbf{a}^d\|_2} > \frac{|\langle \mathbf{a}^0, \mathbf{a}^j \rangle|}{\|\mathbf{a}^j\|_2} \quad \forall\, j \in \mathcal{S} \backslash d \tag{5.46}$$

where $\langle \mathbf{a}, \mathbf{b} \rangle \triangleq (\mathbf{a})^H \mathbf{b}$ is the Euclidean inner product and $\|\mathbf{a}\|_2 \triangleq \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$ is the Euclidean norm.

This condition on the mold means that the angle between the mold and the mixing column of the desired source should be smaller than the angle between the mold and any other mixing column. In other words, the estimate $\mathbf{a}^0$ should be close enough to $\mathbf{a}^d$.

Based on this mold, the reference system should be chosen as the complex conjugate of an orthonormal basis with the first basis vector being the mold, thus $\tilde{\mathbf{t}}_l = (\mathbf{a}^0)^H$ and

$$\tilde{\mathbf{t}}_{l_1} (\tilde{\mathbf{t}}_{l_2})^H = \begin{cases} \left|\left|\mathbf{a}^0\right|\right|_2^2 & \text{for} \quad \{(l_1, l_2) \in \mathcal{L} \times \mathcal{L} |\, l_1 = l_2\} \\ 0 & \text{for} \quad \{(l_1, l_2) \in \mathcal{L} \times \mathcal{L} |\, l_1 \neq l_2\} \end{cases} . \tag{5.47}$$

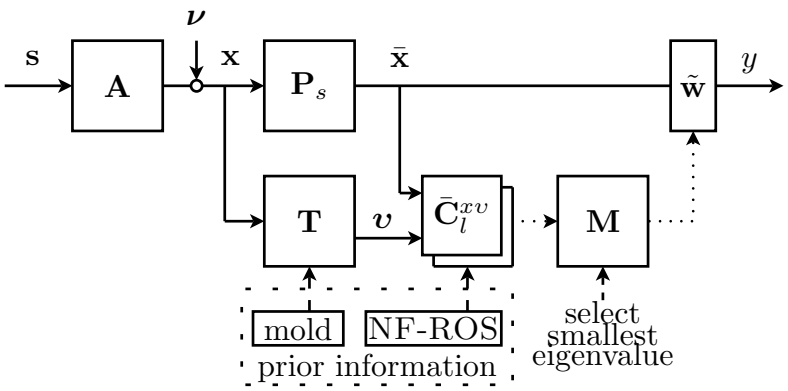


Figure 5.6: Overview of the spatial reference based source extraction algorithm. The *mold* represents the a priori information about the mixing system that is used to design the reference system $\mathbf{T}$.

For the reference system from Figure 5.2 this means that $L = D$ reference filters are required, while for the reference system from Figure 5.3 only $L = D_\upsilon$ reference filters are required. A proof for this design can be found in Section 4.3.2 of Chapter 4. Other designs of reference systems based on array signal processing techniques can be found in Section 4.4 of Chapter 4. These designs lead to a significant reduction in the number of reference signals L and therefore increase the efficiency of the source extraction algorithm.

5.4.2.2 Summary of spatial reference based source extraction algorithm

An overview of the spatial reference based source extraction algorithm is depicted in Figure 5.6. This algorithm can be used to identify the main components $\tilde{\boldsymbol{\mu}}$ and $\mathbf{h}$ of the LCMV and MVDR extraction filter $\tilde{\mathbf{w}}$, respectively. A summary of the steps in the source extraction algorithm is as follows:

1. Design a spatial reference system $\mathbf{T}$ and apply it to the observations, i.e., $\boldsymbol{\upsilon}[n] = \mathbf{T}\mathbf{x}[n]$.
2. Calculate sensor-reference crosscorrelation matrices $\mathbf{C}_l^{x\upsilon}$ and $\mathbf{C}_l^{\upsilon x}$ as in (5.37) and (5.38).
3. Identify subspace matrices $\mathbf{P}_s$ and $\mathbf{Q}^s$ and calculate the reduced size correlation matrices $\bar{\mathbf{C}}_l^{x\upsilon}$ and $\bar{\mathbf{C}}_l^{\upsilon x}$ as in (5.42) and (5.43) or reduce the observations directly.
4. Combine the reduced crosscorrelation matrices in the following way:

$$\mathbf{M} = \sum_{l=2}^{L} \bar{\mathbf{C}}_l^{x\upsilon} \left(\bar{\mathbf{C}}_1^{\upsilon x}\right)^{-1} \bar{\mathbf{C}}_l^{\upsilon x} \left(\bar{\mathbf{C}}_1^{x\upsilon}\right)^{-1}. \tag{5.48}$$

5. Calculate the left or right eigenvector, $\tilde{\boldsymbol{\mu}}$ or $\mathbf{h}$, that corresponds to the smallest eigenvalue of $\mathbf{M}$.
6. Extract the desired source and apply noise reduction if desired.

The underlying structure of $\mathbf{M} \equiv \bar{\mathbf{A}}\mathbf{\Lambda}\left(\bar{\mathbf{A}}\right)^{-1}$ makes that the left eigenvectors of $\mathbf{M}$ correspond to rows from the inverse of the reduced mixing system $\bar{\mathbf{A}}$, while the right eigenvectors correspond to columns from the reduced mixing system. The eigenvalues, i.e., the elements on the diagonal of $\mathbf{\Lambda}$, are positive and have the following structure:

$$\lambda^j = \sum_{l=2}^{L} \frac{|\mathbf{t}_l \mathbf{a}^j|^2}{|\mathbf{t}_1 \mathbf{a}^j|^2} \quad \forall\, j \in \mathcal{S}. \tag{5.49}$$

Each eigenvalue depends on the mixing column of one of the sources and the parameters of the reference system. Consequently, a proper design of the reference system as in Section 5.4.2.1 makes that the left and right eigenvectors that correspond to the smallest eigenvalue can be used to extract the desired source.

For more details on the application of identified extraction filters $\tilde{\boldsymbol{\mu}}$ and $\mathbf{h}$ we refer to Section 5.3.

5.4.3 Source extraction in the presence of spatially uncorrelated noise

So far, we have assumed that a NF-ROS exists and is known. In practice, this information is often difficult to obtain and a NF-ROS does not have to exist. In this section we assume that spatially uncorrelated noise has an unknown contribution in the SOTS of the considered sensor signals.

For the model in Figure 5.2 the presence of spatially uncorrelated signals leads to the following structure in the crosscorrelation functions of sensor and reference signals:

$$r_{il}^{x\upsilon}[n,\kappa] = \sum_{j=1}^{S} a_i^j b_l^{j,*} r_{jj}^{s}[n,\kappa] + t_l^{i,*} r_{ii}^{\nu}[n,\kappa] \tag{5.50}$$

$$r_{li}^{\upsilon x}[n,\kappa] = \sum_{j=1}^{S} b_l^j a_i^{j,*} r_{jj}^{s}[n,\kappa] + t_l^{i} r_{ii}^{\nu}[n,\kappa]. \tag{5.51}$$

It follows that the crosscorrelation functions $r_{il}^{x\upsilon}[n,k]$ and $r_{li}^{\upsilon x}[n,k]$ are only contaminated by noise if the reference signal υ_l is derived from the sensor signal x_i, i.e., if t_l^i is non-zero. Therefore, we split the sensors into two subsets. One set is used for generating reference signals and the other set is used for source extraction, as is depicted in Figure 5.3.

An overview of the algorithm that can be used to extract the desired source for such a scenario is depicted in Figure 5.7, where the concatenations of $\mathbf{A}_x, \mathbf{A}_\upsilon$ and $\boldsymbol{\nu}_x, \boldsymbol{\nu}_\upsilon$ lead to respectively the mixing matrix $\mathbf{A}$ and noise vector $\boldsymbol{\nu}$ in Figure 5.6. The structure of the reduced sensor-reference crosscorrelation matrices is as follows:

$$\bar{\mathbf{C}}_l^{x\upsilon} = \bar{\mathbf{A}}_x \operatorname{diag}\left(\tilde{\mathbf{b}}_l^{\upsilon,*}\right) \bar{\mathbf{C}}^s \quad \forall\, l \in \mathcal{L} \tag{5.52}$$

$$\bar{\mathbf{C}}_l^{\upsilon x} = \bar{\mathbf{A}}_x^* \operatorname{diag}\left(\tilde{\mathbf{b}}_l^{\upsilon}\right) \bar{\mathbf{C}}^s \quad \forall\, l \in \mathcal{L} \tag{5.53}$$

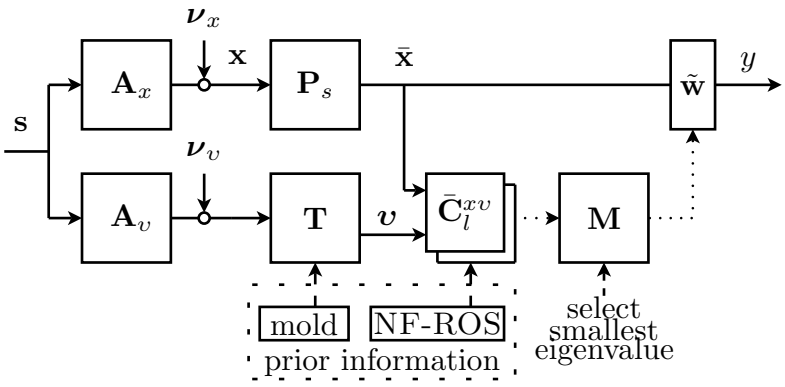


Figure 5.7: Overview of the reference based source extraction algorithm that deals with spatially uncorrelated noise.

where $\bar{\mathbf{A}}_x = \mathbf{P}_s\mathbf{A}_x$, $\mathbf{b}_l^v = \tilde{\mathbf{t}}_l\mathbf{A}_v$, and $\bar{\mathbf{C}}^s = \mathbf{C}^s\mathbf{Q}^s$. If we follow the steps from the algorithm in Section 5.4.2.2, then the matrix $\mathbf{M}$, built from the new type of sensor-reference crosscorrelation matrices, has the following structure:

$$\mathbf{M} = \sum_{l=2}^{L} \bar{\mathbf{C}}_l^{xv} \left(\bar{\mathbf{C}}_1^{vx}\right)^{-1} \bar{\mathbf{C}}_l^{vx} \left(\bar{\mathbf{C}}_1^{xv}\right)^{-1} \equiv \sum_{l=2}^{L} \bar{\mathbf{A}}_x \mathbf{\Lambda} \left(\bar{\mathbf{A}}_x\right)^{-1}. \tag{5.54}$$

Each eigenvalue of the matrix $\mathbf{M}$ corresponds to a unique column of the mixing matrix $\mathbf{A}_x$; and, its left and right eigenvectors are extraction filters with respect to the mixing matrix $\mathbf{A}_x$.

Extraction of the desired source with this new algorithm can be guaranteed by meeting the following two conditions. First, the reference system has to be designed in a similar way as is discussed in Section 5.4.2.1, but now with respect to the mixing matrix $\mathbf{A}_v$ and a priori information about the mixing of the desired source w.r.t. the reference sensors. Second, the mixing matrix $\mathbf{A}_x$ has to be well conditioned, i.e., it has to have rank S. This latter condition, combined with the minimum number of reference signals required for this new approach makes that the total number of sensors D should be larger than $S + 2$. The formal condition on the noise is that the noise observed by a sensor in the reference subset is not correlated with the noise observed by the sensors in the source extraction subset. An example of such noise is spatially uncorrelated sensor noise.

5.5 Simulation results and discussions

In this section we present results from computer simulations. In the first experiment we compare different configurations of the two new reference based source extraction algorithms and verify that source extraction is accomplished. In the second experiment we apply the new algorithms in a wireless acoustic sensor network (WASN)

scenario.

5.5.1 Comparison of several configurations of the source extraction algorithms

In this experiment we compare the performance of the source extraction algorithms for six different configurations while varying the number of samples used in the estimation process. First we describe the simulation setup and performance measures. Finally, we present and discuss the simulation results.

We mixed three complex-valued sources with unit power that were generated by filtering a real-valued Gaussian signal with a filter consisting of a single complex pole. The poles for the three sources are respectively $-0.9\,\mathrm{e}^{3\jmath}$, $0.7\,\mathrm{e}^{2\jmath}$, and $0.8\,\mathrm{e}^{4\jmath}$. Each iteration a new random mixing matrix of size 6×3 was drawn from a complex Gaussian distribution. The columns of the mixing matrix were normalized such that a predictable overall power of the sources is observed at the sensors. Finally, the six sensor signals were contaminated by spatially white, complex Gaussian noise with variance $\sigma_\nu^2 = 0.1$.

The performance is evaluated for six different configurations of the new reference based source extraction algorithms. The first three algorithms use respectively 6, 4, and 3 sensors. These sensors are used for extraction as well as the generation of reference signals. The mold, i.e, the estimate of the mixing column vector corresponding to the desired source was equal to the actual mixing column vector and the remaining reference signals are generated by calculating an orthonormal basis as discussed in Section 5.4.2.1. An ideal mold is chosen to focus on the difference in performance for different sensor configurations. In [61] we have shown that perturbations on the mold lead to little performance degradation. The three remaining algorithms use the first three sensors for applying the extraction filter, while respectively the sensors 1-2, 4-5 and 4-6 are used for generating the reference signals. Again, an ideal mold and an orthonormal basis are used to design the reference system. For the first four algorithms the NF-ROS was chosen as the lags -10 up to 10 excluding lag 0, while the fifth and sixth algorithm included lag 0 in the NF-ROS.

The main objective of the presented algorithms is to identify the extraction filter $\boldsymbol{\mu}$ and the mixing column vector $\mathbf{h}$ that correspond to the desired source. The performance of the extraction filter $\tilde{\boldsymbol{\mu}}$ is evaluated by measuring the SIR in the extracted signal, which has to be maximized. The SIR is obtained by calculating the ratio of the gain for the desired source and for the gain for the interfering sources. The filter $\mathbf{h}$ is an estimate of the mixing column that corresponds to the desired source and can be used to build an MVDR filter if an additional noise suppression step is implemented. Deviation from the actual mixing column vector results in leakage and thus a performance drop. Therefore, we measure the Hermitian angle [67] between the identified mixing column vector $\mathbf{h}$ and the mixing column vector corresponding to the desired source.

In Figures 5.8 and 5.9 we depict the results for a Monte Carlo simulation consisting of 500 simulations per selected number of samples. We observe that in both figures the

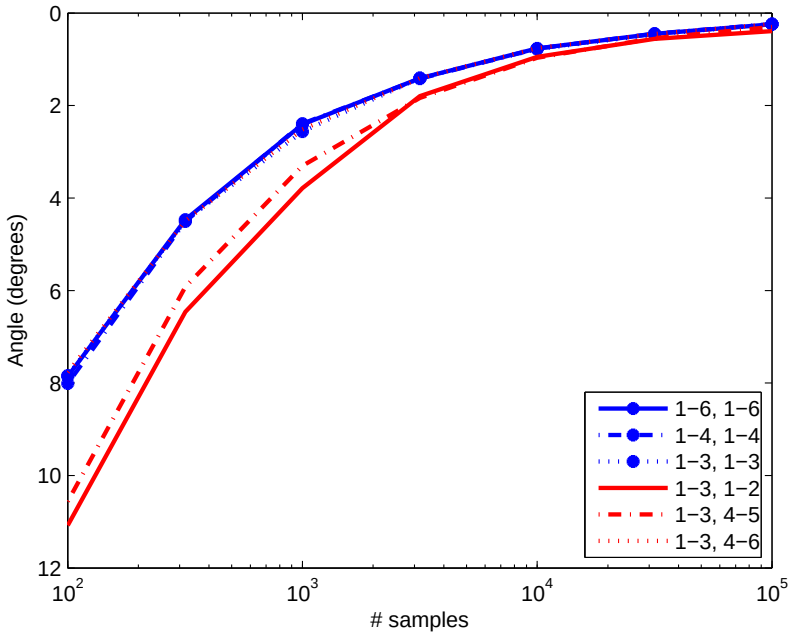


Figure 5.8: Average angle between estimated and actual mixing column corresponding to the desired source for multiple algorithms and different number of samples. The first part in the legend indicates the sensors used for extraction and the second part indicates the sensors used for generating reference signals.

identification performance increase for increasing number of samples and for increasing the number of sensors used for extraction. In Figure 5.8 we observe that the algorithms using only two sensors for generating reference signals have a poorer performance if the used number of samples is relatively low. The third and the sixth algorithm show a similar performance for identification of the mixing column $\mathbf{h}$. An advantage of the latter algorithm is that less a priori information about the noise is required. In Figure 5.9 we observe that the performances of the third and sixth algorithm are similar. Furthermore, we observe that using more sensors increases the extraction performance.

From these simulations we conclude that both reference based source extraction algorithms can be used for extraction of a desired source. As expected, using more sensors and more samples leads to a better performance.

5.5.2 Application in a wireless acoustic sensor network

The objective of this second experiment is to validate the new reference based source extraction algorithm in a more practical scenario.

A WASN consists of spatially distributed sensing nodes that are equipped with

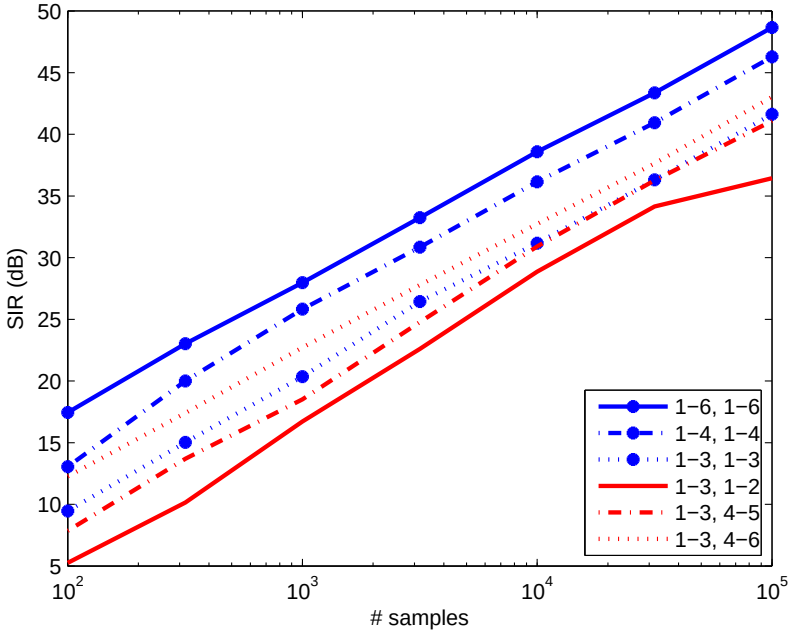


Figure 5.9: SIR of the first experiment for a range of algorithms and different number of samples. The first part in the legend indicates the sensors used for extraction and the second part indicates the sensors used for generating reference signals.

at least one microphone and a radio for wireless communication. In the considered extraction scenario a WASN is deployed near a number of sound sources, e.g., human speakers, and the microphone signals have to be combined in such a way that the desired sound or speech is extracted from the observed mixture. In this experiment we consider the WASN scenario as depicted in Figure 5.10, where a mixture of two sources is observed by three wireless sensor nodes equipped with two microphones used for extraction and a single sensor node having three microphones for generating reference signals.

The source signals are generated by filtering real-valued white Gaussian signals with a comb filter that repeats the poles -0.9 and 0.8 for respectively the desired and interfering source. The room impulse responses (RIR) are calculated with a RIR generator [69] and have a length of 512 samples. We consider a mixing scenario without reflections and a scenario where the reverberation time to 0.15 seconds. The sensor signals are sampled with a sampling frequency of 8 kHz and corrupted with spatially white Gaussian noise such that the source signal to noise ratio per sensor is 20 dB.

We use the method from Section 5.4.3 such that no information about the SOTS of the spatially uncorrelated noise is required. For the mold we use the first 50 samples of the impulse responses from the desired source to the three reference microphones, which model the direct path and some early reflections. The reference system is

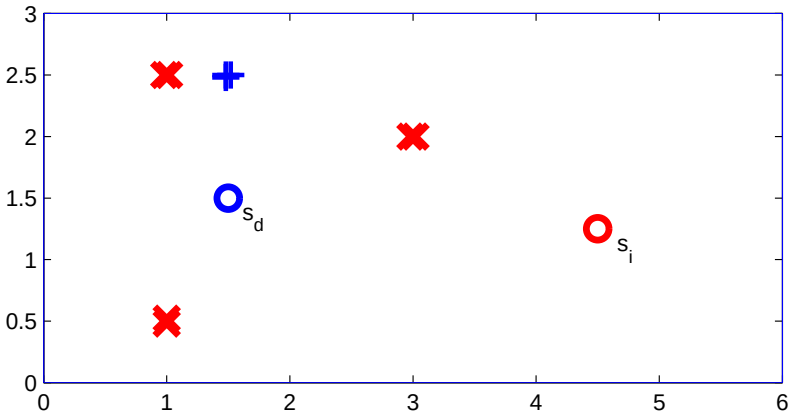


Figure 5.10: Overview of the WASN scenario. The room has a size of 6 by 3 meters. The desired source s_d and the interfering source s_i are indicated with circles, the three sensors used for extraction with cross signs, and the sensors for generating reference signals with plus signs. The sensors on a sensor node have a mutual spacing of 0.06 meters. All sensors are located at a height of 1.2 meters while the height of the room is 3 meters.

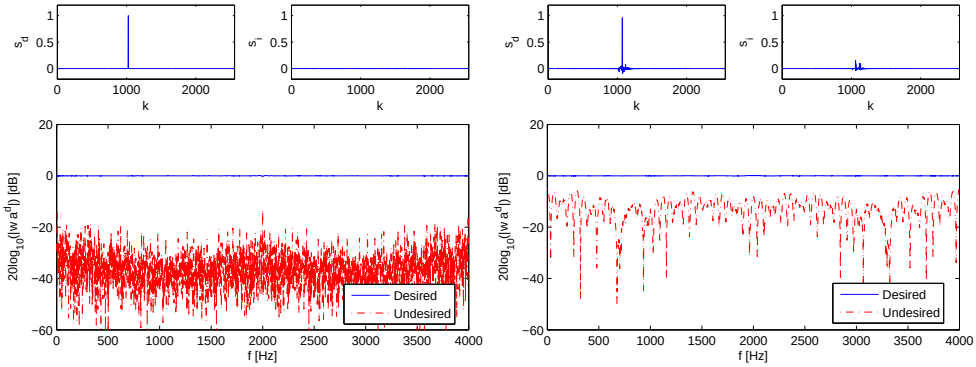
designed by calculating a 2048 points discrete Fourier transform (DFT) of the mold. Subsequently, per frequency bin we calculate an orthonormal basis with the first vector being the mold. The NF-ROS is chosen as the lags -6 up to 6 per frequency bin.

The extraction filters $\tilde{\mathbf{w}}_\mu$ and $\tilde{\mathbf{w}}_h$, are obtained by scaling the vectors $\tilde{\boldsymbol{\mu}}$ and $\mathbf{h}$ using the mold, i.e.,

$$\tilde{\mathbf{w}}_\mu = \frac{\tilde{\boldsymbol{\mu}}}{\tilde{\boldsymbol{\mu}}\mathbf{a}^0} \quad \text{and} \quad \tilde{\mathbf{w}}_h = \frac{(\mathbf{h})^H}{(\mathbf{h})^H\mathbf{a}^0}. \tag{5.55}$$

After scaling and selection of the desired filter we calculate the impulse response of the filter by means of an inverse DFT. The overall transfer function from each source to the extracted filter is calculated by summing the convolution of the extraction filters with the respective RIRs. The resulting overall impulse responses and corresponding transfer functions are depicted in Figure 5.11 and Figure 5.12 for the anechoic and reverberant room, respectively.

We observe that both extraction algorithms are able to extract the desired source from the observations. For the observations in the anechoic room the filter $\tilde{\mathbf{w}}_\mu$ suppresses the interfering source more than the filter $\tilde{\mathbf{w}}_h$; however, both filters suppress the interfering source at all frequencies. For the observations in the reverberant room we observe that both filters suppress the interfering source for most frequencies. The filter $\tilde{\mathbf{w}}_\mu$ has on average a larger suppression of the interfering source than the filter $\tilde{\mathbf{w}}_h$; however, the filter $\tilde{\mathbf{w}}_h$ has fewer frequencies where the interfering source is not suppressed. Furthermore we observe that for the anechoic room the desired source is extracted without distortions while for the reverberant room the extracted desired



(a) Extraction filter $\tilde{\boldsymbol{\mu}}$. (b) Extraction filter $\mathbf{h}$.

Figure 5.11: Results for two source extraction algorithms applied to data from an anechoic room.

source is filtered. Possible reasons for the enhancement of the interfering source by the extraction filters in the reverberant room are that the mixing system is ill conditioned at those frequencies. Furthermore, due to longer RIRs in the reverberant room w.r.t. the anechoic room, the assumption of instantaneous mixtures per frequency bin can become less accurate; therefore, further research on source extraction from convolutive mixtures is desired.

Interesting differences between the spatial and temporal reference based source extraction for spatially uncorrelated noise is in the required number of sensors and the number of time-lag pairs that are required in order to make immediate source extraction of the desired source feasible.

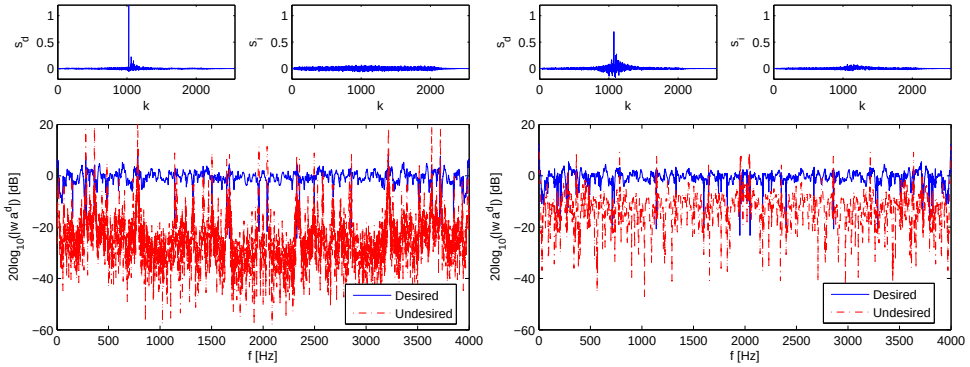
Table 5.1: Comparison of conditions for spatial and temporal reference based source extraction algorithms

Algorithm	# time-lag pairs	# sensors
spatial, regular	$K \geq S$	$D \geq S$
spatial, uncorrelated noise	$K \geq S$	$D \geq S + 2$
temporal, regular	$K \geq 2$	$D \geq S$
temporal, uncorrelated noise	$K \geq 2$	$D \geq 2S$

In the following chapter we show that these conditions can be released further.

5.6 Discussion and conclusions

In this chapter we addressed the source extraction problem. We have shown that the design of a reference system based on physical parameters can be exploited in order



(a) Extraction filter $\tilde{\boldsymbol{\mu}}$. (b) Extraction filter $\mathbf{h}$.

Figure 5.12: Results for two source extraction algorithms applied to data from a reverberant room room.

to guarantee extraction of the desired source. The corresponding source extraction algorithms presented in this work are beneficial over other reference based source extraction algorithms in the sense that only second order statistics are exploited and that extraction of the desired source is guaranteed based on a priori information about physical parameters.

We presented two strategies and associated algorithms to obtain and deal with reference signals. Both algorithms are flexible in the choice of an objective for the source extraction filter. Both the MVDR filter, which minimizes the interference and noise in a balanced way, and the LCMV filter, which minimizes the noise subject to cancelation of the interfering sources, can be identified with the proposed algorithms. The major difference between the two strategies is in the way they deal with spatially uncorrelated noise. The first approach exploits all available sensor signals and identifies the optimal extraction filter using all these signals, at the cost of information about the second order temporal structure of the noise signals. In the second approach, reference signals are derived from a subset of the sensor signals while an extraction filter is identified for the complementary set of sensor signals. This second approach is insensitive to spatially uncorrelated noise; however, the identified extraction filter is sub-optimal since not all available sensors are used for extraction of the desired source.

We validated the algorithms through computer simulations and we have verified that the second approach requires less information about spatially uncorrelated noise at the cost of a sub-optimal extraction filter. Finally, we applied the new source extraction algorithm in a WASN scenario and we have shown that extraction of the desired source is feasible in such a scenario.

A topic for future work is to develop reference system design techniques that are able to incorporate all available information such as information about the location of undesired sources in order to guarantee extraction of the desired source. Furthermore, the development of an adaptive, on-line algorithm and further development and

evaluation of the proposed algorithms for non-stationary signals such as speech have to be considered in the future. Finally, we believe that further research into source extraction from convolutive mixtures can increase the source extraction performance.

6

Source extraction as a polynomial optimization problem

In this chapter we formulate source extraction algorithm design as a polynomial optimization problem. In this optimization problem, homogenous polynomial equations form equality constraints that restrict the feasible set of the optimization problem to extraction filters. Identification of the desired source extraction filter is accomplished by designing a rational objective function using a priori information. The work in this chapter is considered an important step and providing insight and leads for translating the results from the previous chapters to work for convolutive mixtures.

Many similarities exist between the polynomial based and the informed source extraction algorithms presented in the previous chapters. For example, both real-valued and complex-valued mixing scenarios are considered. A priori information about either the mixing parameters or the correlation function of the sources can be exploited and two different types of source extraction filters can be identified. Also the concept of reference signals can be translated to the polynomial optimization based approach. Unique features of the polynomial optimization based methodology presented in this chapter are that identification of extraction filters is feasible for certain under-determined mixtures and that a natural extension towards the use of higher order statistics exists. Finally, we present design techniques for objective function and verify the methodology via computer simulations.

6.1 Introduction

In this chapter we use our knowledge about informed and reference based source extraction to formulate source extraction as a polynomial optimization problem. In such an optimization problem, polynomial equality constraints, formulated in [36], ensure that the feasible set consists of only extraction filters while the objective function can be designed based on a priori information such that the global optimum corresponds to the extraction filter for the desired source. In this way, selection of the desired filter is de-coupled from identifying the actual filter coefficients.

Several techniques are available to solve systems of polynomial equations; however, in general it is difficult to obtain stable algorithms for empirical data. A well known and widely used technique for solving such systems is the homotopy method [71]. The homotopy method uses a known start system and moves from the start system to the unknown system while tracking its roots. Disadvantages of methods based on homotopy are that it is not possible to predict which root will be found. Therefore, in this thesis we focus on a method that uses well known linear algebra techniques such as linear subspaces and (generalized) eigenvalue decompositions to find roots of polynomial systems. This was first recognized by Stetter [72]. Nowadays, Dreesen et al. [73,74] continue working on solving systems of multivariate polyonomial equations using linear algebra techniques and also started to incorporate objective functions to create optimization problems with polynomial constraints.

We show that the polynomial optimization problem can be formulated as a generalized eigenvalue decomposition (GEVD) problem where either the largest or the smallest eigenvalue and its corresponding eigenvector correspond to the optimal solution. This polynomial optimization framework gives insight in the source extraction problem and allows for several generalizations. The objective function of the polynomial optimization problem allows for source extraction algorithm designs based on a priori information. We study the consequences of procedures such as spatial and temporal filtering of the observations. As we have seen in the previous chapters, spatial and temporal filtering procedures can be used to obtain efficient and flexible implementations of source extraction algorithms. Finally, we show that the polynomial optimization framework can be extended in a natural way such that source extraction can be accomplished for under-determined mixtures and by exploiting higher order statistics.

The outline of this chapter is as follows. In Section 6.2 we introduce the mixing model and assumptions on the source and noise signals. The structure in the sensor correlation data is discussed in Section 6.3. Subsequently, we formulate blind system identification as root finding for systems of polynomial equations in Section 6.4 and identification of the desired source extraction filter is presented in Section 6.5. Additional features of the polynomial optimization based source extraction approach are discussed in Section 6.6. Design of polynomial optimization based source extraction algorithms is discussed in Section 6.7. Finally, we evaluate and conclude this chapter in Section 6.8 and Section 6.9, respectively.

6.2 Model and assumptions

As discussed in the previous chapters, several types of source extraction filters can be designed, based on different objectives. Furthermore, many different sorts and amounts of a priori information can be available, either about the desired or undesired sources or their mixing parameters. In the current chapter we focus on the design of source extraction filters using a priori information about the mixing parameters of the desired source. We consider a similar set of assumptions as used in the previous chapters, especially Chapter 4. However, we extend our notation on conjugation in order to include different conjugation schemes and the case for real-valued mixing systems.

6.2.1 Mixing model

In the considered mixing scenario a total of S source signals $s_j[n]$ are mixed by a size $D \times S$ real-valued or complex-valued instantaneous mixing system $\mathbf{A}$. The observations consist of the mixtures that are contaminated by additive noise $\nu_i[n]$. A mathematical model for the structure in these observations is as follows:

$$\mathbf{x}[n] = \sum_{j=1}^{S} \mathbf{a}^j s_j[n] + \boldsymbol{\nu}[n] = \mathbf{A}\mathbf{s}[n] + \boldsymbol{\nu}[n] \tag{6.1}$$

where the vectors containing respectively the sensor, source, and noise signals are defined as follows:

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x_1[n] \\ \vdots \\ x_D[n] \end{bmatrix}, \quad \mathbf{s}[n] \triangleq \begin{bmatrix} s_1[n] \\ \vdots \\ s_S[n] \end{bmatrix}, \quad \boldsymbol{\nu}[n] \triangleq \begin{bmatrix} \nu_1[n] \\ \vdots \\ \nu_D[n] \end{bmatrix} \tag{6.2}$$

and the mixing matrix of size $D \times S$ is defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}^1 & \cdots & \mathbf{a}^S \end{bmatrix} = \begin{bmatrix} a_1^1 & \cdots & a_1^S \\ \vdots & \ddots & \vdots \\ a_D^1 & \cdots & a_D^S \end{bmatrix}. \tag{6.3}$$

Using the model from (6.1) for the observations, our objective is to identify a filter $\tilde{\mathbf{w}}$ that extracts the desired source from the observations, i.e.,

$$y[n] = \tilde{\mathbf{w}}\mathbf{x}[n] = \tilde{\mathbf{w}}\mathbf{A}\mathbf{s}[n] + \tilde{\mathbf{w}}\boldsymbol{\nu}[n] \tag{6.4}$$

where $y[n]$ should only contain the desired source. Due to noise, extraction of only the desired source is impossible and a trade-off has to be made between interference and noise reduction. Two well known filters in which this trade-off is made are the MVDR and LCMV filter. The filters are the solutions to respectively the following

optimization problems:

$$\tilde{\mathbf{w}}_{\mathrm{MVDR}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s. t. } \tilde{\mathbf{w}}\mathbf{a}^d = 1 \tag{6.5}$$

$$\tilde{\mathbf{w}}_{\mathrm{LCMV}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s. t. } \tilde{\mathbf{w}}\mathbf{A} = \tilde{\mathbf{e}}_d \tag{6.6}$$

where d represents the index for the desired source.

The MVDR filter minimizes the energy in $y[n]$ while retaining the desired source in the signal, leading to a balanced reduction of interfering sources and noise. The LCMV filter can be used to actively cancel the undesired sources while preserving the desired source. In order to apply the MVDR filter, identification of the mixing parameters corresponding to the desired source is required. For the LCMV filter, a specific row from the inverse of the mixing matrix has to be identified. This row is orthogonal to the columns in the mixing matrix that correspond to the undesired sources. In this work, identification of these parameters is performed by exploiting second order temporal structure in the data.

6.2.2 Second order temporal structure in the sensor signals

The second order statistical features of the source, sensor, and noise signals are evaluated in the form of auto- and crosscorrelation functions.

Definition 6.2.1 (Auto- and crosscorrelation functions)**:** The auto- and crosscorrelation functions for potentially complex-valued signals $p_i[n]$ and $q_j[n]$ at time-frame n and lag k are defined as follows:

$$r_{ij}^{pq}[n,k] \triangleq \mathbb{E}\{p_i^{c_1}[n] q_j^{c_2}[n-k]\} \tag{6.7}$$

where $\mathbb{E}$ is the mathematical expectation operator and the superscripts c_1 and c_2 indicate the type of conjugation that is applied. The symbol $\circ$ indicates no conjugation while $*$ indicates conjugation. The function is called an autocorrelation function if a signal is correlated with itself. Furthermore, if the two signals carry the same symbol, say p, then we use the following short-hand notation: $r_{ij}^{p}[n,k] = r_{ij}^{pp}[n,k]$.

For completeness it would be more elegant to include the conjugation tuple $(c_1 c_2)$ also into the notation of the correlation functions, e.g., $r_{ij}^{pq,c_1c_2}[n,k]$. However, for conciseness and readability we only include these symbols in the correlation functions if the conjugation does not follow from the context.

Using the model for the mixing system and the corresponding assumptions as in (6.1) and exploiting Definition 6.2.1, leads to the following sets of sensor, source, and noise auto- and crosscorrelation functions:

$$r_{i_1 i_2}^{x}[n,k] \triangleq \mathbb{E}\{x_{i_1}^{c_1}[n] x_{i_2}^{c_2}[n-k]\} \ \forall\, (i_1, i_2) \in \mathcal{D} \times \mathcal{D} \tag{6.8}$$

$$r_{j_1 j_2}^{s}[n,k] \triangleq \mathbb{E}\{s_{j_1}^{c_1}[n] s_{j_2}^{c_2}[n-k]\} \ \forall\, (j_1, j_2) \in \mathcal{S} \times \mathcal{S} \tag{6.9}$$

$$r_{i_1 i_2}^{\nu}[n,k] \triangleq \mathbb{E}\{\nu_{i_1}^{c_1}[n] \nu_{i_2}^{c_2}[n-k]\} \ \forall\, (i_1, i_2) \in \mathcal{D} \times \mathcal{D}. \tag{6.10}$$

Note that the correlation functions are only called autocorrelation functions when their indices are the same, e.g., $i_1 = i_2$. Next to these correlation functions we define the following crosscorrelation functions between the source and noise signals:

$$r_{ij}^{\nu s}[n,k] \triangleq \mathbb{E}\{\nu_i^{c_1}[n] s_j^{c_2}[n-k]\} \; \forall\, (i,j) \in \mathcal{D} \times \mathcal{S} \tag{6.11}$$

$$r_{ji}^{s\nu}[n,k] \triangleq \mathbb{E}\{s_j^{c_1}[n] \nu_i^{c_2}[n-k]\} \; \forall\, (i,j) \in \mathcal{D} \times \mathcal{S}. \tag{6.12}$$

Given the mixing model and the corresponding correlation functions we introduce a set of assumptions that are necessary to enable source extraction filter identification, i.e., without introducing additional assumptions, identification of extraction filters is not feasible.

The first assumption we make is that the sources are independent, or at least mutually uncorrelated, i.e.,

$$r_{j_1 j_2}^{s}[n,k] = 0 \quad \forall\, \{(j_1, j_2) \in \mathcal{S} \times \mathcal{S} |\, j_1 \neq j_2\}. \tag{6.13}$$

Our assumptions on the second order temporal structure of the noise signals are introduced via the definition of a noise-free region of support.

Definition 6.2.2 (Noise-free region of support)**:** The noise-free region of support (NF-ROS), denoted with Ω, consists of a set of time-lag pairs $\Omega_\kappa = (n,k)_\kappa$ with $\kappa \in \mathcal{K}$ for which the following conditions hold:

$$\forall\, \Omega_\kappa \in \Omega : \begin{cases} r_{i_1 i_2}^{\nu}[\Omega_\kappa] = 0 & \forall (i_1, i_2) \in \mathcal{D} \times \mathcal{D} \\ r_{ij}^{\nu s}[\Omega_\kappa] = 0 & \forall (i,j) \in \mathcal{D} \times \mathcal{S} \\ r_{ji}^{s\nu}[\Omega_\kappa] = 0 & \forall (i,j) \in \mathcal{D} \times \mathcal{S} \end{cases} \tag{6.14}$$

where the number of elements in the NF-ROS is indicated with K, i.e., $K = |\Omega|$.

On the other hand, for certain time-lag pairs statistical properties of the noise are often known or can be estimated such that it can be compensated for in the sensor correlation data. In such cases, the time-lag pairs for which noise compensation is applied can also be incorporated into the NF-ROS.

From our assumptions on the source signals and the assumptions on the noise presented in Definition 6.2.2 it follows that in the NF-ROS only the source autocorrelation functions are non-zero. This property is exploited in the remainder of this chapter to formulate polynomial optimization problems from which the desired source extraction filter can be identified for a wide variety in applications.

6.3 Structure in the sensor correlation data

Considering the model for the mixing system and the corresponding assumptions on the second order temporal structure of the source and noise signals, it follows that the sensor auto-and crosscorrelation functions have the following structure in terms

of the mixing parameters and the source autocorrelation functions [36]:

$$r^x_{i_1 i_2}[\Omega_\kappa] = \sum_{j=1}^{S} a^{j,c_1}_{i_1} a^{j,c_2}_{i_2} r^s_{jj}[\Omega_\kappa] \quad \forall\,(i_1, i_2, \kappa) \in \mathcal{D} \times \mathcal{D} \times \mathcal{K}. \tag{6.15}$$

By evaluating the sensor correlation function values for all sensor pairs $(i_1, i_2) \in \mathcal{D} \times \mathcal{D}$ and K time-lag pairs in the NF-ROS $\Omega_\kappa \in \Omega$, we are able to construct the following sensor correlation matrix of size $D^2 \times K$:

$$\mathbf{C}^x \triangleq \begin{bmatrix} r^x_{11}[\Omega_1] & r^x_{11}[\Omega_2] & \cdots & r^x_{11}[\Omega_K] \\ r^x_{12}[\Omega_1] & r^x_{12}[\Omega_2] & \cdots & r^x_{12}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r^x_{DD}[\Omega_1] & r^x_{DD}[\Omega_2] & \cdots & r^x_{DD}[\Omega_K] \end{bmatrix}. \tag{6.16}$$

For real-valued mixtures, this matrix can be obtained by stacking the sensor correlation matrices $\mathbf{C}^x_i$ for $i \in \mathcal{D}$ from (3.5) while for complex-valued mixtures with conjugation tuple $(\circ, *)$ the matrix can be obtained by stacking the sensor correlation matrices $\mathbf{C}^x_{i\bullet}$ for $i \in \mathcal{D}$ from (4.17).

The structure in this sensor correlation matrix follows from the structure in the sensor correlation data, as is presented in (6.15), and is as follows [36]:

$$\mathbf{C}^x \equiv (\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})\mathbf{C}^s \tag{6.17}$$

where $\diamond$ represents the second order Khatri-Rao product, $\mathbf{A}$ is the mixing matrix of size $D \times S$, and $\mathbf{C}^s$ is a source autocorrelation matrix of size $S \times K$ with the following structure:

$$\mathbf{C}^s \triangleq \begin{bmatrix} r^s_{11}[\Omega_1] & r^s_{11}[\Omega_2] & \cdots & r^s_{11}[\Omega_K] \\ r^s_{22}[\Omega_1] & r^s_{22}[\Omega_2] & \cdots & r^s_{22}[\Omega_K] \\ \vdots & \ddots & \ddots & \vdots \\ r^s_{SS}[\Omega_1] & r^s_{SS}[\Omega_2] & \cdots & r^s_{SS}[\Omega_K] \end{bmatrix}. \tag{6.18}$$

In Example 6.3.1 the structure of sensor correlation matrix $\mathbf{C}^x$ is visualized for the scenario where a mixture of two complex sources is observed with two sensors. This example shows how the mixing matrix elements and source autocorrelation function values are related.

Example 6.3.1: Suppose we observe two mixtures of two complex sources, i.e., $D = 2$ and $S = 2$. The sensor correlation matrix $\mathbf{C}^x$ is built from sensor correlation function values with conjugation tuple $(c_1 c_2) = (\circ *)$ for two lags in the NF-ROS, i.e., $K = 2$, which leads to the following structure:

$$\begin{bmatrix} r^x_{11}[\Omega_1] & r^x_{11}[\Omega_2] \\ r^x_{12}[\Omega_1] & r^x_{12}[\Omega_2] \\ r^x_{21}[\Omega_1] & r^x_{21}[\Omega_2] \\ r^x_{22}[\Omega_1] & r^x_{22}[\Omega_2] \end{bmatrix} = \begin{bmatrix} a^1_1 a^{1,*}_1 & a^2_1 a^{2,*}_1 \\ a^1_1 a^{1,*}_2 & a^2_1 a^{2,*}_2 \\ a^1_2 a^{1,*}_1 & a^2_2 a^{2,*}_1 \\ a^1_2 a^{1,*}_2 & a^2_2 a^{2,*}_2 \end{bmatrix} \begin{bmatrix} r^s_{11}[\Omega_1] & r^s_{11}[\Omega_2] \\ r^s_{22}[\Omega_1] & r^s_{22}[\Omega_2] \end{bmatrix}. \tag{6.19}$$

Note that in the example we do not explicitly mention the $\circ$ symbol. This symbol for indicating that no conjugation is applied is only used when it is important to indicate the order of conjugation.

6.3.1 Uniquiefication of sensor correlation functions

For real-valued mixtures and for complex-valued mixtures with conjugation tuples $(\circ\circ)$ or $(**)$, the crosscorrelation function $r^x_{i_1 i_2}[n,k]$ is equal to the crosscorrelation function $r^x_{i_2 i_1}[n,k]$. One way to deal with this symmetry is by means of a process called *uniquiefication* [36].

In the uniquiefication process we select the largest possible subset of unique sensor correlation functions, i.e., duplicates are removed. One way to accomplish this in a structured way for the sensor correlation matrix $\mathbf{C}^x$ is to remove all sensor correlation functions $r^x_{i_1 i_2}$ for the sensor pairs where $i_1 > i_2$ for $(i_1, i_2) \in \mathcal{D} \times \mathcal{D}$. This uniquiefication leads to a significant reduction in the number of rows in the sensor correlation matrix and it removes trivial linear dependencies in the rows of the sensor correlation matrix. The structure of a uniquiefied sensor correlation matrix $\breve{\mathbf{C}}^x$ is denoted as follows:

$$\breve{\mathbf{C}}^x = (\mathbf{A}^c \breve{\diamond} \mathbf{A}^c)\mathbf{C}^s \tag{6.20}$$

where $\breve{\cdot}$ denotes uniquiefication and $c \in \{\circ, *\}$ indicates if conjugation is applied. In order to illustrate the structure in the uniquiefied sensor correlation matrix, an example is presented for a mixture containing real-valued sources.

Example 6.3.2: Suppose we observe two real-valued mixtures of two real-valued sources, i.e., $D = 2$ and $S = 2$. We calculate the sensor correlation function values for two lags in the NF-ROS, i.e., $K = 2$. The uniquiefied sensor correlation matrix then has the following structure:

$$\begin{bmatrix} r^x_{11}[\Omega_1] & r^x_{11}[\Omega_2] \\ r^x_{12}[\Omega_1] & r^x_{12}[\Omega_2] \\ r^x_{22}[\Omega_1] & r^x_{22}[\Omega_2] \end{bmatrix} = \begin{bmatrix} a^1_1 a^1_1 & a^2_1 a^2_1 \\ a^1_1 a^1_2 & a^2_1 a^2_2 \\ a^1_2 a^1_2 & a^2_2 a^2_2 \end{bmatrix} \begin{bmatrix} r^s_{11}[\Omega_1] & r^s_{11}[\Omega_2] \\ r^s_{22}[\Omega_1] & r^s_{22}[\Omega_2] \end{bmatrix}. \tag{6.21}$$

In the uniquiefication process the sensor correlation function values for $r^x_{21}[\Omega_\kappa]$ for $\Omega_\kappa \in \Omega$ have been removed since they are the same as the autocorrelation function values for $r^x_{12}[\Omega_\kappa]$ for $\Omega_\kappa \in \Omega$.

This uniquiefication is not a necessary process here, and it cannot be applied for complex-valued mixtures with conjugation tuples $(\circ *)$ and $(*\circ)$, but it reduces the size of correlation matrices significantly. Determining the number rows in the uniquiefied sensor correlation matrix $\breve{\mathbf{C}}^x$, i.e., the number of unique sensor correlation functions, is a combinatorial problem closely related to the 'stars and bars' problem with non-negative integers [75]. The total number of rows in the uniquiefied sensor correlation matrix $\breve{\mathbf{C}}^x$, i.e., the number of unique index sets (i_1, i_2) for $(i_1, i_2) \in \mathcal{D} \times \mathcal{D}$ while

ignoring the order, is given as follows:

$$D_u = \sum_{i=1}^{D} i = \binom{D+1}{2} = \frac{(D+1)!}{2! \cdot (D+1-2)!} = \frac{D(D+1)}{2}. \tag{6.22}$$

Consequently, whereas the sensor correlation matrix $\mathbf{C}^x$ has a size of $D^2 \times K$, its uniquiefied counterpart $\breve{\mathbf{C}}^x$ has a size of $\frac{1}{2}D(D+1) \times K$.

Later on we show that uniquiefication is a useful tool that can be used to allow for the design of more advanced objective functions for selecting the desired filter based on a priori information and to allow for the identification of mixing column vectors for under-determined mixtures with more sources than sensors.

6.3.2 Subspaces of noise-free sensor correlation matrices

From (6.15) it follows that the sensor correlation functions are linear combinations of the source autocorrelation functions, where the elements $a_{i_1}^{j,c_1} a_{i_2}^{j,c_2}$ form the coefficients. This property is exploited in the current section to identify four linear subspaces of the sensor correlation matrices.

Assuming that the sensor correlation matrix is sufficiently large, four different subspaces can be identified, i.e., two signal subspaces and two noise or null subspaces. One way to perform a decomposition of the sensor correlation matrix into these subspaces is by applying a singular value decomposition, i.e.,

$$\begin{bmatrix} \mathbf{U}^s & \mathbf{U}^\nu \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_\nu \end{bmatrix} \begin{bmatrix} \mathbf{V}_s \\ \mathbf{V}_\nu \end{bmatrix} = \mathrm{SVD}\left(\mathbf{C}^x\right) \tag{6.23}$$

where $\mathbf{U}^s$ and $\mathbf{V}_s$ are called signal subspace matrices and $\mathbf{U}^\nu$ and $\mathbf{V}_\nu$ are called null or noise subspace matrices. Ideally, the singular values on the diagonal of $\mathbf{\Sigma}_\nu$ equal zero; however, in practice they are typically not exactly equal to zero due to estimation errors and influence from noise. Therefore the matrices $\mathbf{\Sigma}_\nu$, $\mathbf{U}^\nu$ and $\mathbf{V}_\nu$ carry the noise symbol ν.

From (6.17) it follows that the dimensions of the subspace matrices depend on properties such as the size and rank of the mixing matrix $\mathbf{A}$, the size and rank of the source autocorrelation matrix $\mathbf{C}^s$, the conjugation tuple that is used, and if uniquiefication is applied. Since the rank of both $(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})$ and $\mathbf{C}^s$ is at most equal to S, it follows immediately that the rank of $\mathbf{C}^x$ is at most equal to S. Consequently, the signal subspaces have at most dimension S. For sensor correlation matrices with a size of at least $S \times S$ and rank S, the sizes of the subspace matrices are as in Table 6.1.

For the source autocorrelation matrix $\mathbf{C}^s$ it follows that its rank is at most S since the number of source autocorrelation functions is at most S. In order to obtain this maximum rank it is required that the source autocorrelation functions are mutually linearly independent for the considered time-lag pairs [36]. A consequence of this requirement is that the number of time-lag pairs in the NF-ROS must be larger than or equal to S, i.e., $K \geq S$.

Table 6.1: Sizes of subspace matrices for both the regular and uniquiefied case.

Subspace matrix	Regular size	Uniquiefied size
$\mathbf{U}^s$	$D^2 \times S$	$\frac{1}{2}D(D+1) \times S$
$\mathbf{U}^\nu$	$D^2 \times D^2 - S$	$\frac{1}{2}D(D+1) \times \frac{1}{2}D(D+1) - S$
$\mathbf{V}_s$	$S \times K$	$S \times K$
$\mathbf{V}_\nu$	$K - S \times K$	$K - S \times K$

The Khatri-Rao product $(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})$ contains only S columns, limiting its rank to a maximum of S. In general, the Khatri-Rao product of two random matrices results in a full rank matrix. However, the Khatri-Rao product of a matrix with itself results in trivial linear dependencies, as we have seen in Section 6.3.1. For scenarios where uniquiefication is applied the matrix $(\mathbf{A}^c \bar{\diamond} \mathbf{A}^c)$ has a size of $\frac{1}{2}D(D+1) \times S$ and it typically has full rank in case the mixing matrix $\mathbf{A}$ is chosen randomly. Consequently, the matrix $\mathbf{C}^x$ has rank S as long as $D^2 \geq S$, or $\frac{1}{2}D(D+1) \geq S$ in case equal correlation functions exist.

From this analysis it follows that when the sensor correlation matrix has rank S, then the column space of $\mathbf{C}^x$ equals the column space of $(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})$ and it equals the columns space of $\mathbf{U}^s$, i.e., $\text{range}(\mathbf{U}^s) = \text{range}(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})$. This property creates opportunities for building systems of homogeneous polynomial equations from which the mixing parameters and source extraction filters can be identified.

6.4 Blind system identification as root finding for systems of homogeneous polynomial equations

Since the column vectors in subspace matrix $\mathbf{U}^\nu$ are orthogonal to the column vectors in subspace matrix $\mathbf{U}^s$, and thus the columns in sensor correlation matrix $\mathbf{C}^x$, they must also be orthogonal to the Khatri-Rao product $(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})$, i.e.,

$$(\mathbf{U}^\nu)^T \mathbf{C}^x = \mathbf{0}_{D^2 - S \times K} \quad \Longrightarrow \quad (\mathbf{U}^\nu)^T (\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2}) = \mathbf{0}_{D^2 - S \times S}. \tag{6.24}$$

A similar equation holds in case uniquiefication is applied; however, from now on we only discuss results after uniquiefication in case significant differences occur.

In [36, 70] it is shown that (6.24) can be used to formulate the blind mixing matrix identification problem as a root finding problem for a system of homogeneous polynomial equations. More specifically, it is shown that each root corresponds to the elements in one column of the mixing matrix. In this section we discuss some important properties of homogeneous polynomial equations and we present the system of homogeneous polynomial equations from which mixing column vectors can be identified. This system of polynomial equations is used later on in optimization problems to identify specific source extraction filters.

6.4.1 Homogeneous polynomial equations

First the definition of a homogeneous polynomial equation is given. Subsequently we present a scaling and multiplicative property of homogeneous polynomial equations that are used later on to solve the source extraction problem via a polynomial optimization problem.

Definition 6.4.1 (Homogenous polynomial equation)**:** Given a polynomial equation $f(\mathbf{z}) = 0$ of order p with and $\mathbf{z}$ containing the variables z_i for $i \in \mathcal{D}$. The polynomial equation $f(\mathbf{z}) = 0$ is called a p'th order *homogeneous polynomial equation* if for all non-zero coefficients the corresponding monomials, i.e., products of the variables z_i, have order p.

Example 6.4.1: An example second order homogeneous polynomial equation with variables z_i and coefficients $\phi^{i_1 i_2}$ is as follows:

$$f(\mathbf{z}) = \phi^{11} z_1 z_1 + \phi^{12} z_1 z_2 + \phi^{22} z_2 z_2 = 0 \tag{6.25}$$

where it follows that the monomials with non-zero coefficients, i.e., $z_1 z_1$, $z_1 z_2$, and $z_2 z_2$, have order two.

In case some of the variables in the monomials are conjugated, such as in the following equation:

$$f(\mathbf{z}) = \phi^{11} z_1 z_1^* + \phi^{12} z_1 z_2^* + \phi^{21} z_2 z_1^* + \phi^{22} z_2 z_2^* = 0 \tag{6.26}$$

then the equation is not a polynomial equation anymore. In [36] these equations are called *polyconjugal* equations; however, in this work we keep calling such equations polynomial equations. This approach is justified by treating the variables and their conjugated counterpart as independent variables, i.e., the conjugation relationship is not exploited. Note that under these conditions, (6.26) becomes a homogeneous polynomial equation with four second order monomials $z_1 z_1^*$, $z_1 z_2^*$, $z_2 z_1^*$, $z_2 z_2^*$ in four independent variables z_1, z_2, z_1^*, and z_2^*.

Theorem 6.4.1 (Scaling property of homogeneous polynomials)**:** Homogeneous polynomial $f(\mathbf{z})$ of degree p with q conjugations for all monomials having non-zero coefficients have the following structure:

$$f(\mathbf{z}) = \sum_{i_1, \cdots, i_p \in \mathcal{D} \times \cdots \times \mathcal{D}} \phi^{i_1 \cdots i_p} z_{i_1} \cdots z_{i_{p-q}} z^*_{i_{p-q+1}} \cdots z^*_{i_p}. \tag{6.27}$$

Here we used the property that the order of variables in a monomial can be chosen arbitrarily, i.e., $z_1 z_2^* = z_2^* z_1$. Homogeneous polynomials as in (6.27) have the following scaling property when conjugated variables are treated as independent variables:

$$f(\eta \mathbf{z}) = (\eta)^p f(\mathbf{z}) \qquad \forall\, \eta \in \mathbb{C}, \qquad \forall\, \mathbf{z} \in \mathbb{C}_D. \tag{6.28}$$

In case the conjugation property is taken into account then this scaling property is as follows:

$$f(\eta\mathbf{z}) = (\eta)^{p-q}(\eta^*)^q f(\mathbf{z}) \qquad \forall\, \eta \in \mathbb{C}, \qquad \forall\, \mathbf{z} \in \mathbb{C}_D. \tag{6.29}$$

Proof. When the conjugated variables are treated as independent variables then we must replace the variables z_i by ηz_i and z_i^* by ηz_i^*, which leads to the following:

$$\begin{aligned} f(\eta\mathbf{z}) &= \sum_{i_1\cdots i_p} \phi^{i_1\cdots i_p} \eta z_{i_1} \cdots \eta z_{i_{p-q}} \eta z^*_{i_{p-q+1}} \cdots \eta z^*_{i_p} \\ &= (\eta)^p \sum_{i_1\cdots i_p} \phi^{i_1\cdots i_p} z_{i_1} \cdots z_{i_{p-q}} z^*_{i_{p-q+1}} \cdots z^*_{i_p} = (\eta)^p f(\mathbf{z}). \end{aligned} \tag{6.30}$$

When the relationship $z_i = (z_i^*)^*$ is exploited, then we must replace the variables z_i by ηz_i and z_i^* by $\eta^* z_i^*$, which leads to the following:

$$\begin{aligned} f(\eta\mathbf{z}) &= \sum_{i_1\cdots i_p} \phi^{i_1\cdots i_p} \eta z_{i_1} \cdots \eta z_{i_{p-q}} \eta^* z^*_{i_{p-q+1}} \cdots \eta^* z^*_{i_p} \\ &= (\eta)^{p-q}(\eta^*)^q \sum_{i_1\cdots i_p} \phi^{i_1\cdots i_p} z_{i_1} \cdots z_{i_{p-q}} z^*_{i_{p-q+1}} \cdots z^*_{i_p} \\ &= (\eta)^{p-q}(\eta^*)^q f(\mathbf{z}). \end{aligned} \tag{6.31}$$

□

Note that when conjugated variables are considered to be independent variables, then the condition of having q conjugated variables for each monomial is not required for the scaling property to hold. However, from now on we assume that this condition holds for all homogenous polynomial equations, except when it is explicitly mentioned.

A direct result of the scaling properties is that the following holds for homogeneous polynomial equations:

$$f(\mathbf{z}) = 0 \qquad \Longleftrightarrow \qquad f(\eta\mathbf{z}) = 0 \qquad \forall\, \eta \in \mathbb{C}\backslash 0, \qquad \forall\, \mathbf{z} \in \mathbb{C}_D. \tag{6.32}$$

Thus, if $\mathbf{z}$ is a solution of the system of polynomial equations, then also $\eta\mathbf{z}$ solves the system. Since we do not allow for $\eta = 0$, the converse also holds. This property is related to the scaling indeterminacy in blind signal processing and cannot be solved without exploiting additional information [8, 11].

Finally, the following multiplicative property holds for polynomial equations:

$$f(\mathbf{z}) = 0 \quad \Longrightarrow \quad z_i^c f(\mathbf{z}) = 0 \qquad \forall\, i \in \mathcal{D},\ c \in \{\circ, *\},\ \mathbf{z} \in \mathbb{C}_D\backslash\mathbf{0}_D. \tag{6.33}$$

Notice that the converse of this multiplicative property is not true since the right system is also zero if $z_i^c = 0$. This multiplicative property of polynomials is used later in two ways. First, it is used to design optimization problems such that identification of the mixing column that corresponds to the desired source can be guaranteed. Second, it is used to solve under-determined source extraction problems.

6.4.2 Constructing a system of homogeneous polynomial equations

The system of homogeneous polynomial equations for blind system identification is constructed as follows. By considering (6.24) and the structure of the Khatri-Rao product $(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})$ it follows that the following set of equations must hold for the known subspace matrix $\mathbf{U}^{\nu}$ and the unknown mixing column vectors $\mathbf{a}^j$ for $j \in \mathcal{S}$:

$$(\mathbf{U}^{\nu})^H (\mathbf{a}^{j,c_1} \otimes \mathbf{a}^{j,c_2}) = \mathbf{0}_{D^2} \quad \forall\, j \in \mathcal{S}. \tag{6.34}$$

Replacing the unknown mixing column vectors in (6.34) by the vector $\mathbf{z}$ containing variables z_i for $i \in \mathcal{D}$ leads to the following system of second order homogeneous polynomial equations in vector-matrix notation [36]:

$$(\mathbf{U}^{\nu})^H (\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) = \mathbf{0}_{D^2}. \tag{6.35}$$

Another way to formulate (6.35) is in the following form of a D-variate system of second order homogeneous polynomial equations [36]:

$$\{f_q(\mathbf{z})\}_{q\in\mathcal{Q}} = \left\{ \sum_{(i_1,i_2)\in\mathcal{D}\times\mathcal{D}} \phi_q^{i_1 i_2} z_{i_1}^{c_1} z_{i_2}^{c_2} = 0 \right\}_{q\in\mathcal{Q}} \tag{6.36}$$

where coefficient $\phi_q^{i_1 i_2}$ corresponds to the conjugate of the element at the q'th column and $(i_1 D + i_2)$'th row of $\mathbf{U}^{\nu}$ and $\mathcal{Q} = \{q \in \mathbb{N} | \, 1 \leq q \leq Q\}$, with Q being the number of columns in $\mathbf{U}^{\nu}$, i.e., $Q = D^2 - S$. Note that with uniquiefication the system contains less equations and monomials; however, the principle remains the same.

From (6.34) and the scaling property of homogeneous polynomials it follows that $\mathbf{z} = \eta \mathbf{a}^j$ for $j \in \mathcal{S}$ and arbitrary scaling $\eta \in \mathbb{C}\backslash 0$ form S essentially unique solutions of (6.35) and (6.36). Furthermore, from linear algebra it follows that the nullspace of nullspace matrix $\mathbf{U}^{\nu}$ has dimension S such that at most S linearly independent solution vectors $(\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2})$ exist. In [36, 70] it is shown via geometry that the S solutions $\mathbf{z} = \eta \mathbf{a}^j$, for $j \in \mathcal{S}$ and $\eta \in \mathbb{C}\backslash 0$ an arbitrary scaling, are the only solutions for the considered system of homogeneous polynomial equations.

Later on, systems of homogeneous polynomial equations as in (6.35) and (6.36) are used to formulate the source extraction problem as a polynomial optimization problem; however, first we show how solutions of the system of polynomial equations can be found using linear algebra techniques.

6.4.3 Solving systems of homogeneous polynomial equations by means of a multi-matrix GEVD

Several techniques are available to solve systems of (homogeneous) polynomial equations; however, in general it is difficult to obtain stable algorithms for empirical data. A well known and widely used technique for solving such systems is the homotopy method [71]. The homotopy method uses a known start system and moves from the start system to the unknown system while tracking its roots. Disadvantages of

methods based on homotopy are that it is impossible to predict which root will be found. Furthermore, it is not clear if all roots are found and the same root can be found multiple times. Therefore, in this thesis we focus on a method that uses well known linear algebra techniques such as linear subspaces and (generalized) eigenvalue decompositions to find roots of polynomial systems [72].

As shown in (6.35), the system of polynomial equations $\{f_q(\mathbf{z}) = 0\}_{q\in\mathcal{Q}}$ from (6.36) can be written in the following vector-matrix form:

$$\mathbf{\Phi}_2(\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) = \mathbf{0}_Q \quad \text{with} \quad \mathbf{\Phi_2} = \begin{bmatrix} \phi_1^{11} & \phi_1^{12} & \cdots & \phi_1^{DD} \\ \vdots & \ddots & \ddots & \vdots \\ \phi_Q^{11} & \phi_Q^{12} & \cdots & \phi_Q^{DD} \end{bmatrix} \tag{6.37}$$

where $\mathbf{\Phi}_2$ is a polynomial coefficient matrix and $\mathbf{0}_Q$ is a length Q column vector containing only zeros. The subscript in the polynomial coefficient matrix $\mathbf{\Phi}_2$ is used for indicating the degree of the considered homogeneous polynomials. In this specific case the coefficient matrix is given as $\mathbf{\Phi}_2 = (\mathbf{U}^\nu)^T$; however, later on coefficient matrices with different structures are presented.

From (6.37) it follows immediately that the solutions $(\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2})$ must be found in the linear subspace orthogonal to the polynomial coefficient matrix $\mathbf{\Phi}_2$. A representation of this orthogonal subspace for this specific coefficient matrix is found in the form of signal subspace matrix $\mathbf{U}^s$, which has the same number of columns as the number of solutions for the system of polynomial equations. Although the signal subspace matrix $\mathbf{U}^s$ represents the solution space, its column vectors do not yet have the Kronecker structure $(\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2})$. However, there must exist an invertible matrix $\mathbf{V}$ such that the following holds [36]:

$$\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2} = \mathbf{U}^s\mathbf{V} \tag{6.38}$$

where $\mathbf{V}$ is a matrix that restores the Kronecker structure $(\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2})$ in the signal subspace matrix $\mathbf{U}^s$. This property can also be understood from the structure in the sensor correlation matrix $\mathbf{C}^x$ as is presented in (6.17). Without applying subspace decompositions it follows that the (pseudo-)inverse of the source autocorrelation matrix $\mathbf{C}^s$ has to be applied on the right hand side of the sensor correlation matrix, i.e.,

$$\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2} = \mathbf{C}^x \left(\mathbf{C}^s\right)^\dagger \equiv (\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2})\mathbf{C}^s \left(\mathbf{C}^s\right)^\dagger. \tag{6.39}$$

One way to restore the structure in the signal subspace matrix $\mathbf{U}^s$ is by means of a multi-matrix GEVD. If we split the subspace matrix $\mathbf{U}^s$ of size $D^2 \times S$ into D non-overlapping submatrices $\mathbf{U}_1^s, \cdots, \mathbf{U}_D^s$ of size $D \times S$, with submatrix $\mathbf{U}_i^s$ consisting of the i'th set of D rows from $\mathbf{U}^s$, then the following multi-matrix GEVD can be used to reconstruct the structure in $\mathbf{U}^s$ and to identify the mixing column vectors [36]:

$$\mathbf{z}^{c_2} = \frac{1}{z_1^{c_1}}\mathbf{U}_1^s\boldsymbol{\mu} = \frac{1}{z_2^{c_1}}\mathbf{U}_2^s\boldsymbol{\mu} = \cdots = \frac{1}{z_D^{c_1}}\mathbf{U}_D^s\boldsymbol{\mu}. \tag{6.40}$$

The structure of the selected submatrices $\mathbf{U}_i^s$ is as follows:

$$\mathbf{U}_i^s \equiv \mathbf{A}^{c_2} \operatorname{diag}\left(\tilde{\mathbf{a}}_i^{c_1}\right) (\mathbf{V})^{-1} \quad \forall\, i \in \mathcal{D} \tag{6.41}$$

where $\tilde{\mathbf{a}}_i^{c_1}$ is the i'th row of $\mathbf{A}^{c_1}$ and $\mathbf{V}$ is the matrix that restores the Kronecker structure in $\mathbf{U}^s$.

For well conditioned mixing systems with $D \geq S$ it follows that the multi-matrix GEVD can be solved by calculating the eigenvalue decomposition of the product $\left(\mathbf{U}_{i_2}^s\right)^\dagger \mathbf{U}_{i_1}^s$ for $\{(i_1, i_2) \in \mathcal{D} \times \mathcal{D} | \, i_1 \neq i_2\}$, i.e.,

$$\left(\mathbf{U}_{i_2}^s\right)^\dagger \mathbf{U}_{i_1}^s \boldsymbol{\mu} = \lambda \boldsymbol{\mu} \tag{6.42}$$

leads to S eigenvalue-eigenvector pairs $(\lambda^j, \boldsymbol{\mu}^j)$ for $j \in \mathcal{S}$. As long as the eigenvalues are different and finite, then all the eigenvectors can be calculated. If these conditions are not met, then a different set or linear combinations of submatrices can be used, as will be shown later on.

Suppose we create the matrix $\hat{\mathbf{V}}$ of size $S \times S$ by stacking all eigenvectors $\boldsymbol{\mu}^1, \cdots, \boldsymbol{\mu}^S$. By applying the matrix $\hat{\mathbf{V}}$ to the subspace matrix $\mathbf{U}^s$ we reconstruct the Kronecker structure, i.e.,

$$\mathbf{U}^s \hat{\mathbf{V}} = \left(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2}\right) \left(\mathbf{V}\right)^{-1} \hat{\mathbf{V}} = \left(\mathbf{A}^{c_1} \diamond \mathbf{A}^{c_2}\right) \mathbf{D}\mathbf{P} \tag{6.43}$$

where $\mathbf{D}$ and $\mathbf{P}$ are an unknown scaling and permutation matrices. The unknown, arbitrary scaling $\mathbf{D}$ comes from the fact that $\eta\boldsymbol{\mu}$ for all $\eta \in \mathbb{C} \backslash 0$ is an eigenvector if $\boldsymbol{\mu}$ is an eigenvector. The permutation matrix $\mathbf{P}$ follows from the fact that the order of the eigenvectors is arbitrary. Both these problems are well known in the field of blind signal processing [8] and cannot be solved without using more a priori information.

Now that we have restored the Kronecker structure in the matrix $\mathbf{U}^s \hat{\mathbf{V}}$, the mixing column vectors can be found in arbitrary order and with an arbitrary scaling in the first D rows of $\mathbf{U}^s \hat{\mathbf{V}}$. An alternative way to calculate these estimated mixing column vectors $\hat{\mathbf{a}}^j$ is from the products $\hat{\mathbf{a}}^j = \mathbf{U}_i^s \boldsymbol{\mu}^j$ for $j \in \mathcal{S}$, where $\mathbf{U}_i^s$ is an arbitrary submatrix.

From previous chapters we know that in case $D \geq S$ the source extraction filters $\tilde{\mathbf{w}}_j$ for $j \in \mathcal{S}$ can also be identified from the GEVD of submatrices $\mathbf{U}_i^s$. In order to identify these filters the left instead of the right eigenvectors have to be calculated from the following multi-matrix GEVD, i.e.,

$$\frac{1}{z_1} \tilde{\boldsymbol{\mu}} \mathbf{U}_1^s = \frac{1}{z_2} \tilde{\boldsymbol{\mu}} \mathbf{U}_2^s = \cdots = \frac{1}{z_D} \tilde{\boldsymbol{\mu}} \mathbf{U}_D^s. \tag{6.44}$$

If these filters exist, i.e., if $D \geq S$, then the generalized eigenvectors $\tilde{\boldsymbol{\mu}}_j$ can be calculated as the eigenvectors corresponding to the S non-zero eigenvalues of the following eigenvalue decomposition problem:

$$\tilde{\boldsymbol{\mu}} \mathbf{U}_{i_1}^s \left(\mathbf{U}_{i_2}^s\right)^\dagger = \lambda \tilde{\boldsymbol{\mu}} \tag{6.45}$$

where both $\mathbf{U}_{i_1}^s$ and $\mathbf{U}_{i_1}^s$ must have rank S.

From the structure in the submatrices $\mathbf{U}_i^s$, see (6.41), it follows that left eigenvectors are scaled rows from the inverse of $\mathbf{A}^{c_2}$. Consequently, the source extraction

filters $\tilde{\mathbf{w}}_j$ are calculated, in arbitrary order, as $\tilde{\mathbf{w}}_j = \tilde{\boldsymbol{\mu}}_j^{c_2}$. Applying these filters to the mixing system leads to $\tilde{\mathbf{w}}_j \mathbf{A} = \eta \tilde{\mathbf{e}}_j$ for all $j \in \mathcal{S}$ and $\eta \in \mathbb{C} \backslash 0$.

Finally, an alternative method to identify the S mixing column vectors is via the following eigenvalue decomposition:

$$\mathbf{U}_{i_1}^s \left(\mathbf{U}_{i_2}^s\right)^\dagger \boldsymbol{\mu} = \lambda \boldsymbol{\mu} \tag{6.46}$$

where again the eigenvectors corresponding to the S non-zero eigenvalues have to be selected. The estimated mixing column vectors are then found, in arbitrary order, as $\mathbf{a}^j = \eta \boldsymbol{\mu}^{j,c_2}$ for all $j \in \mathcal{S}$ and $\eta \in \mathbb{C} \backslash 0$. This means that a single matrix product can be used to identify both the mixing column vectors $\mathbf{a}^j$ and the source extraction filters $\tilde{\mathbf{w}}_j$.

Note that the following relationship holds between the submatrices $\mathbf{U}_i^s$ and sensor correlation matrices $\mathbf{C}_i^x$ are built from the corresponding rows of $\mathbf{C}^x$:

$$\mathbf{C}_i^x = \mathbf{U}_i^s \boldsymbol{\Sigma}_s \mathbf{V}_s \quad \Longleftrightarrow \quad \mathbf{U}_i^s = \mathbf{C}_i^x \left(\boldsymbol{\Sigma}_s \mathbf{V}_s\right)^\dagger . \tag{6.47}$$

Consequently, for the matrix product in (6.46) we have the following:

$$\mathbf{U}_{i_1}^s \left(\mathbf{U}_{i_2}^s\right)^\dagger \quad = \quad \mathbf{C}_{i_1}^s \left(\mathbf{C}_{i_2}^s\right)^\dagger \tag{6.48}$$

where the matrices $\mathbf{U}_i$ have size $D \times S$ and the matrices $\mathbf{C}_i^x$ have size $D \times K$. Note that the sensor correlation matrices $\mathbf{C}_i^x$ have the same structure as in (3.5), (4.16), or (4.17), depending on type of conjugation. In this chapter we consider the source extraction problem in the context of systems of polynomial equations and linear subspaces of the matrices representing them. Therefore we keep using the signal subspace matrix $\mathbf{U}^s$ and its submatrices instead of the sensor correlation matrices $\mathbf{C}_i^x$.

Summary of a BSI and BSS algorithm Identification of all the mixing columns $\mathbf{a}^j$ or source extraction filters $\tilde{\mathbf{w}}_j$ corresponds to solving the BSI and BSS problem, respectively. Given the number of sources S, knowledge about a NF-ROS, and the constraint $D \geq S$, the BSI and BSS problems can be solved via the following steps:

1. Calculate the sensor correlation matrix $\mathbf{C}^x$ as defined in (6.16).
2. Extract two submatrices $\mathbf{U}_{i_1}^s$ and $\mathbf{U}_{i_2}^s$ from the signal subspace matrix $\mathbf{U}^s$, e.g., obtained via SVD $(\mathbf{C}^x)$.
3. For BSI: Calculate the eigenvectors corresponding to the S non-zero eigenvalues of (6.46), leading to $\mathbf{a}^j = \eta \boldsymbol{\mu}^{j,c_2}$.
4. For BSS: Calculate the eigenvectors corresponding to the S non-zero eigenvalues of (6.45), leading to $\tilde{\mathbf{w}}_j = \eta \tilde{\boldsymbol{\mu}}_j^{c_2}$.

When the eigenvectors are calculated one by one, then these steps constitute a (sequential) BSE algorithm.

In the following section we present an example where we calculate both the S mixing column vectors $\mathbf{a}^i$ and the S source extraction filters $\tilde{\mathbf{w}}_j$. Subsequently, we show that a priori information can be incorporated into the eigenvalues to enable identification of only the desired mixing column vector and source extraction filter.

6.4.4 Blind identification example for $D = 2$ and $S = 2$

For a mixing system consisting of two sources and two sensors, the sensor correlation matrix $\mathbf{C}^x$ with conjugation tuple $(\circ*)$ has the following structure when considering $K = 2$ time-lag pairs in the NF-ROS:

$$\mathbf{C}^x = \begin{bmatrix} r_{11}^x[\Omega_1] & r_{11}^x[\Omega_2] \\ r_{12}^x[\Omega_1] & r_{12}^x[\Omega_2] \\ r_{21}^x[\Omega_1] & r_{21}^x[\Omega_2] \\ r_{11}^x[\Omega_1] & r_{22}^x[\Omega_2] \end{bmatrix} = \begin{bmatrix} \mathbf{U}^s & \mathbf{U}^\nu \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_\nu \end{bmatrix} \begin{bmatrix} \mathbf{V}_s \\ \mathbf{V}_\nu \end{bmatrix}. \tag{6.49}$$

Since the rank of the uniquiefied sensor correlation matrix is S, the nullspace matrix $\mathbf{U}^\nu$ has size 2×4, leading to the following system of polynomial equations:

$$\mathbf{\Phi}_2(\mathbf{z} \otimes \mathbf{z}^*) = \begin{bmatrix} \phi_1^{11} & \phi_1^{12} & \phi_1^{21} & \phi_1^{22} \\ \phi_2^{11} & \phi_2^{12} & \phi_2^{21} & \phi_2^{22} \end{bmatrix} \begin{bmatrix} z_1 z_1^* \\ z_1 z_2^* \\ z_2 z_1^* \\ z_2 z_2^* \end{bmatrix} = 0. \tag{6.50}$$

The solutions of this system of polynomial equations are embedded in the nullspace of coefficient matrix $\mathbf{\Phi}_2$, i.e., the size 4×2 matrix $\mathbf{U}^s$, which contains the following two submatrices of size 2×2:

$$\mathbf{U}^s = \begin{bmatrix} \mathbf{U}_1^s \\ \mathbf{U}_2^s \end{bmatrix}. \tag{6.51}$$

By solving the following GEVDs of submatrices $\mathbf{U}_1^s$ and $\mathbf{U}_2^s$ the mixing column vectors and source extraction filters can be calculated, i.e.,

$$\mathbf{z}^* = \lambda_d \mathbf{U}_1^s \boldsymbol{\mu} = \lambda_n \mathbf{U}_2^s \boldsymbol{\mu} \qquad \text{and} \qquad \lambda_d \tilde{\boldsymbol{\mu}} \mathbf{U}_1^s = \lambda_n \tilde{\boldsymbol{\mu}} \mathbf{U}_2^s \tag{6.52}$$

where for the solutions it holds that $\mathbf{a}^j = \eta_a \mathbf{z}^j$ and $\tilde{\mathbf{w}}_j = \eta_w \tilde{\boldsymbol{\mu}}_j^*$ for $j \in \mathcal{S}$, $\eta_a \in \mathbb{C} \backslash 0$, and $\eta_w \in \mathbb{C} \backslash 0$. Note that the solutions are found in arbitrary order and with arbitrary scaling.

6.5 Identification of the desired source extraction filter

In this section we show that immediate identification of the desired mixing column vector or source extraction filter can be accomplished by exploiting a priori information about the mixing parameters. The a priori information is incorporated into the eigenvalues such that the largest eigenvalue corresponds to the desired source. Subsequently, we exploit this insight in order to formulate an optimization problem for source extraction filter identification. In this optimization problem a system of polynomial equations is used to restrict the feasible set to mixing column vectors and source extraction filters while a fractional objective function contains a priori information to ensure selection of the desired mixing column or extraction filter.

6.5.1 Linear fractional structure in the eigenvalues

In the previous section we have seen that identification of mixing column vectors and source extraction filters can be accomplished by finding the roots of a system of polynomial equations. Furthermore, we have shown that these roots can be calculated via a multi-matrix GEVD. Another way to formulate the multi-matrix GEVD of submatrices as in (6.40) is as follows:

$$\mathbf{z}^{c_2} = \frac{1}{z_1^{c_1}}\mathbf{S}_1\mathbf{U}^s\boldsymbol{\mu} = \frac{1}{z_2^{c_1}}\mathbf{S}_2\mathbf{U}^s\boldsymbol{\mu} = \cdots = \frac{1}{z_D^{c_1}}\mathbf{S}_D\mathbf{U}^s\boldsymbol{\mu} \tag{6.53}$$

where the matrices $\mathbf{S}_i$ for $i \in \mathcal{D}$ are *selection matrices* of size $D \times D^2$ that have the following structure:

$$\mathbf{S}_i = \tilde{\mathbf{e}}_i \otimes \mathbf{I}_D \quad \forall\, i \in \mathcal{D} \tag{6.54}$$

with $\tilde{\mathbf{e}}_i$ a length D row vector with a one on the i'th column and zeros elsewhere and $\mathbf{I}_D$ the size $D \times D$ identity matrix. For example, when $D = 3$ then we have that $S_2 = \begin{bmatrix}\mathbf{0}_D & \mathbf{I}_D & \mathbf{0}_D\end{bmatrix}$.

With the design from (6.54), a selection matrix $\mathbf{S}_i$ collects a subset of rows in $\mathbf{U}^s$. Consequently, the products $\mathbf{S}_i\mathbf{U}^s$ equal the submatrices $\mathbf{U}_i$ such that the multi-matrix GEVD in (6.53) is the same as the multi-matrix GEVD in (6.40). As a result, the eigenvalue decompositions in (6.42), (6.45), and (6.46) have the following structure in their eigenvalues:

$$\lambda = \frac{\langle \tilde{\mathbf{e}}_{i_1}, \mathbf{z}^{c_1}\rangle}{\langle \tilde{\mathbf{e}}_{i_2}, \mathbf{z}^{c_1}\rangle} = \frac{z_{i_1}^{c_1}}{z_{i_2}^{c_1}}. \tag{6.55}$$

When the D vectors $\tilde{\mathbf{e}}_i$ to create the D selection matrices $\mathbf{S}_i$ for $i \in \mathcal{D}$ are replaced by L arbitrary vectors $(\boldsymbol{\xi}^l)^H$ to form L selection matrices $\mathbf{S}_l$, then the following multi-matrix GEVD of linear combinations of submatrices is formed, which is first formulated in (5.6.21) of [36]:

$$\mathbf{z}^{c_2} = \frac{1}{\langle \boldsymbol{\xi}^1, \mathbf{z}^{c_1}\rangle}\boldsymbol{\Gamma}_1\boldsymbol{\mu} = \frac{1}{\langle \boldsymbol{\xi}^2, \mathbf{z}^{c_1}\rangle}\boldsymbol{\Gamma}_2\boldsymbol{\mu} = \cdots = \frac{1}{\langle \boldsymbol{\xi}^L, \mathbf{z}^{c_1}\rangle}\boldsymbol{\Gamma}_L\boldsymbol{\mu} \tag{6.56}$$

where the matrices $\boldsymbol{\Gamma}_l = \mathbf{S}_l\mathbf{U}^s \equiv \sum_{i\in\mathcal{D}} \xi_i^{l,*}\mathbf{U}_i^s$ have size $D \times S$.

For well-determined mixtures with $D \geq S$, eigenvalue decompositions of two linear combinations of submatrices $\boldsymbol{\Gamma}_{l_1}$ and $\boldsymbol{\Gamma}_{l_2}$ instead of using two submatrices $\mathbf{U}_{i_1}^s$ and $\mathbf{U}_{i_2}^s$ results in the same eigenvectors as in (6.42), (6.45), and (6.46). However, the corresponding eigenvalues now have the following linear fractional structure:

$$\lambda = \frac{\langle \boldsymbol{\xi}^{l_1}, \mathbf{z}^{c_1}\rangle}{\langle \boldsymbol{\xi}^{l_2}, \mathbf{z}^{c_1}\rangle} = \frac{(\boldsymbol{\xi}^{l_1})^H\mathbf{z}^{c_1}}{(\boldsymbol{\xi}^{l_2})^H\mathbf{z}^{c_1}}. \tag{6.57}$$

When the vectors $\boldsymbol{\xi}^{l_1}$ and $\boldsymbol{\xi}^{l_2}$ are chosen based on a priori information such that the following condition holds:

$$\left|\frac{\langle \boldsymbol{\xi}^{l_1}, \mathbf{a}^d\rangle}{\langle \boldsymbol{\xi}^{l_2}, \mathbf{a}^d\rangle}\right| > \left|\frac{\langle \boldsymbol{\xi}^{l_1}, \mathbf{a}^j\rangle}{\langle \boldsymbol{\xi}^{l_2}, \mathbf{a}^j\rangle}\right| \quad \forall\{j \in \mathcal{S} |\, j \neq d\} \tag{6.58}$$

with d the index corresponding to the desired source, then identification of the desired mixing column vector and source extraction filter is guaranteed when the eigenvectors corresponding to the absolute largest eigenvalue is selected. This de-coupling of selection via eigenvalues and identification via eigenvectors ensures that the identified extraction filter is independent from the selection mechanism that is used, i.e., no bias is introduced due to the selection mechanism. Furthermore, by selecting the largest eigenvalue, always the solution with the best match with respect to the a priori information is selected.

Exploiting the structure in the eigenvalues allows for the formulation of desired mixing column vector or source extraction filter identification algorithm design as a design problem for the following polynomial optimization problem:

$$\begin{aligned} \max_{\mathbf{z}\in\mathbb{C}_D\backslash\mathbf{0}_D} \quad & \frac{\langle \boldsymbol{\xi}^1, \mathbf{z}^{c_1}\rangle}{\langle \boldsymbol{\xi}^2, \mathbf{z}^{c_1}\rangle} \\ \text{s. t.} \quad & f_q(\mathbf{z}) = 0 \quad \forall\, q \in \mathcal{Q}. \end{aligned} \tag{6.59}$$

In this optimization problem the polynomial constraints ensure that the feasible set consists of only mixing column vectors and source extraction filters, which was already formulated in [36]. The objective function can be designed based on a priori information by choosing appropriate weight vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ in order to ensure selection of the extraction filter corresponding to the desired source. As we have seen, solutions of such optimization problems can be found via largest eigenvalue decompositions.

In the current form, the amount of a priori information that can be incorporated into the eigenvalue decomposition problem is limited, which will be discuss in detail in Section 6.7. In the remainder of this section we show how the multiplicative property of polynomials can be used to allow for incorporating more advanced a priori information, leading to more flexible and robust source extraction algorithms.

6.5.2 Exploiting the polynomial multiplicative property to generate second order fractional objective functions.

The polynomial optimization problem in (6.59) and the corresponding method to solve the problem only allow for the design of linear fractional objective functions. Here we exploit the multiplicative property of polynomials as in (6.33) in order to design higher order fractional objective functions such that identification of the desired source extraction filter can be guaranteed based on more advanced a priori information.

From the multiplication property we know that for a single homogeneous polynomial equation $f_q(\mathbf{z}) = 0$ for $q \in \mathcal{Q}$ it holds that:

$$f_q(\mathbf{z}) = 0 \quad \Longrightarrow \quad z_i^c f_q(\mathbf{z}) = 0 \qquad \text{for } i \in \mathcal{D}, c \in \{\circ, *\},\ \forall \mathbf{z} \in \mathbb{C}_D\backslash\mathbf{0}_D. \tag{6.60}$$

The right hand side of (6.60) holds if $f_q(\mathbf{z})$, z_i^c, or both $f_q(\mathbf{z})$ and z_i^c equal zero. Since we do not allow that all z_i^c with $c \in \{\circ, *\}$ equal zero for all $i \in \mathcal{D}$ simultaneously, i.e.,

$\mathbf{z}^c \neq \mathbf{0}_D$, we know that the following relation holds between a polynomial equation $f_q(\mathbf{z}) = 0$ and a system of D polynomial equations as on the right hand side of (6.60):

$$f_q(\mathbf{z}) = 0 \quad \Longleftrightarrow \quad \{z_i^c f_q(\mathbf{z}) = 0\}_{i\in\mathcal{D}} \qquad \forall\, \mathbf{z} \in \mathbb{C}_D \backslash \mathbf{0}_D. \tag{6.61}$$

Consequently, given a system of polynomial equations $f_q(\mathbf{z})$ for all $q \in \mathcal{Q}$ we obtain the following relationship:

$$\{f_q(\mathbf{z}) = 0\}_{q\in\mathcal{Q}} \quad \Longleftrightarrow \quad \{z_i^c f_q(\mathbf{z}) = 0\}_{(i,q)\in\mathcal{D}\times\mathcal{Q}} \qquad \forall\, \mathbf{z} \in \mathbb{C}_D \backslash \mathbf{0}_D. \tag{6.62}$$

Note that these systems of polynomial equations are independent of the choice for the conjugation symbol $c \in \{\circ, *\}$. Furthermore, treating the variables $\mathbf{z}^*$ as independent variables with respect to $\mathbf{z}$, as we discussed in Section 6.4.1, leads to the additional condition that $\mathbf{z}^* \in \mathbb{C}\backslash 0$.

The property in (6.62) means that for the system of second order polynomials in (6.37) a system of third order polynomials exists that has the same roots. In vector-matrix notation this system of third order polynomial equations can be formulated as follows:

$$\mathbf{\Phi}_3(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) = \mathbf{0}_{DQ} \tag{6.63}$$

where $\mathbf{\Phi}_3$ is the third order coefficient matrix of size $DQ \times D^3$ with $Q = D^2 - S$. This coefficient matrix depends strongly on the coefficient matrix for the system of second order polynomial equations $\mathbf{\Phi}_2$ and is well structured, i.e.,

$$\mathbf{\Phi}_3(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) \equiv \bigoplus_{i\in\mathcal{D}} \mathbf{\Phi}_2(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}). \tag{6.64}$$

Note that this structure only holds because we inserted the vectors with variables $\mathbf{z}^{c_0}$ on the left hand side in (6.63). A change in the order of the vectors $\mathbf{z}^{c_0}$, $\mathbf{z}^{c_1}$, and $\mathbf{z}^{c_2}$ results in a shuffling of the rows of coefficient matrix $\mathbf{\Phi}_3$. Consequently, the matrix would still be well structured and lead to the same results; however, it is harder to derive a concise mathematical description in terms of the coefficient matrix $\mathbf{\Phi}_2$.

A similar system of third order polynomial equations can be derived for the uniquiefied system of second order polynomial equations. This system has the following structure:

$$\mathbf{\Phi}_3(\mathbf{z}^{c_0} \otimes (\mathbf{z}^{c_1} \breve{\otimes} \mathbf{z}^{c_1})) = \mathbf{0}_{DQ} \tag{6.65}$$

where $\mathbf{\Phi}_3$ has size $DQ \times \frac{1}{2}D^2(D+1)$ with $Q = \frac{1}{2}D(D+1) - S$ and $\breve{\otimes}$ denotes the uniquiefied Kronecker product. The structure in the corresponding coefficient matrix for the system of third order polynomial equations is as follows:

$$\mathbf{\Phi}_3(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \breve{\otimes} \mathbf{z}^{c_1}) \equiv \bigoplus_{i\in\mathcal{D}} \mathbf{\Phi}_2(\mathbf{z}^{c_0} \otimes (\mathbf{z}^{c_1} \breve{\otimes} \mathbf{z}^{c_1})) \tag{6.66}$$

where $\bigoplus$ is the direct sum operator.

For a mixture with two sources and two sensors, this results in the following form for coefficient matrix $\mathbf{\Phi}_3$:

$$\mathbf{\Phi}_3 = \begin{bmatrix} \mathbf{\Phi}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}_2 \end{bmatrix} \tag{6.67}$$

where only the size of $\mathbf{\Phi}_3$ is depending on uniquiefication for the system of second order polynomial equations and not its structure.

From (6.64) and (6.66) it follows that currently the system of third order polynomial equations consists of D orthogonal subsystems that can be solved independently. In order to benefit from building a system of third order polynomial equations the orthogonal subsystems have to be mixed. In the remainder of this section we show that applying uniquiefication from Section 6.3.1 to the system of third order polynomial equations leads the desired mix of subsystems such that the source extraction problem can be solved via a polynomial optimization problem with a quadratic fractional objective function.

In order to be able to apply uniquiefication to the systems in (6.64) and (6.66) it is required that the conjugation applied to the first vector of variables $\mathbf{z}^{c_0}$ must also be applied to one of the other two vectors of variables such that at least $c_0 = c_1$ or $c_0 = c_2$ holds. Note that for the uniquiefied case this automatically means that all conjugations are the same, i.e., $c_0 = c_1 = c_2$. Without loss of generality, we assume in the remainder of this section that the conjugations c_0 and c_1 are the same, i.e., $c_0 = c_1$. Applying the uniquiefication with respect to variables $\mathbf{z}^{c_0}$ and $\mathbf{z}^{c_1}$ leads to the following uniquiefied system of polynomials in vector-matrix format:

$$\breve{\mathbf{\Phi}}_3((\mathbf{z}^{c_1} \breve{\otimes} \mathbf{z}^{c_1}) \otimes \mathbf{z}^{c_2}) = 0 \tag{6.68}$$

and a similar system in case uniquiefication was applied to a system of second order polynomial equations, i.e.,

$$\breve{\mathbf{\Phi}}_3(\mathbf{z}^{c_1} \breve{\otimes} \mathbf{z}^{c_1} \breve{\otimes} \mathbf{z}^{c_1}) = 0. \tag{6.69}$$

Applying uniquiefication to the system of third order polynomial equations leads to a reduction in the number of columns in $\mathbf{\Phi}_3$. This reduction is from D^3 to $\frac{1}{2}D^2(D+1)$ columns; and, in case of uniquiefication for the system of second order polynomial equations, from $\frac{1}{2}D(D+1)$ to $\frac{1}{6}D(D+1)(D+2)$ columns. If now DQ is larger than or equal to the number of columns in $\breve{\mathbf{\Phi}}_3$, then the number of equations is sufficient to restrict to S solutions; however, this only holds if sufficient equations are linearly independent, i.e., if the rank of $\breve{\mathbf{\Phi}}_3$ is large enough. In simulations with randomly chosen mixing systems this condition on the rank was met every time.

By calculating the signal subspace of the uniquiefied coefficient matrix $\breve{\mathbf{\Phi}}_3$, again denoted by $\mathbf{U}^s$, we are able to obtain submatrices $\mathbf{U}^s_{i_1 i_2}$ of size $D \times S$ for $\{(i_1, i_2) \in \mathcal{D} \times \mathcal{D}\}$, or $\{(i_1, i_2) \in \mathcal{D} \times \mathcal{D} | \, i_1 \leq i_2\}$ in case of uniquiefication, where submatrix $\mathbf{U}^s_{i_1 i_2}$ is obtained by collecting rows from $\mathbf{U}^s$ corresponding to the monomials $z_{i_1}^{c_1} z_{i_2}^{c_2} \otimes \mathbf{z}^{c_1}$. The structure in these submatrices is as follows:

$$\mathbf{U}^s_{i_1 i_2} = \mathbf{A}^{c_1} \operatorname{diag}\left(\tilde{\mathbf{a}}_{i_1}^{c_1} \odot \tilde{\mathbf{a}}_{i_2}^{c_2}\right) (\mathbf{V})^{-1} \tag{6.70}$$

where $\odot$ represents element wise multiplication.

Following the strategy from the previous section it follows that the following multi-matrix GEVD holds for the submatrices $\mathbf{U}^s_{i_1 i_2}$:

$$\mathbf{z}^{c_1} = \frac{1}{z_1^{c_1} z_1^{c_2}} \mathbf{U}^s_{11} \boldsymbol{\mu} = \frac{1}{z_1^{c_1} z_2^{c_2}} \mathbf{U}^s_{12} \boldsymbol{\mu} = \cdots = \frac{1}{z_D^{c_1} z_D^{c_2}} \mathbf{U}^s_{DD} \boldsymbol{\mu}. \tag{6.71}$$

Consequently, for well determined mixtures with $D \geq S$ the following eigenvalue decompositions can be used to identify mixing column vectors and source extraction filters:

$$\left(\mathbf{U}^s_{i_3 i_4}\right)^\dagger \mathbf{U}^s_{i_1 i_2} \boldsymbol{\mu} = \lambda \boldsymbol{\mu} \quad \longrightarrow \quad \mathbf{a}^j = \eta \mathbf{U}_{i_1 i_2} \boldsymbol{\mu}^{j,c_1} \tag{6.72}$$

$$\mathbf{U}^s_{i_1 i_2} \left(\mathbf{U}^s_{i_3 i_4}\right)^\dagger \boldsymbol{\mu} = \lambda \boldsymbol{\mu} \quad \longrightarrow \quad \mathbf{a}^j = \eta \boldsymbol{\mu}^{j,c_1} \tag{6.73}$$

$$\tilde{\boldsymbol{\mu}} \mathbf{U}^s_{i_1 i_2} \left(\mathbf{U}^s_{i_3 i_4}\right)^\dagger = \lambda \tilde{\boldsymbol{\mu}} \quad \longrightarrow \quad \tilde{\mathbf{w}}_j = \eta \tilde{\boldsymbol{\mu}}_j^{c_1} \tag{6.74}$$

where $\{(i_1, i_2, i_3, i_4) \in \mathcal{D} \times \mathcal{D} \times \mathcal{D} \times \mathcal{D}\}$ and for the last two eigenvalue decompositions the eigenvectors corresponding to the S non-zero eigenvalues have to be selected. The S non-zero eigenvalues of the eigenvalue decomposition problems in (6.72), (6.73), and (6.74) have the following quadratic fractional structure:

$$\lambda = \frac{z_1^{c_1} z_2^{c_2}}{z_3^{c_1} z_4^{c_2}}. \tag{6.75}$$

Again, taking linear combinations of the submatrices $\mathbf{U}^s_{i_1 i_2}$ allows for the incorporation of a priori information into the eigenvalues. The structure in the eigenvalues for an eigenvalue decomposition of linear combinations of submatrices $\mathbf{U}^s_{i_1 i_2}$ is as follows:

$$\lambda = \frac{\langle \boldsymbol{\xi}^1, \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2} \rangle}{\langle \boldsymbol{\xi}^2, \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2} \rangle}. \tag{6.76}$$

Proper design of the vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ based on a priori information about the mixing parameters allows for identification of the desired mixing column vector or source extraction filter via a largest eigenvalue decomposition problem.

In a generalized form, the identification problem can be formulated as the following polynomial optimization problem with quadratic fractional objective function:

$$\begin{aligned} \max_{\mathbf{z} \in \mathbb{C}_D \setminus \mathbf{0}_D} \quad & \frac{(\mathbf{z}^{c_1})^T \boldsymbol{\Xi}_1 \mathbf{z}^{c_2}}{(\mathbf{z}^{c_1})^T \boldsymbol{\Xi}_2 \mathbf{z}^{c_2}} \\ \text{s. t.} \quad & f_q(\mathbf{z}) = 0 \quad \forall\, q \in \mathcal{Q}. \end{aligned} \tag{6.77}$$

The matrices $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$ represent the coefficients for taking linear combinations of submatrices $\mathbf{U}^s_{i_1 i_2}$ in such a way that the following equation holds:

$$(\mathbf{z}^{c_1})^T \boldsymbol{\Xi}_l \mathbf{z}^{c_2} = \sum_{(i_1 i_2) \in \mathcal{D} \times \mathcal{D}} \xi^{l,*}_{i_1 i_2} z_{i_1}^{c_1} z_{i_2}^{c_2} \tag{6.78}$$

where $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ are the length D^2 vectors for taking linear combinations of submatrices $\mathbf{U}^s_{i_1 i_2}$. Note that in case uniquiefication is applied to the system of second order polynomial equations then the vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ have a length of $\frac{1}{2} D(D+1)$.

Summary of the source extraction algorithm Suppose the number of sources S is known, knowledge about a NF-ROS is available, and the constraint $D \geq S$ holds. Given a priori information about the mixing parameters of the sources such that a quadratic fractional objective function as in (6.77), and thus the matrices $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$, can be designed such that it is maximum for the mixing parameters corresponding to the desired source s_d, then identification of the mixing column vector $\mathbf{a}^d$ and source extraction filter $\tilde{\mathbf{w}}_d$ can be accomplished via the following steps:

1. Calculate sensor correlation matrix $\mathbf{C}^x$.
2. Calculate null/noise subspace matrix $\mathbf{U}^\nu$, e.g., via an SVD.
3. Apply the polynomial multiplication by shifting rows of $\boldsymbol{\Phi}_2 = (\mathbf{U}^\nu)^H$ and apply uniquiefication, leading to the uniquiefied coefficient matrix $\breve{\boldsymbol{\Phi}}_3$.
4. Calculate signal subspace matrix $\mathbf{U}^s$ of $\breve{\boldsymbol{\Phi}}_3$, e.g., via an SVD.
5. Derive vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ from the objective function and take two linear combinations of submatrices $\boldsymbol{\Gamma}_l = \sum_{i_1 i_2} \xi^{l,*}_{i_1 i_2} \mathbf{U}^s_{i_1 i_2}$ of $\mathbf{U}^s$.
6. Calculate the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{\Gamma}_1 (\boldsymbol{\Gamma}_2)^\dagger$.

The approach in this chapter has several benefits over the algorithms presented in the previous chapters. First, a largest eigenvalue problem is formulated instead of a smallest eigenvalue problem, which is beneficial in terms of numerical stability. Furthermore, applying reduction matrices to move from D to S are optional instead of required. Finally, the design method for the objective function, and therefore the structure in the eigenvalues, is more flexible.

6.6 Advanced source extraction algorithms based on polynomial optimization

The developed source extraction method from the previous sections and knowledge from other chapters can be combined in order to obtain more flexible and efficient source extraction algorithms. First we show that identification of the mixing column corresponding to the desired source is also possible for certain under-determined mixtures, i.e., mixtures consisting of more sources than sensors. Subsequently, we show that a reference based approach can be exploited in order to obtain a much more efficient source extraction algorithm. Finally, we show that mixing column vectors and source extraction filters can be identified by exploiting a priori information about the autocorrelation function of the desired source.

6.6.1 Solving under-determined blind source extraction problems

For under-determined mixtures, i.e., a mixture with more sources than sensors, source extraction using the LCMV filter $\tilde{\mathbf{w}}$ that cancels all interfering sources is not possible because the inverse of the mixing matrix does not exist. However, when the mixing parameters of the desired source are available, source extraction via an MVDR filter remains possible. In this section we use the multiplicative property of polynomials in

order to identify the mixing parameters that correspond to the desired source via an eigenvalue decomposition.

For under-determined mixtures the system of polynomial equations has the same structure as for well-determined mixtures, i.e.,

$$\mathbf{\Phi}_2(\mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) = 0. \tag{6.79}$$

The only difference is in the number of equations.

When uniquiefication can be applied, i.e., when $c_1 = c_2$, then the total number of equations is $\frac{1}{2}D(D+1) - S$. Due to the scaling indeterminacy, the solutions $\mathbf{z}$ represents lines instead of points in a D dimensional space. Consequently, $D-1$ equations are required in $\mathbf{\Phi}_2$ [36, 70], leading to the following maximum number of sources that can be estimated with this approach for real-valued mixtures:

$$\frac{1}{2}D(D+1) - S \geq D-1 \quad \longrightarrow \quad S_{\max} = \frac{D^2 - D + 2}{2}. \tag{6.80}$$

For complex-valued mixtures the total number of equations in $\mathbf{\Phi}_2$ is $D^2 - S$. Since the solution vectors $\mathbf{z}$ and $\mathbf{z}^*$ are treated as independent vectors, i.e., the fact that the identity $\mathbf{z}^* = (\mathbf{z})^*$ is never used. Consequently, we are searching for solutions in a 2D dimensional space. Furthermore, since the vectors $\mathbf{z}$ and $\mathbf{z}^*$ may have their own arbitrary scaling, the solutions are two dimensional intersections in a $2D$ dimensional space. Consequently, at least $2D-2$ equations are required in $\mathbf{\Phi}_2$. From this analysis it follows that the maximum number of sources that can be estimated with this approach is:

$$D^2 - S \geq 2D - 2 \quad \longrightarrow \quad S_{\max} = D^2 - 2D + 2 = (D-1)^2 + 1. \tag{6.81}$$

In Table 6.2 an overview of the maximum number of sources that can be identified for the different conjugation tuples is depicted.

Table 6.2: Maximum number of sources $S_{\max}$ given the number of sensors for different conjugation tuples.

D	2	3	4	5	6
$(\circ\circ), (**)$	2	4	7	11	16
$(\circ*), (*\circ)$	2	5	10	17	26

In [36] it is shown that homotopy methods can be used in order to identify all mixing columns. Here we show that immediate identification of a specific mixing column vector can be achieved by exploiting the multiplicative property of polynomials.

As we have seen in (6.64), the following system of third order polynomial equations can be obtained when the multiplicative property of polynomials is applied a system

of homogeneous second order polynomial equations:

$$\mathbf{\Phi}_3(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) = \bigoplus_{i \in \mathcal{D}} \mathbf{\Phi}_2(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2}) = 0. \tag{6.82}$$

Note that this system of polynomial equations for under-determined mixtures is equivalent to the system of polynomial equations constructed for well-determined mixtures. The only difference is in the number of equations. Again, when the conjugation c_0 is also applied to one of the other sets of variables, i.e., $c_0 = c_1$ or $c_0 = c_2$, then uniquiefication can be used to mix the orthogonal subsystems.

Applying uniquiefication means that a uniquiefied coefficient matrix $\breve{\mathbf{\Phi}}_3$ with $\frac{1}{2}D^2(D+1)$ instead of D^3 columns is obtained. The nullspace matrix of $\breve{\mathbf{\Phi}}_3$, again indicated with the symbol $\mathbf{U}^s$, has a size $D^2(D+1)/2 \times S$ and has the uniquiefied Kronecker structure $(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} \otimes \mathbf{z}^{c_2})$ embedded. This means that vectors $\boldsymbol{\mu}$ exist such that $\mathbf{U}^s\boldsymbol{\mu}$ restores this structure. These vectors can be identified again using a multi-matrix generalized eigenvalue decomposition, i.e.,

$$\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1} = \frac{1}{z_1^{c_2}}\mathbf{U}_1^s\boldsymbol{\mu} = \cdots = \frac{1}{z_D^{c_2}}\mathbf{U}_D^s\boldsymbol{\mu} \tag{6.83}$$

where the submatrices $\mathbf{U}_i^s$ are found by collecting all rows of the matrix $\mathbf{U}^s$ where the monomials are $(\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1})z_i^{c_2}$.

The mixing parameters are found, up to an arbitrary scaling factor and permutation, as the first D elements in $\mathbf{z}^{c_0} \otimes \mathbf{z}^{c_1}$. Note that the eigenvalues can now be designed as linear fractional functions by taking linear combinations of the submatrices $\mathbf{U}_i^s$. Furthermore, applying the multiplicative property of polynomials for a second time makes identification of a specific mixing column vector based on a quadratic fractional objective function feasible.

6.6.1.1 Example for a complex-valued mixture with $D = 3$ and $S = 4$

For an under-determined, complex-valued mixture consisting of three sensors and four sources, the sensor correlation matrix $\mathbf{C}^x$ has size 9×4 and has rank 4. Using subspace techniques we can find a size 5×9 coefficient matrix $\mathbf{\Phi}_2$ that leads to the following system of second order polynomial equations in vector-matrix notation:

$$\mathbf{\Phi}_2(\mathbf{z} \otimes \mathbf{z}^*) =$$

$$\begin{bmatrix} \phi_1^{11} & \phi_1^{12} & \phi_1^{13} & \phi_1^{21} & \phi_1^{22} & \phi_1^{23} & \phi_1^{23} & \phi_1^{31} & \phi_1^{32} & \phi_1^{33} \\ \phi_2^{11} & \phi_2^{12} & \phi_2^{13} & \phi_2^{21} & \phi_2^{22} & \phi_2^{23} & \phi_2^{23} & \phi_2^{31} & \phi_2^{32} & \phi_2^{33} \\ \phi_3^{11} & \phi_3^{12} & \phi_3^{13} & \phi_3^{21} & \phi_3^{22} & \phi_3^{23} & \phi_3^{23} & \phi_3^{31} & \phi_3^{32} & \phi_3^{33} \\ \phi_4^{11} & \phi_4^{12} & \phi_4^{13} & \phi_4^{21} & \phi_4^{22} & \phi_4^{23} & \phi_4^{23} & \phi_4^{31} & \phi_4^{32} & \phi_4^{33} \\ \phi_5^{11} & \phi_5^{12} & \phi_5^{13} & \phi_5^{21} & \phi_5^{22} & \phi_5^{23} & \phi_5^{23} & \phi_5^{31} & \phi_5^{32} & \phi_5^{33} \end{bmatrix} \begin{bmatrix} z_1 z_1^* \\ z_1 z_2^* \\ z_1 z_3^* \\ z_2 z_1^* \\ z_2 z_2^* \\ z_2 z_3^* \\ z_3 z_1^* \\ z_3 z_2^* \\ z_3 z_3^* \end{bmatrix} = \mathbf{0}_Q. \tag{6.84}$$

The considered system of third order polynomial equations without uniquiefication is as follows:

$$\mathbf{\Phi}_3(\mathbf{z}^* \otimes \mathbf{z} \otimes \mathbf{z}^*) = \bigoplus_{i \in \mathcal{D}} \mathbf{\Phi}_2(\mathbf{z}^* \otimes \mathbf{z} \otimes \mathbf{z}^*) = \mathbf{0}_{DQ}. \tag{6.85}$$

This system of polynomial equations consists of $DQ = 3 \cdot 5 = 15$ equations and $D^3 = 27$ monomials of which $D^2(D+1)/2 = 18$ are unique. Due to the size of the matrix it is not printed here; however, its structure is logical and follows the structure of earlier examples. Applying the uniquiefication leads to a uniquiefied coefficient matrix $\breve{\mathbf{\Phi}}_3$ of size 15×18. The singular value decomposition (SVD) of this coefficient matrix $\breve{\mathbf{\Phi}}_3$ is used to identify an orthogonal basis for the nullspace of $\breve{\mathbf{\Phi}}_3$ in the form of the matrix $\mathbf{U}^s$.

By selecting appropriate rows from the matrix $\mathbf{U}^s$, three submatrices $\mathbf{U}_i$ of size 9×4 are built. The submatrices $\mathbf{U}_1^s$, $\mathbf{U}_2^s$, and $\mathbf{U}_3^s$ are created by stacking the rows corresponding to respectively the following monomials:

$$z_1^*(\mathbf{z} \otimes \mathbf{z}^*), \quad z_2^*(\mathbf{z} \otimes \mathbf{z}^*), \quad z_3^*(\mathbf{z} \otimes \mathbf{z}^*). \tag{6.86}$$

The following multi-matrix GEVD can be used to identify mixing columns of the under-determined system:

$$\mathbf{z} \otimes \mathbf{z}^* = \frac{1}{z_1^*}\mathbf{U}_1^s\boldsymbol{\mu} = \frac{1}{z_2^*}\mathbf{U}_2^s\boldsymbol{\mu} = \frac{1}{z_3^*}\mathbf{U}_3^s\boldsymbol{\mu}. \tag{6.87}$$

Notice that the GEVD of two out of the three matrices has exactly four solutions and the vector $\mathbf{z}$ can be found up to an arbitrary scaling by calculating the first D elements of $\mathbf{U}_i^s\boldsymbol{\mu}$. Again, linear combinations of the submatrices $\mathbf{U}_1^s$, $\mathbf{U}_2^s$, and $\mathbf{U}_3^s$ can be taken in order to incorporate a priori information into the eigenvalues such that a specific mixing column vector can be identified. Furthermore, exploiting the multiplicative property for a second time allows for designing second order fractional objective functions for identifying a specific mixing column vector.

6.6.2 Reference systems and signals in a polynomial optimization problem

In Chapter 5 we have seen that spatial reference systems can be designed in such a way that a priori information can be incorporated into the source extraction algorithm without taking linear combinations of submatrices $\mathbf{U}_i^s$. In this chapter we investigate the consequence of using spatial reference systems for source extraction using a polynomial optimization approach.

In Chapter 5 two types of spatial reference systems are considered. In the first system, reference signals are obtained by filtering the observations $\mathbf{x}[n]$ by a reference filter. In the second system, reference signals are obtained from separate observations, i.e., from different sensors. In this chapter, the origin of the reference signals is ignored and we assume to have D observations $\mathbf{x}$ and L reference signals $\boldsymbol{\upsilon}$, i.e.,

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] + \dot{\boldsymbol{\nu}}[n] \tag{6.88}$$

$$\boldsymbol{\upsilon}[n] = \mathbf{B}\mathbf{s}[n] + \ddot{\boldsymbol{\nu}}[n] \tag{6.89}$$

where possibly the transformation $\mathbf{B} \triangleq \mathbf{TA}$ is known and $\dot{\boldsymbol{\nu}}$ and $\ddot{\boldsymbol{\nu}}$ are noise signals that are potentially related via $\ddot{\boldsymbol{\nu}} = \mathbf{T}\dot{\boldsymbol{\nu}}$. More details on the construction and structure of these signals can be found in Chapter 5. Here we focus on the second order statistics of these signals under the assumption of a NF-ROS. The crosscorrelation functions between the sensor signals and the reference signals have the following structure in the NF-ROS:

$$r_{il}^{x\upsilon}[\Omega_\kappa] = \sum_{j\in\mathcal{S}} a_i^{j,c_1} b_l^{j,c_2} r_{jj}^{s}[\Omega_\kappa] \quad \text{and} \quad r_{li}^{\upsilon x}[\Omega_\kappa] = \sum_{j\in\mathcal{S}} b_l^{j,c_1} a_i^{j,c_2} r_{jj}^{s}[\Omega_\kappa]. \tag{6.90}$$

In matrix form, this leads to the following reference-sensor crosscorrelation matrix:

$$\mathbf{C}^{\upsilon x} = \begin{bmatrix} r_{11}^{\upsilon x}[\Omega_1] & \cdots & r_{11}^{\upsilon x}[\Omega_K] \\ r_{12}^{\upsilon x}[\Omega_1] & \cdots & r_{12}^{\upsilon x}[\Omega_K] \\ \vdots & \ddots & \vdots \\ r_{LD}^{\upsilon x}[\Omega_1] & \cdots & r_{DL}^{\upsilon x}[\Omega_K] \end{bmatrix} \equiv (\mathbf{B}^{c_1} \diamond \mathbf{A}^{c_2})\mathbf{C}^{s}. \tag{6.91}$$

The size of crosscorrelation matrix $\mathbf{C}^{\upsilon x}$ is $LD \times K$. From the reasoning that the Khatri-Rao product of two randomly chosen matrices has full rank, it follows that the rank of $\mathbf{C}^{\upsilon x}$ is S as long as $\operatorname{rank}(\mathbf{A}^{c_2}) \cdot \operatorname{rank}(\mathbf{B}^{c_1}) \geq S$, i.e., with full rank matrices $\mathbf{A}$ and $\mathbf{B}$ we assume that $DL \geq S$.

Note that the Khatri-Rao product of two random matrices does not lead to trivial linear dependencies. Consequently, the number of unique rows in the crosscorrelation matrix $\mathbf{C}^{\upsilon x}$ is LD. Applying subspace techniques to the crosscorrelation matrix $\mathbf{C}^{\upsilon x}$ leads to the following system of polynomial equations:

$$\mathbf{\Phi}_2(\mathbf{v}^{c_1} \otimes \mathbf{z}^{c_2}) = 0 \tag{6.92}$$

where $\mathbf{v} \in \mathbb{C}_L \backslash \mathbf{0}_L$ contains L variables v_i for $i \in \mathcal{L}$.

The sets of variables $\mathbf{v}$ and $\mathbf{z}$ of length L and D, respectively, could be merged into a single set of variables of length $L + D$. However, since only crosscorrelation functions between the sensor signals $x_i[n]$ and reference signals $\upsilon_l[n]$ are considered, information about the relation between $\mathbf{v}$ and $\mathbf{z}$ is missing in the system of equations. The consequence of this lack of information is that identification of the combined mixing column vector $(\mathbf{a}^j, \mathbf{b}^j)$ for $j \in \mathcal{S}$ is only possible up to an arbitrary mutual scaling, i.e., $(\eta_1\mathbf{a}^j, \eta_2\mathbf{b}^j)$ for $\eta_1 \in \mathbb{C}\backslash 0$ and $\eta_2 \in \mathbb{C}\backslash 0$. This mutual scaling can be resolved when correlation functions related to $r_{i_1 i_2}^{x}[\Omega_\kappa]$ and/or $r_{i_1 i_2}^{\upsilon}[\Omega_\kappa]$ are considered. Alternatively, if the transformation matrix $\mathbf{T}$ used to generate the reference signals as $\boldsymbol{\upsilon}[n] = \mathbf{T}\mathbf{x}[n]$ is available, then this matrix can also be used to resolve the mutual scaling indeterminacy; however, solving the problem for this specific scenario does not lead to better source extraction filters since then the observations $\mathbf{x}$ and the reference signals $\boldsymbol{\upsilon}$ are linearly dependent.

The system of homogeneous polynomial equations in (6.92) contains $LD - S$ equations in a $D + L$ dimensional space. Due to the independent scaling indeterminacy for the two sets of variables $\mathbf{v}$ and $\mathbf{z}$, a total of $D + L - 2$ equations is required in

order to enable identification of mixing column vectors. Consequently, we have the following condition for the required number of equations:

$$LD - S \geq D + L - 2. \tag{6.93}$$

From this condition it follows that the maximum number of mixing column vectors that can be identified as function of the number of sensors D and the number of reference signals L is as follows:

$$S_{\max} = (L-1)(D-1) + 1. \tag{6.94}$$

Note that this condition is independent from the conjugation tuple as long as the transformation matrix $\mathbf{T}$ is sufficiently different from the identity matrix, i.e., no trivial linear dependencies are allowed in crosscorrelation matrix $\mathbf{C}^{vx}$.

In Table 6.3 an overview is given for the maximum number of sources that is allowed given the number of sensors en number of reference signals in order to identify their corresponding mixing column vectors. Note that the constraint is symmetric in D and L. Furthermore, the diagonal elements, i.e., where $L = D$, correspond to the results for mixtures with conjugation tuples $(\circ*)$ and $(*\circ)$ in Section 6.6.1.

Table 6.3: Maximum number of mixing column vectors $S_{\max}$ given the number of sensors D and number of reference signals L.

$L \backslash D$	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	2	3	4	5	6
3	1	3	5	7	9	11
4	1	4	7	10	13	16
5	1	5	9	13	17	21
6	1	6	11	16	21	26

Like in the previous sections, the system of polynomial equations as in (6.92) can be solved using a GEVD. The signal space of $\mathbf{\Phi}_2$, represented by the matrix $\mathbf{U}^s$, has a size $LD \times S$. Taking appropriate submatrices $\mathbf{U}_l^s$ of size $D \times S$ leads to the following multi-matrix GEVD:

$$\mathbf{z}^{c_2} = \frac{1}{v_1^{c_1}} \mathbf{U}_1^s \boldsymbol{\mu} = \cdots = \frac{1}{v_L^{c_1}} \mathbf{U}_L^s \boldsymbol{\mu}. \tag{6.95}$$

From this GEVD it follows that the reference based source extraction problem can be written as the following polynomial optimization problem:

$$\begin{aligned} &\max_{\mathbf{z},\mathbf{v}} \quad \lambda(\mathbf{v}) \\ &\text{s.t.} \quad f_q(\mathbf{v}, \mathbf{z}) = 0 \quad \forall\, q \in \mathcal{Q} \end{aligned} \tag{6.96}$$

where $\lambda(\mathbf{v})$ is a linear fractional function in the variables $\mathbf{v}$.

In this system of polynomial equations, the (known) transformation $\mathbf{B} = \mathbf{TA}$ can be interpreted as the following change in variables: $\mathbf{v} = \mathbf{Tz}$. As we have already seen in Chapter 5, design of the transformation matrix $\mathbf{T}$ can be used in order to incorporate a priori information in the source extraction algorithm. In the polynomial optimization problem the a priori information returns solely in the objective function, which in turn translates to eigenvalue in a GEVD.

For example, if the GEVD of the submatrices $\mathbf{U}_1^s$ and $\mathbf{U}_2^s$ is used to identify source extraction filters, then the objective function has the following structure:

$$\lambda(\mathbf{v}^{c_1}) = \frac{v_1^{c_1}}{v_2^{c_1}} = \frac{\tilde{\mathbf{t}}_1^{c_1}\mathbf{z}^{c_1}}{\tilde{\mathbf{t}}_2^{c_1}\mathbf{z}^{c_1}}. \tag{6.97}$$

This construction of the problem leads to the following S eigenvalues:

$$\lambda^j = \frac{v_1^{c_1}}{v_2^{c_1}} = \frac{\tilde{\mathbf{t}}_1^{c_1}\mathbf{a}^{j,c_1}}{\tilde{\mathbf{t}}_2^{c_1}\mathbf{a}^{j,c_1}}. \tag{6.98}$$

Applying the multiplicative property w.r.t. the variables $\mathbf{v}$ allows for the design of reference based polynomial optimization problems with quadratic objective functions in terms of the variables $\mathbf{v}$. However, note that this requires that the conjugation for the variables is the same.

6.6.3 Exploiting a priori information about the source autocorrelation function

In previous sections we have seen that identification of the desired source extraction filter can be accomplished by exploiting a priori information about the mixing parameters of this desired source. In this section we show that source extraction filters can be identified as well via information about the source autocorrelation function by exploiting a different sensor correlation matrix structure.

The structure of the sensor correlation matrix exploited in this section is as follows:

$$\mathbf{R}^x = \begin{bmatrix} \mathbf{R}_1^x \\ \mathbf{R}_2^x \\ \vdots \\ \mathbf{R}_K^x \end{bmatrix} \equiv \left((\mathbf{C}^s)^T \diamond \mathbf{A}^{c_1}\right)(\mathbf{A}^{c_2})^T \tag{6.99}$$

where $\mathbf{R}_\kappa^x$ for $\kappa \in \mathcal{K}$ is the following widely used sensor correlation matrix:

$$\mathbf{R}_\kappa^x \triangleq \begin{bmatrix} r_{11}^x[\Omega_\kappa] & r_{12}^x[\Omega_\kappa] & \cdots & r_{1D}^x[\Omega_\kappa] \\ r_{21}^x[\Omega_\kappa] & r_{22}^x[\Omega_\kappa] & \cdots & r_{2D}^x[\Omega_\kappa] \\ \vdots & \ddots & \ddots & \vdots \\ r_{D1}^x[\Omega_\kappa] & r_{D2}^x[\Omega_\kappa] & \cdots & r_{DD}^x[\Omega_\kappa] \end{bmatrix}. \tag{6.100}$$

The size of $\mathbf{R}^x$ is $KD \times D$ and the size of $\mathbf{R}^x_\kappa$ is $D \times D$.

Applying subspace techniques to the sensor correlation matrix $\mathbf{R}^x$ leads to the following system of polynomial equations:

$$\mathbf{\Phi}_2(\mathbf{v} \otimes \mathbf{z}^{c_1}) = 0 \tag{6.101}$$

where the variables in $\mathbf{v}$ correspond to the source autocorrelation function $\tilde{\mathbf{r}}^s_{jj}$ and the variables in $\mathbf{z}^{c_1}$ correspond to the mixing column vector $\mathbf{a}^j$ with conjugation c_1.

Following the reasoning that source extraction filters can be identified by taking linear combinations of submatrices $\mathbf{U}^s_\kappa$ leads to the following polynomial optimization problem:

$$\begin{aligned} \max_{\mathbf{z},\mathbf{v}} \quad & \frac{\langle \boldsymbol{\xi}^1, \mathbf{v} \rangle}{\langle \boldsymbol{\xi}^2, \mathbf{v} \rangle} \\ \text{s.t.} \quad & f_q(\mathbf{v}, \mathbf{z}) = 0 \quad \forall\, q \in \mathcal{Q}. \end{aligned} \tag{6.102}$$

The following conditions have to hold in order to be able to identify the desired source extraction filter. The mixing system must be well-determined, i.e., $D \geq S$ with rank S. If this is not the case, then the matrix $\mathbf{A}^{c_2}$ on the right hand side of (6.99) is not invertible. The number of time lag pairs in the NF-ROS must be at least two. Since the matrices $\mathbf{A}$ and $\mathbf{C}^s$ are independent from each other, their Khatri-Rao product is expected to have maximum rank in general. In total we have $KD - S$ equations while $K + D - 2$ equations are required, leading to the following condition:

$$S_{\max} = (K-1)(D-1) + 1. \tag{6.103}$$

On the other hand, we have $D \geq S$, leading to the following:

$$S_{\max} = \min\{D, (K-1)(D-1) + 1\}. \tag{6.104}$$

6.6.4 Connection to higher order statistics

When considering higher order statistics, assumptions are made that are very similar to the second order statistics scenario. In this section we focus on third and fourth order statistics. Instead of sensor correlation functions, cumulant functions are obtained. These cumulant functions have the following structure for third order and fourth order statistics, respectively:

$$r^x_{i_1 i_2 i_3}[\Omega_\kappa] = \sum_{j \in \mathcal{S}} a^{j,*}_{i_1} a^j_{i_2} a^{j,*}_{i_3} r^s_{jjj}[\Omega_\kappa] \tag{6.105}$$

$$r^x_{i_1 i_2 i_3 i_3}[\Omega_\kappa] = \sum_{j \in \mathcal{S}} a^j_{i_1} a^{j,*}_{i_2} a^j_{i_3} a^{j,*}_{i_4} r^s_{jjjj}[\Omega_\kappa] \tag{6.106}$$

where $r_{jjj}[\Omega_\kappa]$ and $r^s_{jjjj}[\Omega_\kappa]$ are the third and fourth order source autocumulant functions evaluated in the noise-free region of support. We assume that the source are mutually independent, which means that their cross statistics equal zero. Furthermore, we assume that the source autocumulant functions are linearly independent [36].

The higher order equivalent of the sensor correlation matrix $\mathbf{C}^x$ is called the sensor cumulant matrix $\mathbf{C}^x$. The structure of this sensor cumulant matrix for third and fourth order statistics is as follows:

$$\mathbf{C}^x = (\mathbf{A}^* \diamond \mathbf{A} \diamond \mathbf{A}^*)\mathbf{C}^s \tag{6.107}$$

and

$$\mathbf{C}^x = (\mathbf{A} \diamond \mathbf{A}^* \diamond \mathbf{A} \diamond \mathbf{A}^*)\mathbf{C}^s. \tag{6.108}$$

Applying subspace techniques and forming systems of polynomial equations leads to the following systems:

$$\mathbf{\Phi}_3(\mathbf{z}^* \otimes \mathbf{z} \otimes \mathbf{z}^*) = \mathbf{0} \tag{6.109}$$

and

$$\mathbf{\Phi}_4(\mathbf{z} \otimes \mathbf{z}^* \otimes \mathbf{z} \otimes \mathbf{z}^*) = \mathbf{0}. \tag{6.110}$$

It follows immediately that the systems of polynomial equations are very similar to the systems of polynomial equations where the multiplicative property of polynomials is exploited either to incorporate more advanced a priori information or to enable identification of mixing column vectors for under-determined mixtures. The main difference is that the multiplicative property of polynomials leads to a Macaulay like coefficient matrix $\mathbf{\Phi}_3$ or $\mathbf{\Phi}_4$ [74]. When using solely higher order statistics, this structure is not present and therefore it cannot be exploited as a feature for efficient algorithms.

6.7 Design of polynomial optimization based source extraction algorithms

In a source extraction algorithm based on polynomial optimization the objective function $\lambda(\mathbf{z})$ is evaluated at the points where $\mathbf{z} \propto \mathbf{a}^j$ for $j \in \mathcal{S}$, leading to S eigenvalues λ^j. In order to guarantee identification of the extraction filter corresponding to the desired source, the objective function must be designed in such a way that the eigenvalue corresponding to the desired source is distinguishable from the other eigenvalues. In this chapter we present design strategies that lead to largest or absolute largest eigenvalue problems. For such problems efficient and numerically stable algorithms, e.g., the power method [46], are available; however, other design criteria can be used as well.

The first designs are based on the source extraction algorithm design procedure from Section 6.5.1 where a linear fractional objective function is used to ensure identification of the desired source. Subsequently, source extraction algorithm designs using a second order fractional objective function, as discussed in Section 6.5.2, are presented. Finally, more advanced design strategies are discussed, including design strategies for the source extraction algorithms from Section 6.6.

6.7.1 Linear fractional objective functions

The linear fractional objective function as function of the D unknown mixing parameters $\mathbf{z}$ from (6.57) is repeated here for convenience, i.e.,

$$\lambda(\mathbf{z}^{c_1}) = \frac{\langle \boldsymbol{\xi}^1, \mathbf{z}^{c_1} \rangle}{\langle \boldsymbol{\xi}^2, \mathbf{z}^{c_1} \rangle} \simeq \frac{\|\boldsymbol{\xi}^1\|_2}{\|\boldsymbol{\xi}^2\|_2} \frac{\cos\theta_1}{\cos\theta_2} \tag{6.111}$$

where θ_l is the, potentially complex-valued, angle between $\boldsymbol{\xi}^l$ and $\mathbf{z}^{c_1}$ and $\|\boldsymbol{\xi}^l\|_2$ represents the length of $\boldsymbol{\xi}^l$. We use the symbol $\simeq$ to indicate that the right hand side is a differently parameterized version of the left hand side of the equation.

Suppose we have a guess or estimate of the actual mixing parameters corresponding to the desired source in the form of a mold $\mathbf{a}^0$ such that $\mathbf{a}^0 \approx \mathbf{a}^d$. Given this a priori information, our goal is to design an objective function $\lambda(\mathbf{z})$ that is largest when $\mathbf{z} = \mathbf{a}^0$ and decreases when $\mathbf{z}$ moves away from this direction. It follows that these conditions are met when we choose $\boldsymbol{\xi}^1$ in the direction of $\mathbf{a}^0$ and $\boldsymbol{\xi}^2$ orthogonal with respect to $\boldsymbol{\xi}^1$, i.e., $\boldsymbol{\xi}^1 = \mathbf{a}^{0,c_1}$ and $\boldsymbol{\xi}^1 \perp \boldsymbol{\xi}^2$. The corresponding linear fractional objective function then has the following form:

$$\lambda(\theta_0) = \frac{\|\boldsymbol{\xi}^1\|_2}{\|\boldsymbol{\xi}^2\|_2} \frac{1}{\tan\theta_0} = \frac{\|\boldsymbol{\xi}^1\|_2}{\|\boldsymbol{\xi}^2\|_2} \cot\theta_0 \tag{6.112}$$

where θ_0 is the angle between the mold $\mathbf{a}^0$ and vector $\mathbf{z}$ when $\mathbf{z}$ is projected onto the two dimensional subspace spanned by $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$. In case only two sensors are considered for the source selection task, then this projection can be ignored. Note that the lengths of the vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ only define a constant scaling for all eigenvalues simultaneously. Therefore, without loss of generality, we assume for the remainder of this section that the vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ have equal lengths.

For a real-valued mixing system, the actual and absolute linear fractional objective function is depicted as a blue, solid line as function of the angle θ_0 in the upper and lower plots in Figure 6.1, respectively. Since the objective function, and therefore the eigenvalues, can become infinitely large the inverse tangent function is used to map the objective function to finite numbers. Figures such as Figure 6.1 can be interpreted as the selection patterns from Chapter 4. The selection patterns provide insight in the source that will be extracted, given the actual mixing parameters.

The following condition on the mold is required in order to guarantee extraction of the desired source when the absolute largest eigenvalue is selected:

$$\frac{|\langle \mathbf{a}^0, \mathbf{a}^d \rangle|}{|\langle \boldsymbol{\xi}^2, \mathbf{a}^d \rangle|} > \frac{|\langle \mathbf{a}^0, \mathbf{a}^i \rangle|}{|\langle \boldsymbol{\xi}^2, \mathbf{a}^i \rangle|} \quad \forall\, j \in \mathcal{S}\backslash d \tag{6.113}$$

with $\boldsymbol{\xi}^2 \perp \mathbf{a}^0$ and where index d corresponds to the desired source. In other terms, we require that the angle between the mixing column vector projected onto the space spanned by $\boldsymbol{\xi}^1 = \mathbf{a}^0$ and $\boldsymbol{\xi}^2 \perp \mathbf{a}^0$ is smallest for the desired source, i.e., $\theta_0^d < \theta_0^j$ for $j \in \mathcal{S}\backslash d$. If only two sensors are considered, then this condition on the mold implies that the mold has to be *close enough* to the actual mixing parameters of the

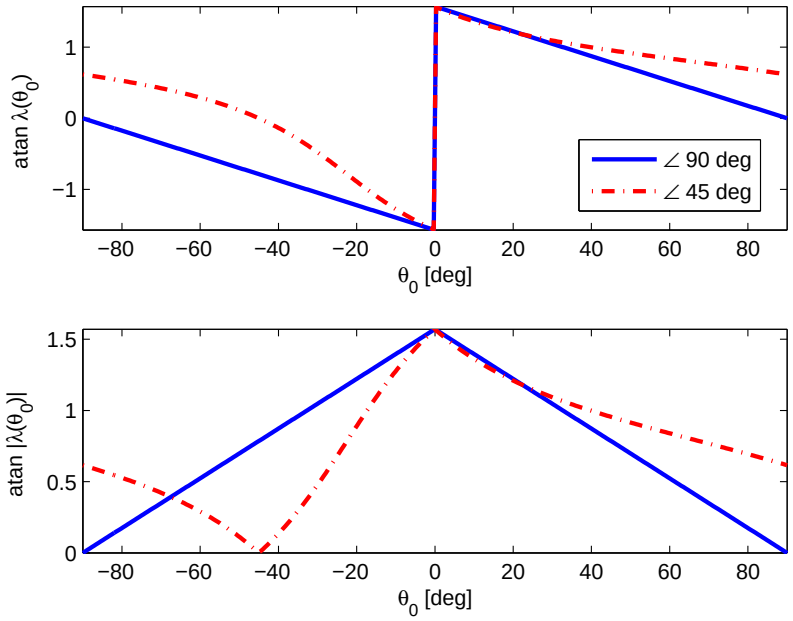


Figure 6.1: Linear fractional objective functions as function of the angle between the mold and the actual mixing column vector, projected onto the two dimensional subspace spanned by $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$. In the upper and lower graph respectively the actual and absolute value of the functions are depicted for angles of 90 and 45 degrees between the vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$. The inverse tangent function is used in order to compress the y-axis.

desired source. If more than two sensors are considered, then the projections of the mixing column vectors are arbitrary and can lead to small projected angles while the angle between the mold and the actual mixing column vector can be relatively large. This property is a fundamental limitation of working with linear fractional objective functions.

Choosing non-orthogonal vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ allows for the incorporation of additional a priori information into the linear fractional objective function or selection pattern. In Figure 6.1 the red, dashed line shows the linear fractional objective function where $\boldsymbol{\xi}^2$ is chosen orthogonal to the mold and $\boldsymbol{\xi}^1$ is chosen at an angle of 45 degrees with respect to $\boldsymbol{\xi}^2$, in the two dimensional subspace spanned by $\mathbf{a}^0$ and $\boldsymbol{\xi}^2$. This design strategy allows to design extraction algorithms that incorporate information about a certain preferred direction of the error with respect to the mold. For example, by choosing $\boldsymbol{\xi}^1$ orthogonal to an estimate or guess of the mixing column corresponding to a single interfering source, the objective function will become zero in this direction. One way to calculate these vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ based on estimates $\hat{\mathbf{a}}^0$ and $\hat{\mathbf{a}}^i$ of the mixing column vectors corresponding to the desired and a single

interfering source, respectively, is as follows:

$$\boldsymbol{\xi}^1 = \hat{\mathbf{a}}^0 - \frac{\langle \hat{\mathbf{a}}^i, \hat{\mathbf{a}}^0 \rangle}{\langle \hat{\mathbf{a}}^i, \hat{\mathbf{a}}^i \rangle} \hat{\mathbf{a}}^i \quad \text{and} \quad \boldsymbol{\xi}^2 = \hat{\mathbf{a}}^i - \frac{\langle \hat{\mathbf{a}}^0, \hat{\mathbf{a}}^i \rangle}{\langle \hat{\mathbf{a}}^0, \hat{\mathbf{a}}^0 \rangle} \hat{\mathbf{a}}^0 \tag{6.114}$$

where we leave out the conjugation c_1 for notational convenience.

For specific scenarios, such as the ones described above, source extraction based on polynomial optimization problems with linear fractional objective functions can be practical. However, due to the linear characteristic of the objective function and the corresponding projection onto a two dimensional subspace it is fundamentally impossible to generalize the design strategies towards multiple sensors and sources without requiring significantly more a priori information. Furthermore, due to the presence of both positive and negative eigenvalues it is difficult to apply regularization in order to limit the maximally attainable eigenvalue. In the following section we show that these generalizations and improvements are feasible when considering second order fractional objective functions.

6.7.2 Second order fractional objective functions

The structure of a second order fractional objective function is originally presented in (6.76) and (6.77) and is repeated here for convenience, i.e., the second order fractional objective function in vector-matrix form is as follows:

$$\lambda(\mathbf{z}) = \frac{(\mathbf{z}^{c_1})^T \boldsymbol{\Xi}_1 \mathbf{z}^{c_2}}{(\mathbf{z}^{c_1})^T \boldsymbol{\Xi}_2 \mathbf{z}^{c_2}} \tag{6.115}$$

where the matrices $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$ can be chosen arbitrarily or designed based on a priori information. Depending on the conjugation, either D^2 or $D(D+1)/2$ degrees of freedom exist in the matrices $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$. In the remainder of this section we assume that the mixture is real-valued. Consequently, the conjugation symbols c_1 and c_2 are not taken into account. The results for real-valued mixtures easily extend to mixtures with conjugation tuples $(c_1, c_2) = (\circ, *)$ and $(c_1, c_2) = (*, \circ)$. For complex-valued mixtures with conjugation tuples $(c_1, c_2) = (\circ, \circ)$ and $(c_1, c_2) = (*, *)$ an objective function designer has to take into account how to deal with the phase in the objective function.

The following design strategies are based on the assumption that a priori information in the form of a guess or estimate of the mixing parameters corresponding to the desired source is available. This a priori information is represented by the vector $\mathbf{a}^0$, which we call a mold. In the design strategies a projection matrix $\mathbf{P}_0$ that projects vectors onto the space spanned by the mold is regularly used. Therefore we present this matrix in advance, i.e.,

$$\mathbf{P}_0 \triangleq \frac{\mathbf{a}^0 (\mathbf{a}^0)^T}{(\mathbf{a}^0)^T \mathbf{a}^0}. \tag{6.116}$$

From the following design strategies it follows which conditions have to hold on this mold in order to guarantee extraction of the desired source by selecting the extraction filter corresponding to the largest eigenvalue.

Design strategy I If the mold $\mathbf{a}^0$ is available in a source extraction scenario, then the following *matched filter* design strategy can be used in order to guarantee extraction of the desired source:

$$\mathbf{\Xi}_1 = \mathbf{P}_0 \quad \text{and} \quad \mathbf{\Xi}_2 = \mathbf{I} \tag{6.117}$$

We refer to this design in terms of a matched filter due to the direct application of the mold $\mathbf{a}^0$. Using this design, the second order fractional objective function has the following form:

$$\lambda(\mathbf{z}) = \frac{(\mathbf{z})^T \mathbf{P}_0 \mathbf{z}}{(\mathbf{z})^T \mathbf{z}} = \frac{1}{\|\mathbf{a}^0\|_2^2} \frac{(\mathbf{z})^T \mathbf{a}^0 (\mathbf{a}^0)^T \mathbf{z}}{(\mathbf{z})^T \mathbf{z}} \simeq \cos^2 \theta_0 \tag{6.118}$$

where θ_0 is the angle between $\mathbf{z}$ and the mold $\mathbf{a}^0$. The symbol $\simeq$ is again used to denote that the left hand side and right hand side are equal up to a different parametrization. Note that this angle θ_0 is different from the angle considered in Section 6.7.1. In the current design we are not concerned with a projection onto a two dimensional subspace and the design can be used for any number of sensors larger than or equal to two.

The matched filter design strategy leads to the following structure for the eigenvalues:

$$\lambda^j = \frac{1}{\|\mathbf{a}^0\|_2^2} \frac{|\langle \mathbf{a}^0, \mathbf{a}^j \rangle|^2}{\|\mathbf{a}^j\|_2^2} \simeq \cos^2 \theta_0^j \tag{6.119}$$

where θ_0^j is the angle between the mold $\mathbf{a}^0$ and actual mixing column vector $\mathbf{a}^j$. In order to guarantee that the largest eigenvalue corresponds to the desired source, the following condition has to hold for the mold:

$$\frac{|\langle \mathbf{a}^0, \mathbf{a}^d \rangle|}{\|\mathbf{a}^d\|_2} > \frac{|\langle \mathbf{a}^0, \mathbf{a}^j \rangle|}{\|\mathbf{a}^j\|_2} \quad \forall\, j \in \mathcal{S} \backslash d \tag{6.120}$$

where index d corresponds to the desired source. In other terms, we require that $\theta_0^d < \theta_0^j$ for $j \in \mathcal{S} \backslash d$. In practice this condition on the mold implies that the mold is *close enough* to the actual mixing parameters of the desired source.

An interesting property of this design is that the eigenvalues are always positive and in the range $[0, 1]$. Applying an arbitrary gain to $\mathbf{\Xi}_1$ leads to a scaling of all the eigenvalues by the same amount and allows for an arbitrary maximum eigenvalue. A disadvantage of the matched filter design strategy is that all eigenvalues are relatively close to each other in case the mixing system is not orthogonal. The following design strategy leads to a second order fractional objective function with a sharp peak, resulting in large eigenvalues for mixing column vectors close to the mold. Such a separation of eigenvalues is beneficial for numerical algorithms that identify eigenvectors corresponding to the largest eigenvalue.

Design strategy II In order to obtain a larger separation of the largest eigenvalues from the other eigenvalues, the following design strategy based on the same a priori

information as used for the matched filter design strategy can be used:

$$\boldsymbol{\Xi}_1 = \mathbf{P}_0 \quad \text{and} \quad \boldsymbol{\Xi}_2 = \mathbf{I} - \mathbf{P}_0. \tag{6.121}$$

This design leads to the following fractional objective function:

$$\lambda(\mathbf{z}) = \frac{(\mathbf{z})^T \mathbf{P}_0 \mathbf{z}}{(\mathbf{z})^T \left[\mathbf{I} - \mathbf{P}_0\right] \mathbf{z}} \simeq \frac{\cos^2 \theta_0}{\sin^2 \theta_0} = \cot^2 \theta_0. \tag{6.122}$$

For this peaked design, the same condition on the mold has to hold as for the matched filter design from (6.117), i.e., we assume that (6.120) holds for the mold. A benefit of this design strategy with respect to the first design strategy is that the spread of eigenvalues leads to faster converging algorithms; however, since the eigenvalues become infinitely large in case the mold exactly matches the observed mixing column vector, the current design strategy also leads to numerically instable algorithms. In order to deal with this numerical problem the following design strategy with a shape parameter can be exploited.

Design strategy III In order to benefit from both the peakedness of the design in (6.121) for fast convergence and from the limited range for the eigenvalues for the design in (6.117) for numerical stability we present the following design strategy with a shape parameter:

$$\boldsymbol{\Xi}_1 = \mathbf{P}_0 \quad \text{and} \quad \boldsymbol{\Xi}_2 = \gamma \left[\mathbf{I} - \mathbf{P}_0\right] + \mathbf{P}_0 \tag{6.123}$$

where γ is a shape parameter that can be used to tune the peakedness of the objective function. The second order fractional objective function corresponding to this design strategy is as follows:

$$\lambda(\mathbf{z}) = \frac{(\mathbf{z})^T \mathbf{P}_0 \mathbf{z}}{(\mathbf{z})^T \left[\gamma \mathbf{I} + (1-\gamma)\mathbf{P}_0\right] \mathbf{z}} \simeq \frac{1}{\gamma \tan^2 \theta_0 + 1}. \tag{6.124}$$

As for the design in (6.117), the eigenvalues are again all positive and in the range $[0, 1]$. Furthermore, applying an arbitrary gain to $\boldsymbol{\Xi}_1$ leads to a scaling of all eigenvalues by the same amount and allows for an arbitrary maximally attainable eigenvalue. The peakedness parameter γ can be tuned in order to separate the largest eigenvalue from the other eigenvalues. For $\gamma = 1$ we obtain the matched filter design from (6.117) while towards $\gamma = 0$ the objective function becomes more and more flat. For increasing γ the objective function becomes sharper, leading to a Dirac delta function for infinitely large γ. In Figure 6.2a the second order fractional objective function is depicted for a range of peakedness parameter values.

In the last part of this section we present a design strategy that incorporates a priori information about mixing parameters corresponding to interfering sources in order to design extraction algorithms for which the eigenvalue corresponding to the desired source is maximally separated from the other eigenvalues.

Design strategy IV In case not only information about the mixing parameters corresponding to the desired source, but also information about the mixing parameters corresponding to the interfering sources is available, then this information can be incorporated into the design of the second order fractional objective function.

Suppose we have an estimate $\hat{\mathbf{A}}$ of the mixing column vectors corresponding to the interfering sources. The projection matrix that projects vectors onto the space spanned by these mixing column vectors is defined as follows:

$$\mathbf{P}_i = \hat{\mathbf{A}} \left((\hat{\mathbf{A}})^T \hat{\mathbf{A}} \right)^\dagger (\hat{\mathbf{A}})^T. \tag{6.125}$$

Based on this a priori information $\hat{\mathbf{A}}$ and the original mold $\mathbf{a}^0$ the following design strategy leads to large eigenvalues for mixing column vectors near $\mathbf{a}^0$ and small, positive eigenvalues near the estimated or guessed mixing column vectors corresponding to the interfering sources:

$$\mathbf{\Xi}_1 = \mathbf{I} - \mathbf{P}_i \quad \text{and} \quad \mathbf{\Xi}_2 = \gamma \left[\mathbf{I} - \mathbf{P}_0 \right] + \mathbf{I} - \mathbf{P}_i. \tag{6.126}$$

This design strategy leads to the following second order fractional objective function:

$$\lambda(\mathbf{z}) = \frac{(\mathbf{z})^T \left[\mathbf{I} - \mathbf{P}_i \right] \mathbf{z}}{(\mathbf{z})^T \left[\gamma \left(\mathbf{I} - \mathbf{P}_0 \right) + \left(\mathbf{I} - \mathbf{P}_i \right) \right] \mathbf{z}}. \tag{6.127}$$

For this design strategy, the eigenvalue corresponding to the desired source ideally becomes equal to one, while the eigenvalues corresponding to the interfering sources ideally become equal to zero. In order to be able to use this design the mixing column vectors for which a priori information is available, including the mold, must be linearly independent. If this condition is met, then the peakedness parameter γ can be used to control the peakedness of the objective function. In Figures 6.2b and 6.2c second order fractional objective functions are depicted as function of θ_0 for guessed interfering source angles of 70 and 30 degrees with respect $\theta_0 = 0$, respectively, for different peakedness parameter values.

6.7.3 Advanced source extraction algorithms

The design strategies presented for the linear and second order fractional objective functions form a basis for the design and development of source extraction algorithms exploiting a priori information about the application at hand. In the remainder of this section we discuss optional extensions or alternative design strategies and we discuss design strategies for the advanced source extraction algorithms from Section 6.6. Only a high level discussion is presented since the optimal design strongly depends on specific properties of the application.

- If an estimate or guess of the mixing parameters corresponding to the desired source is only available for a subset of the sensors, i.e., if the mold is only partially known, then the presented design strategies can also be applied to only those sensors. In such case the coefficients for the sensors for which no

information is available must be set to zero for both $\boldsymbol{\Xi}_1$ and $\boldsymbol{\Xi}_2$. The condition on the mold should now hold for the mixing parameters corresponding to the sensors used for selection. Finally, note that for such scenarios the extraction filters are still identified for all available sensors.

- Another way to use the peakedness parameter γ from *Design strategy III* is to design source extraction algorithms that detect and/or extract a source within a certain region of interest. Given the maximally allowed angle θ_{BW} between the actual mixing column vector and the mold $\mathbf{a}^0$ and a threshold parameter $0 < \lambda_T < 1$, then choosing the peakedness parameter γ as follows:

$$\gamma = \frac{1 - \lambda_T}{\lambda_T} \cot^2 \theta_{BW} \tag{6.128}$$

  ensures that eigenvalues that are larger than the threshold λ_T correspond to mixing column vectors in the region of interest, i.e., their angle with resect to the mold is smaller than θ_{BW}.

- If initially only the mold is available then a blind identification algorithm, for example based on *Design strategy III*, can be used to provide information about the mixing column vectors corresponding to the interfering sources. Subsequently, the information about the interfering sources can be exploited to track slow variations in the mixing system using *Design strategy IV*.

- If a parametrization of the mixing column vectors is available, then design of the source extraction algorithm based on a priori information in this parameterized domain can be used as well. More specifically, all selection beamformer designs from Chapter 4 can be used in this polynomial optimization approach as well.

- For under-determined mixtures where the strategy from Section 6.6.1 is applied, i.e., with exploiting the multiplicative property of polynomial only once, the polynomial optimization problem has a linear fractional objective function. Consequently, the same design techniques can be used as presented in Section 6.7.1. However, due to the generality of the approach, the multiplicative property could be exploited for a second time such that also second order fractional objective functions can be used for selecting the desired solution for under-determined mixtures. Note that in this latter case *Design strategy IV* cannot be applied when estimates of all $(S > D)$ mixing column vectors are considered since these vectors are linearly dependent.

- Reference based source extraction based on the linear fractional objective function can be applied in a straightforward way by choosing the transformation matrix $\mathbf{T}$ as the concatenation of the two vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$. For real-valued mixtures and complex-valued mixtures with conjugation tuples $(\circ, \circ)$ and $(*, *)$ the following second order fractional objective function can be considered:

$$\lambda(\mathbf{z}) = \frac{(\mathbf{z})^T (\mathbf{T})^T \boldsymbol{\Xi}_1 \mathbf{T}\mathbf{z}}{(\mathbf{z})^T (\mathbf{T})^T \boldsymbol{\Xi}_2 \mathbf{T}\mathbf{z}} \tag{6.129}$$

  where we left out the conjugation symbols for notational convenience. The transformation matrix $\mathbf{T}$ can be used for several purposes, e.g., dimensionality

reduction, incorporating a priori information, and conditioning. For complex-valued mixtures with conjugation tuples $(\circ, *)$ and $(*, \circ)$ there is no straightforward way to formulate polynomial optimization based source extraction as reference based source extraction.

- Source extraction algorithms exploiting information about the autocorrelation functions of the sources can be designed in a similar way as source extraction algorithms based on information about the mixing parameters of the sources. Both linear and second order fractional objective functions can be designed based on a priori information about the desired source. An example design based on information about the autocorrelation function of a single interfering source, provided by a remote wireless microphone, is presented in [**?**]. The design is used there for informed source extraction; however, it can also be used for polynomial optimization based source extraction.
- Source extraction algorithms based on higher order statistics contain the higher order structure for second order fractional objective functions in a natural way. Consequently, the design techniques from Section 6.7.2 can be used for source extraction algorithms based on higher order statistics as well.

6.8 Validation of polynomial optimization based source extraction

In the remainder of this chapter we present results for two scenarios where the polynomial optimization based source extraction algorithm is applied. The results are obtained from artificial mixtures via computer simulations. In the first scenario, four sensors observe an over-determined mixture of three sources. In the second scenario three sensors observe an under-determined mixture of four sources. The objective of this exercise is on the one hand to validate the theoretical results from this chapter and on the other hand to provide examples that illustrate the working of polynomial optimization based source extraction algorithms.

For both experiments stationary source signals with variance $\sigma_s^2 = 1$ are created by filtering a real-valued Gaussian signal with a filter consisting of a single pole. The poles for the four sources are respectively 0.9, 0.6, -0.7 and -0.9. For the first scenario, only the first three sources are used. After mixing the sources, the sensor signals are contaminated by temporally and spatially white Gaussian noise with a variance of $\sigma_\nu^2 = 0.1$.

An over-determined mixture with $D = 4$ and $S = 3$: The over-determined mixing matrix for the first scenario is given as follows:

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} & \frac{1}{\sqrt{10}}\begin{bmatrix}-2\\2\\3\\-1\end{bmatrix} & \frac{1}{2\sqrt{10}}\begin{bmatrix}-2\\-1\\1\\2\end{bmatrix} \end{bmatrix}. \tag{6.130}$$

Using normalized mixing column vectors and source signals is useful for evaluation purposes. When a different, unknown scaling of the sources and the mixing column vectors is used, then the extraction filters and mixing column vectors can be identified up to an arbitrary scaling.

For the polynomial optimization based source extraction algorithm we use design strategy III from Section 6.7.2, with shape parameter $\gamma = 10$, i.e.,

$$\mathbf{\Xi}_1 = \mathbf{P}_0 \quad \text{and} \quad \mathbf{\Xi}_2 = \gamma(\mathbf{I} - \mathbf{P}_0) + \mathbf{P}_0 \tag{6.131}$$

where $\mathbf{P}_0 = \frac{\mathbf{a}^0(\mathbf{a}^0)^T}{(\mathbf{a}^0)^T\mathbf{a}^0}$. We choose the mold $\mathbf{a}^0$ in the plane spanned by the mixing column vectors $\mathbf{a}^1$ and $\mathbf{a}^3$ by means of the following parametrization:

$$\mathbf{a}^0 = \cos(\theta)\mathbf{a}^1 + \sin(\theta)\mathbf{a}^3. \tag{6.132}$$

Since $\mathbf{a}^1$ and $\mathbf{a}^3$ are orthogonal vectors, the angle θ corresponds to the angle between the mold $\mathbf{a}^0$ and the mixing column vector corresponding to the first source. In our performance measures we assume that the first source is the desired source. For both the extraction filter and the identified mixing column vector we measure performance via the signal to interference ratio, which is defined as follows:

$$SIR_{\tilde{\mathbf{w}}} = \frac{(\tilde{\mathbf{w}}\mathbf{a}^1)^2}{\sum_{j=2}^{S}(\tilde{\mathbf{w}}\mathbf{a}^j)^2} \quad \text{and} \quad SIR_{\mathbf{h}} = \frac{((\mathbf{h})^T\mathbf{a}^1)^2}{\sum_{j=2}^{S}((\mathbf{h})^T\mathbf{a}^j)^2} \tag{6.133}$$

where $\tilde{\mathbf{w}}$ and $\mathbf{h}$ are the identified source extraction filter and mixing column vector, respectively.

Since we created the mixing system, we can use (6.124) to calculate the expected value of the eigenvalue corresponding to the desired source, i.e,

$$\lambda(\theta) = \frac{1}{10\tan^2\theta + 1}. \tag{6.134}$$

In Figure 6.3 the $SIR_{\tilde{\mathbf{w}}}$, $SIR_{\mathbf{h}}$, and the expected eigenvalue λ are depicted as function of the angle θ between the mold $\mathbf{a}^0$ and the mixing column vector corresponding to the desired source $\mathbf{a}^1$.

The results in Figure 6.3 confirm that the polynomial optimization based source extraction algorithm can be used to identify the source extraction filter and the mixing column vector corresponding to a desired source, for which a priori information about the mixing parameters is available. The black vertical lines show the expected angle of ± 45 degrees for which the selection procedure breaks down, i.e., when the angle between the mold and the mixing column vector corresponding to the desired source becomes larger than the angle between the mold and one of the other mixing column vectors. In this case, the third source is selected when θ becomes larger than 45 degrees. For the current parametrization, the angle between the mold and the second mixing column vector is always larger than the angle between the mold and one of the other mixing column vectors. Therefore, selection of this second source does not occur in this example. For the extraction filter $\tilde{\mathbf{w}}$, the optimal SIR is infinite since

the interfering sources can be fully cancelled. However, due to estimation errors, a SIR of 35 dB is reached. For the identified mixing column vector a SIR of 12.5 dB is expected, as is indicated by the black horizontal line. Note that both filters reach and hold their maximum SIR for all molds, as long as the angle between the mold and the mixing column vectors is smallest for the desired source. Consequently, errors in the mold do not translate to errors in the identified extraction filters or mixing column vectors as long as the desired source is selected.

An under-determined mixture with $D = 3$ and $S = 4$: For the under-determined mixing scenario the mixing matrix is given as follows:

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix} & \frac{1}{\sqrt{19}}\begin{bmatrix}1\\-3\\-3\end{bmatrix} & \frac{1}{\sqrt{21}}\begin{bmatrix}4\\-1\\-2\end{bmatrix} & \frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix} \end{bmatrix}. \tag{6.135}$$

Since the mixing system is under-determined, source extraction filters that cancel all interfering sources do not exist, i.e., only identification of the mixing column vector corresponding to the desired source is possible. In order to identify this desired mixing column vector we use the algorithm presented in Section 6.6.1, where selection of the desired solution is performed via a linear fractional objective function. The vectors $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$ are chosen orthogonal to each other in the plane spanned by $\mathbf{a}^1$ and $\mathbf{a}^4$. Since the vectors $\mathbf{a}^1$ and $\mathbf{a}^4$ are orthogonal we use the following parametrization:

$$\boldsymbol{\xi}^1 = \cos(\theta)\mathbf{a}^1 + \sin(\theta)\mathbf{a}^4 \tag{6.136}$$

$$\boldsymbol{\xi}^2 = \sin(\theta)\mathbf{a}^1 - \cos(\theta)\mathbf{a}^4 \tag{6.137}$$

where the angle θ forms the angle between the mold $\boldsymbol{\xi}^1 = \mathbf{a}^0$ and the mixing column vector corresponding to the desired source. The eigenvalue corresponding to this desired source can be calculated by using (6.112), i.e.,

$$\lambda(\theta) = \cot\theta. \tag{6.138}$$

Results for the under-determined mixing scenario are depicted in Figure 6.4. In the top figure, the $\text{SIR}_{\mathbf{h}}$ is depicted as function of the angle θ. The horizontal black line indicates again the expected SIR, which is 3.4 dB for the current scenario. In the bottom graph the absolute largest eigenvalue is depicted as well as the expected absolute eigenvalue for the desired source as function of the angle θ. Since the eigenvalues can take values in the range from zero up to infinity, the function $\frac{\pi}{2}\arctan(\cdot)$ is used in the graph to compress the eigenvalues to a range from zero to one. An additional benefit of this compression function is that the compressed (absolute) eigenvalues become linear as function of θ. The black vertical lines indicate where the selection procedure is expected to go wrong. These points are obtained by calculating the angles between the desired mixing column vector and the other mixing column vectors projected onto the space spanned by $\mathbf{a}^1$ and $\mathbf{a}^4$. This leads to the angles of -29.5, 78.6 and 90.0 degrees, for the second up to the third source, respectively. Selection

of the desired source will fail when θ becomes closest to another source. For negative and positive angles this means that selection of the desired source fails when $\theta < -29.5/2 = -14.75$ degrees and $\theta > 78.6/2 = 39.3$ degrees, respectively. However, as long as the desired source is selected, maximum SIR is obtained.

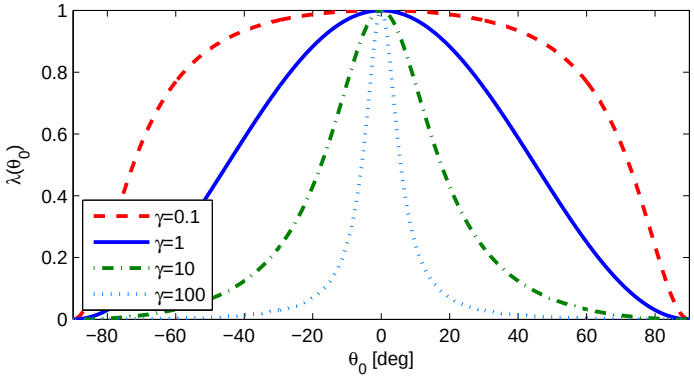
6.9 Conclusions

In the current chapter we presented a way to develop and design source extraction algorithms by formulating source extraction as a polynomial optimization problem. In this optimization problem, homogenous polynomial equations form equality constraints that restrict the feasible set of the optimization problem to extraction filters. A first or second order fractional objective function is used to identify the desired source extraction filter.

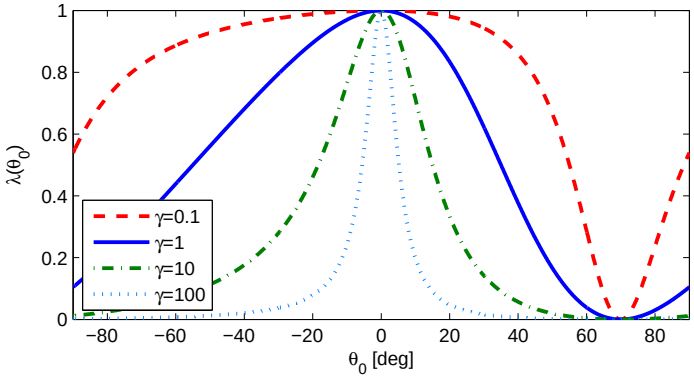
The presented polynomial optimization based approach can be considered a generalization of the methods presented in previous chapters. It works for both real-valued and complex-valued mixtures and it can be used to identify source extraction filters according both the MVDR and LCMV principle. A priori information about either the mixing parameters or the autocorrelation functions can be considered and in certain cases the source extraction algorithm can be formulated as a reference based source extraction algorithm.

Next to these similarities with previously presented algorithms, the polynomial optimization based approach has several beneficial properties. First, a largest eigenvalue problem is formulated instead of a smallest eigenvalue problem, which is beneficial in terms of numerical stability. Furthermore, applying reduction matrices to move from D to S are optional instead of required. We have shown that the polynomial optimization based approach can be used to identify extraction filters for under-determined mixtures. Finally, we have indicated that a natural extension towards using higher order statistics exists.

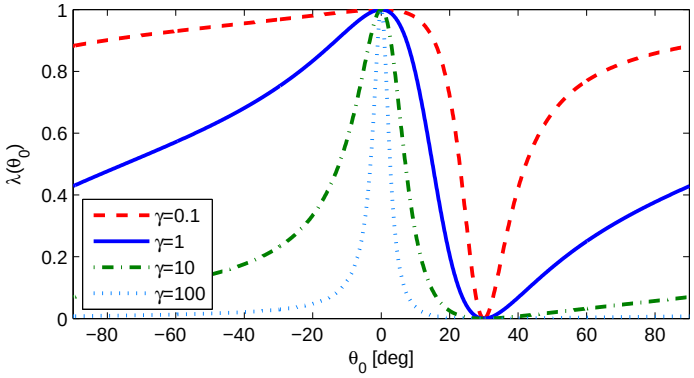
In the final part of this chapter we have presented several design strategies for the linear and second order fractional objective functions, based on several types and amounts of a priori information, and we validated the methodology for an over-determined and under-determined mixture via computer simulations.



(a) Regularized design



(b) Interferer at 70 degrees



(c) Interferer at 30 degrees

Figure 6.2: Second order fractional objective functions for several peakedness parameter values. In (6.2a) objective functions based on Design Strategy I are depicted. In (6.2b) and (6.2c) objective functions based on Design Strategy IV are depicted for interferers expected at 70 and 30 degrees, respectively.

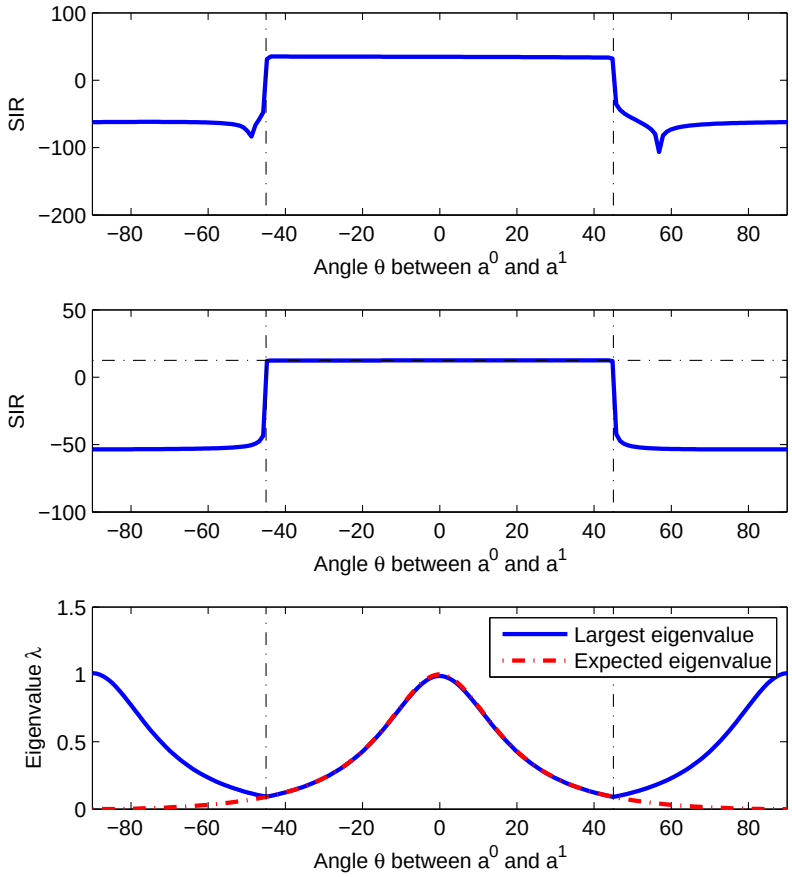


Figure 6.3: Results for the polynomial optimization based source extraction algorithm as a function of the angle θ between the mold and the mixing column vector corresponding to the desired source for an over-determined mixture. In the top and middle figure the SIR is depicted for the calculated extraction filter and identified mixing column vector, respectively. In the bottom figure the largest eigenvalue and the expected eigenvalue corresponding to the desired source is depicted.

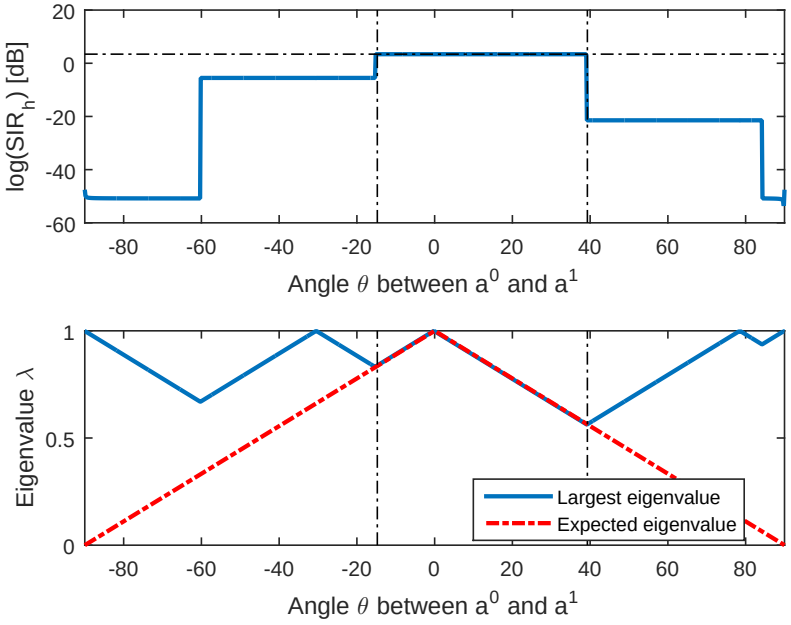


Figure 6.4: Results for the polynomial optimization based source extraction algorithm as a function of the angle θ between the mold and the mixing column vector corresponding to the desired source for an under-determined mixture. In the top figure the SIR_h is depicted for the identified mixing column vector. In the bottom figure the absolute largest eigenvalue and the expected absolute eigenvalue corresponding to the desired source is depicted through the compression function $\frac{\pi}{2}\arctan(\cdot)$.

7

A first study towards source extraction from convolutive mixtures

In this thesis the focus has been on the extraction of a desired source signal from an instantaneous mixture of multiple sources. However, in many other applications the sensor signals consist of more complex mixtures such as convolutive mixtures. In literature, both time and frequency domain domain methods have been proposed to solve problems such as blind system identification, blind source separation or extraction, and blind deconvolution for convolutive mixtures. By considering the convolutive mixing system in the frequency domain, informed source extraction can be solved by applying approaches for instantaneous mixtures; however, due to permutations in each frequency bin, this frequency domain approach puts high demands on the available a priori information.

In this chapter we use our insights on informed source extraction for instantaneous mixtures in order to provide insight and tools that support the design and development of informed source extraction algorithms for convolutive mixtures. In our approach we apply subspace techniques to statistical features of the observations in order to formulate polynomial optimization problems. In such an optimization problem a system of homogeneous polynomial equations restricts the feasible set to extraction filters and a priori information is used in the objective function in order to identify only the extraction filter for the desired source. In this chapter we show results of a first study where we apply our approach to a convolutive mixing model with short finite impulse responses. The foreseen benefit and objective of this approach is that requirements on available a priori information can be lowered.

7.1 Introduction

The objective in this chapter is to use our insights on informed source extraction for instantaneous mixtures in order to provide insight and tools that support the design and development of informed source extraction algorithms for convolutive mixtures. Convolutive mixtures can be considered a generalization of instantaneous mixtures and are useful in a wide range of applications such as speech enhancement, biomedical signal processing and telecommunication. In a convolutive mixing scenario the relation between a source and sensor is not modeled by a, potentially complex, gain, but by a linear, time-invariant impulse response. Models that describe such impulse responses are for example finite impulse responses (FIR) and infinite impulse responses (IIR) filters, which in turn can be dense or sparse [9, 11, 76]. Next to the different types of mixing filters, many different assumptions on the source and noise signals can be made.

With the introduction of a convolutive mixing system, also a new type of source extraction objective becomes available. In the literature, difference is made between blind source separation (BSS) and blind deconvolution (BD) objectives [11]. The difference between BSS and BD is that a BD algorithm tries to invert the mixing system whereas a BSS algorithm solely tries to separate the source signals by cancelling the interfering source signals, i.e., the original source signals are obtained up to an arbitrary filtering. Next to the objective for performing source extraction, the structure of the source extraction filters have to be chosen. Also for these filters a large variety of filter structures is available. Despite the many potential structures of convolutive mixing systems, in this chapter our main focus is on convolutive mixing models and source extraction filters that are represented by FIR filters. These types of filters are widely used and very practical in applications.

Convolutive mixtures can be expressed either in time domain or frequency domain [9]. An advantage of frequency domain approaches is that, under the condition that the number of frequency bins is sufficient, the convolutive mixing system reduces to independent, complex-valued, instantaneous mixing problems per frequency bin. Therefore, algorithms that work for instantaneous mixtures can be reused. Another advantage of the frequency domain approach is that due to the independence, convergence of frequency domain adaptive algorithms can typically be faster than convergence of time domain algorithms. However, frequency domain approaches also suffer from some disadvantages. In case the mixing filters are relatively long, many frequency bins are required to prevent circular convolution artifacts and it may become difficult to obtain sufficient samples per frequency bin in order to estimate statistical properties. Furthermore, the scaling and permutation indeterminacy that come with instantaneous mixing problems now arise in every frequency bin. Consequently, when our informed source extraction algorithm is applied in such a frequency domain approach, a priori information that ensures extraction of only the desired source has to be available for each frequency bin, which may be challenging in many practical applications. For example, in a speech enhancement scenario, a rough estimate of the direction of arrival of a desired source can be available; however, due to reverberation and long filter lengths, this information may be too much deteriorated for usage per

frequency bin. Finally, the arbitrary scalings per frequency bin translate into arbitrary, single channel filters for the extracted signal. Next to the time and frequency domain, also subband approaches have been applied [77]. Applying our informed source extraction approach combined with such a subband technique could potentially lead to informed source extraction algorithms for convolutive mixtures where a priori information is only required per subband. However, in each subband the source signals are still mixed by a convolutive mixing system such that our informed source extraction algorithms for instantaneous mixtures cannot be applied directly.

In this chapter we study the transformation of our informed source extraction approach for instantaneous mixtures to convolutive mixtures. In the previous Chapter 6 we presented a generic formulation of the source extraction problem for instantaneous mixtures in the form of a polynomial optimization problem. We have also shown methods on how to compute the solutions to these optimization problems. In order to derive informed source extraction algorithms for convolutive mixtures, based on our approach for instantaneous mixtures, we first need to be able to formulate a system of polynomial equations from which the roots correspond to the mixing parameters or source extraction filters. Subsequently, we should be able to use a priori information in order to formulate an objective function that ensures that directly the root corresponding to the desired source is computed. Generic mathematical tools to solve such polynomial optimization problems are being developed [73, 74]. Consequently, our focus is on formulating informed source extraction for convolutive mixtures as a polynomial optimization problem by applying subspace techniques to the statistical features of the convolutive mixtures. In this chapter we do not come to a generic formulation, but we are able to show that short mixing filters can be identified from the roots of a system of homogenous polynomial equations.

The outline of this chapter is as follows. In Section 7.2 we discuss different types of convolutive mixing models and present the assumptions on the source and noise signals that we use in the remainder of this chapter. The structure in sensor correlation data from convolutive mixtures is exploited in Section 7.3 in order to formulate a system of polynomial equations from which the roots correspond to the convolutive mixing parameters. Finally, we conclude this chapter in Section 7.4.

7.2 Model and assumptions

7.2.1 Mixing model

In a convolutive mixing scenario we assume that a total of S source signals $\mathbf{s}_j$ are mixed by a size $D \times S$ convolutive mixing system $\mathbf{H}$, consisting of FIR impulse responses or filters. The observations consist of the mixtures contaminated by additive noise $\nu_i[n]$. Mathematically, this model can be formulated as follows:

$$\mathbf{x}[n] = \sum_{j=1}^{S} (\mathbf{h}^j * s_j)[n] + \boldsymbol{\nu}[n] = (\mathbf{H} * \mathbf{s})[n] + \boldsymbol{\nu}[n] \tag{7.1}$$

where $*$ represents convolution and the vectors containing respectively the sensor, source, and noise signals are given as follows:

$$\mathbf{x}[n] \triangleq \begin{bmatrix} x_1[n] \\ \vdots \\ x_D[n] \end{bmatrix}, \quad \mathbf{s}[n] \triangleq \begin{bmatrix} s_1[n] \\ \vdots \\ s_S[n] \end{bmatrix}, \quad \boldsymbol{\nu}[n] \triangleq \begin{bmatrix} \nu_1[n] \\ \vdots \\ \nu_D[n] \end{bmatrix}. \tag{7.2}$$

The convolutive mixing system of size $D \times S$ is defined as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}^1 & \cdots & \mathbf{h}^S \end{bmatrix} = \begin{bmatrix} h_1^1 & \cdots & h_1^S \\ \vdots & \ddots & \vdots \\ h_D^1 & \cdots & h_D^S \end{bmatrix}. \tag{7.3}$$

In the convolutive mixing model, each mixing component is considered an impulse response or filter. Evaluating the mixing component h_i^j at sample index p is formulated as $h_i^j[p]$ and corresponds to the impulse response value at time pT_s, where T_s is the sample time. Many different structures can be underlying the finite impulse response. For example, the impulse responses can have many or almost no zeros, or have a wideband or narrowband characteristic. If knowledge about these properties is known from the application, it should be incorporated in source extraction algorithms.

We assume that similar properties as presented in the previous chapter hold for the source signals. To summarize these properties, we assume that all source signals are mutually statistically independent, or at least uncorrelated. Furthermore, we assume that the source signals contain second order temporal structure (SOTS). In the current chapter we focus on SOTS in the form of non-whiteness. Furthermore, we assume that the statistical properties of the sources, i.e., the source autocorrelation functions, are sufficiently different from each other. Finally, we assume that the noise is uncorrelated from the source signals and for now we assume that we can ignore the contribution of the noise in the sensor correlation data. However, this is only allowed if we can extend the concept of a NF-ROS to convolutive mixtures. A straightforward approach to define a NF-ROS is to estimate the noise statistics and compensate for it in the sensor correlation data.

In the remainder of this chapter we discuss several topics related to convolutive mixtures based on the generic convolutive mixing model in (7.1); however, for certain topics we consider a specific mixing system consisting of causal, linear, and finite impulse responses of maximum order P. For such impulse responses the convolutive mixture can be written as follows:

$$\mathbf{x}[n] = \sum_{j=1}^{S} \sum_{p=0}^{P} \mathbf{h}^j[p] s_j[n-p] + \boldsymbol{\nu}[n]. \tag{7.4}$$

Next we consider different objectives and structures for source extraction filters.

7.2.2 Source extraction filters

Given the convolutive mixing models in (7.1) and (7.4), our objective is to identify a filter $\tilde{\mathbf{w}}$ that extracts the desired source from the observed mixture, i.e.,

$$y[n] = (\tilde{\mathbf{w}} * \mathbf{x})[n] = (\tilde{\mathbf{w}} * \mathbf{H} * \mathbf{s})[n] + (\tilde{\mathbf{w}} * \boldsymbol{\nu})[n] \tag{7.5}$$

where $y[n]$ ideally only contains the desired source signal. The extraction filter components are considered to be impulse responses or filters. Same as for the filters that represent the mixing system, the source extraction filters can have many different forms such as FIR and IIR filters. In the remainder of this thesis we assume that the extraction filters are FIR filters of maximum order Q. Consequently, the output signal $y[n]$ can be written as follows:

$$y[n] = \sum_{q=0}^{Q} \tilde{\mathbf{w}}[q]\mathbf{x}[n-q] = \sum_{i=1}^{D}\sum_{q=0}^{Q} w^i[q]x_i[n-q]. \tag{7.6}$$

For the FIR source extraction filters, their optimal filter lengths Q depend on the mixing system at hand and the objective of the source extraction filter. In this section we consider three different objectives for the source extraction filters and the corresponding required filter lengths.

First of all, it is known that a matched filter, where the extraction filters are chosen as the time-reversed impulse responses of the desired source, increases the signal to interference ratio. Consequently, identifying the mixing filters or impulse responses that correspond to the desired source can be an objective for constructing a source extraction filter. For optimal performance in terms of signal to noise ratio (SNR), the extraction filter length of a matched filter should match the length of the impulse responses in the mixing system. When the number of sensors is increased the SNR increases as well. However, in order to maintain a certain desired SNR at the output of the extraction filter the extraction filters may be shortened when the number of sensors is increased. Alternatively, if the source extraction filters are longer than required, additional noise reduction leads to the following objective for an MVDR extraction filter:

$$\tilde{\mathbf{w}}_{\text{MVDR}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{h}^d = g_\delta \tag{7.7}$$

where $\mathbf{h}^d$ is a vector containing the impulse responses from the desired source to the sensors and g_δ represents a, potentially delayed, unit response. When the MVDR objective is applied, the filter lengths should be chosen at least equal or larger than the filter lengths of the mixing system for optimal performance.

In case the objective is to separate the desired source from the other sources, i.e., to maximize the signal to interference ratio, then we can derive optimal filter lengths as function of the number of sensors based on the following analysis. The number of constraints for this objective are $(S-1)(P+Q+1)$ zero constraints for the interference cancelation and a single constraint for not cancelling the desired source. The degrees of freedom for filters of length $(Q+1)$ applied to D sensors is given as $D(Q+1)$.

Consequently, the source separation objective leads to the following constraints on the number of filter coefficients $(Q+1)$ per sensor:

$$Q+1 \geq \frac{P(S-1)+1}{D-S+1}. \tag{7.8}$$

The corresponding optimization problem for this source separation objective, with additional noise reduction, is as follows:

$$\tilde{\mathbf{w}}_{\text{separation}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{h}^i = g_0 \quad \text{and} \quad \tilde{\mathbf{w}}\mathbf{h}^d \neq g_0 \tag{7.9}$$

where $\mathbf{h}^i$ corresponds to the mixing filters or impulse responses of the interfering sources, $\mathbf{h}^d$ corresponds to the mixing filters of the desired source, and g_0 represents an impulse response consisting of only zeroes. Again, increasing the number of extraction filter coefficients allows for additional noise reduction.

Finally, in case the objective is deconvolution, which means that besides separation, also the impulse response applied to the desired source is constraint, then the number of constraints increases to $S(P+Q+1)$ while there are still $D(Q+1)$ degrees of freedom. This deconvolution objective leads to the following constraints on the minimally required number of filter coefficients:

$$Q+1 \geq \frac{PS}{D-S} \tag{7.10}$$

From this constraint we observe immediately the constraint that the number of sensors should be strictly larger than the number of sources, i.e., $D > S$. The corresponding optimization problem for the deconvolution problem is mathematically formulated as follows:

$$\tilde{\mathbf{w}}_{\text{deconvolution}} = \arg\min_{\tilde{\mathbf{w}}} \mathbb{E}\{|y[n]|^2\} \quad \text{s.t. } \tilde{\mathbf{w}}\mathbf{H} = \tilde{\mathbf{g}}_d \tag{7.11}$$

where $\tilde{\mathbf{g}}_d$ is a vector of filters with the, potentially delayed, unit response g_δ for the desired source and g_0 for the remaining filters.

From these three different objectives we observe that increasing the number of sensors reduces the required number of filter coefficients per sensor. In the remainder of this chapter our focus will be on the matched filter as a source extraction filter and therefore on the identification of the mixing system parameters that correspond to the desired source. However, we realize that also the other objectives can be very relevant for applications and consider the matched filter approach a first step. In our approach we apply subspace techniques to the sensor correlation data. Therefore, in the next section we investigate the structure in the sensor correlation data for convolutive mixtures.

7.3 Structure in the sensor correlation data

When the source extraction problem is considered as a number of independent instantaneous mixing systems in the frequency domain, then a priori information must

be available that ensures extraction of the desired source for each separate frequency bin. In case the convolutive source extraction problem is tackled as a wideband or subband problem, then the requirements on the available a priori information used for selection of the desired filter are less restrictive. However, before we can incorporate a priori information about the desired source into a source extraction algorithm for convolutive mixtures we must first be able to identify source extraction filters for such mixtures. Therefore, in the remainder of this chapter we apply subspace techniques to sensor correlation data for convolutive mixing systems where sources are mixed by short FIR filters. This approach can be a first step towards the development of informed source extraction algorithms for convolutive mixtures.

Based on the assumptions in the previous section, the structure in the noise-free sensor correlation data follows from the convolutive mixing model in (7.1) and is given as follows (see Appendix 7.A for a derivation):

$$r^x_{i_1 i_2}[k] \triangleq \mathbb{E}\{x_{i_1}[n]x_{i_2}[n-k]\} = \sum_{j=1}^{S}\sum_{p_1=0}^{P}\sum_{p_2=0}^{P} h^j_{i_1}[p_1]h^j_{i_2}[p_2]r^s_{jj}[k+p_2-p_1] \quad (7.12)$$

where we exploit again that the source signals are mutually uncorrelated.

Next we consider this structure in vector matrix format such that we can identify linear subspaces and formulate systems of homogeneous polynomial equations from which the roots correspond to mixing filter parameters. In order to introduce the notation and structure of the problem at hand we start with a single source use-case and present our approach to exploit the temporal structure in the mixing system.

7.3.1 Structure for a single source

Although the work in this thesis is concerned with the extraction of one source signal from a mixture of multiple sources, here we use a single source model in order to introduce the notation of and structure in sensor correlation data for the convolutive mixing scenario.

For $D = 2$, $P = 1$ and $S = 1$, the structure in the sensor correlation data from (7.12) can be represented in the following vector-matrix format:

$$\begin{bmatrix} r^x_{11}[k] \\ r^x_{12}[k] \\ r^x_{21}[k] \\ r^x_{22}[k] \end{bmatrix} = \begin{bmatrix} h^1_1[0]h^1_1[1] & h^1_1[0]h^1_1[0]+h^1_1[1]h^1_1[1] & h^1_1[1]h^1_1[0] \\ h^1_1[0]h^1_2[1] & h^1_1[0]h^1_2[0]+h^1_1[1]h^1_2[1] & h^1_1[1]h^1_2[0] \\ h^1_2[0]h^1_1[1] & h^1_2[0]h^1_1[0]+h^1_2[1]h^1_2[1] & h^1_2[1]h^1_1[0] \\ h^1_2[0]h^1_2[1] & h^1_2[0]h^1_2[0]+h^1_2[1]h^1_1[1] & h^1_2[1]h^1_2[0] \end{bmatrix} \begin{bmatrix} r^s_{11}[k+1] \\ r^s_{11}[k] \\ r^s_{11}[k-1] \end{bmatrix}. \quad (7.13)$$

This structure in the sensor correlation data is exploited by evaluating a sensor correlation matrix $\mathbf{C}^x$ that is built by stacking sensor correlation vectors from (7.13) for K lags next to each other, which leads to the following sensor correlation matrix

structure of size $D^2 \times K$:

$$\mathbf{C}^x \triangleq \begin{bmatrix} r^x_{11}[k_1] & \cdots & r^x_{11}[k_K] \\ r^x_{12}[k_1] & \cdots & r^x_{12}[k_K] \\ r^x_{21}[k_1] & \cdots & r^x_{21}[k_K] \\ r^x_{22}[k_1] & \cdots & r^x_{22}[k_K] \end{bmatrix}. \tag{7.14}$$

We use the following short-hand notation for the structure in the sensor correlation matrix $\mathbf{C}^x$:

$$\mathbf{C}^x \triangleq \mathbf{G}\mathbf{C}^s \tag{7.15}$$

where

$$\mathbf{G} \triangleq \begin{bmatrix} h^1_1[0]h^1_1[1] & h^1_1[0]h^1_1[0] + h^1_1[1]h^1_1[1] & h^1_1[1]h^1_1[0] \\ h^1_1[0]h^1_2[1] & h^1_1[0]h^1_2[0] + h^1_1[1]h^1_2[1] & h^1_1[1]h^1_2[0] \\ h^1_2[0]h^1_1[1] & h^1_2[0]h^1_1[0] + h^1_2[1]h^1_1[1] & h^1_2[1]h^1_1[0] \\ h^1_2[0]h^1_2[1] & h^1_2[0]h^1_2[0] + h^1_2[1]h^1_1[1] & h^1_2[1]h^1_2[0] \end{bmatrix} \tag{7.16}$$

$$\mathbf{C}^s \triangleq \begin{bmatrix} r^s[k_1+1] & \cdots & r^s[k_K+1] \\ r^s[k_1] & \cdots & r^s[k_K] \\ r^s[k_1-1] & \cdots & r^s[k_K-1] \end{bmatrix} \tag{7.17}$$

with $\mathbf{G}$ a size $D^2 \times 2P+1$ matrix containing auto- and crosscorrelations of the impulse responses of the mixing system and $\mathbf{C}^s$ a size $2P+1 \times K$ source autocorrelation matrix.

When the sensor correlation matrix is built for K consecutive lags, then the corresponding source autocorrelation matrix $\mathbf{C}^s$ obtains a Toeplitz matrix structure. Selecting different lags lowers redundancy in the sensor correlation matrix, but it allows for incorporating sensor correlation data from different time instances such that non-stationarity of source signals can be exploited.

Note that the columns of $\mathbf{G}$ and the rows of $\mathbf{C}^s$ can be permuted simultaneously. In case the order of the columns of $\mathbf{G}$ and the rows of $\mathbf{C}^s$ is reversed, then for consecutive lags the Toeplitz matrix structure of $\mathbf{C}^s$ becomes a Hankel matrix structure. We do not explicitly exploit these observations on the structure and therefore the different orderings or permutations are a matter of convention and here we choose, without loss of generality, for the structure from (7.13), which leads to the Toeplitz matrix structure.

By applying subspace techniques to a sensor correlation matrix $\mathbf{C}^x$ information about the mixing parameters can be obtained. We first investigate the structure of the sensor correlation matrix $\mathbf{C}^x$. The rows in the matrix $\mathbf{G}$ form the auto- and crosscorrelation functions of the impulse responses of the mixing system. From this observation it follows that the impulse response matrix $\mathbf{G}$ is rank deficient in case the impulse responses have so called common zeroes in the frequency domain. This property is also widely recognized in the literature. Although common zeroes and rank deficiency of $\mathbf{G}$ is recognized as a relevant problem, for brevity and focus, here we assume that the matrix $\mathbf{G}$ has full rank. For the source autocorrelation matrix $\mathbf{C}^s$ it is in general likely to have full rank in case the source autocorrelation functions are

mutually different. A class of signals that form an exception to this assumption are signals that are generated from an autoregressive model with a single pole. When for these signals either positive or negative lags are not taken into account, then columns of the source autocorrelation matrix become linearly dependent such that the rank of the source autocorrelation matrix reduces. In this thesis we assume that the source autocorrelation functions are linearly independent and that the source autocorrelation matrix has maximum rank, which is a likely scenario in many applications.

The sensor correlation matrix $\mathbf{C}^x$ has size $D^2 \times K$ and due to its structure a maximum rank of $\min(D^2, K, 2P+1)$. Due to our assumptions that the matrices $\mathbf{G}$ and $\mathbf{C}^s$ have full rank, this maximum rank of $\min(D^2, K, 2P+1)$ of the sensor correlation matrix is also realized. Our subspace approach relies on the fact that the sensor correlation matrix has a left nullspace. Consequently, with $K \geq 2P+1$ we require at least that $D^2 > 2P+1$. For the scenario where $D = 2$ and $P = 0$, i.e., the instantaneous mixing scenario, it follows that a left nullspace exists; however, it follows that for longer filters no left nullspace exists in the sensor correlation matrix when using only two sensors. However, temporal structure in the source signals can be exploited in order to create a larger system of equations, allowing for the identification of the mixing parameters. Instead of considering one lag per column of sensor correlation data, multiple consecutive lags can be considered. This leads to the following structure for $N = 2$ lags:

$$\begin{bmatrix} r^x_{11}[k] \\ r^x_{12}[k] \\ r^x_{21}[k] \\ r^x_{22}[k] \\ r^x_{11}[k-1] \\ r^x_{12}[k-1] \\ r^x_{21}[k-1] \\ r^x_{22}[k-1] \end{bmatrix} = \begin{bmatrix} h^1_1[0]h^1_1[1] & h^1_1[0]h^1_1[0]+h^1_1[1]h^1_1[1] & h^1_1[1]h^1_1[0] & 0 \\ h^1_1[0]h^1_2[1] & h^1_1[0]h^1_2[0]+h^1_1[1]h^1_2[1] & h^1_1[1]h^1_2[0] & 0 \\ h^1_2[0]h^1_1[1] & h^1_2[0]h^1_1[0]+h^1_2[1]h^1_1[1] & h^1_2[1]h^1_1[0] & 0 \\ h^1_2[0]h^1_2[1] & h^1_2[0]h^1_2[0]+h^1_2[1]h^1_2[1] & h^1_2[1]h^1_2[0] & 0 \\ 0 & h^1_1[0]h^1_1[1] & h^1_1[0]h^1_1[0]+h^1_1[1]h^1_1[1] & h^1_1[1]h^1_1[0] \\ 0 & h^1_1[0]h^1_2[1] & h^1_1[0]h^1_2[0]+h^1_1[1]h^1_2[1] & h^1_1[1]h^1_2[0] \\ 0 & h^1_2[0]h^1_1[1] & h^1_2[0]h^1_1[0]+h^1_2[1]h^1_1[1] & h^1_2[1]h^1_1[0] \\ 0 & h^1_2[0]h^1_2[1] & h^1_2[0]h^1_2[0]+h^1_2[1]h^1_2[1] & h^1_2[1]h^1_2[0] \end{bmatrix} \begin{bmatrix} r^s_{11}[k+1] \\ r^s_{11}[k] \\ r^s_{11}[k-1] \\ r^s_{11}[k-2] \end{bmatrix}. \tag{7.18}$$

This approach to exploit multiple consecutive lags can generalized. For sensor correlation matrices that are built from N lags in the vertical direction we use the following notation:

$$\mathbf{C}^x_N = \mathbf{G}_N \mathbf{C}^s_N. \tag{7.19}$$

An example matrix $\mathbf{G}_N$ for $N = 3$ is given in Appendix 7.A.

Note that the dimensions of the extended sensor correlation matrix $\mathbf{C}^x_N$ increase to $ND^2 \times K$, whereas its maximum rank is given as $\min(ND^2, K, 2P+N)$. In order to obtain a left null space, with $K \geq 2P+N$, we require at least that $ND^2 > 2P+N$. Consequently, without increasing the number of sensors, for $D = 2$ a left nullspace is obtained in case $N > \frac{2P}{3}$.

The nullspace of expanded sensor correlation matrices is used in the following sections in order to identify impulse responses of the mixing system. Again, we start with the single source use-case in order to show the principle and later we present results for a mixture of two sources.

7.3.2 System identification for $D = 2$, $P = 1$ and $S = 1$

In this section we consider a system where a single source signal is filtered by short impulse responses of order $P = 1$ and the filtered signals are captured by two sensors. Even though this scenario does not correspond to a convolutive mixing scenario since there is only one source, we use this scenario to obtain insight in the structure of the data. Furthermore, we use this scenario in order to present the main steps of our approach to formulate systems of homogeneous polynomial equations from which impulse responses of convolutive mixing systems can be identified.

For $D = 2$, $P = 1$, $S = 1$, and $K \geq 2P + N$, it follows that with $N = 2$ the sensor correlation matrix $\mathbf{C}_2^x = \mathbf{G}_2\mathbf{C}_2^s$ has a left nullspace. The correlation matrix has eight rows, while its maximum rank, $2 \cdot P + N$, equals four. When the matrices $\mathbf{G}_2$ and $\mathbf{C}_2^s$ are full rank, the nullspace of the sensor correlation matrix $\mathbf{C}_2^x$ has a dimension of four. Consequently, we are able to find a size 4×8 nullspace matrix $\mathbf{\Phi}$ such that

$$\mathbf{\Phi}\mathbf{C}_2^x = \mathbf{0} \tag{7.20}$$

where $\mathbf{0}$ is a size $4 \times K$ matrix of zeroes. Note that for $N = 1$ also a nullspace can be identified; however, with the corresponding nullspace we are unable to demonstrate our approach when the number of sources is increased.

Since the matrix $\mathbf{C}_2^s$ is assumed to have full rank, we are able to exploit the structure in $\mathbf{G}_2$ in order to form a system of homogeneous second order polynomials from which the roots correspond to the parameters of the convolutive mixing system. In order to capture the structure in $\mathbf{G}_2$ we substitute the filter coefficients $h_i^j[p]$ by the variables $z_{i,p}$ and for short filters with single digit filter orders, i.e., $P < 10$, we use the short hand notation z_{ip}. This leads to the following substitutions for a convolutive system with $D = 2$ sensors and an impulse response order of $P = 1$:

$$h_1^j[0] \rightarrow z_{10}, \quad h_1^j[1] \rightarrow z_{11}, \quad h_2^j[0] \rightarrow z_{20}, \quad \text{and}, \quad h_2^j[1] \rightarrow z_{21}. \tag{7.21}$$

Combining (7.20), the structure in $\mathbf{G}_2$, and the representation of the filters by the variables z_{ip}, we obtain the following two systems of four homogeneous second order polynomial equations:

$$\phi_q^1 z_{10}z_{11} + \phi_q^2 z_{10}z_{21} + \phi_q^3 z_{20}z_{11} + \phi_q^4 z_{20}z_{21} = 0 \qquad \forall\, 1 \leq q \leq 4 \tag{7.22}$$

and

$$\begin{aligned}\phi_q^1 z_{10}z_{11} + \phi_q^3 z_{10}z_{21} + \phi_q^2 z_{20}z_{11} + \phi_q^4 z_{20}z_{21} + \phi_q^5(z_{10}z_{10} + \\ z_{11}z_{11}) + (\phi_q^6 + \phi_q^7)(z_{10}z_{20} + z_{11}z_{21}) + \phi_q^8(z_{20}z_{20} + z_{21}z_{21}) = 0\end{aligned} \tag{7.23}$$

where ϕ_q^i is the element in the i'th column on the q'th row of the nullspace matrix $\mathbf{\Phi}$. These equations are based on the first and third column of the matrix structure $\mathbf{G}_2$. Solving these equations allows for identification of the impulse responses of the mixing system. Next we present two ways to compute the roots of these equations.

7.3.2.1 Computing roots of a system of polynomial equations: first approach

The system of homogeneous polynomial equations from (7.22) and (7.23) can be considered in the following vector-matrix relation:

$$\begin{bmatrix} \phi_1^1 & \phi_1^2 & \phi_1^3 & \phi_1^4 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_4^1 & \phi_4^2 & \phi_4^3 & \phi_4^4 & 0 & 0 & 0 \\ \phi_1^1 & \phi_1^2 & \phi_1^3 & \phi_1^4 & \phi_1^5 & \phi_1^6 + \phi_1^7 & \phi_1^8 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_4^1 & \phi_4^2 & \phi_4^3 & \phi_4^4 & \phi_4^5 & \phi_4^6 + \phi_4^7 & \phi_4^8 \end{bmatrix} \begin{bmatrix} z_{10}z_{11} \\ z_{10}z_{21} \\ z_{20}z_{11} \\ z_{20}z_{21} \\ z_{10}z_{10} + z_{11}z_{11} \\ z_{10}z_{20} + z_{11}z_{21} \\ z_{20}z_{20} + z_{21}z_{21} \end{bmatrix} = \mathbf{0}. \tag{7.24}$$

The matrix on the left has size 8×7. Although a reasoning and proof is still missing, from Matlab experiments we learned that the matrix typically has rank 6 for arbitrarily generated mixing filter coefficients. This means that the right nullspace $\mathbf{u}$ has size 7×1, which can be used to formulate the following system of homogeneous polynomial equations:

$$\begin{bmatrix} z_{10}z_{11} \\ z_{10}z_{21} \\ z_{20}z_{11} \\ z_{20}z_{21} \\ z_{10}z_{10} + z_{11}z_{11} \\ z_{10}z_{20} + z_{11}z_{21} \\ z_{20}z_{20} + z_{21}z_{21} \end{bmatrix} = \alpha\mathbf{u}. \tag{7.25}$$

From these equations the following relations can be derived:

$$\frac{z_{20}}{z_{10}} = \frac{u_3}{u_1} \quad \text{and} \quad \frac{z_{21}}{z_{11}} = \frac{u_2}{u_1}. \tag{7.26}$$

Which in turn leads to the following equations:

$$z_{10}z_{10} + z_{11}z_{11} = \alpha u_5 \tag{7.27}$$

$$\frac{u_3}{u_1} z_{10}z_{10} + \frac{u_2}{u_1} z_{11}z_{11} = \alpha u_6 \tag{7.28}$$

$$\left(\frac{u_3}{u_1}\right)^2 z_{10}z_{10} + \left(\frac{u_2}{u_1}\right)^2 z_{11}z_{11} = \alpha u_7. \tag{7.29}$$

Due to the scaling indeterminacy, which we know from the instantaneous mixing scenario, we know that we can choose $z_{10} = 1$ as long as $h_1^j[0] \neq 0$ holds. Using this assumption and by substituting (7.27) into (7.28) we obtain the following result for α:

$$u_3 + u_2(\alpha u_5 - 1) = \alpha u_1 u_6 \qquad \Rightarrow \qquad \alpha = \frac{u_3 - u_2}{u_1 u_6 - u_2 u_5}. \tag{7.30}$$

From this it follows that the scaled impulse response coefficients can be computed by:

$$z_{10} = h_1^j[0]/h_1^j[0] = 1 \tag{7.31}$$
$$z_{20} = h_2^j[0]/h_1^j[0] = u_3/u_1 \tag{7.32}$$
$$z_{11} = h_1^j[1]/h_1^j[0] = \alpha u_1 \tag{7.33}$$
$$z_{21} = h_2^j[1]/h_1^j[0] = \alpha u_2. \tag{7.34}$$

Next an alternative approach to compute the roots of the system of polynomial equation from (7.22) and (7.23) is presented.

7.3.2.2 Computing roots of a system of polynomial equations: second approach

In our second approach we first identify the columns of $\mathbf{H}^j$ from the nullspace of the matrix generated by the equations from (7.22):

$$\phi_q^1 z_{10} z_{11} + \phi_q^2 z_{10} z_{21} + \phi_q^3 z_{20} z_{11} + \phi_q^4 z_{20} z_{21} = 0 \qquad \forall\, 1 \leq q \leq 4. \tag{7.35}$$

Thus

$$\mathbf{\Phi} \begin{bmatrix} z_{10}z_{11} \\ z_{10}z_{21} \\ z_{20}z_{11} \\ z_{20}z_{21} \end{bmatrix} = \begin{bmatrix} \phi_1^1 & \phi_1^2 & \phi_1^3 & \phi_1^4 \\ \phi_2^1 & \phi_2^2 & \phi_2^3 & \phi_2^4 \\ \phi_3^1 & \phi_3^2 & \phi_3^3 & \phi_3^4 \\ \phi_4^1 & \phi_4^2 & \phi_4^3 & \phi_4^4 \end{bmatrix} \begin{bmatrix} z_{10}z_{11} \\ z_{10}z_{21} \\ z_{20}z_{11} \\ z_{20}z_{21} \end{bmatrix} = \mathbf{0}. \tag{7.36}$$

From experiments with randomly generated impulse responses it follows that the matrix $\mathbf{\Phi}$ has a one dimensional nullspace $\mathbf{u}$. Exploiting this nullspace leads to the following system of homogeneous polynomial equations:

$$\begin{bmatrix} z_{10}z_{11} \\ z_{10}z_{21} \\ z_{20}z_{11} \\ z_{20}z_{21} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}. \tag{7.37}$$

From this system we are able to identify the ratios z_{10}/z_{20} and z_{11}/z_{21}, i.e., $z_{20} = u_3/u_1 z_{10} = \alpha_1 z_{10}$ and $z_{11} = u_1/u_2 z_{21} = \alpha_2 z_{21}$, where we use α_1 and α_2 as shorthand notation. Substituting these ratios in the system of homogeneous polynomial equations of (7.23) leads to the following system of four homogeneous polynomial equations:

$$\begin{aligned} \left(\phi_q^5 + (\phi_q^6 + \phi_q^7)\alpha_1 + \phi_q^8(\alpha_1)^2\right) z_{10}z_{10} \\ + (\phi_q^1 + \phi_q^2\alpha_1 + \phi_q^3\alpha_2 + \phi_q^4\alpha_1\alpha_2) z_{10}z_{11} \\ + \left(\phi_q^5 + (\phi_q^6 + \phi_q^7)\alpha_2 + \phi_q^8(\alpha_2)^2\right) z_{11}z_{11} = 0. \end{aligned} \tag{7.38}$$

Transforming this system of polynomial equations into vector notation for the variables $z_{10}z_{10}$, $z_{10}z_{11}$, and $z_{11}z_{11}$ shows that a system of four equations in three vari-

ables, i.e.,

$$\mathbf{\Psi} \begin{bmatrix} z_{10}z_{10} \\ z_{10}z_{11} \\ z_{11}z_{11} \end{bmatrix} = \mathbf{0} \tag{7.39}$$

where the elements on the q'th row of the matrix $\mathbf{\Psi}$ have the following value:

$$\psi_q^1 = \phi_q^5 + (\phi_q^6 + \phi_q^7)\alpha_1 + \phi_q^8(\alpha_1)^2 \tag{7.40}$$

$$\psi_q^2 = \phi_q^1 + \phi_q^2\alpha_1 + \phi_q^3\alpha_2 + \phi_q^4\alpha_1\alpha_2 \tag{7.41}$$

$$\psi_q^3 = \phi_q^5 + (\phi_q^6 + \phi_q^7)\alpha_2 + \phi_q^8(\alpha_2)^2. \tag{7.42}$$

Since the first column of $\mathbf{\Psi}$ is equal to the third column of $\mathbf{\Psi}$, i.e., $\psi_q^1 = \psi_q^3$ for $1 \leq q \leq 4$, a left nullspace $\mathbf{v}$ of the matrix $\mathbf{\Psi}$ with a size of 3×1 exists. If we assume again that $h_1^j[0] \neq 0$, we can normalize the nullspace vector $\mathbf{v}$ by scaling such that its first element becomes equal to one. Consequently, we know that $z_{10} = 1$ and it follows that $z_{11} = v_2$ holds. Finally, we use again the ratios $z_{20} = \alpha_1 z_{10}$ and $z_{21} = \alpha_2 z_{11}$ such that the impulse response filters are successfully identified.

We realize that many steps have to be taken in order to proof the concept and to generalize these results to more practical scenarios; however, now that we have demonstrated the principle of the mixing filter identification strategy, we apply this technique to sensor correlation data for a mixture of two sources in order to demonstrate feasibility by formulating the system identification problem for convolutive mixtures as a root finding problem for a system of homogeneous polynomial equations.

7.3.3 System identification for $D = 2$, $P = 1$ and $S = 2$

In this section we consider a system where two source signals are filtered by short impulse responses of order $P = 1$ and the filtered signals are captured by two sensors. We use this scenario to present our results on a study where we apply our insights on informed source extraction from instantaneous mixtures in order develop informed source extraction algorithm for convolutive mixtures.

For $D = 2$, $P = 1$, $S = 2$, the structure in a sensor correlation matrix $\mathbf{C}^x$ is changed in the sense that the number of columns in a matrix $\mathbf{G}$ and the number of rows in a matrix $\mathbf{C}^s$ is doubled due to the presence of an extra source. Consequently, the rank of the matrix $\mathbf{C}_N^x$ is given by $\min(ND^2, K, (2P+N)S)$. Consequently, in case $K \geq (2P+N)S$, the condition for having a left nullspace is $N > \frac{(2PS}{D^2-S} = \frac{4}{2} = 2$.

For $N = 3$, the rank of the sensor correlation matrix is at most $(2P+N)S = (2+3)2 = 10$, while we have $ND^2 = 12$ rows of sensor correlation data. Therefore, the nullspace $\mathbf{\Phi}$ is expected to have a size of 2×12, i.e.,

$$\mathbf{\Phi}\mathbf{C}_3^x = \mathbf{0}. \tag{7.43}$$

We use again the variables z_{10}, z_{11}, z_{20}, and z_{21} in order to form systems of homogenous polynomial equations using the nullspace matrix $\mathbf{\Phi}$ and the structure

of $\mathbf{G}_3$. Combining the nullspace matrix $\mathbf{\Phi}$ with the structure in $\mathbf{G}_3$ leads to three different types of polynomial equations, which are in total the following six unique equations:

$$\phi_q^1 z_{10} z_{11} + \phi_q^2 z_{10} z_{21} + \phi_q^3 z_{20} z_{11} + \phi_q^4 z_{20} z_{21} = 0 \qquad \forall\, 1 \leq q \leq 2 \tag{7.44}$$

and

$$\begin{aligned}\phi_q^5 z_{10} z_{11} + \phi_q^6 z_{10} z_{21} + \phi_q^7 z_{20} z_{11} + \phi_q^8 z_{20} z_{21} + \phi_q^1 (z_{10} z_{10} + z_{11} z_{11}) \\ + (\phi_q^2 + \phi_q^3)(z_{10} z_{20} + z_{11} z_{21}) + \phi_q^4 (z_{20} z_{20} + z_{21} z_{21}) = 0 \qquad \forall\, 1 \leq q \leq 2\end{aligned} \tag{7.45}$$

and

$$\begin{aligned}(\phi_q^1 + \phi_q^9) z_{10} z_{11} + (\phi_q^3 + \phi_q^{11}) z_{10} z_{21} + (\phi_q^2 + \phi_q^{10}) z_{20} z_{11} \\ + (\phi_q^4 + \phi_q^{12}) z_{20} z_{21} + \phi_q^5 (z_{10} z_{10} + z_{11} z_{11}) + (\phi_q^6 + \phi_q^7)(z_{10} z_{20} + z_{11} z_{21}) \\ + \phi_q^8 (z_{20} z_{20} + z_{21} z_{21}) = 0 \qquad \forall\, 1 \leq q \leq 2.\end{aligned} \tag{7.46}$$

These equations lead to the following vector-matrix equation:

$$\begin{bmatrix} \phi_1^1 & \phi_1^2 & \phi_1^3 & \phi_1^4 & 0 & 0 & 0 \\ \phi_2^1 & \phi_2^2 & \phi_2^3 & \phi_2^4 & 0 & 0 & 0 \\ \phi_1^5 & \phi_1^6 & \phi_1^7 & \phi_1^8 & \phi_1^1 & \phi_1^2 + \phi_1^3 & \phi_1^4 \\ \phi_2^5 & \phi_2^6 & \phi_2^7 & \phi_2^8 & \phi_2^1 & \phi_2^2 + \phi_2^3 & \phi_1^4 \\ \phi_1^1 + \phi_1^9 & \phi_1^2 + \phi_1^{11} & \phi_1^3 + \phi_1^{10} & \phi_1^4 + \phi_1^{12} & \phi_1^5 & \phi_1^6 + \phi_1^7 & \phi_1^8 \\ \phi_2^1 + \phi_2^9 & \phi_2^2 + \phi_2^{11} & \phi_2^3 + \phi_2^{10} & \phi_2^4 + \phi_2^{12} & \phi_2^5 & \phi_2^6 + \phi_2^7 & \phi_2^8 \end{bmatrix} \begin{bmatrix} z_{10} z_{11} \\ z_{10} z_{21} \\ z_{20} z_{11} \\ z_{20} z_{21} \\ z_{10} z_{10} + z_{11} z_{11} \\ z_{10} z_{20} + z_{11} z_{21} \\ z_{20} z_{20} + z_{21} z_{21} \end{bmatrix} = \mathbf{0}. \tag{7.47}$$

The matrix on the left has size 6×7; however, from experiments with randomly generated impulse responses it follows that its rank is typically 5. More specifically, in our experiments we found that the two equations of third type, i.e., (7.46) are equivalent. In those cases, the matrix has a two dimensional right nullspace $\mathbf{Z}$, which corresponds exactly with the number of sources.

From (7.47) it follows that the mixing filter parameters are embedded in the right nullspace $\mathbf{Z}$ of $\mathbf{\Phi}$; however, the nullspace matrix $\mathbf{Z}$ does not necessarily contain the polynomial structure in terms of the variables z_{10}, z_{11}, z_{20}, and z_{21}. However, with a full rank matrix $\mathbf{V}$ we can create a matrix $\mathbf{K} = \mathbf{ZV}$ in canonical form that has the polynomial structure restored since we know that the following relation must hold for the mixing parameters:

$$\begin{bmatrix} a_1^1[0] a_1^1[1] & a_1^2[0] a_1^2[1] \\ a_1^1[0] a_2^1[1] & a_1^2[0] a_2^2[1] \\ a_2^1[0] a_1^1[1] & a_2^2[0] a_1^2[1] \\ a_2^1[0] a_2^1[1] & a_2^2[0] a_2^2[1] \\ a_1^1[0] a_1^1[0] + a_1^1[1] a_1^1[1] & a_1^2[0] a_1^2[0] + a_1^2[1] a_1^2[1] \\ a_1^1[0] a_2^1[0] + a_1^1[1] a_2^1[1] & a_1^2[0] a_2^2[0] + a_1^2[1] a_2^2[1] \\ a_2^1[0] a_2^1[0] + a_2^1[1] a_2^1[1] & a_2^2[0] a_2^2[0] + a_2^2[1] a_2^2[1] \end{bmatrix} = \mathbf{ZV} \tag{7.48}$$

In the previous chapter we have seen that this polynomial structure can be restored by formulating and solving a generalized eigenvalue decomposition (GEVD) problem. In (6.50) up to (6.52) an example is shown where eigenvectors are used to restore the polynomial structure in the nullspace matrix. Furthermore, in Section 6.5 we formulate polynomial optimization problems as GEVD problems in a more generalized way.

If we follow the steps from the previous chapter, then we are looking for matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ such that the following generalized eigenvalue decomposition holds:

$$\mathbf{S}_1\mathbf{Z}\mathbf{V}\mathbf{D} = \mathbf{S}_2\mathbf{Z}\mathbf{V} \tag{7.49}$$

where $\mathbf{V}$ is a full rank matrix containing generalized eigenvectors and $\mathbf{D}$ is a diagonal matrix containing generalized eigenvalues.

The first four rows of the nullspace matrix $\mathbf{Z}$ can be used to construct such a generalized eigenvalue decomposition. If we choose the following selection matrices:

$$\mathbf{S}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.50}$$

$$\mathbf{S}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{7.51}$$

then we know that for the elements on the diagonal of $\mathbf{D}$ the generalized eigenvalues $d_j^j = \frac{z_{10}^j}{z_{20}^j}$ form solutions. The generalized eigenvector matrix $\mathbf{V}$ that restores the nullspace matrix $\mathbf{Z}$ into canonical form, i.e., $\mathbf{K} = \mathbf{Z}\mathbf{V}$, can be computed from the following eigenvalue decomposition:

$$\mathbf{V}\mathbf{D}\left(\mathbf{V}\right)^{-1} = \left(\mathbf{S}_1\mathbf{Z}\right)^{-1}\mathbf{S}_2\mathbf{Z} \tag{7.52}$$

Now that we have restored the canonical structure in the nullspace matrix via $\mathbf{K} = \mathbf{Z}\mathbf{V}$ we are able to identify relations between mixing filter coefficients and eventually the actual mixing filter coefficients.

For the first four rows in each column of $\mathbf{K}$ we obtain the following structure:

$$\begin{bmatrix} z_{10}z_{11} \\ z_{10}z_{21} \\ z_{20}z_{11} \\ z_{20}z_{21} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}. \tag{7.53}$$

From this structure we are able to identify the following relations:

$$z_{20} = \frac{u_3}{u_1}z_{10} = \alpha_1 z_{10} \tag{7.54}$$

$$z_{21} = \frac{u_2}{u_1}z_{11} = \alpha_2 z_{11}. \tag{7.55}$$

The relation for z_{10} and z_{20} is also embedded in the corresponding generalized eigenvalue, i.e., $\alpha_1^j = \frac{1}{d_j^j}$. Note that the index j is related to the j'th column in the canonical

matrix $\mathbf{K}$ and eigenvalue matrix $\mathbf{D}$ and it does not reveal information about the source index. However, each eigenvalue and each corresponding column of $\mathbf{K}$ is related to the mixing parameters of only one of the sources. Consequently, when a priori information about the mixing parameters is available, one can use this information in order to select which columns of $\mathbf{K}$ are most interesting to identify since they correspond to interesting sources.

When one of the columns is selected, the relations in (7.54) and (7.55) combined with the second type of equations in (7.45) leads to the following constraints in vector-matrix format:

$$\begin{bmatrix} \phi_g^1 + (\phi_q^2 + \phi_q^3)\alpha_1 + \phi_q^4(\alpha_1)^2 \\ \phi_q^5 + \phi_q^7\alpha_1 + \phi_q^6\alpha_2 + \phi_q^8\alpha_1\alpha_2 \\ \phi_q^1 + (\phi_q^2 + \phi_q^3)\alpha_2 + \phi_q^4(\alpha_2)^2 \end{bmatrix}^T \begin{bmatrix} z_{10}z_{10} \\ z_{10}z_{11} \\ z_{11}z_{11} \end{bmatrix} = \mathbf{0}. \tag{7.56}$$

This matrix has a right nullspace $\mathbf{u}$ of size 3×1 and matches the following structure:

$$\begin{bmatrix} z_{10}z_{10} \\ z_{10}z_{11} \\ z_{11}z_{11} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{7.57}$$

from which it follows that $z_{11} = \frac{u_2}{u_1} z_{10}$.

Using (7.54) and (7.55) allows for identification of all coefficients. Note that due to the fact that each column in $\mathbf{K}$ is related to the impulse response coefficients of a single source, we do not suffer from local permutations as is the case in a frequency domain approach. Furthermore, note that by selecting the first filter coefficient z_{10} as 1, we apply a restriction on the solution which is a manifestation of the scaling indeterminacy that is obtained for instantaneous mixtures. Finally, it follows that in case the length of the impulse responses of the mixing system are known, identification of the mixing system becomes possible up to only an arbitrary gain, i.e., without additional filtering.

7.4 Discussion and conclusions

Convolutive mixing systems can be characterized by many parameters, such as sparsity and finiteness of impulse responses. Furthermore, we have shown that the source extraction problem can be formulated in different ways by means of source extraction filter objectives. Instead of only suppressing the interfering sources, which is performed with source separation algorithms, an additional requirement to perform deconvolution may be desired. In this chapter we focussed on convolutive mixtures that can be modeled with finite impulse response filters and we concentrated on the identification of the impulse response parameters that correspond to the desire source.

Our objective in this chapter was to apply our insights on informed source extraction for instantaneous mixtures onto the source extraction problem for convolutive problems. In order to do so, it was required to go back to the basics of the mixing system and study the structure in the sensor correlation matrices. We have shown for

a small mixing system consisting of length two mixing filters for two sources that it is possible to identify the mixing parameters. The mixing filters or impulse responses can be identified by formulating a system of second order homogenous polynomial equations from which the roots correspond to the mixing parameters. Furthermore, we have indicated that the approach allows for incorporating a priori information about the impulse responses in order to select and therefore identify only the impulse response coefficients that correspond to the desired source. The identified impulse response can be used to formulate a matched filter that extracts the desired source.

In order to obtain a widely applicable approach to design informed source extraction algorithms much additional research is required. For instance, we considered up to only two sources and two sensors. Furthermore, we considered only FIR filters of length two. Once appropriate systems of homogeneous polynomials are formulated, mathematical tools such as the ones presented in [73, 74] can help to compute the desired solutions. Next to generalizing the approach to long filters and more sensors and sources, it is also very useful to have methods for detecting the lengths of the impulse responses and the number of sources that are active. Finally, a challenging topic that remains is the translation of the noise-free region of support approach such that it can be applied for convolutive mixtures containing noise.

7.A Derivations of selected equations

Derivation of (7.12) The structure in the sensor correlation matrix is found as follows:

$$r^x_{i_1 i_2}[k] = \mathbb{E}\{x_{i_1}[n]x_{i_2}[n-k]\} \tag{7.58}$$

$$= \mathbb{E}\left\{\sum_{j_1=1}^{S}\sum_{p_1=0}^{P} h^{j_1}_{i_1}[p_1]s_{j_1}[n-p_1]\sum_{j_2=1}^{S}\sum_{p_2=0}^{P} h^{j_2}_{i_1}[p_2]s_{j_2}[n-p_2-k]\right\} \tag{7.59}$$

$$= \sum_{j_1=1}^{S}\sum_{j_2=1}^{S}\sum_{p_1=0}^{P}\sum_{p_2=0}^{P} h^{j_1}_{i_1}[p_1]h^{j_2}_{i_1}[p_2]\mathbb{E}\left\{s_{j_1}[n-p_1]s_{j_2}[n-p_2-k]\right\}. \tag{7.60}$$

Due to the assumption that the sources are indepdentent, or at least mutually uncorrelated, we have that $\mathbb{E}\{s_{j_1}[n]s_{j_2}[n-k]\} = 0$ for all n, k when $j_1 \neq j_2$. Taking this assumption into account results into the following structure in the sensor correlation data:

$$r^x_{i_1 i_2}[k] = \sum_{j=1}^{S}\sum_{p_1=0}^{P}\sum_{p_2=0}^{P} h^{j}_{i_1}[p_1]h^{j}_{i_1}[p_2]r^s_{jj}[k+p_2-p_1]. \tag{7.61}$$

Example matrix G_3 from (7.19) for $N = 3$ consecutive lags Here we present an example correlated impulse response matrix G_N for $N = 3$ to illustrate the structure in sensor correlation matrices when consecutive lags are considered. The struc-

ture of $\mathbf{G}_3$ is as follows:

$$\begin{bmatrix}
h_1^1[0]h_1^1[1] & h_1^1[0]h_1^1[0]+h_1^1[1]h_1^1[1] & h_1^1[1]h_1^1[0] & 0 & 0 \\
h_1^1[0]h_2^1[1] & h_1^1[0]h_2^1[0]+h_1^1[1]h_2^1[1] & h_1^1[1]h_2^1[0] & 0 & 0 \\
h_2^1[0]h_1^1[1] & h_2^1[0]h_1^1[0]+h_2^1[1]h_1^1[1] & h_2^1[1]h_1^1[0] & 0 & 0 \\
h_2^1[0]h_2^1[1] & h_2^1[0]h_2^1[0]+h_2^1[1]h_2^1[1] & h_2^1[1]h_2^1[0] & 0 & 0 \\
0 & h_1^1[0]h_1^1[1] & h_1^1[0]h_1^1[0]+h_1^1[1]h_1^1[1] & h_1^1[1]h_1^1[0] & 0 \\
0 & h_1^1[0]h_2^1[1] & h_1^1[0]h_2^1[0]+h_1^1[1]h_2^1[1] & h_1^1[1]h_2^1[0] & 0 \\
0 & h_2^1[0]h_1^1[1] & h_2^1[0]h_1^1[0]+h_2^1[1]h_1^1[1] & h_2^1[1]h_1^1[0] & 0 \\
0 & h_2^1[0]h_2^1[1] & h_2^1[0]h_2^1[0]+h_2^1[1]h_2^1[1] & h_2^1[1]h_2^1[0] & 0 \\
0 & 0 & h_1^1[0]h_1^1[1] & h_1^1[0]h_1^1[0]+h_1^1[1]h_1^1[1] & h_1^1[1]h_1^1[0] \\
0 & 0 & h_1^1[0]h_2^1[1] & h_1^1[0]h_2^1[0]+h_1^1[1]h_2^1[1] & h_1^1[1]h_2^1[0] \\
0 & 0 & h_2^1[0]h_1^1[1] & h_2^1[0]h_1^1[0]+h_2^1[1]h_1^1[1] & h_2^1[1]h_1^1[0] \\
0 & 0 & h_2^1[0]h_2^1[1] & h_2^1[0]h_2^1[0]+h_2^1[1]h_2^1[1] & h_2^1[1]h_2^1[0]
\end{bmatrix}.$$

8

Conclusions and future research

In this thesis we have developed methods for designing algorithms that extract one desired source signal from an instantaneous mixture of multiple sources where only the mixtures are observed. Extraction of a source signal from the observations is accomplished by exploiting assumptions on the statistical features of the source signals and selection of the desired source signal is accomplished by incorporating additional a priori information directly into the source extraction algorithm. This additional information can be information about either the mixing parameters or the statistical features of the source signals or both. Algorithms that use this information to ensure immediate extraction of only the desired source are called informed source extraction algorithms.

In the literature only two generic methods for informed source extraction were found, both having their own limitations and disadvantages. By studying the structure in the second order statistical features of the sensor signals and by taking widely used assumptions on the second order temporal structure of source signals into account we obtained insight in the informed source extraction problem. Based on this insight we showed that for a large variety of applications the informed source extraction problem can be formulated and solved as smallest or largest eigenvalue problems and as polynomial optimization problems. Finally, we made a first attempt to apply our subspace based approach to convolutive mixtures. In this chapter we summarize the results and highlight the contributions of this work. Furthermore, we present suggestions for future work.

8.1 Introduction

In a source extraction problem, a number of mutually statistically independent source signals are mixed by a MIMO instantaneous mixing system and only the mixtures are observed. From these observations one particular source signal has to be recovered or extracted. Data driven methods that solve closely related problems by exploiting only the observed mixtures and certain assumptions on statistical properties of the sources, such as system identification and source separation, are called blind signal processing (BSP) methods in the literature. A result of this limited set of assumptions is that BSP algorithms suffer from a permutation indeterminacy, which entails that the sources are obtained in an arbitrary order. In order to solve the source extraction problem, a BSP algorithm can be used in a multi-stage approach. Such a multistage source extraction algorithm consists of a BSP algorithm, which extracts randomly one the sources or separates all sources, followed by a classifier that selects the desired source. However, such multi-stage algorithms are naturally inefficient and utilize more resources than strictly required. Furthermore, methods based on BSP algorithms can only be used in batch mode in order to remove undesired extracted sources with a deflation algorithm. By incorporating the a priori information from the classifier directly into the source extraction algorithm, more efficient, informed source extraction algorithms can be obtained.

Only two generic methods for informed source extraction were found in the literature. In the first method non-Gaussianity of the source signals is combined with a priori information and exploited in a constrained ICA framework in order to guarantee extraction of only the desired source. The a priori information about the desired source is assumed to be available in the form of reference signals; and extraction of the desired source is guaranteed by requiring that a distance between this reference signal and the extracted signal is below a threshold value. If the threshold value cannot be determined based on a priori information, an iterative approach is proposed where the threshold is increased until the source that is closest to the reference signal is extracted. In the second method linear prediction based source extraction algorithms are used in order to extract the desired source. In such an algorithm a linear prediction filter is applied to the output of an extraction filter. By adapting the extraction filter parameters such that the linear prediction filter output has minimum energy, the source with the smallest prediction error is extracted. For this algorithm, immediate extraction of the desired source can be guaranteed by designing the linear prediction filters based on a priori information about the second order temporal structure of the sources.

The two informed source extraction methods found in the literature have their own limitations and disadvantages. The first method relies on higher than second order statistical features, which can be difficult to estimate robustly in practice. Furthermore, the proposed algorithm depends on a threshold value, which can be difficult to determine a priori. If the threshold value cannot be determined a priori, then the algorithm looses efficiency since a proper threshold value has to be determined using an iterative approach. Finally, in many applications it may be difficult to obtain a proper reference signal. The algorithms based on linear prediction exploit second

order statistical features, which are in general easier to estimate than higher order statistical features. However, linear prediction filters contain information about only the autocorrelation or the power spectrum of the source signals. Consequently, these methods do not provide a solution in case a priori information about the mixing parameters is available. Therefore, this approach is restricted and can only be used in a limited number of applications. A generic approach for designing source extraction algorithms that are flexible in their use of a priori information, efficient in their use of resources, and efficient in finding the desired source extraction filters is missing.

8.2 Results and contributions

The lack of suitable methods for designing informed source extraction algorithms was our motivation for studying the informed source extraction problem. In this work we have shown that the informed source extraction problem can be formulated and solved for a large variety of applications as smallest or largest eigenvalue problems and as polynomial optimization problems. For such problems, efficient and robust numerical algorithms are available.

In Chapter 3 we considered the informed source extraction problem for real-valued instantaneous mixtures. By studying the second order temporal structure of the sensor signals, i.e., auto- and crosscorrelation data, we provided insight in the informed source extraction problem. By combining specifically arranged matrices that contain the sensor correlation data we were able to formulate the informed source extraction problem as an eigenvalue decomposition problem and a priori information can be incorporated by taking specifically weighted combinations of the sensor correlation matrices. The eigenvalues contain information about the mismatch in a priori information w.r.t. the information that is represented while left and right eigenvectors correspond to two different types of source extraction filters. The source extraction filters can be identified in case a priori an estimate of the mixing parameters and/or the autocorrelation function of the desired source is available. For one of the source extraction filters the focus is on suppression of all interfering source signals while the objective for the other filter is to minimize the desired signal to interference plus noise ratio. This latter filter is especially useful in case a mixing system is ill-conditioned. A strong advantage of the proposed approach is that the identified extraction filter coefficients are not biased by an error or mismatch in the available a priori information, as long as the a priori information ensures that the extraction filters for the desired source are selected.

In Chapter 4 we extended the work for real-valued instantaneous mixtures from Chapter 3 to work also for complex-valued instantaneous mixtures. Extending the informed source extraction principles for real-valued mixtures to complex-valued mixtures is non-trivial and requires a different organization of the sensor correlation matrices. Furthermore, we presented a measure that can be used to indicate the conditioning of a mixing matrix related to the informed source extraction problem. For source extraction problems, the mixing system is only considered to be ill-conditioned if the mixing parameters corresponding to the desired, extracted source are relatively

close to the mixing parameters corresponding to any other source. In addition, we considered a parametrization of the mixing system that is very common in the fields of telecommunication and array signal processing and we presented tools in the form of selection patterns and selection beamformers that help the design of informed source extraction algorithms. These tools can be used in order to ensure immediate extraction of the desired source based on the available a priori information. Finally, through simulations we have shown that minimimum variance distortionless response (MVDR) and linearly constrained minimum variance (LCMV) filters based on the extraction algorithms are much more robust w.r.t. errors in the a priori information than an MVDR filter based on the same information. The only condition is that the smallest eigenvalue has to correspond to the desired source.

In Chapter 5 the source extraction problem was considered for both real-valued and complex-valued mixtures in case reference signals are available. We derived conditions that have to hold for these reference signals such that immediate extraction of the desired source is guaranteed. Furthermore, we presented different design procedures for creating reference signals in such a way that immediate extraction of the desired source can be guaranteed. These design procedures describe how to choose temporal or spatial filters based on a priori information about either the autocorrelation function or the mixing parameters of the sources. The reference signals can be derived from the regular observations or even obtained from separate additional systems such as a dedicated reference system. Finally, we have shown that using separate sensors for obtaining reference signals can be beneficial in the presence of noise.

In Chapter 6 we formulated the informed source extraction problem as a polynomial optimization problem. In the optimization problem the polynomial equality constraints limit the feasible set to source extraction filters, while a priori information about the mixing parameters or the autocorrelation functions of the sources can be used to design fractional objective functions that ensure identification of only the filter coefficients that extract the desired source. We also showed that through the use of subspace techniques such polynomial optimization problems can be formulated and solved as generalized eigenvalue decomposition problems. We showed that source extraction can be accomplished for certain under-determined mixtures with more source than sensors. Furthermore, we indicated that the approach can be extended in a natural way towards the use of higher order statistical features instead of second order statistical features. Finally, we have demonstrated that the polynomial optimization problem formulation is a more natural and efficient approach than the ICA with reference based approach that depends on a threshold parameter.

Finally, in Chapter 7 we discussed the convolutive mixing scenario and applied our subspace based approach to a specific convolutive mixing system consisting of short finite impulse response filters. We formulated systems of homogenous polynomial equations from which the mixing parameters can be identified up to a global scaling factor. Although the results were obtained for a very specific mixing scenario, we consider them a first step towards the design of application specific informed source extraction algorithms for convolutive mixtures.

8.3 Future work

The work in this thesis forms a basis for the design of informed source extraction algorithms for a large variety of practical applications, allowing for the use of second order temporal structure, exploiting different types of a priori information, and identifying two kinds of source extraction filters. Although the algorithms presented in this work can be implemented using standard numerical algorithms, derivation of real-time, efficient, and robust source extraction algorithms from the results in this thesis is considered a very relevant step for bringing the insights, design techniques, and algorithms from this thesis into practice. In the remainder of this section we focus on two relevant topics that are closely related to the work in this thesis and that deserve further research.

First, the methods for performing informed source extraction presented in this thesis assumed that the sources are more or less continuously active. However, in many applications signals are temporally sparse, i.e., the signals are only active for a limited amount of time. The currently proposed methods do not provide an optimal solution for such signals. On the one hand, temporal sparsity leads to rank deficient correlation matrices such that proposed source extraction filter identification algorithms fail. On the other hand, for temporally sparse source signals the optimal source extraction filter parameters become time dependent. In some use cases the extraction filters should remain constant in order to suppress interfering sources at an onset, while in other use cases the extraction filters should adapt to the mixture that is currently at hand, which can be achieved for example with adaptive filters. Since the desired configuration of the source extraction filter is application dependent, algorithms should be developed that allow for the identification of source extraction filters for both cases. Taking this temporal sparseness of sources properly into account, allows for solving the informed source extraction problem for many more applications and is therefore considered a relevant topic for further research.

Second, the work in this thesis was initiated in order to translate the results for informed source extraction from instantaneous mixtures into solutions for the more complex convolutive mixing model. By means of a time-frequency analysis, the convolutive mixing problem is typically translated into multiple independent instantaneous mixing problems per frequency band. In order to solve the source extraction problem for such a system, the permutation problem has to be dealt with in each frequency band. Furthermore, the resolution of the time-frequency analysis should be sufficiently high. In Chapter 6 we have already shown that the formulation of the source extraction problem as a polynomial optimization problem can be used to solve source extraction problems for mixtures that cannot be dealt with when focussing on solutions only involving linear algebra. We believe that by extending the work from Chapter 6 and Chapter 7 and by formulating systems of polynomial equations, source separation and extraction problems can be solved for a wider range of convolutive mixtures.

8.4 Conclusions

In this thesis we provided insight in the source extraction problem and we provided tools that support the design and development of application specific informed source extraction algorithms for a wide variety of applications. The source extraction problem has been considered for various mixing models, different types of a priori information, and multiple objectives for the source extraction filter. We presented conditions that have to hold on available a priori information and techniques for the design of source extraction algorithms such that immediate extraction of the desired source from instantaneous mixtures is guaranteed. Furthermore, we made a first attempt to translate the approach for instantaneous mixtures towards convolutive mixtures. Finally, we discussed the results from this work and provided two relevant topics for future research that are closely related to the work in this thesis.

Bibliography

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393 – 422, 2002.

[2] Jennifer Yick, Biswanath Mukherjee, and Dipak Ghosal. Wireless sensor network survey. *Computer Networks*, 52(12):2292 – 2330, 2008.

[3] Alexander Bertrand. *Signal Processing Algorithms for Wireless Acoustic Sensor Networks*. PhD thesis, Katholieke Universiteit Leuven, Leuven, Belgium, May 2011.

[4] Don H. Johnson and Dan E. Dudgeon. *Array signal processing : concepts and techniques*. P T R Prentice Hall Englewood Cliffs, NJ, 1993.

[5] Harry L. van Trees. *Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing*. John Wiley & Sons, 2002.

[6] Jacob Benesty, Jingdong Chen, and Yiteng Huang. *Microhpone array signal processing*, volume 1 of *Springer topics in signal processing*. Springer, 2008.

[7] Aapo Hyvrinen, Juha Karhunen, and Erkki Oja. *Independent component analysis*. John Wiley & Sons, 2001.

[8] Andrzej Cichocki and S. Amari. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[9] M. S. Pedersen, J. Larsen, U. Kjems, and L. C. Parra. A survey of convolutive blind source separation methods. In *Springer Handbook of Speech Processing*. Springer Press, nov 2007.

[10] S. Makino, T.-W. Lee, and H. *Blind Speech Separation*. Springer, 2007.

[11] Pierre Comon and Christian Jutten. *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. Academic Press, 2010.

[12] S. Doclo and M. Moonen. GSVD-based optimal filtering for single and multimicrophone speech enhancement. *Signal Processing, IEEE Transactions on*, 50(9):2230 – 2244, sep 2002.

[13] C.J. James and O.J. Gibson. Temporally constrained ICA: an application to artifact rejection in electromagnetic brain signal analysis. *Biomedical Engineering, IEEE Transactions on*, 50(9):1108–1116, 2003.

[14] Wei Liu, D.P. Mandic, and A. Cichocki. A class of novel blind source extraction algorithms based on a linear predictor. *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, 4:3599–3602 Vol. 4, May 2005.

[15] Wei Liu, D.P. Mandic, and A. Cichocki. Blind source extraction of instantaneous noisy mixtures using a linear predictor. In *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*, pages 4202–4206, 0-0 2006.

[16] Xi-Lin Li. Sequential blind extraction adopting second-order statistics. *IEEE Signal Process. Lett.*, 14(1):58–61, 2007.

[17] R. Dubroca, C. De Luigi, Marc Castella, and E. Moreau. A general algebraic algorithm for blind extraction of one source in a mimo convolutive mixture. *Signal Processing, IEEE Transactions on*, 58(5):2484–2493, 2010.

[18] R. Phlypo, V. Zarzoso, and I. Lemahieu. Source extraction by maximizing the variance in the conditional distribution tails. *Signal Processing, IEEE Transactions on*, 58(1):305–316, 2010.

[19] S. Markovich, S. Gannot, and I. Cohen. Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals. *Audio, Speech, and Language Processing, IEEE Transactions on*, 17(6):1071–1086, 2009.

[20] S. Applebaum and D. Chapman. Adaptive arrays with main beam constraints. *Antennas and Propagation, IEEE Transactions on*, 24(5):650 – 662, sep. 1976.

[21] L. Griffiths and C. Jim. An alternative approach to linearly constrained adaptive beamforming. *Antennas and Propagation, IEEE Transactions on*, 30(1):27 – 34, jan. 1982.

[22] O. Hoshuyama, A. Sugiyama, and A. Hirano. A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters. *Signal Processing, IEEE Transactions on*, 47(10):2677–2684, Oct 1999.

[23] L. C. Parra and C. V. Alvino. Geometric source separation: merging convolutive source separation with geometric beamforming. *IEEE Transactions on Speech and Audio Processing*, 10(6):352–362, Sept 2002.

[24] Affan H. Khan, Maja Taseska, and Emanuël A. P. Habets. A geometrically constrained independent vector analysis algorithm for online source extraction. In Emmanuel Vincent, Arie Yeredor, Zbyněk Koldovský, and Petr Tichavský, editors, *Latent Variable Analysis and Signal Separation*, pages 396–403, Cham, 2015. Springer International Publishing.

[25] T. Kim, I. Lee, and T. Lee. Independent vector analysis: Definition and algorithms. In *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, pages 1393–1396, Oct 2006.

[26] T. Kim, H. T. Attias, S. Lee, and T. Lee. Blind source separation exploiting

higher-order frequency dependencies. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(1):70–79, Jan 2007.

[27] Maja Taseska and Emanuël A. P. Habets. Doa-informed source extraction in the presence of competing talkers and background noise. *EURASIP Journal on Advances in Signal Processing*, 2017(1):60, Aug 2017.

[28] M. Taseska and E. A. P. Habets. Nonstationary noise psd matrix estimation for multichannel blind speech extraction. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(11):2223–2236, Nov 2017.

[29] Boaz Schwartz, Sharon Gannot, Emanuel A. P. Habets, Boaz Schwartz, Sharon Gannot, and Emanuel A. P. Habets. Two model-based em algorithms for blind source separation in noisy environments. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, 25(11):2209–2222, November 2017.

[30] M. Souden, S. Araki, K. Kinoshita, T. Nakatani, and H. Sawada. A multichannel mmse-based framework for speech source separation and noise reduction. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9):1913–1928, Sept 2013.

[31] Allan Kardec Barros and Andrzej Cichocki. Extraction of specific signals with temporal structure. *Neural Computation*, 13(9):1995–2003, 2001.

[32] Pierre Comon. Independent component analysis, a new concept? *Signal Process.*, 36(3):287–314, April 1994.

[33] Marc Castella, S. Rhioui, E. Moreau, and J. Pesquet. Quadratic higher order criteria for iterative blind separation of a mimo convolutive mixture of sources. *Signal Processing, IEEE Transactions on*, 55(1):218–232, 2007.

[34] Wei Lu and J.C. Rajapakse. Approach and applications of constrained ICA. *Neural Networks, IEEE Transactions on*, 16(1):203–212, 2005.

[35] Simon Haykin, editor. *Unsupervised adaptive filtering - volume 1 blind source separation*. John Wiley & Sons, New York, NY, USA, 2000.

[36] J. van de Laar. *MIMO Instantaneous Blind Identification and Separation based on Arbitrary Order Temporal Structure in the Data*. PhD thesis, Technische Universiteit Eindhoven, 2007.

[37] H. Buchner, R. Aichner, and W. Kellermann. A generalization of blind source separation algorithms for convolutive mixtures based on second order statistics. *IEEE Trans on Speech and Audio Processing*, 13(1):120–134, January 2005.

[38] R. Aichner, H. Buchner, and W. Kellermann. Exploiting narrowband efficiency for broadband convolutive blind source separation. *EURASIP Journal on Advances in Signal Processing*, 2007:1–9, September 2006.

[39] Jean-Franois Cardoso and Antoine Souloumiac. Blind beamforming for non gaussian signals. *IEE Proceedings-F*, 140:362–370, 1993.

[40] A. Hyvrinen. Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 10(3):626–634, May 1999.

[41] Wei Lu and Jagath C. Rajapakse. ICA with reference. In *Proceedings of 3rd International Conference on Independent Component Analysis and Blind Signal Separation (ICA01)*, pages 120–125, 2001.

[42] Wei Lu and Jagath C. Rajapakse. ICA with reference. *Neurocomputing*, 69(1618):2244 – 2257, 2006.

[43] Dimitri P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods (Optimization and Neural Computation Series)*. Athena Scientific, 1 edition, 1996.

[44] L. Tong, V.C. Soon, Y. Huang, and R. Liu. Amuse: a new blind identification algorithm. In *Circuits and Systems, 1990., IEEE International Symposium on*, pages 1784–1787 vol.3, May 1990.

[45] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *Signal Processing, IEEE Transactions on*, 45(2):434 –444, feb 1997.

[46] Gene H. Golub and Charles F. Van Loan. *Matrix Computations (3rd Ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[47] L. Molgedey and H. G. Schuster. Separation of a mixture of independent signals using time delayed correlations. *Phys. Rev. Lett.*, 72:3634–3637, Jun 1994.

[48] R. A. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis. *UCLA Working Papers in Phonetics*, 16(1):84, 1970.

[49] Adel Belouchrani and Andrzej Cichocki. Robust whitening procedure in blind source separation context. *Electronics Letters*, 36:2050–2051, 2000.

[50] Dinh Tuan Pham. Joint approximate diagonalization of positive definite hermitian matrices. *SIAM Journal on Matrix Analysis and Applications*, 22:1136–1152, 2001.

[51] Wei Liu, D.P. Mandic, and A. Cichocki. Blind second-order source extraction of instantaneous noisy mixtures. *Circuits and Systems II: Express Briefs, IEEE Transactions on*, 53(9):931–935, Sept. 2006.

[52] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. Robust blind extraction of a signal with the best match to a prescribed autocorrelation. In *Proceedings of EUSIPCO 2010*, 2010.

[53] S. Javidi, B. Jelfs, and D. P. Mandic. Blind extraction of noncircular complex signals using a widely linear predictor. In *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, pages 501–504, Aug 2009.

[54] S. Javidi, D. P. Mandic, and A. Cichocki. Complex blind source extraction from noisy mixtures using second-order statistics. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(7):1404–1416, July 2010.

[55] CG Khatri and C Radhakrishna Rao. Solutions to some functional equations and their applications to characterization of probability distributions. *Sankhyā:*

The Indian Journal of Statistics, Series A, pages 167–180, 1968.

[56] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. A single stage approach to blind source extraction based on second order statistics. *Signal Processing*, 93(2):432 – 444, 2013.

[57] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. Blind extraction algorithm with direct desired signal selection. In *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2010 IEEE*, pages 13 –16, October 2010.

[58] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. Overdetermined blind source extraction exploiting a generalized sidelobe canceller structure. In *Proceedings of the 12th International Workshop on Acoustic Echo and Noise Control 2010*, 2010.

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

[60] Zhi-Lin Zhang and Zhang Yi. Robust extraction of specific signals with temporal structure. *Neurocomputing*, 69(7-9):888–893, March 2006.

[61] Brian Bloemendal, Jakob van de Laar, and Piet Sommen. Design of signal extraction algorithms based on second order statistics exploiting beamforming techniques. *Signal Processing*, 96, Part B(0):125 – 137, 2014.

[62] S. Gorlow and S. Marchand. Informed audio source separation using linearly constrained spatial filters. *Audio, Speech, and Language Processing, IEEE Transactions on*, 21(1):3 –13, jan. 2013.

[63] Kevin H. Knuth. Informed source separation: A bayesian tutorial. In *Proceedings of the 13th European Signal Processing Conference (EUSIPCO 2005*, 2005.

[64] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. Beamformer design exploiting blind source extraction techniques. In *Proceedings of the 2012 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 2589–2592, 2012.

[65] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. Blind source extraction for a combined fixed and wireless sensor network. In *Proceedings of the 20th European Signal Processing Conference (EUSIPCO 2012)*, 2012.

[66] B.B.A.J. Bloemendal, J. van de Laar, and P.C.W. Sommen. Multichannel signal enhancement using a remote wireless microphone. In *Proceedings of the International Workshop on Acoustic Signal Enhancement 2012*, 2012.

[67] K. Scharnhorst. Angles in complex vector spaces. *Acta Applicandae Mathematicae*, 69:95–103, 2001. 10.1023/A:1012692601098.

[68] B.L.R. De Moor. Daisy: Database for the identification of systems, department of electrical engineering, esat/sista, k.u.leuven, belgium, http://homes.esat.kuleuven.be/~smc/daisy/, Visited: January 2013.

[69] E. Habets. Room impulse response (rir) generator. http://home.tiscali.nl/

ehabets/rir_generator.html, July 2006.

[70] J. van de Laar, M. Moonen, and P. Sommen. Mimo instantaneous blind identification based on second-order temporal structure. *Signal Processing, IEEE Transactions on*, 56(9):4354–4364, Sept. 2008.

[71] Jan Verschelde. *Homotopy continuation methods for solving polynomial systems.* PhD thesis, Department of Computer Science, K.U.Leuven, Leuven, Belgium, May 1996. Haegemans, Anny and Van de Vel, Hugo (supervisors).

[72] Hans J. Stetter. *Numerical polynomial algebra.* SIAM, 2004.

[73] Philippe Dreesen, Kim Batselier, and Bart De Moor. Back to the roots: Polynomial system solving, linear algebra, systems theory. *Proc 16th IFAC Symposium on System Identification (SYSID 2012)*, pages 1203–1208, 2012.

[74] Philippe Dreesen. *Back to the Roots: Polyonomial System Solving Using Linear Algebra.* PhD thesis, KU Leuven, September 2013.

[75] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, January 1968.

[76] Emmanuel Vincent, Nancy Bertin, Rémi Gribonval, and Frédéric Bimbot. From blind to guided audio source separation: How models and side information can improve the separation of sound. *IEEE Signal Processing Magazine*, 31(3):107–115, May 2014.

[77] H. Attias. New em algorithms for source separation and deconvolution with a microphone array. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, volume 5, pages V–297–300 vol.5, April 2003.

Acknowledgments

It has taken ten years from the start of my PhD project till the defense of this thesis. During the first four years I worked on my PhD research at the Eindhoven University of Technology. After this period I started working at Philips Research and continued writing this thesis in my own time. In these ten years many people have directly or indirectly contributed to the work in this thesis.

First of all I would like to thank my co-promoter Piet Sommen for his support, patience and confidence that I would ultimately finish my thesis. We have had many discussions where you learned me to always look for the simplest representation, such as a drawing, that grasps the essence of a problem. Typically, these simple representations lead to new valuable insights and this approach is still very valuable to me. I would also like to thank my co-promoter Jakob van de Laar, for having fruitful contributions when discussing my progress and next steps and for extensively reviewing my papers and this thesis. Furthermore, I thank my promoter Jan Bergmans, as well as Piet Sommen and Kees Janse, for giving me the opportunity to work on my PhD project. In addition, I would like to extend my gratitude to the members of my defense committee, Marc Moonen from KU Leuven, Emmanuël Habets from International Audio Laboratories Erlangen, Kees Janse from Philips Research Eindhoven, and Siep Weiland from Eindhoven University of Technology, for participating in the committee and reading and reviewing my thesis.

My PhD research has been conducted at the Signal Processing Systems (SPS) group at the Eindhoven University of Technology. I would like to thank my colleagues from the SPS group for providing a pleasant work environment and I particularly appreciate our collaborations in various teaching tasks. Furthermore, I've enjoyed being together with a large group of fellow PhD candidates. At the university we worked together by reviewing each other's work and by sharing our thoughts, questions, and challenges during whiteboard sessions and joint lunches. I've also enjoyed being together with you during dinner, movie and, board game evenings where everybody was welcome. Thank you, Michiel, Aline, Joep, Maarten, Iman, Hooman, Agis, Paola, Lily, Marija, Máté, and many more. Many thanks go to my colleagues at Philips Research, who let me experience a very warm welcome when I joined Philips.

I really appreciate working together with you and the attractive working atmosphere has definitely extended the time span for finishing my thesis.

Also people from outside my professional environment have had a key role in finishing this work. Thank you to all my friends and family for your interest and patience. It is impossible to address everyone personally, but I would like to sincerely thank you all. A very special thank you is in place for my parents, Joke en Wim, for their support and patience, for example during dinner where I regularly left the table to write down some idea, but most importantly for providing the opportunity to develop myself into who I am now. And finally, thank you Chantal for your support and understanding, and thank you for having a beautiful life together.

Curriculum Vitae

Brian Bloemendal was born in Deurne, The Netherlands on October 24, 1983. He grew up in Helenaveen, which is a small village near De Peel, a Dutch national park known for its peat. He started his study in 2002 and received both his B.Sc and M.Sc degree in Electrical Engineering from Eindhoven University of Technology (TU/e), The Netherlands. In 2008 he graduated within the Signal Processing Systems group on blind source extraction.

From 2008 to 2013 he continued his graduation work at the TU/e in the form of a PhD project of which the results are presented in this dissertation. In 2013 he started as a Scientist at Philips Research, working primarily on speech enhancement and interventional ultrasound topics, but also shares his knowledge and experience on informed source extraction to assist research on other topics. In 2018 he continued as a Senior Scientist at Philips Research. His main research interest is real-time digital signal processing applied to raw signals from one or more sensors such as microphones or ultrasonic transducers. Research activities related to this field concern real-time, embedded signal processing and statistical signal processing, including blind signal processing, adaptive filters, array signal processing, and compressive sensing.