ARRAY SIGNAL PROCESSING WITH ALPHA-STABLE DISTRIBUTIONS

by

Panagiotis Tsakalides

A Dissertation Presented to the

FACULTY OF THE GRADUATE SCHOOL

UNIVERSITY OF SOUTHERN CALIFORNIA

In Partial Fulfillment of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

(Electrical Engineering)

December 1995

Acknowledgments

While the completion of a doctorate is looked upon as an individual accomplishment, it is rarely achieved without the help of a significant number of people. Such is the case here, for it has been the friendship, guidance, and support of many people that have helped make the last five years some of the most exciting and stimulating of my life and have helped me through the times when things got tough.

Over the last four years, Professor Max Nikias has been a source of inspiration and stimulation, and has provided the constant encouragement and critical evaluation which I needed. His guidance and friendship have helped me grow both professionally and personally. For his concern, caring and effort, I am eternally grateful.

I would also like to thank Professor Irving Reed and Professor Alan Nadel for participating in my dissertation committee as well as for providing valuable academic and career guidance. Irving Reed has taken considerable time to provide constant feedback on radar issues and to hear my recollections of Greek mythology readings. I am also grateful to Professor Robert Scholtz and Professor Zhen Zhang for their participation in my guidance committee.

All my colleagues in the Signal and Image Processing Institute (SIPI), past and present, have been instrumental in making the last five years fun and challenging. Particular among these have been Professor George Tsihrintzis and Ioannis Katsavounidis with whom I shared a common language over lunch and coffee breaks. Special thanks to George Tsihrintzis for being a worthy and occasionally dangerous opponent on the racketball court. I would also like to thank my officemates who came from the four corners of the world to Los Angeles to share the roller coaster ride of graduate studies with me: Sam Heidari who taught me my first steps in racketball, Dae Shin for sharing his cool LaTeX bits of knowledge with me, Jun Shen and Jack Ma the American football aficionados, and Mehmet Gurelli. I would also like to acknowledge the help and support of the staff at SIPI, Linda Varilla and Mercedes Morente. I am especially grateful to Mayitta Penoliar and Dawn Ernst at the Center for Research on Applied Signal Processing for their constant

help and support. My thanks go to Allan Weber and Toy Mayeda for keeping the computer facilities running and for answering so many of my computer-related questions.

With Ellen Strain, I shared the last five years, each of us helping the other over the hurdles of graduate studies. Outside of school, we discovered the sun-kissed city of Angels together and kept back the threatening clouds of noir.

The unconditional love and understanding of my family has been the pillar of support which has kept me upright over the years. Without them none of this would have happened. My brother Costas and my sister Effie have been my role models throughout my adolescent years as well as constant sources of love. No amount of gratitude can express the debt I owe to my parents for they have always believed in me and provided for me. In all my life I have just been trying to follow their example of hard work, determination, and honesty. With love, this dissertation is dedicated to them.

Contents

List of Tables

List of Figures

Abstract

The importance of extending the statistical signal processing methodology to the so-called alpha-stable framework is apparent. First, scientists and engineers have started to appreciate alpha-spectra and the elegant scaling and self-similarity properties of stable distributions. Additionally, real life applications exist in which impulsive channels tend to produce large-amplitude, short-duration interferences more frequently than Gaussian channels do. The stable law has been shown to successfully model noise over certain impulsive channels. In this dissertation, we propose new robust techniques for source detection and localization in the presence of signals and/or noise modeled as complex isotropic stable processes.

First, we present optimal, maximum likelihood-based approaches to the direction-of-arrival estimation problem and we introduce the Cauchy Beamformer. We show that the Cauchy Beamformer provides better bearing estimates than the Gaussian Beamformer in a wide range of impulsive noise environments and for very low signal-to-noise ratios. In addition, we calculate the Cramér-Rao bound on the estimation error covariance for the case of deterministic incoming signals retrieved in the presence of additive complex Cauchy noise.

In the second part of the dissertation, we develop subspace methods based on fractional lower-order statistics, for applications where reduced computational cost is a crucial design parameter. Since symmetric alpha-stable processes do not possess finite pth order moments for $p \geq \alpha$, traditional subspace techniques employing second- and higher-order moments cannot be applied in impulsive noise environments modeled under the stable law. Instead, we use properties of fractional lower-order moments and covariations. We define the spatial covariation matrix of the observation vector process and employ subspace-based bearing estimation techniques to the sample covariation matrix resulting in improved direction-of-arrival estimates in impulsive noise environments. We name the introduced technique Robust Covariation-Based Multiple Signal Classification or ROC-MUSIC. In addition, we present consistent estimators for the marginals of the covariation matrix and we study their asymptotic performance through both theory and simulations.

Finally, in the last part of the dissertation, we investigate the problem of localizing wideband sources in the presence of noise modeled as a complex isotropic stable process. We consider the frequency-domain representation of the sensor outputs and show that the spectral density of complex stable processes plays a role in array processing problems analogous to that played by the power spectral density of second-order processes.

Chapter 1

Introduction

Statistical array processing based on the linear theory of random processes with finite second-order moments has been the focus of considerable academic research. Critical problems such as high-resolution direction finding, null- and beam-steering, and detection of the number of sources illuminating an array of sensors have been studied under the assumption of a Gaussian or second-order model. Many different classes of methods, compromising optimality for the sake of computational efficiency, have been proposed under the aforementioned statistical framework [34].

Looking toward real world applications, we are interested in developing array processing methods for a larger class of random processes which include the Gaussian processes as special elements. The availability of such methods would make it possible to operate in environments which, while sharing many characteristics also differ from Gaussian environments in significant ways.

The class of stable distributions, a natural generalization of the Gaussian distribution, has some important characteristics which make it very attractive for modeling impulsive signals. Stable processes satisfy the stability property which states that linear combinations of jointly stable variables are indeed stable. They arise as limiting processes of sums of independent, identically-distributed random variables via the generalized central limit theorem. That is, if the observed randomness is the result of many cumulative effects, and these effects follow a heavy-tailed distribution, then a stable model may be appropriate. Stable processes are described by their characteristic exponent α, taking values $0 < \alpha \leq 2$. Gaussian processes are stable processes with $\alpha = 2$. Stable distributions have heavier tails than the normal distribution, possess finite pth order moments only for $p < \alpha$, and are appropriate for modeling signals with outliers. Hence, non-Gaussian stable deviates have infinite variance and in some cases infinite first moment. Unlike the Gaussian distribution

which is symmetric about its mean, stable distributions may be assymetric, i.e., they admit skewness. Therefore, in certain applications where a heavy-tailed and assymetric model is called for, the stable model may be a viable alternative to the Gaussian distribution. The difficulty in developing signal processing methods based on stable processes is due to the fact that the linear space of a stable process is not a Hilbert space, as in the case of Gaussian processes, but either a Banach ($1 < \alpha < 2$) or a metric space ($0 < \alpha \leq 1$), both of which are more unyielding in their structure.

This dissertation addresses the solution of the signal parameter estimation problem through the use of sensor array data retrieved in the presence of impulsive interference. One of the most interesting problems in this area is the estimation of the direction-of-arrival (DOA) of narrow-band source signals having the same known center frequency. A related problem is the detection of the number of sources impinging on a sensor array. In the past, these problems have been studied extensively under the assumption of Gaussian distributed signals and/or noise, and a variety of methods for their solution have been proposed. As a result of the Gaussianity assumption, most methods are based on the second- or higher-order statistics of the signals [51]. The following section gives a brief review of these methods.

1.1 Literature Review

Despite the different characteristics of several existing multisensor data systems, there exists a common methodology of processing data from multiple sensors in order to make inferences about a physical event [26]. A block diagram of a typical multisensor system is shown in Figure 1.1. The sensors detect incoming energy and observe a combination of intentional energy from a target as well as environmental energy such as noise, multipath signals, and interference. An additional source of incoming energy may include jamming signals, which are intentionally generated by an enemy to deceive the observer by decreasing the signal-to-noise-ratio (SNR) of the observed environment. The ensuing signal conditioning process does not alter the information content of the signal but it performs translations (e.g. frequency shifts, heterodyning, A/D or D/A changes) that facilitate subsequent processing. A data alignment function transforms the raw sensor observations into a standard set of units. An association process groups the observations into meaningful groups with every group representing observations of a single physical entity.

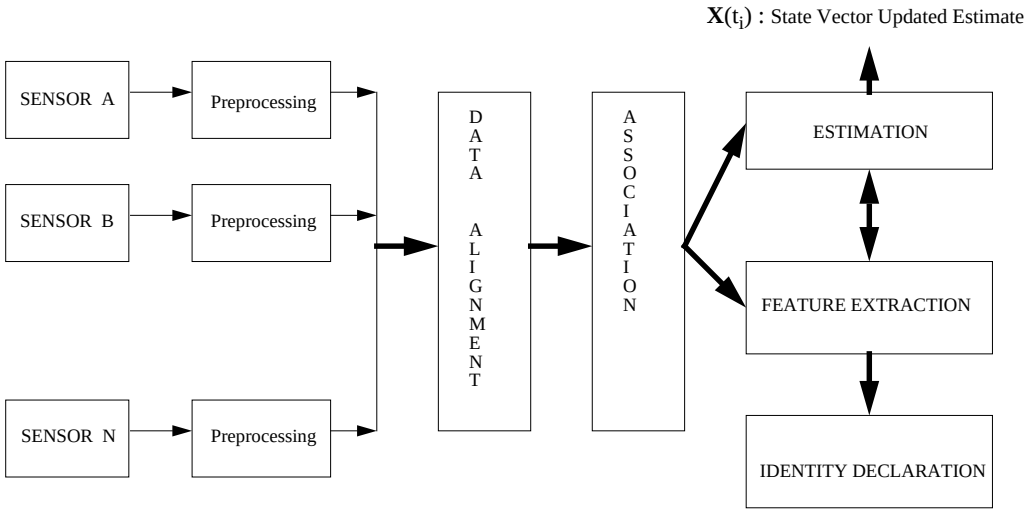


Figure 1.1: Typical multisensor system.

Our work focuses on development of methods pertaining to the estimation part of the system shown in Figure 1.1. The estimation process combines the observations to obtain an estimate of a state vector $\mathbf{X}(t)$ which best fits the observed data. Basic inferences about the observed entity's azimuth, elevation, range, or velocity are made at this level of processing. The estimation processor is essentially an algorithm which optimizes an appropriate chosen cost function with respect to the observation vectors. The cost function is formed by assuming certain statistics about the data based on the physics of the problem and by using an optimization criterion. There exist several optimization criteria such as Least-Squares (LS), Weighted LS, Maximum Likelihood, and Constrained (Bayesian) which, until recently, were based on a second-order moments statistical model.

Maximum Likelihood (ML) was one of the first methods to be applied in the area of array processing [34]. When applying the ML technique to the source localization problem, two different assumptions for the signal waveforms result into two different methods. According to the *Stochastic ML* (SML), the signals are modeled as Gaussian random processes. This is often motivated by the Central Limit Theorem and results in mathematically convenient expressions. On the other hand, in the *Deterministic ML* (DML) the signals are considered as unknown, deterministic quantities that need to be estimated

in conjunction with the direction of arrival. This is a natural model for digital communication applications where the signals are far from being normal random variables, and where estimation of the signal is of equal interest.

The importance of the ML technique comes from the mathematical property that, under certain regularity conditions, the ML estimator is known to be asymptotically efficient, i.e., it achieves the Cramér-Rao bound (CRB) for the estimation error variance. In this sense, ML has the best possible asymptotic properties.

Many researchers have studied the ML technique as a means of approaching the source localization problem in the presence of Gaussian additive noise. Stoica and Nehorai examined the ML performance and its asymptotic properties, derived expressions for the CRB and established some of its properties [71]. Additionally, they investigated the relationship between the ML and large sample approximation methods such as MUSIC. With high computational cost constituting the main shortcoming of the ML technique, Ziskind and Wax introduced a computationally attractive method for calculating the ML estimator [97]. Their method was based on an iterative technique named "Alternating Projection" (AP) that transformed the multivariate optimization problem into a sequence of simpler one-dimensional optimization problems.

However, due to the high computational load of the multivariate nonlinear maximization problem involved in the ML estimator, sub-optimal methods have also been developed [13, 30, 41, 57, 61, 66, 80, 81, 84, 85]. The better known ones are cited here: Minimum Variance Distortionless method of Capon [13] and the so-called eigenvector-based methods including the MUSIC [66], Minimum Norm [41, 57], and the ESPRIT method [61]. The performance of the aforementioned methods is inferior to that of the ML method, especially for low SNR values or when the number of observation snapshots is small.

The MUSIC method, a generalization of Pisarenko's harmonic retrieval method, has received the most attention and triggered the development of a large number of algorithms referred to as *eigenvector* or *subspace* techniques. Besides offering a new geometric interpretation of the array processing problem, MUSIC uses concepts from complex vector spaces and well-known tools from linear algebra, such as the singular value decomposition (SVD), in order to achieve high resolution while keeping the computational complexity relatively low compared to that of the ML methods.

Source localization can be considered as an important initial step in adaptive array processing. Once the desired signal and interference directions are determined, adaptive

statistically optimum beamforming maximizes the signal-to-noise ratio of the array output for optimal signal detection. The multiple sidelobe canceller (MSC) is perhaps the earliest statistically optimum beamformer introduced by Applebaum *et al.* [1]. Widrow and associates studied the adaptive antenna problem by minimizing the mean square error between the beamformer output and a reference signal [92]. Reed *et al.* concentrated on adaptive radar applications. They developed an adaptive processor which maximizes the probability of detection for a fixed false-alarm rate [4]. They also introduced a direct method of adaptive weight computation, based on the sample covariance matrix of the noise field, which provides rapid convergence [59].

Recently, special interest has been shown in relaxing some of the assumptions concerning the statistical nature of the noise in the bearing estimation problem [19, 20, 39]. One such method is the bispectrum beamformer introduced by Forster and Nikias [19]. It was demonstrated that, for the case of spatially correlated Gaussian additive noise with unknown cross-spectral matrix (CSM), the bispectrum beamformer may provide asymptotically better bearing estimates than the stochastic ML method with known CSM.

Often in actual applications the Gaussian noise assumption proves inadequate, as systems designed under this assumption exhibit a significant performance degradation. There exist physical processes generating interferences which contain *noise components that are impulsive in nature*. These processes can be natural, as well as man-made, and include radar clutter, underwater acoustic signals, lightning in the atmosphere, and transients in power lines and car ignitions. In modeling this type of signals, the *stable distribution law* provides a very attractive theoretical tool. It was proven that under broad conditions, a general class of impulsive noise follows the stable law [68]. As a result, considerable research interest has been shown in designing robust signal processing algorithms for detection, direction finding, and equalization, that can do well not only in the presence of Gaussian noise, but also, more importantly, in the presence of impulsive noise environments [75, 78, 79].

1.2 Dissertation Organization and Contribution

This dissertation is devoted to the detection and localization problem of multiple sources in the presence of impulsive additive interference which can be modeled as a complex isotropic α-stable process. The dissertation is organized as follows: In Chapter 2, we formulate the direction-of-arrival estimation problem. In Chapter 3, we introduce the statistical model,

based on the class of bivariate symmetric α-stable ($S\alpha S$) distributions. This model is well-suited for describing signal and/or noise processes that are impulsive in nature and contains the Gaussian process as a special case. In Chapter 4, we develop the deterministic ML estimator and derive the Cramér-Rao bound for the Cauchy case. We also discuss the application of the ML method for the general case of $S\alpha S$ processes and we present Monte-Carlo simulation results. In Chapter 5, we develop large-sample, subspace-based bearing estimation techniques based on the eigendecomposition of the array covariation matrix. In Chapter 6, we extend our methods to handle the case of wide-band source signals. Finally, in Chapter 7, we propose future research directions.

1.3 Abbreviations

The following abbreviations are used in this dissertation:

CRB	Cramér-Rao bound
DML	deterministic maximum likelihood
DOA	direction-of-arrival
FLOM	fractional lower-order moment
GSNR	generalized signal-to-noise ratio
IM	incoherent MUSIC
IRM	incoherent ROC-MUSIC
ML	maximum likelihood
MLC	maximum likelihood Cauchy
MLG	maximum likelihood Gaussian
MSE	mean-square error
MUSIC	multiple signal classification
ROC-MUSIC	robust covariation-based MUSIC
PSNR	pseudo signal-to-noise ratio
$S\alpha S$	symmetric α-stable
SML	stochastic maximum likelihood
SNR	signal-to-noise ratio
SSC	steered spectral covariation
STMV	steered minimum variance

Chapter 2

Array Signal Processing Fundamentals and Current Approaches

The classical problem in array signal processing is to determine the location of an energy-radiating planar source relative to the location of the array. We are interested in estimating the direction-of-arrival of a signal in the presence of noise and interfering signals. The array antenna attacks this problem by exploiting the spatial separation in the locations where the desired and interfering signals originate. It filters data collected over a spatial aperture much in the same manner that an FIR filter processes temporally sampled data. Thus, the key issue is to recognize the equivalence of the direction estimation problem with that of estimating the spatial spectrum of the radiation field. Then, spectrum estimation techniques can be viewed in the framework of array processing applications.

To facilitate the discussion, some notation must be developed for referring to the array processing system parameters. First we define some mathematical conventions: Scalar quantities are denoted with italic typeface. Vectors are denoted with lower case boldface letters while matrices are denoted by uppercase boldface letters. The operations of transposition, complex cojugation, and conjugate transposition are denoted by superscripts T, $*$, and H, respectively. The notation $\mathbf{T} = \text{diag}(\mathbf{t})$ denotes a diagonal matrix whose main diagonal is the vector $\mathbf{t}$. For q $r \times 1$ vectors $\mathbf{a}_1, \ldots, \mathbf{a}_q$, $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_q]$ is an $r \times q$ matrix. Finally, $E\{\cdot\}$ denotes the expected value of a random quantity. For reference, Table 2.1 provides a list of the various array processing system parameters which will be used in the following sections.

Table 2.1: Symbols for array processing system parameters.

r	number of sensor elements in the array
q	number of source signals
M	number of snapshots

2.1 Problem Formulation

Consider an array of r sensors with arbitrary locations and arbitrary directional characteristics, which receive signals generated by q narrow-band sources with known center frequency ω and locations $\theta_1, \theta_2, \ldots, \theta_q$. Since the signals are narrow-band, the propagation delay across the array is much smaller than the reciprocal of the signal bandwidth, and it follows that by using a complex envelop representation, the array output can be expressed as

$$\mathbf{x}(t) = \sum_{k=1}^{q} \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t) \tag{2.1}$$

where we have the following:

- $\mathbf{x}(t) = [x_1(t), \ldots, x_r(t)]^T$ is the vector of the signals received by the array sensors;
- $s_k(t)$ is the signal emitted by the kth source as received at the reference sensor 1 of the array;
- $\mathbf{a}(\theta_k) = [1, e^{-\jmath\omega\tau_2(\theta_k)}, \ldots, e^{-\jmath\omega\tau_r(\theta_k)}]^T$ is the steering vector of the array toward direction θ_k;
- $\tau_i(\theta_k)$ is the propagation delay between the first and the ith sensor for a waveform coming from direction θ_k;
- $\mathbf{n}(t) = [n_1(t), \ldots, n_r(t)]^T$ is the noise vector.

Equation (2.1) can be expressed in a compact form as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{2.2}$$

where $\mathbf{A}(\boldsymbol{\theta})$ is the $r \times q$ matrix of the array steering vectors

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_q)] \tag{2.3}$$

and $\mathbf{s}(t)$ is the $q \times 1$ vector of the signals

$$\mathbf{s}(t) = [s_1(t), \ldots, s_q(t)]^T. \tag{2.4}$$

Assuming that M snapshots are taken at time instants $t_1, \ldots, t_M$, the data can be expressed as

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N} \tag{2.5}$$

where $\mathbf{X}$ and $\mathbf{N}$ are the $r \times M$ matrices

$$\mathbf{X} = [\mathbf{x}(t_1), \ldots, \mathbf{x}(t_M)] \tag{2.6}$$

$$\mathbf{N} = [\mathbf{n}(t_1), \ldots, \mathbf{n}(t_M)] \tag{2.7}$$

and $\mathbf{S}$ is the $q \times M$ matrix

$$\mathbf{S} = [\mathbf{s}(t_1), \ldots, \mathbf{s}(t_M)]. \tag{2.8}$$

Our objective is to estimate the DOA's $\theta_1, \ldots, \theta_q$ of the sources from the M snapshots of the array $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_M)$.

Toward this goal, we are going to make the following assumptions regarding the array, the signals, and the noise:

A.1 The number of signals is known and is smaller than the number of sensors, i.e., $q < r$.

A.2 The set of any q steering vectors is linearly independent.

A.3 The noise samples $n_i(t_j)$; $i = 1, \ldots, r$; $j = 1, \ldots, M$, come from a complex (bivariate) *isotropic stable distribution.*

A.4 The noise samples $n_i(t_j)$ are statistically independent from one another both along the array sensors, namely, along index i, and along time, namely, along index j.

Assumptions A.1 and A.2 guarantee the uniqueness of the solution. Assumption A.3 draws a new element into our analysis as we deviate from the conventional assumption that the noise in sensor arrays is a complex-valued Gaussian process. Our assumption incorporates a wide range of noise environments which are impulsive in nature and includes the Gaussian noise as a special case.

2.2 Maximum Likelihood DOA Estimation with Gaussian Distributions

In this section, we review the maximum likelihood methodology for the case of additive Gaussian noise of zero mean and variance matrix $\sigma^2\mathbf{I}$, where σ^2 is the unknown noise power, and $\mathbf{I}$ is the identity matrix. When applying the ML technique to the array processing problem, two main methods have been considered depending on the signal data model assumption. According to the *Stochastic ML*, the signals are modeled as Gaussian random processes. On the other hand, in the *Deterministic ML* the signals are considered as unknown, deterministic quantities that need to be estimated in conjunction with the direction of arrival. In the following, we give a brief presentation of the two methods together with bounds on the estimation accuracy.

2.2.1 The Stochastic Maximum Likelihood Method

The *Stochastic ML* (SML) method models the signals as stationary jointly Gaussian stochastic processes with covariance matrix given by $\mathbf{\Sigma} = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$. Under the assumptions of Section 2.1, the observation process $\mathbf{x}(t)$ is a stationary, zero-mean, Gaussian process completely described by its second-order covariance matrix:

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^H + \sigma^2\mathbf{I}, \tag{2.9}$$

where we write $\mathbf{A}$ instead of $\mathbf{A}(\boldsymbol{\theta})$ for notational convenience. Since the snapshots are independent and identically distributed, the density function of the complete data set $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_M)$ is given by:

$$f(\mathbf{X}) = \prod_{t=1}^{M} \frac{1}{\pi^r |\mathbf{R}|} \exp\left(-\mathbf{x}^H(t)\mathbf{R}^{-1}\mathbf{x}(t)\right). \tag{2.10}$$

Ignoring constant terms and after some simplifications, the SML estimator is obtained by solving the following optimization problem:

$$[\hat{\sigma}^2, \hat{\mathbf{\Sigma}}, \hat{\boldsymbol{\theta}}] = \arg \min_{\sigma^2, \mathbf{\Sigma}, \boldsymbol{\theta}} \{\log |\mathbf{R}| + \mathrm{Tr}\{\mathbf{R}^{-1}\hat{\mathbf{R}}\}\}, \tag{2.11}$$

where $\mathrm{Tr}\{\cdot\}$ is the trace operator, and

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{t=1}^{M}\mathbf{x}(t)\mathbf{x}^H(t) \tag{2.12}$$

is the sample covariance matrix. The dimensionality of the problem can be reduced by substituting the SML estimates for σ^2 and $\mathbf{\Sigma}$ [2, 31] in (2.11):

$$\hat{\mathbf{\Sigma}}(\boldsymbol{\theta}) = \mathbf{A}^{\dagger}(\boldsymbol{\theta})\left[\hat{\mathbf{R}} - \hat{\sigma}^2(\boldsymbol{\theta})\mathbf{I}\right]\mathbf{A}^{\dagger H}(\boldsymbol{\theta}), \tag{2.13}$$

and

$$\hat{\sigma}^2(\boldsymbol{\theta}) = \frac{1}{r-q}\mathrm{Tr}\left\{\mathbf{P}_{\mathbf{A}}^{\perp}(\boldsymbol{\theta})\hat{\mathbf{R}}\right\}, \tag{2.14}$$

where $\mathbf{A}^{\dagger} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ is the pseudo-inverse of $\mathbf{A}$ and $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger}$ is the orthogonal projector onto the null space of $\mathbf{A}^H$. The SML signal parameter estimates $\hat{\boldsymbol{\theta}}_{SML}$ are now obtained by solving the following optimization problem:

$$\hat{\boldsymbol{\theta}}_{SML} = \arg\min_{\boldsymbol{\theta}} \log|\mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{\Sigma}}(\boldsymbol{\theta})\mathbf{A}^H(\boldsymbol{\theta}) + \hat{\sigma}^2(\boldsymbol{\theta})\mathbf{I}|. \tag{2.15}$$

In general, the SML estimate $\hat{\boldsymbol{\theta}}_{SML}$ cannot be found analytically. Hence, numerical procedures must be employed to carry out the required optimization. Several optimization methods have been proposed in the literature, including the Alternating Projection method [97], several Newton-type techniques [23, 70, 82, 87], and the Expected Maximization (EM) method [18, 49]. The SML likelihood function is regular and the SML estimator is consistent and asymptotically efficient, i.e., the covariance of the estimates asymptotically attains the stochastic Cramér-Rao Bound [42].

2.2.2 The Deterministic Maximum Likelihood Method

According to the *Deterministic ML* (DML), the signals are considered as unknown, deterministic quantities that need to be estimated in conjunction with the direction of arrival. This is a natural model for digital communication applications where the signals are far from being normal random variables, and where estimation of the signal is of equal interest.

Under assumption A.4 of Section 2.1, it follows from (2.2) that the density function of the sampled data is given by

$$f(\mathbf{X}) = \prod_{t=1}^{M} \frac{1}{\pi \sigma^{2r}} \exp\left(-\frac{1}{\sigma^2} \left| \mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) \right|^2 \right). \tag{2.16}$$

Hence, by ignoring constant terms, the log likelihood function $L(\mathbf{X}; \sigma^2, \mathbf{S}, \boldsymbol{\theta})$ can be expressed as:

$$L(\mathbf{X}; \sigma^2, \mathbf{S}, \boldsymbol{\theta}) = -Mr \log(\sigma^2) - \frac{1}{\sigma^2} \sum_{t=1}^{M} \left| \mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) \right|^2. \tag{2.17}$$

The ML estimator is obtained by maximizing $L(\mathbf{X}; \sigma^2, \mathbf{S}, \boldsymbol{\theta})$ with respect to σ^2, $\mathbf{S}$, and $\boldsymbol{\theta}$. Fixing $\mathbf{S}$ and $\boldsymbol{\theta}$, and maximizing with respect to σ^2 we get

$$\hat{\sigma}^2 = \frac{1}{Mr} \sum_{t=1}^{M} \left| \mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) \right|^2. \tag{2.18}$$

Substituting (2.18) back into (2.17) and maximizing with respect to $\mathbf{S}$ we get a Least–Squares estimate for the signal

$$\hat{\mathbf{s}}(t) = (\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta}))^{-1}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{x}(t). \tag{2.19}$$

Thus, the dimension of the required optimization is reduced, and the maximum likelihood estimator for $\boldsymbol{\theta}$ is given by the following minimization problem:

$$\hat{\boldsymbol{\theta}}_{DML} = \arg\min_{\boldsymbol{\theta}} \operatorname{Tr}\left\{ \mathbf{P}^{\perp}_{\mathbf{A}(\boldsymbol{\theta})} \hat{\mathbf{R}} \right\}. \tag{2.20}$$

The solution of (2.20) requires a q-dimensional optimization for computing the directions of arrival. As in the case of the SML estimator, analytical solutions for this problem are not available in general, and one has to resort to numerical methods, such as Newton-type algorithms. A drawback of the DML method is that the dimension of the parameter vector increases without bound as the number of incoming snapshots M grows. For this reason, the DML function is non-regular and although the DML estimate of $\boldsymbol{\theta}$ is consistent, the DML estimate of the signal $\mathbf{S}$ is inconsistent. Furthermore, $\hat{\boldsymbol{\theta}}_{DML}$ is asymptotically efficient as M goes to infinity only when the number of sensors r also goes to infinity [71]. Hence, the requirement from the DML method to estimate the source signals results in poorer source bearing estimates, unless r is large.

2.2.3 The Deterministic Cramér-Rao Bound for Gaussian Noise

Under the assumptions stated in Section 2.1, and for the case of deterministic incoming signal waveforms in the presence of complex Gaussian noise, the following theorem holds [28]:

Theorem 2.1 *The CRB for* $\boldsymbol{\theta}$*, and* σ^2 *is given by*

$$CRB(\boldsymbol{\theta}) = \frac{\sigma^2}{2}\left\{\sum_{t=1}^{M} \Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\left[I - \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\right]\mathbf{D}\mathbf{S}(t)\right\}\right\}^{-1} \tag{2.21}$$

and

$$CRB(\sigma^2) = \frac{\sigma^4}{Mr}, \tag{2.22}$$

where

$$\mathbf{S}(t) = \begin{bmatrix} s_1(t) & 0 & \cdots & 0 & 0 \\ 0 & s_2(t) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{q-1}(t) & 0 \\ 0 & 0 & \cdots & 0 & s_q(t) \end{bmatrix}; \quad t = 1, \ldots, M, \tag{2.23}$$

$$\mathbf{D} = [\mathbf{d}(\theta_1), \ldots, \mathbf{d}(\theta_q)], \tag{2.24}$$

and $\mathbf{d}(\theta_i) = \partial\mathbf{a}(\theta_i)/\partial\theta_i;\ \ i = 1, \ldots, q.$

For the case of a single source ($q = 1$) impinging from direction θ in a linear array whose sensors are spaced a half-wavelength apart we have

$$\mathbf{A} = [1, e^{-\jmath\pi\sin(\theta)}, \ldots, e^{-\jmath(r-1)\pi\sin(\theta)}]^T, \tag{2.25}$$

and

$$\mathbf{D} = [0, -\jmath\pi\cos(\theta)e^{-\jmath\pi\sin(\theta)}, \ldots, -\jmath(r-1)\pi\cos(\theta)e^{-\jmath(r-1)\pi\sin(\theta)}]^T. \tag{2.26}$$

So, it holds

$$\mathbf{A}^H\mathbf{A} = r, \tag{2.27}$$

$$\mathbf{D}^H\mathbf{D} = \pi^2\frac{r(r-1)(2r-1)}{6}\cos^2(\theta), \tag{2.28}$$

and

$$\mathbf{D}^H\mathbf{A} = \jmath\pi\frac{r(r-1)}{2}\cos(\theta). \tag{2.29}$$

Then,

$$CRB(\theta) = \frac{6}{\pi^2} \cdot \frac{1}{r(r^2-1)\cos^2(\theta)} \cdot \frac{\sigma^2}{\sum_{t=1}^{M} |s(t)|^2}. \tag{2.30}$$

As we can clearly see, the larger the signal-to-noise ratio and the bigger the array, the higher the CRB.

2.3 Subspace-Based DOA Estimation Using Second-Order Statistics

Due to the high computational load of the multivariate nonlinear maximization problem involved in the ML estimator, sub-optimal high-resolution methods have also been developed, and the most important of them are presented here. In this section, we will assume that the signal waveforms are noncoherent Gaussian processes and we will present the algorithms based in this assumption. Means of reducing the susceptibility of the methods to coherent signals have been introduced in the literature for special array structures [54, 55, 58, 67, 91, 93].

The covariance matrix of the observation vector $\mathbf{x}(t)$ is given by

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^H + \sigma^2\mathbf{I}. \tag{2.31}$$

The $q \times q$ matrix $\mathbf{\Sigma} = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$ has full rank since we assumed noncoherence of the q incoming plane waves. Assumption A.3 of Section 2.1 implies that matrix $\mathbf{A}$ has full rank q.

A class of spatial spectral estimation techniques is based on the eigenvalue decomposition of the spatial covariance matrix. The rationale behind this approach is that one wants to emphasize the choices for the steering vector $\mathbf{a}(\theta)$, which correspond to signal directions. The method exploits the property that the directions of arrival determine the eigenstructure of the matrix.

Let $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_r$ denote the eigenvalues of matrix $\mathbf{R}$, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$ denote the eigenvalues of matrix $\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H$, respectively. Then, from (2.31) it follows that

$$\rho_i = \lambda_i + \sigma^2 \qquad i = 1, 2, \ldots, r. \tag{2.32}$$

Since matrix $\mathbf{A}$ is of full column rank q, the $(r-q)$ smallest eigenvalues of the matrix $\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H$ are equal to zero. Correspondingly, the $(r-q)$ eigenvalues of $\mathbf{R}$ are equal to σ^2:

$$\rho_i = \begin{cases} \lambda_i + \sigma^2 & \text{i=1,...,q} \\ \sigma^2 & \text{i=q+1,...,r.} \end{cases} \tag{2.33}$$

Hence, the eigenvalue decomposition of the spatial covariance matrix $\mathbf{R}$ can be written as

$$\mathbf{R} = \sum_{i=1}^{q} [\lambda_i + \sigma^2] \mathbf{v}_i \mathbf{v}_i^H + \sum_{i=q+1}^{r} \sigma^2 \mathbf{v}_i \mathbf{v}_i^H, \tag{2.34}$$

where

$$\mathbf{v}_i^H \mathbf{v}_j = \delta_{i,j}, \tag{2.35}$$

are the orthogonal eigenvectors of the matrix $\mathbf{R}$, i.e., $\mathbf{R}\mathbf{v}_i = \rho_i \mathbf{v}_i, \quad i = 1, 2, \ldots, r$. It holds that

$$\mathbf{R}\mathbf{v}_i = \sigma^2 \mathbf{v}_i, \qquad i = q+1, \ldots, r \tag{2.36}$$

or, equivalently,

$$(\mathbf{R} - \sigma^2 \mathbf{I})\mathbf{v}_i = \mathbf{0}, \qquad i = q+1, \ldots, r \tag{2.37}$$

Using (2.31), we get

$$\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H \mathbf{v}_i = \mathbf{0}, \qquad i = q+1, \ldots, r, \tag{2.38}$$

from which it follows that

$$\mathbf{A}^H \mathbf{v}_i = \mathbf{0}, \qquad i = q+1, \ldots, r. \tag{2.39}$$

Equation (2.39) readily implies that the subspace spanned by the eigenvectors $\{\mathbf{v}_{q+1}, \mathbf{v}_{q+2}, \ldots, \mathbf{v}_r\}$ is the orthogonal complement of the subspace spanned by the steering vectors $\{\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_q)\}$. Since the eigenvectors of the covariance matrix $\mathbf{R}$ are orthogonal to each other we can also conclude that the subspace spanned by the eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q\}$ is exactly the same as the subspace spanned by the vectors $\{\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_q)\}$.

The previous analysis leads to the following observations. If the propagation field contains q distinct noncoherent propagating signals in a spatially white noise environment, then the eigenvalue decomposition of the spatial covariance matrix $\mathbf{R}$ results in the formation of two disjoint subspaces that are the orthogonal complement of each other. The first

one, called the *signal plus noise subspace,* is spanned by the eigenvectors corresponding to the q largest eigenvalues of $\mathbf{R}$. The second, called the *noise subspace* is spanned by the eigenvectors corresponding to the $r-q$ smallest eigenvalues of $\mathbf{R}$. Given the eigenvectors of $\mathbf{R}$, we may determine the signal directions of arrival by searching for those steering vectors $\mathbf{a}(\theta)$ that are orthogonal to the noise subspace.

In the following, we present four spatial spectra estimation methods based on the eigenvalue decomposition of the spatial covariance matrix.

2.3.1 The MUSIC Algorithm

As we already saw, given the eigenvectors of the spatial covariance matrix $\mathbf{R}$, we may determine the signal directions of arrival by searching for those steering vectors $\mathbf{a}(\theta)$ that are orthogonal to the noise subspace which is spanned by the eigenvectors $\{\mathbf{v}_{q+1}, \mathbf{v}_{q+2}, \ldots, \mathbf{v}_r\}$ corresponding to the $r-q$ smallest eigenvalues of $\mathbf{R}$. In practice, $\mathbf{R}$ is unknown but can be consistently estimated from the available data as

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{t=1}^{M}\mathbf{x}(t)\mathbf{x}(t)^H, \tag{2.40}$$

Because of the uncertainty in the eigenvector estimates $\{\hat{\mathbf{v}}_{q+1}, \hat{\mathbf{v}}_{q+2}, \ldots, \hat{\mathbf{v}}_r\}$ introduced by the way we estimate the matrix $\mathbf{R}$, we can only search for the steering vectors that are most closely orthogonal to the noise subspace. The MUSIC algorithm estimates the signal directions as the peaks of the MUSIC spatial spectrum estimate given by

$$P_{MUSIC}(\theta) = \frac{1}{\sum_{i=q+1}^{r}|\mathbf{a}(\theta)^H\hat{\mathbf{v}}_i|^2}. \tag{2.41}$$

Note that the MUSIC spectrum is based on a single realization of the stochastic process represented by the snapshots $\mathbf{x}(t), \quad t = 1, \ldots, M$. MUSIC estimates are consistent and they converge to the true source bearings as the number of snapshots grows to infinity [38]. Furthermore, when the signal covariance matrix $\mathbf{\Sigma}$ is diagonal, MUSIC is a large sample realization of the DML method described in Section 2.2.2. The asymptotic equivalence of MUSIC and the DML method for uncorrelated signals was proven in [71]. A drawback of MUSIC is its sensitivity to model errors [21, 73]. MUSIC requires the computationally costly procedure of calibration and it is very sensitive to errors in the calibration procedure. The cost of calibration increases as the number of parameters that define the array manifold increases. Friedlander approached the calibration problem in a "blind" way, i.e.,

without knowledge of the direction of the calibrating source and with minimal knowledge about the array manifold [22].

2.3.2 The Minimum Norm Spatial Spectrum Estimator

The Minimum Norm (MN) method is also based on the eigenvalue decomposition of the spatial covariance matrix $\mathbf{R}$. We form the two sets of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q\}$ and $\{\mathbf{v}_{q+1}, \mathbf{v}_{q+2}, \ldots, \mathbf{v}_r\}$ into two matrices $\mathbf{V}_s$ and $\mathbf{V}_n$ respectively:

$$\mathbf{V}_s = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q] = \begin{bmatrix} \mathbf{g}_s^T \\ \mathbf{G}_s \end{bmatrix}, \tag{2.42}$$

and

$$\mathbf{V}_n = [\mathbf{v}_{q+1}, \mathbf{v}_{q+2}, \ldots, \mathbf{v}_r] = \begin{bmatrix} \mathbf{g}_n^T \\ \mathbf{G}_n \end{bmatrix}, \tag{2.43}$$

where $\mathbf{g}_s^T$ and $\mathbf{g}_n^T$ are $1 \times q$ and $1 \times (r-q)$ row vectors, respectively. The matrices $\mathbf{G}_s$ and $\mathbf{G}_n$ have dimensions $(r-1) \times q$ and $(r-1) \times (r-q)$, respectively.

The Minimum Norm method aims in finding an $r \times 1$ vector $\mathbf{W}$ that satisfies the following requirements:

- The vector $\mathbf{W}$ lies in the range of $\mathbf{V}_n$, and therefore, it is orthogonal to the columns of $\mathbf{V}_s$, i.e.,

$$\mathbf{V}_s^H \mathbf{W} = \mathbf{0} \tag{2.44}$$

- The first element of $\mathbf{W}$ is equal to 1, i.e.,

$$\mathbf{W} = \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} \tag{2.45}$$

- The Euclidean Norm of $\mathbf{W}$ is minimum.

By using (2.42), (2.44), and (2.45), we can write

$$\mathbf{G}_s^H \mathbf{w} = \mathbf{g}_s. \tag{2.46}$$

This is a linear system of q equations in $r-1$ unknowns. Therefore, with $q < r$, it is an undetermined system with no unique solution. The minimum norm solution to this system is given by

$$\mathbf{w} = (1 - \mathbf{g}_s^H \mathbf{g}_s)^{-1} \mathbf{G}_s^H \mathbf{g}_s. \tag{2.47}$$

Correspondingly, the minimum norm solution for $\mathbf{W}$ is given by

$$\mathbf{W} = \begin{bmatrix} 1 \\ -(1 - \mathbf{g}_s^H \mathbf{g}_s)^{-1} \mathbf{G}_s^H \mathbf{g}_s \end{bmatrix} \tag{2.48}$$

An analogous expression for $\mathbf{W}$ in terms of the elements constituting the matrix $\mathbf{V}n$ is

$$\mathbf{W} = \begin{bmatrix} 1 \\ (\mathbf{g}_n^H \mathbf{g}_n)^{-1} \mathbf{G}_n^* \mathbf{g}_n \end{bmatrix} \tag{2.49}$$

Having computed the vector $\mathbf{W}$, the Minimum Norm spatial spectrum estimate is given by

$$P_{MN}(\theta) = \frac{1}{|\mathbf{a}(\theta)^H \mathbf{W}|^2} \tag{2.50}$$

Note that, as in the case of the MUSIC spectrum, the Minimum Norm spectrum is based on a single realization of the underlying stochastic process.

2.3.3 The Minimum Variance Distortionless Method

The Minimum Variance Distortionless (MVD) method is developed as a constrained optimization problem. The sensor outputs are weighted by a vector $\mathbf{W}$ to produce a beamformer output $y(t)$ given by

$$y(t) = \mathbf{W}^H \mathbf{X}(t). \tag{2.51}$$

The spectral estimate is derived by finding the vector $\mathbf{W}$ which yields the minimum beam energy $E\{|y(t)|^2\} = \mathbf{W}^H \mathbf{R} \mathbf{W}$ subject to the constraint that $\mathbf{W}^H \mathbf{a}(\theta) = 1$, where $\mathbf{a}(\theta)$ represents an ideal plane wave corresponding to the direction of interest θ. Thus, the MVD estimator ensures that any signal from angle θ and frequency ω_0 is passed to the output undistorted, while at the same time contributions to the output due to noise and interfering signals arriving from directions other than θ are minimized.

Then, the linearly constrained minimization problem is written as:

$$\min_{\mathbf{W}} \mathbf{W}^H \mathbf{R} \mathbf{W} \quad \text{subject to} \quad \mathbf{W}^H \mathbf{a}(\theta) = 1. \tag{2.52}$$

The solution technique is to use the Lagrange multipliers to minimize the expression:

$$\mathbf{W}^H \mathbf{R} \mathbf{W} + \alpha[\mathbf{W}^H \mathbf{a}(\theta) - 1]. \tag{2.53}$$

The resulting solution for $\mathbf{W}$ is

$$\mathbf{W} = \frac{\mathbf{R}^{-1}\mathbf{a}(\theta)}{\mathbf{a}(\theta)^H \mathbf{R}^{-1}\mathbf{a}(\theta)}. \tag{2.54}$$

The power in the beam when steered in the direction of interest determined by $\mathbf{a}(\theta)$ becomes

$$P_{MVD}(\theta) = \frac{1}{\mathbf{a}(\theta)^H \mathbf{R}^{-1}\mathbf{a}(\theta)}, \tag{2.55}$$

which is the MVD spatial spectrum estimate.

Expressing the spatial covariance matrix in terms of its eigenvectors and eigenvalues, we can write

$$\mathbf{R} = \sum_{i=1}^{r} \rho_i \mathbf{v}_i \mathbf{v}_i^H. \tag{2.56}$$

The inverse $\mathbf{R}^{-1}$ has the same eigenvectors as $\mathbf{R}$, but its eigenvalues are the reciprocal of those of $\mathbf{R}$. Therefore,

$$\mathbf{R}^{-1} = \sum_{i=1}^{r} \frac{1}{\rho_i} \mathbf{v}_i \mathbf{v}_i^H. \tag{2.57}$$

The eigenvector expansion of $\mathbf{R}^{-1}$ is truncated to include only the terms corresponding to the signal plus noise subspace

$$\hat{\mathbf{R}}^{-1} = \sum_{i=1}^{q} \frac{1}{\rho_i - \sigma^2} \mathbf{v}_i \mathbf{v}_i^H. \tag{2.58}$$

This expression for $\hat{\mathbf{R}}^{-1}$ can be substituted in (2.55) to give an alternative MVD spatial spectrum estimate.

2.3.4 The Autoregressive Spatial Spectrum Estimator

Using the same idea as in linear prediction for time series, let x_{m_0} be the total signal measured at the m_0 sensor. We assume that x_{m_0} is estimated by a weighted linear combination of the signals at the other $r-1$ sensors:

$$x_{m_0} = -\sum_{m \neq m_0} w_m x_m. \tag{2.59}$$

The method is based on finding the weights w_m which minimize the mean squared prediction error

$$E\{|x_{m_0} + \sum_{m \neq m_0} w_m x_m|^2\} \tag{2.60}$$

We define the weights column vector $\mathbf{W} = [w_1, \ldots, w_{m_0-1}, w_{m_0}, w_{m_0+1}, \ldots, w_r]^T$, and $\mathbf{U}_{m_0}$ as the vector having its m_0th element equal to one and the other elements equal to zero. Then, the linear prediction problem can be restated as:

$$\min_{\mathbf{W}} \mathbf{W}^H \mathbf{R} \mathbf{W} \text{ subject to } \mathbf{W}^H \mathbf{U}_{m_0} = 1. \tag{2.61}$$

The above constrained optimization problem is very similar to the MVD problem as described in (2.52), the only difference being the presence of the vector $\mathbf{U}_{m_0}$ in the constraint equation for the linear prediction problem instead of the steering vector $\mathbf{a}(\theta)$ in the constraint equation for the MVD problem. Therefore, the solution is:

$$\mathbf{W} = \frac{\mathbf{R}^{-1}\mathbf{U}_{m_0}}{\mathbf{U}_{m_0}^T \mathbf{R}^{-1} \mathbf{U}_{m_0}}. \tag{2.62}$$

Using the ideas found in time series, the linear prediction method corresponds to an autoregressive model for the signal. Thus, the power spectrum is given by the mean-squared prediction error divided by the magnitude squared of the spectrum of the predictor coefficients

$$P(\theta) = \frac{\mathbf{W}^H \mathbf{R} \mathbf{W}}{|\mathbf{W}^H \mathbf{a}(\theta)|^2} \tag{2.63}$$

Substituting $\mathbf{W}$ from (2.62) we get

$$P(\theta) = \frac{\mathbf{U}_{m_0}^T \mathbf{R}^{-1} \mathbf{U}_{m_0}}{|\mathbf{U}_{m_0}^T \mathbf{R}^{-1} \mathbf{a}(\theta)|^2} \tag{2.64}$$

By discarding the constant factor $\mathbf{U}_{m_0}^T \mathbf{R}^{-1} \mathbf{U}_{m_0}$ in the nominator and using the vector $\mathbf{U} = [1, 0, \ldots, 0]^T$ we get the Autoregressive (AR) spatial spectrum estimate

$$P_{AR}(\theta) = \frac{1}{|\mathbf{U}^T \mathbf{R}^{-1} \mathbf{a}(\theta)|^2} \tag{2.65}$$

Again, the expression for $\hat{\mathbf{R}}^{-1}$ given in (2.58) can be used in the estimate of the spatial spectrum.

2.3.5 Relationships Between the MVD, AR, MUSIC and MN Spectra

The spatial spectra estimators presented in the previous sections are interrelated in an asymptotic sense. Here, we assume that the model we introduced for our physical problem is exact: We have a space series of superimposed uncorrelated signals in an additive white noise environment, which consist of uniformly spaced spatial samples and whose spatial correlation matrix is exactly known. Then, we can state the following about the relationships among the different spatial spectra:

- As shown by Burg [5], the MVD and AR spectra are related to each other analytically in the case of an equally spaced linear array by

$$\frac{1}{P_{MVD}(\theta)} = \frac{1}{r} \sum_{m=1}^{r} \frac{1}{P_{AR}^{(m)}(\theta)}, \tag{2.66}$$

  where $P_{AR}^{(m)}(\theta)$ is the AR spectrum obtained with an mth order model. This result suggests that the MVD spectrum is a smoothed version of the AR spectrum as it is a harmonic average of lower order AR spectra. This explains the reason for the greater fluctuations observed in the AR spectrum compared to the MVD spectrum.

- Nickel [50] showed that if the spatial covariance matrix $\mathbf{R}$ is known exactly, the noise subspace projection matrix $\mathbf{V}_n \mathbf{V}_n^H$ used to compute the MUSIC spectrum is the limiting value of the inverse covariance matrix $\mathbf{R}^{-1}$ when the signal to noise power approaches infinity. Therefore, the MUSIC spectrum can be interpreted as the limit of an MVD type of spectrum when the signal-to-noise ratio is infinite. This observation explains the better resolution of the MUSIC spectrum compared to that of the MVD spectrum.

Figure 2.1: Relationships between the four high resolution spatial spectra

- The MN spectrum is the limiting value of the AR spectrum for infinite SNR, hence the superior resolution of the MN spectrum compared to the AR spectrum.
- The MUSIC spectrum is a weighted harmonic average of the MN spectra for array sizes 1 to M, i.e., the MUSIC spectrum is a smoothed version of the MN spectrum.

A graphical depiction of the above relationships is given in Figure 2.1.

2.4 High Resolution Methods Based on Higher-Order Statistics

The orthogonal decomposition methods described in the previous section have been developed by using the assumption that the additive noise is white, i.e., the noise covariance matrix is diagonal. When the white noise assumption does not hold, as it is often the case in sonar or radar applications, the harmonic decomposition methods based on second-order statistics will not give reliable spatial spectrum estimates. Array processing problems in the presence of spatially correlated noise have been studied by assuming known spatial noise autocovariance or by estimating it from secondary inputs [57, 63].

Pan and Nikias [74] studied the bearing estimation problem in the presence of spatially correlated Gaussian noise by employing fourth-order cumulants of the array snapshots.

The motivation behind the use of higher-order statistics lies in their ability to suppress the Gaussian noise. Indeed, if the additive noise follows the Gaussian probability density function, all its cumulants of order greater than two are identically zero.

The spatial spectrum estimates defined in the previous section can be reformulated by introducing the fourth-order cumulant matrix of the array snapshots as

$$\mathbf{C} \stackrel{\text{def}}{=} \text{Cum}\left\{ \begin{pmatrix} x_1(t) & x_1^*(t) & x_1(t) \\ x_2(t) & x_2^*(t) & x_2(t) \\ \vdots & \vdots & \vdots \\ x_M(t) & x_M^*(t) & x_M(t) \end{pmatrix} \cdot [x_1^*(t) x_2^*(t) \cdots x_M^*(t)] \right\}, \tag{2.67}$$

where Cum denotes the cumulant operator for a zero mean random process $x(k)$, defined as

$$\begin{aligned} &\text{Cum}\left\{x(k), x^*(k+\tau_1), x(k+\tau_2), x^*(k+\tau_3)\right\} = c_4(\tau_1, \tau_2, \tau_3) \\ &= E\left\{x(k)x^*(k+\tau_1)x(k+\tau_2)x^*(k+\tau_3)\right\} \\ &\quad -E\left\{x(k), x^*(k+\tau_1)\right\} \cdot E\left\{x(k+\tau_2)x^*(k+\tau_3)\right\} \\ &\quad -E\left\{x(k), x(k+\tau_2)\right\} \cdot E\left\{x^*(k+\tau_1)x^*(k+\tau_3)\right\} \\ &\quad -E\left\{x(k), x^*(k+\tau_3)\right\} \cdot E\left\{x^*(k+\tau_1)x(k+\tau_2)\right\} \end{aligned} \tag{2.68}$$

As we can see, to generate the fourth-order cumulant sequence, we need knowledge of the fourth-order moment and autocorrelation sequences.

The snapshot vector $\mathbf{x}(t)$ as defined in (2.2) is

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{2.69}$$

Substituting (2.69) into (2.67) we get that

$$\mathbf{C} = \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^H, \tag{2.70}$$

where $\boldsymbol{\Gamma} = diag(\gamma_1, \gamma_2, \ldots, \gamma_M)$, and γ_k is the fourth-order cumulant (kurtosis) of the kth source signal defined as

$$\gamma_k \stackrel{\text{def}}{=} \text{Cum}\left\{s_k(t), s_k^*(t), s_k(t), s_k^*(t)\right\} \tag{2.71}$$

The important point is that (2.70) holds even for the case of additive spatially correlated Gaussian noise with unknown covariance function. Based on the cumulant matrix $\mathbf{C}$, the high resolution methods introduced in Section 2.3 can be reformulated by decomposing matrix $\mathbf{C}$ in terms of its eigenvectors $\mathbf{\Upsilon}_i$:

$$\mathbf{C} = \sum_{i=1}^{q} [\mu_i + \epsilon_i] \mathbf{\Upsilon}_i \mathbf{\Upsilon}_i^H + \sum_{i=q+1}^{r} \epsilon_i \mathbf{\Upsilon}_i \mathbf{\Upsilon}_i^H, \tag{2.72}$$

In theory, it holds that $\epsilon_i = 0 \quad i = 1, \ldots, r$. In practice however, ϵ_i will be very small compared to μ_i if we estimate the cumulant matrix $\mathbf{C}$ properly. The inverse of $\mathbf{C}$ expressed in terms corresponding to the signal subspace is

$$\hat{\mathbf{C}}^{-1} = \sum_{i=1}^{q} \frac{1}{\mu_i} \mathbf{\Upsilon}_i \mathbf{\Upsilon}_i^H. \tag{2.73}$$

Then, the four spatial spectrum estimates become as follows:

MVD Spectrum Estimator:

$$P_{MVD}(\theta) = \frac{1}{\mathbf{a}(\theta)^H \hat{\mathbf{C}}^{-1} \mathbf{a}(\theta)} \tag{2.74}$$

AR Spectrum Estimator:

$$P_{AR}(\theta) = \frac{1}{|\mathbf{U}^T \hat{\mathbf{C}}^{-1} \mathbf{a}(\theta)|^2} \tag{2.75}$$

MUSIC Spectrum Estimator:

$$P_{MUSIC}(\theta) = \frac{1}{\sum_{i=q+1}^{r} |\mathbf{a}(\theta)^H \mathbf{\Upsilon}_i|^2} \tag{2.76}$$

MN Spectrum Estimator:

$$P_{MN}(\theta) = \frac{1}{|\mathbf{a}(\theta)^H \mathbf{W}|^2} \tag{2.77}$$

where the vector $\mathbf{W}$ follows from an analysis analogous to the one described in Section 2.3.2.

In addition to the aforementioned techniques, a bearing estimation method based on the cross-bispectrum of the array output data was introduced by Forster and Nikias [19]. The method was based on the asymptotic distribution of the cross-bispectrum estimates and employed maximum likelihood theory. Finally, another interpretation for the use of cumulants in array processing was proposed by Doğan and Mendel [15, 16]. It was shown that fourth-order cumulants of multichannel observations increase the directional information compared with second-order statistics, and suppress spatially colored non-Gaussian noise with the use of an additional remote sensor.

Chapter 3

Alpha-Stable Random Variables and Processes

Gaussian distributions and processes have long been accepted as useful tools for stochastic modeling. In this section, we introduce a statistical model based on the class of *symmetric α-stable ($S\alpha S$) distributions* which is well-suited for describing signals that are impulsive in nature. A review of the state of the art on stable processes from a statistical point of view is provided by a collection of papers edited by Cambanis, Samorodnitsky and Taqqu [11]. Several statisticians including Cambanis, Zolotarev, Weron, *et al.* have published extensively on the theory and applications of stable processes. They studied the properties of stable processes [6, 8, 29, 33, 36, 43, 46, 48, 65, 90, 98], their spectral representation [7, 27, 47], as well as prediction and linear filtering problems [9, 10, 35, 37, 62, 94]. Textbooks in the area were written by Samorodnitsky and Taqqu [64] and by Janicki and Weron [32].

An extensive review of stable processes from a signal processing point of view can be found in a tutorial paper by Shao and Nikias [69] as well as in a monogram written by the same authors [52].

3.1 The Class of Real $S\alpha S$ Distributions

The symmetric α-stable ($S\alpha S$) distribution is best defined by its characteristic function

$$\varphi(\omega) = \exp(\jmath\delta\omega - \gamma|\omega|^\alpha), \tag{3.1}$$

where α is the *characteristic exponent* restricted to the values $0 < \alpha \leq 2$, δ ($-\infty < \delta < \infty$) is the *location parameter*, and γ ($\gamma > 0$) is the *dispersion* of the distribution. For values of α in the interval $(1, 2]$, the location parameter δ corresponds to the mean of the $S\alpha S$

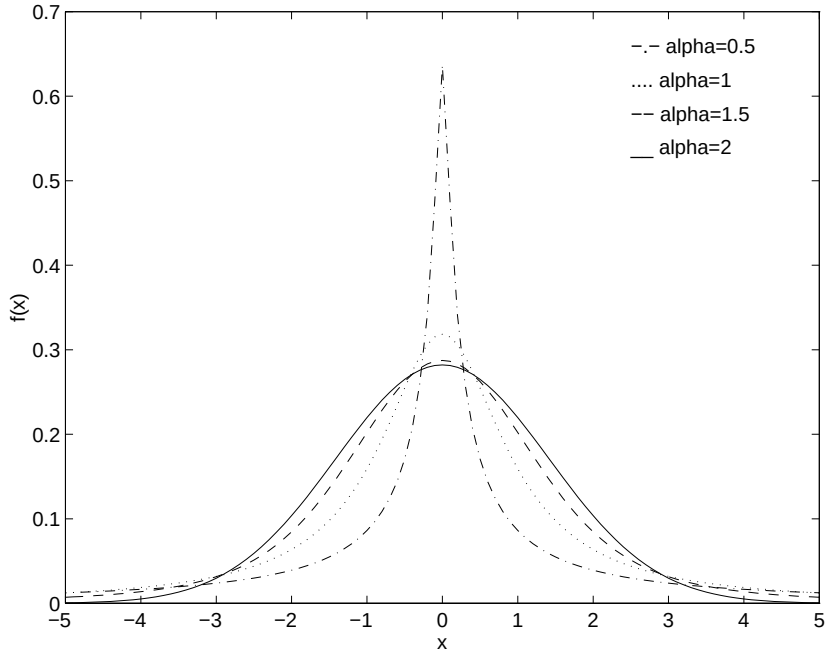


Figure 3.1: Standard $S\alpha S$ densities.

distribution, while for $0 < \alpha \leq 1$, δ corresponds to its median. The dispersion parameter γ determines the spread of the distribution around its location parameter δ, similar to the variance of the Gaussian distribution. The characteristic exponent α is the most important parameter of the $S\alpha S$ distribution and it determines the shape of the distribution.

A stable distribution is called *standard* if $\delta = 0$ and $\gamma = 1$. Clearly, if a random variable X is stable with parameters α, γ, δ, then $(X - \delta)/\gamma^{1/\alpha}$ is standard with characteristic exponent α. The standard $S\alpha S$ density functions for a few values of the characteristic exponent α are shown in Figure 3.1.

By letting α take the values 1 and 2, we get two important special cases of $S\alpha S$ distributions, namely, the *Cauchy* ($\alpha = 1$), and the *Gaussian* ($\alpha = 2$):

Cauchy

$$f_1(\gamma, \delta; x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \delta)^2} \tag{3.2}$$

Gaussian

$$f_2(\gamma, \delta; x) = \frac{1}{\sqrt{4\pi\gamma}} \exp\left[-\frac{(x - \delta)^2}{4\gamma}\right]. \tag{3.3}$$

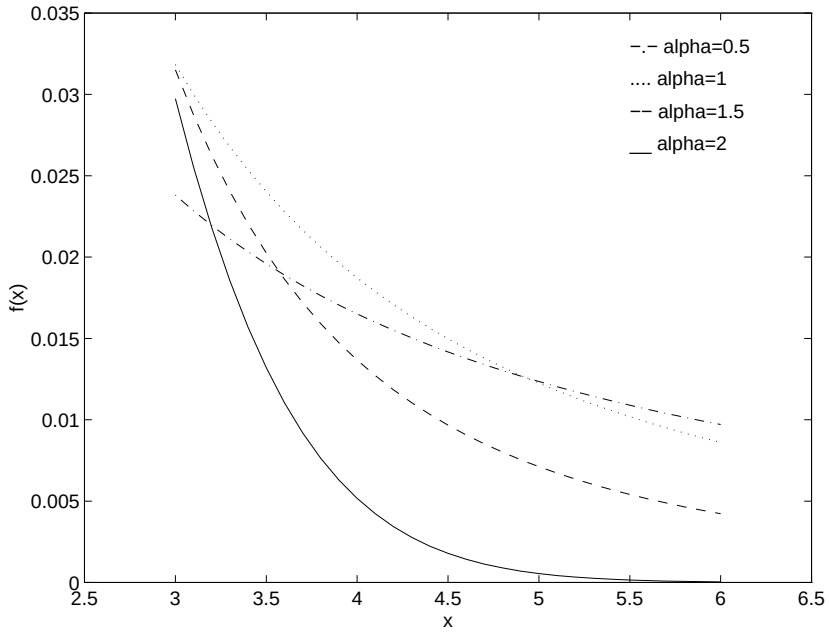


Figure 3.2: A close-up view of the tails of the densities in Figure 3.1.

Unfortunately, no closed form expressions exist for general $S\alpha S$ distributions other than the Cauchy and the Gaussian. However, power series expansions can be derived for $f_\alpha(\gamma, \delta; x)$. In the following, we shall assume that all $S\alpha S$ distributions are centered at the origin, i.e., the location parameter $\delta = 0$. This is equivalent to the zero-mean assumption for Gaussian distributions. Then, the standard $S\alpha S$ density function is given by [69]

$$f_\alpha(x) = \begin{cases} \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) x^{-\alpha k} \sin(\frac{k\alpha\pi}{2}) & \text{for } 0 < \alpha < 1 \\ \frac{1}{\pi(x^2+1)} & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k!} \Gamma(\frac{2k+1}{\alpha}) x^{2k} & \text{for } 1 < \alpha < 2 \\ \frac{1}{2\sqrt{\pi}} \exp\left[-\frac{x^2}{4}\right] & \text{for } \alpha = 2. \end{cases} \tag{3.4}$$

Although the $S\alpha S$ density behaves approximately like a Gaussian density near the origin, its tails decay at a lower rate than the Gaussian density tails. While the Gaussian density has exponential tails, the stable densities have algebraic tails (cf. Figure 3.2). *The smaller the characteristic exponent α is, the heavier the tails of the $S\alpha S$ density.* This implies that random variables following $S\alpha S$ distributions with small characteristic exponents are *highly impulsive*. It is this heavy-tail characteristic that makes the $S\alpha S$ densities appropriate for modeling signals and noise or interference which are impulsive in

nature. Figure 3.3 depicts some representative time series of $S\alpha S$ deviates for comparison purposes. Clearly, several outliers can be observed in the data series when the characteristic exponent α takes values other than 2. As α decreases both the occurrence rate and the strength of the outliers increase, resulting to very impulsive processes.

$S\alpha S$ densities obey two important properties which further justify their role in data modeling:

- The *stability property*, which states that the random variables $X_1, \ldots, X_n$ are independent and symmetrically stable with the same characteristic exponent α if and only if for any constants $a_1, \ldots, a_n$, the linear combination $\sum_{i=1}^{n} a_i X_i$ is also $S\alpha S$;
- The *generalized central limit theorem*, which states that the family of stable distributions contains all limiting distributions of sums of i.i.d. random variables.

An important difference between the Gaussian and the other distributions of the $S\alpha S$ family is that only moments of order less than α exist for the non-Gaussian $S\alpha S$ family members. The *fractional lower order moments* (FLOM's) of a $S\alpha S$ random variable with zero location parameter and dispersion γ are given by:

$$E|X|^p = C(p,\alpha)\gamma^{\frac{p}{\alpha}} \qquad \text{for } 0 < p < \alpha \tag{3.5}$$

where

$$C(p,\alpha) = \frac{2^{p+1}\Gamma(\frac{p+1}{2})\Gamma(-\frac{p}{\alpha})}{\alpha\sqrt{\pi}\Gamma(-\frac{p}{2})} \tag{3.6}$$

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

3.2 Bivariate Isotropic Stable Distributions

Multivariate stable distributions, much like the univariate stable distributions, are characterized by the stability property and the generalized central limit theorem. However, they are much more difficult to describe because they form a nonparametric set [69]. An exception is the family of multidimensional isotropic stable distributions. Here, we concentrate on the two dimensional (bivariate) case which is appropriate for modeling signals and noise in the bearing estimation problem.

The characteristic function of a bivariate isotropic α-stable distribution has the form

$$\varphi(\omega_1, \omega_2) = \exp(\jmath(\delta_1\omega_1 + \delta_2\omega_2) - \gamma|\boldsymbol{\omega}|^\alpha), \tag{3.7}$$

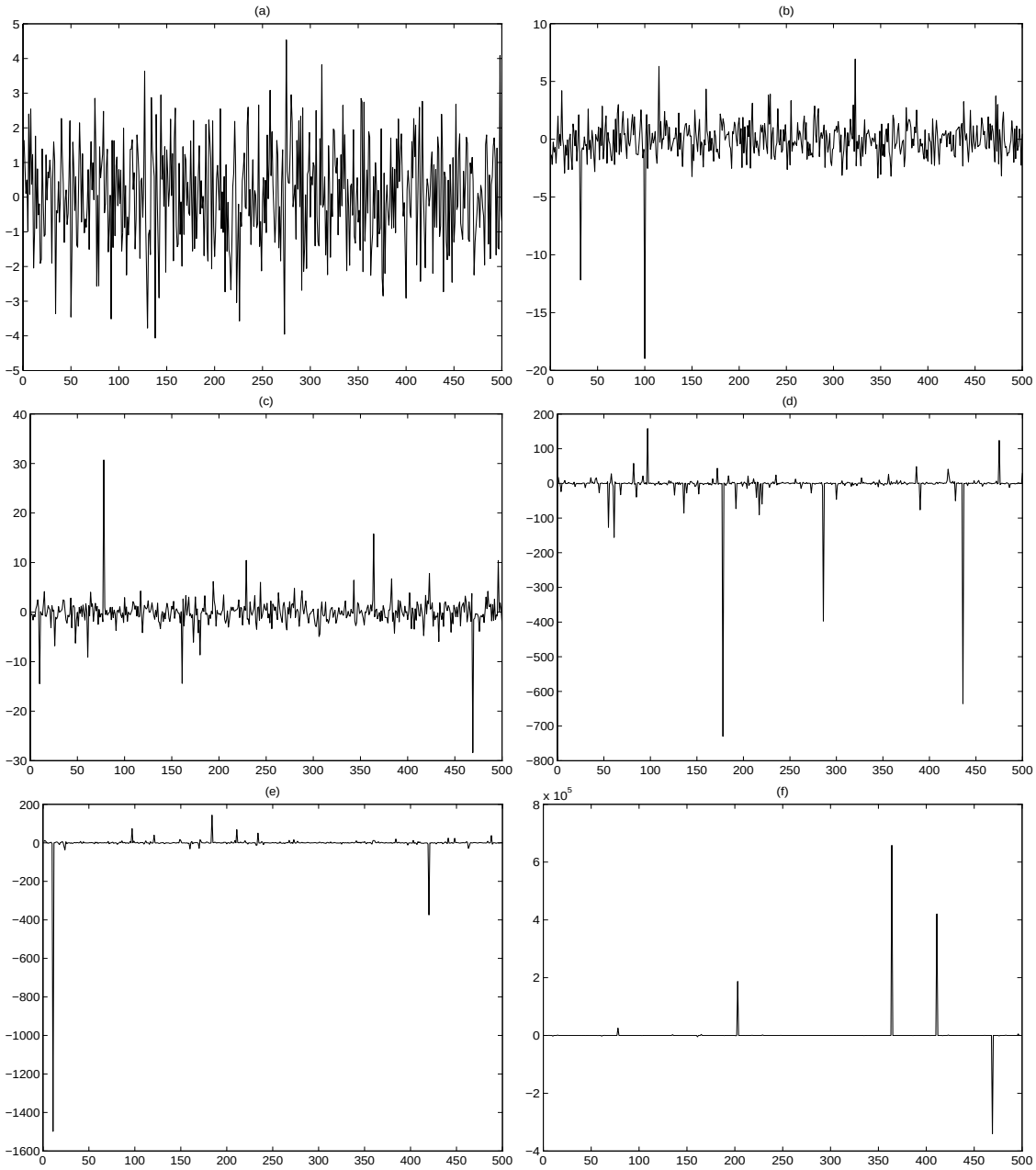


Figure 3.3: $S\alpha S$ time series. (a): $\alpha = 2.0$ (Gaussian), (b): $\alpha = 1.95$, (c): $\alpha = 1.5$, (d): $\alpha = 1.0$ (Cauchy), (e): $\alpha = 0.85$, (f): $\alpha = 0.45$.

where $\boldsymbol{\omega} = (\omega_1, \omega_2)$ and $|\boldsymbol{\omega}| = \sqrt{\omega_1^2 + \omega_2^2}$.

Again here, the parameters α and γ are the characteristic exponent and the dispersion, respectively. The parameters δ_1, δ_2 are the location parameters. The distribution is isotropic with respect to the point (δ_1, δ_2). Note that the two marginal distributions of the isotropic stable distribution are $S\alpha S$ with parameters $(\delta_1, \gamma, \alpha)$ and $(\delta_2, \gamma, \alpha)$. In the following, we will assume that $(\delta_1, \delta_2) = (0, 0)$. The bivariate *isotropic Cauchy* and *Gaussian* distributions are special cases for $\alpha = 1$ and $\alpha = 2$, respectively.

As in the case of the univariate $S\alpha S$ density function, when $\alpha \neq 1$ or $\alpha \neq 2$, no closed form expressions exist for the density function of the bivariate stable random variable. By using the polar coordinate $r = |\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2}$, the density function can be written as $f_{\alpha,\gamma}(x_1, x_2) = \chi_{\alpha,\gamma}(r)$, and can be expressed in a power series expansion form:

$$\chi_{\alpha,\gamma}(r) = \begin{cases} \frac{1}{\pi^2 \gamma^{2/\alpha}} \sum_{k=1}^{\infty} \frac{2^{\alpha k}(-1)^{k-1}}{k!} (\Gamma(\alpha k/2 + 1))^2 \sin(\frac{k\alpha\pi}{2}) (\frac{r}{\gamma^{1/\alpha}})^{-\alpha k - 2} & \text{for } 0 < \alpha < 1 \\ \frac{\gamma}{2\pi(r^2+\gamma^2)^{3/2}} & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha\gamma^{2/\alpha}} \sum_{k=0}^{\infty} \frac{(-1)^k}{2^{2k+1}(k!)^2} \Gamma(\frac{2k+2}{\alpha}) (\frac{r}{\gamma^{1/\alpha}})^{2k} & \text{for } 1 < \alpha < 2 \\ \frac{1}{4\pi\gamma} \exp\left[-\frac{r^2}{4\gamma}\right] & \text{for } \alpha = 2. \end{cases} \tag{3.8}$$

The density function $\chi_{\alpha,\gamma}(r)$ described above is also a heavy-tailed function. An expression for the FLOM's, similar to the one for the single-dimensional case, can be found in [98]. If X is a random vector in $\mathbf{R}^n$ having the isotropic stable distribution with dispersion γ, then

$$E|\mathbf{X}|^p = C_n(p, \alpha)\gamma^{\frac{p}{\alpha}} \qquad \text{for } 0 < p < \alpha \tag{3.9}$$

where

$$C_n(p, \alpha) = \frac{2^{p+1}\Gamma(\frac{p+n}{2})\Gamma(-\frac{p}{\alpha})}{\alpha\Gamma(\frac{n}{2})\Gamma(-\frac{p}{2})}. \tag{3.10}$$

3.3 Symmetric Alpha-Stable Processes

A collection of random variables $\{X(t),\ t \in T\}$ where T is an arbitrary index set, is said to be a $S\alpha S$ stochastic process if for all combinations of distinct indices $t_1, \ldots, t_n \in T$, the random variables $X(t_1), \ldots, X(t_n)$ are jointly $S\alpha S$ with the same characteristic exponent α. The stochastic process $\{X(t),\ t \in T\}$ is *stationary* if the random vectors $(X(t_1), \ldots, X(t_n))$ and $(X(t_1+s), \ldots, X(t_n+s))$ are identically distributed for each choice of $s, t_1, \ldots, t_n \in T$. The family of stable processes has many members with mutually

exclusive properties. In the following, we present three important types of stable processes that are commonly used.

1) Sub-Gaussian Processes: A stable process $\{X(t),\ t \in T\}$ is said to be an *α-sub-Gaussian* process (α-SG(R)) if for distinct indices $t_1, \ldots, t_n \in T$, $(X(t_1), \ldots, X(t_n))$ has characteristic function given by

$$\varphi(\boldsymbol{\omega}) = \exp(-[\frac{1}{2}\sum_{l,m=1}^{n} \omega_l \omega_m R(t_l, t_m)]^{\alpha/2}), \tag{3.11}$$

where $R(t,s)$ is a positive definite function, $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_n]^T$ and α takes values in $(1,2]$. When $\alpha = 2$, $X(t)$ is a Gaussian process with zero mean and covariance function $R(t,s)$. A sub-Gaussian process is stationary if and only if $R(t,s) = R(t-s) = R(s-t)$.

Sub-Gaussian processes share many common features with Gaussian processes. In fact, sub-Gaussian processes are variance mixtures of Gaussian processes [10]. Specifically, if $X(t)$ is α-SG(R), then

$$X(t) = S^{1/2} Y(t) \tag{3.12}$$

where S is a positive $\frac{\alpha}{2}$-stable random variable and independent of $Y(t)$ which is a Gaussian process with zero-mean and covariance function $R(t,s)$. An important distinction between Gaussian and sub-Gaussian processes is that, while linear spaces of Gaussian random variables may contain non-degenerate independent elements, sub-Gaussian random variables cannot be independent [12]. A sub-Gaussian process $X(t) = S^{1/2}Y(t)$ is stationary if and only if the Gaussian process $Y(t)$ is stationary.

2) Linear Stable Processes: Let $\{U(n),\ n = 0, \pm 1, \pm 2, \ldots\}$ be a family of i.i.d. $S\alpha S$ random variables. Then,

$$X(n) = \sum_{i=-\infty}^{+\infty} a_i U(n-i) \tag{3.13}$$

defines a stationary $S\alpha S$ random process if $\sum_{-\infty}^{+\infty} |a_i|^{\alpha-\delta} < \infty$ for some $0 < \delta < \alpha$ when $0 < \alpha < 1$, or if $\sum_{-\infty}^{+\infty} |a_i| < \infty$ when $\alpha \geq 1$. These processes are called *linear* stable processes or stable processes with moving-average representation. Examples of linear stable processes include finite-order autoregressive (AR), moving-average (MA) and autoregressive moving-average ($ARMA$) processes [69].

3) Harmonizable Stable Processes: A complex valued $S\alpha S$ process $X(t)$ is called *harmonizable* if it can be expressed as:

$$X(t) = \int_{-\infty}^{\infty} e^{\jmath t\omega}\, d\xi(\omega); \quad -\infty < t < \infty, \tag{3.14}$$

where $d\xi(\omega)$ is a $S\alpha S$ process with independent increments satisfying

$$\{E|d\xi(\omega)|^p\}^{\alpha/p} = C(p,\alpha)\phi(\omega)\, d\omega \quad \text{for all } \ 0 < p < \alpha, \tag{3.15}$$

and $C(p,\alpha)$ is a constant depending on p and α and $\phi(\omega)$ is a nonnegative function called the *spectral density* of $X(t)$ [27]. The spectral density $\phi(\omega)$ describes fully the distribution of the process $X(t)$. In another sharp contrast with the Gaussian case, the class of $S\alpha S$ harmonizable processes is disjoint from the class of linear processes. In addition, sub-Gaussian processes are neither linear nor harmonizable [12].

In modeling the signals and/or noise for the parameter estimation problem, we need a complex model for the noise samples. We also need to define quantities which describe correlations between random variables. In the following section, we present the family of isotropic complex $S\alpha S$ distributions and describe their correlation properties by means of the *covariation* quantity.

3.4 Complex $S\alpha S$ Random Variables and Covariations

A complex random variable (r.v.) $X = X_1 + \jmath X_2$ is symmetric α-stable ($S\alpha S$) if X_1 and X_2 are jointly $S\alpha S$ and its characteristic function is written as

$$\begin{aligned} \varphi(\omega) &= E\exp[\jmath\Re(\omega X^*)] = E\exp[\jmath(\omega_1 X_1 + \omega_2 X_2)] \\ &= \exp\left[-\int_{S_2} |\omega_1 x_1 + \omega_2 x_2|^\alpha d\Gamma_{X_1,X_2}(x_1,x_2)\right], \end{aligned} \tag{3.16}$$

where $\omega = \omega_1 + \jmath\omega_2$, $\Re[\cdot]$ is the real part operator, and Γ_{X_1,X_2} is a symmetric measure on the unit sphere S_2, called the *spectral measure* of the random variable X.

A complex random variable $X = X_1 + \jmath X_2$ is *isotropic* if and only if (X_1, X_2) has a uniform spectral measure. In this case, the characteristic function of X can be written as

$$\varphi(\omega) = E\exp(\jmath\Re[\omega X^*]) = \exp(-\gamma|\omega|^\alpha), \tag{3.17}$$

where γ ($\gamma > 0$) is the *dispersion* of the distribution.

In the theory of second-order processes, the concept of *covariance* plays an important role in problems of linear prediction, filtering and smoothing. Since $S\alpha S$ processes do not possess finite pth order moments for $p \geq \alpha$, covariances do not exist on the space of $S\alpha S$ random variables. Instead, a quantity called *covariation* plays an analogous role for statistical signal processing problems involving $S\alpha S$ processes to the role played by covariance in the case of second-order processes.

Several complex r.v.'s are jointly $S\alpha S$ if their real and imaginary parts are jointly $S\alpha S$. When $X = X_1 + \jmath X_2$ and $Y = Y_1 + \jmath Y_2$ are jointly $S\alpha S$ with $1 < \alpha \leq 2$, the *covariation* of X and Y is defined by

$$[X, Y]_\alpha = \int_{S_4} (x_1 + \jmath x_2)(y_1 + \jmath y_2)^{<\alpha-1>} d\Gamma_{X_1,X_2,Y_1,Y_2}(x_1, x_2, y_1, y_2), \tag{3.18}$$

where we use throughout the convention

$$Y^{<\beta>} = |Y|^{\beta-1} Y^*. \tag{3.19}$$

It can be shown that for every $1 \leq p < \alpha$, the covariation can be expressed as a function of moments [9]

$$[X, Y]_\alpha = \frac{E X Y^{<p-1>}}{E|Y|^p} \gamma_Y, \tag{3.20}$$

where γ_Y is the dispersion of the r.v. Y given by

$$\gamma_Y^{p/\alpha} = \frac{E|Y|^p}{C(p, \alpha)} \qquad \text{for } 0 < p < \alpha, \tag{3.21}$$

with

$$C(p, \alpha) = \frac{2^{p+1} \Gamma(\frac{p+2}{2}) \Gamma(-\frac{p}{\alpha})}{\alpha \Gamma(\frac{1}{2}) \Gamma(-\frac{p}{2})}. \tag{3.22}$$

Obviously, from (3.20) it holds that

$$[X, X]_\alpha = \gamma_X. \tag{3.23}$$

Also, the *covariation coefficient* of X and Y is defined by

$$\lambda_{X,Y} = \frac{[X, Y]_\alpha}{[Y, Y]_\alpha}, \tag{3.24}$$

and by using (3.20), it can be expressed as

$$\lambda_{X,Y} = \frac{EXY^{<p-1>}}{E|Y|^p} \qquad \text{for } 1 \leq p < \alpha. \tag{3.25}$$

The covariation of complex jointly $S\alpha S$ r.v.'s is not generally symmetric and has the following properties [6]:

P1 If X_1, X_2 and Y are jointly $S\alpha S$, then

$$[aX_1 + bX_2, Y]_\alpha = a[X_1, Y]_\alpha + b[X_2, Y]_\alpha \tag{3.26}$$

for any complex constants a and b.

P2 If Y_1 and Y_2 are independent and X_1, X_2 and Y are jointly $S\alpha S$, then

$$[aX_1, bY_1 + cY_2]_\alpha = ab^{<\alpha-1>}[X_1, Y_1]_\alpha + ac^{<\alpha-1>}[X_1, Y_2]_\alpha \tag{3.27}$$

for any complex constants a, b and c.

P3 If X and Y are independent $S\alpha S$, then $[X, Y]_\alpha = 0$

P4 Let $\{U_i;\ i = 1, \ldots, n\}$ be independent complex $S\alpha S$ r.v.'s with dispersions γ_i. For any complex numbers $\{a_i, b_i;\ i = 1, \ldots, n\}$ form

$$X = a_1U_1 + \cdots + a_nU_n,$$

$$Y = b_1U_1 + \cdots + b_nU_n.$$

Then

$$[X, X]_\alpha = \gamma_1|a_1|^\alpha + \cdots + \gamma_n|a_n|^\alpha,$$

$$[Y, Y]_\alpha = \gamma_1|b_1|^\alpha + \cdots + \gamma_n|b_n|^\alpha,$$

$$[X, Y]_\alpha = \gamma_1 a_1 b_1^{<\alpha-1>} + \cdots + \gamma_n a_n b_n^{<\alpha-1>}. \tag{3.28}$$

3.5 Generation of Complex Isotropic $S\alpha S$ Random Variables

The generation of complex isotropic $S\alpha S$ deviates of characteristic exponent α is based on the following proposition found in [64]:

Proposition 3.1 *A complex $S\alpha S$ ($\alpha < 2$) random variable $X = X_1 + \jmath X_2$ is isotropic if and only if there exist two i.i.d. zero-mean Gaussian variables G_1 and G_2 and a real stable random variable A of characteristic exponent $\alpha/2$, dispersion $\cos^2(\pi\alpha/4)$ and skewness $\beta = 1$ (we write $A \sim S_{\alpha/2}(\cos^2(\pi\alpha/4), 1)$), independent of (G_1, G_2) such that $(X_1, X_2) \stackrel{d}{=} (A^{1/2}G_1, A^{1/2}G_2)$.*

We say that the vector (X_1, X_2) is *sub-Gaussian* with underlying vector (G_1, G_2). It can be shown that the real and imaginary parts of X are always dependent, unless G_1 and G_2 are degenerate. Hence, every complex isotropic $S\alpha S$ random variable with $\alpha < 2$ can be expressed as

$$X = A^{1/2}(G_1 + \jmath G_2), \tag{3.29}$$

and its generation involves the generation of a real, totally skewed stable random variable. The problem of generating a real stable deviate is studied in [14] and [32]. Here, we present the result for easy reference.

To generate a real standard stable random variable $A \sim S_\alpha(1, \beta)$ of characteristic exponent α, skewness β and unit dispersion $\gamma = 1$, the following representations can be deduced:

$$S(\alpha, \beta, 1) = D_{\alpha,\beta}\frac{\sin\alpha(U - U_0)}{(\cos U)^{1/\alpha}}\left(\frac{\cos(U - \alpha(U - U_0))}{W}\right)^{\frac{1-\alpha}{\alpha}}, \quad \text{for } \alpha \neq 1, \tag{3.30}$$

and

$$S(1, \beta, 1) = \frac{2}{\pi}\left[(\frac{\pi}{2} + \beta U)\tan U - \beta\ln(\frac{\frac{\pi}{2}W\cos U}{\frac{\pi}{2} + \beta U})\right], \tag{3.31}$$

where W is standard exponential with $Pr\{W > w\} = e^{-w}$, $w > 0$, and U is uniform on $(-\frac{\pi}{2}, \frac{\pi}{2})$. Also, $D_{\alpha,\beta} = [\cos(\arctan(\beta\tan(\pi\alpha/2)))]^{-1/\alpha}$, and $U_0 = -\frac{\pi}{2}\beta[k(\alpha)/\alpha]$ with $k(\alpha) = 1 - |1 - \alpha|$. Then, a stable variate, A_1, of dispersion γ can be obtained from A by $A_1 = \gamma^{1/\alpha}A$.

To conclude, the following proposition gives the relationship between the dispersion γ of the complex r.v. $X = X_1 + \jmath X_2$ and the variance σ of the complex Gaussian r.v. $G = G_1 + \jmath G_2$.

Proposition 3.2 *The dispersion γ of the complex r.v. $X = X_1 + \jmath X_2$ generated according to Proposition 3.1 is given by $\gamma = (\sigma/2)^\alpha$, where σ is the variance of the underlying complex Gaussian random variable.*

Proof The Laplace transform of the r.v. $A \sim S_{\alpha/2}(\cos^2(\pi\alpha/4), 1)$ is given by [64]:

$$E\{\exp(-sA)\} = \exp(-s^{\alpha/2}), \quad s > 0. \tag{3.32}$$

Also, since $G = G_1 + \jmath G_2 \sim N_C(\sigma, 0)$, its characteristic function is given by

$$\varphi_G(\omega) = \exp(-\sigma^2|\omega|^2/4), \tag{3.33}$$

where $\omega = \omega_1 + \jmath\omega_2$. Then, the characteristic function of X can be expressed as

$$\begin{aligned}
\varphi_X(\omega) &= E\{\exp(\jmath\Re[\omega X^*])\} \\
&= E\{\exp(\jmath(\omega_1 A^{1/2} G_1 + \omega_2 A^{1/2} G_2))\} \\
&= E\{E\{\exp(\jmath(\omega_1 A^{1/2} G_1 + \omega_2 A^{1/2} G_2))|A\}\} \\
&= E\{\exp(-\sigma^2(\omega_1^2 + \omega_2^2)A/4)\} \quad \text{[by use of 3.33]} \\
&= \exp((\sigma/2)^\alpha |\omega|^\alpha) \quad \text{[by use of 3.32].}
\end{aligned} \tag{3.34}$$

Comparing (3.34) with (3.7) we conclude that $\gamma = (\sigma/2)^\alpha$. ∎

Chapter 4

Maximum Likelihood DOA Estimation in Alpha-Stable Noise

In this chapter, we develop the Maximum Likelihood estimator of the source locations in the presence of noise modeled as a *complex isotropic α-stable process* with dispersion γ. Initially, in Section 4.1, we concentrate on the case where $\alpha = 1$, i.e., we consider the additive complex Cauchy noise case. There are two reasons for doing this: First, the Cauchy distribution has a closed-form expression for its density function. This results in a straight-forward implementation of the maximum likelihood estimation, with closed form expressions for the Cramér-Rao bound derived in Section 4.2. Second, it is shown through simulations in Section 4.4 that the Cauchy beamformer is very robust in different impulsive noise environments, i.e., its performance does not change significantly when the parameter α of the $S\alpha S$ noise varies in the interval $[1, 2]$. When the noise follows the $S\alpha S$ distribution with $a \neq 1$, the maximum likelihood estimation of the direction of arrival is approximated by exploring the simple expression for the characteristic function of the noise given by (3.7). This approach is followed in Section 4.3.

In the following, in a similar approach as in [97], we do not regard the source signals as sample functions of random processes but rather as *unknown deterministic sequences*. We will see that under the unknown signal assumption, the ML estimation of the source bearings cannot be decoupled from the ML estimation of the signal waveforms as it can be in the case of Gaussian additive noise. Hence, in order to reduce the dimensionality of the optimization procedure, we will use suboptimal estimates of the signals, refering to the resulting processor as the pseudo maximum likelihood processor.

4.1 Pseudo ML Bearing Estimation of Sources in Cauchy Noise

In this section, we assume that the noise present at the array sensors is modeled as a *complex isotropic Cauchy process* with pdf given by (3.8, $\alpha = 1$). We are most interested in estimating the bearings of the sources, and therefore we consider the signals themselves, as well as the noise dispersion, γ, as nuisance parameters in the estimation problem.

Under assumption A.4, it follows from (2.1) and (3.8) that the joint density function of the sampled data is given by

$$f(\mathbf{X}) = \prod_{t=1}^{M} \prod_{i=1}^{r} \chi_{1,\gamma} \left(\left| x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t) \right| \right) \tag{4.1}$$

or

$$f(\mathbf{X}) = \prod_{t=1}^{M} \prod_{i=1}^{r} \frac{1}{2\pi} \frac{\gamma}{\left(\gamma^2 + \left| x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t) \right|^2 \right)^{3/2}}, \tag{4.2}$$

where $a_1(\theta_k) = 1$ and $a_i(\theta_k) = e^{-\jmath\omega\tau_i(\theta_k)}$; $i = 2, \ldots, r$. Hence the log likelihood function $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$, ignoring constant terms, is expressed as

$$L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = Mr \log(\gamma) - \frac{3}{2} \sum_{t=1}^{M} \sum_{i=1}^{r} \log \left(\gamma^2 + \left| x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t) \right|^2 \right). \tag{4.3}$$

The ML estimator is obtained by maximizing $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to γ, $\mathbf{S}$, and $\boldsymbol{\theta}$, i.e.,

$$\max_{\gamma, \mathbf{S}, \boldsymbol{\theta}} L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}). \tag{4.4}$$

To reduce the dimension of this optimization problem, we first fix γ and $\boldsymbol{\theta}$, and minimize $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to the signal $\mathbf{S}$. For fixed t, we take the derivative of $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to $s_k(t)$:

$$\frac{\partial L}{\partial s_k(t)} = -3 \sum_{i=1}^{r} \frac{a_i(\theta_k) \left[x_i(t) - a_i(\theta_k) s_k(t) \right]^*}{\gamma^2 + |x_i(t) - a_i(\theta_k) s_k(t)|^2}. \tag{4.5}$$

Unfortunately, no explicit solution of (4.5) is possible. In order to be able to obtain closed form expressions for the signals, we resort to the application of the pseudo-maximum likelihood (PML) estimation. PML estimation is an important method in applications

where probability models abound for which the analytical derivation of the maximum likelihood estimate for all the parameters is virtually impossible. The problem formulated in [24] can be stated as follows:

> *Let $X_1, \ldots, X_n$ be i.i.d. random variables with probability distribution $f(X; \boldsymbol{\theta}, \mathbf{S})$ indexed by two sets of parameters. Let $\hat{\mathbf{S}} = \hat{\mathbf{S}}(X_1, \ldots, X_n)$ be an estimate of $\mathbf{S}$ other than the maximum likelihood estimate, and let $\hat{\boldsymbol{\theta}}$ be the solution of the likelihood equation $\partial/\partial\boldsymbol{\theta}\ \log L(\mathbf{X}; \boldsymbol{\theta}, \hat{\mathbf{S}}) = 0$ which maximizes the likelihood. Then, $\hat{\boldsymbol{\theta}}$ is called a pseudo maximum likelihood estimate of $\boldsymbol{\theta}$, and under certain conditions it is consistent and asymptotically normal.*

The PML estimator $\hat{\boldsymbol{\theta}}$ has good large sample properties when $\hat{\mathbf{S}}$ does. In general, the asymptotic analysis for $\hat{\boldsymbol{\theta}}$ will depend on the asymptotic characteristics of $\hat{\mathbf{S}}$.

Returning to the optimization problem described in (4.4), we observe that maximizing $L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ with respect to the signal $\mathbf{S}$ is equivalent to the following minimization problem:

$$\min_{\mathbf{S}} \mathcal{L}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = \min_{\mathbf{S}} \left\{ \sum_{t=1}^{M} \sum_{i=1}^{r} \log(1 + \frac{|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)|^2}{\gamma^2}) \right\}. \tag{4.6}$$

As we can see, (4.6) involves minimizing a double sum expression of logarithmic functions of the form $\log(1+z)$. In the unit disc $B_1(0) = \{z \in \mathcal{C} : |z| < 1\}$, the function $\log(1+z)$ can be expressed as an infinite series:

$$\log(1+z) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} z^n = z - \frac{z^2}{2} + - \cdots; \quad |z| < 1 \tag{4.7}$$

Hence, for $|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)| < \gamma$ the functional $\mathcal{L}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ can be written in the form

$$\mathcal{L}(\mathbf{X}; \boldsymbol{\theta}, \mathbf{S}) = \sum_{t=1}^{M} \sum_{i=1}^{r} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n\gamma^{2n}} |x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)|^{2n}. \tag{4.8}$$

A first order approximation of the above expression results in the following $\mathcal{L}^{(1)}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})$ functional:

$$\mathcal{L}^{(1)}(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta}) = \frac{1}{\gamma^2} \sum_{t=1}^{M} \sum_{i=1}^{r} |x_i(t) - \sum_{k=1}^{q} a_i(\theta_k) s_k(t)|^2 \tag{4.9}$$

which, by using (2.2), can be written in a more compact form as:

$$\mathcal{L}^{(1)}(\mathbf{X};\gamma,\mathbf{S},\boldsymbol{\theta}) = \frac{1}{\gamma^2}\sum_{t=1}^{M}|\mathbf{x}(t)-\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)|^2. \tag{4.10}$$

Hence, the minimization of $\mathcal{L}^{(1)}(\mathbf{X};\gamma,\mathbf{S},\boldsymbol{\theta})$ with respect to $\mathbf{S}$ is equivalent to the Least-Squares (LS) estimation of $\mathbf{S}$. This problem has a well-known solution:

$$\hat{\mathbf{s}}(t) = (\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta}))^{-1}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{x}(t). \tag{4.11}$$

The dispersion γ can be estimated by using the method of moments. Namely, in the expression for the FLOM of the noise given in (3.9), $E|X|^p$ can be approximated by an average sum:

$$\hat{\gamma} = \frac{\left[\frac{1}{M}\sum_{t=1}^{M}\sum_{i=1}^{r}|x_i(t)-\sum_{k=1}^{q}a_i(\theta_k)\hat{s}_k(t)|^p\right]^{\frac{1}{p}}}{[C_2(p,1)]^{\frac{1}{p}}}, \tag{4.12}$$

where $p<1$, and $C_2(p,1)$ is given by:

$$C_2(p,1) = p2^p\frac{\Gamma(\frac{p}{2})\Gamma(-p)}{\Gamma(-\frac{p}{2})}. \tag{4.13}$$

By using the above estimates for the signal $\mathbf{S}$ and the noise dispersion γ, we obtain the following reduced optimization problem:

$$\max_{\boldsymbol{\theta}} L(\mathbf{X};\hat{\gamma},\hat{\mathbf{S}},\boldsymbol{\theta}) = \max_{\boldsymbol{\theta}}\left\{Mr\log(\hat{\gamma}) - \frac{3}{2}\sum_{t=1}^{M}\sum_{i=1}^{r}\log\left(\hat{\gamma}^2+\left|x_i(t)-\sum_{k=1}^{q}a_i(\theta_k)\hat{s}_k(t)\right|^2\right)\right\}. \tag{4.14}$$

An iterative procedure based on the gradient descent principle can be applied in order to solve for $\boldsymbol{\theta}$. In general, the cost function described in (4.14) is non-convex and the optimization procedure has to be initialized sufficiently close to the global extremum. Sub-optimal DOA estimators such as the MUSIC algorithm or the ROC-MUSIC method [76] introduced in the next section, can be used to obtain initial bearing estimates. Concluding this section we point out, once again, the two main assumptions made in order to obtain a closed-form expression for the signal estimate, $\hat{\mathbf{S}}$:

B.1 Assumption $|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k)s_k(t)| < \gamma$ (holds with probability $(\sqrt{2}-1)/\sqrt{2}$) enabled us to express the logarithmic function as an infinite series;

B.2 Assumption $|x_i(t) - \sum_{k=1}^{q} a_i(\theta_k)s_k(t)| \ll \gamma$ enabled us to take a first order approximation of the infinite series, and thus to obtain a least-squares closed form estimate for the signal.

Obviously, when assumptions B.1 and B.2 are not satisfied, the pseudo maximum likelihood processor will suffer from suboptimal signal estimation. For the general application case, when the processor does not possess enough information about the transmitted signals, the two assumptions B.1 and B.2 provide a closed form expression for the signal estimates. The degree that these assumptions affect the performance of the processor is studied in the simulated experiments, in Section 4.4.

4.2 The Cramér-Rao Bound for Cauchy Noise

Under the assumptions stated in Section 2.1 and for the case of complex isotropic Cauchy noise, the following theorem holds:

Theorem 4.1 *The CRB for $\boldsymbol{\theta}$, and γ is given by*

$$CRB(\boldsymbol{\theta}) = \frac{5\gamma^2}{3}\left\{\sum_{t=1}^{M} \Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\left[I - \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\right]\mathbf{D}\mathbf{S}(t)\right\}\right\}^{-1} \tag{4.15}$$

and

$$CRB(\gamma) = \frac{5}{4}\frac{\gamma^2}{Mr}, \tag{4.16}$$

where

$$\mathbf{S}(t) = \begin{bmatrix} s_1(t) & 0 & \cdots & 0 & 0 \\ 0 & s_2(t) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{q-1}(t) & 0 \\ 0 & 0 & \cdots & 0 & s_q(t) \end{bmatrix}; \quad t = 1, \ldots, M, \tag{4.17}$$

$$\mathbf{D} = [\mathbf{d}(\theta_1), \ldots, \mathbf{d}(\theta_q)], \tag{4.18}$$

and $\mathbf{d}(\theta_i) = \partial\mathbf{a}(\theta_i)/\partial\theta_i;\ \ i = 1, \ldots, q.$

Proof See Appendix A.

The above expression for the CRB is very similar to the one presented in Section 2.2.3 for the case of additive Gaussian noise. We should note that the above bound can be achieved only when there exist unbiased estimators of all the model parameters γ, $\mathbf{S}$, and $\boldsymbol{\theta}$. A useful insight on the CRB can be gained if we consider the case of a single source ($q = 1$) impinging from direction θ in a linear array whose sensors are spaced a half-wavelength apart. In this case,

$$\mathbf{A} = [1, e^{-\jmath\pi \sin(\theta)}, \ldots, e^{-\jmath(r-1)\pi \sin(\theta)}]^T, \tag{4.19}$$

and

$$\mathbf{D} = [0, -\jmath\pi \cos(\theta) e^{-\jmath\pi \sin(\theta)}, \ldots, -\jmath(r-1)\pi \cos(\theta) e^{-\jmath(r-1)\pi \sin(\theta)}]^T. \tag{4.20}$$

So, it holds

$$\mathbf{A}^H \mathbf{A} = r, \tag{4.21}$$

$$\mathbf{D}^H \mathbf{D} = \pi^2 \frac{r(r-1)(2r-1)}{6} \cos^2(\theta), \tag{4.22}$$

and

$$\mathbf{D}^H \mathbf{A} = \jmath\pi \frac{r(r-1)}{2} \cos(\theta). \tag{4.23}$$

Then,

$$CRB(\theta) = \frac{20}{\pi^2} \cdot \frac{1}{r(r^2-1)\cos^2(\theta)} \cdot \frac{\gamma^2}{\sum_{t=1}^{M} |s(t)|^2}. \tag{4.24}$$

The term $\frac{\gamma^2}{\sum_{t=1}^{M} |s(t)|^2}$ in the above expression for the CRB can be viewed as the inverse of a quantity analogous to the signal-to-noise ratio (SNR) for the Gaussian case, i.e., a generalized SNR, so to speak. The larger the dispersion γ of the noise, the higher the CRB.

4.3 ML Estimation for General α-Stable Noise

In this section, we discuss the maximum likelihood estimation problem in the presence of additive stable noise with characteristic exponent α in the interval $(1, 2)$. In this case, there are no closed-form expressions but only power series expansions for the noise distributions as shown in (3.8). Here, we present an approximate solution to the ML estimation problem based on the characteristic function of the bivariate α-stable process given by (3.7).

The problem of approximating likelihood calculations has been studied in the past for applications where probability distributions are only conveniently represented by their transforms, like the characteristic function (cf. [3] and the references contained therein). A natural approach for estimating density ($p(x;\theta)$) based functions, like the likelihood function $L(x;\theta) = \log p(x;\theta)$, the score function $S(x;\theta) = \frac{\partial \log p(x;\theta)}{\partial \theta}$, and the Fisher information $\mathcal{I}(\theta) = \mathrm{var}\{S(x;\theta)\}$, is based on prediction theory: The best linear approximation of a function $h(x)$ of a random variable in terms of a kernel class of functions $G = \{g(\omega_i, x);\ \ \omega_i \in (\omega_1, \ldots, \omega_k)\}$ is defined to be $\hat{h}(x) = \sum_{i=1}^{k} d_i g(\omega_i, x)$ minimizing the mean-square error $E\{(\hat{h}(x) - h(x))^2\}$. The well-known solution for this problem is given by:

$$\hat{h}(x) = E\{h(x)\} + \boldsymbol{\lambda}^H \boldsymbol{\Sigma}^{-1}[\mathbf{g}(x) - \boldsymbol{\gamma}], \tag{4.25}$$

where $\mathbf{g}(x) = [g(\omega_1, x), \ldots, g(\omega_k, x)]^T$, $\boldsymbol{\lambda} = \mathrm{cov}\{h(x), \mathbf{g}(x)\}$, $\boldsymbol{\Sigma} = \mathrm{var}\{\mathbf{g}(x)\}$, and $\boldsymbol{\gamma} = E\{\mathbf{g}(x)\}$.

Approximations of the score function $S(x;\theta)$ and the Fisher information $\mathcal{I}(\theta)$ based on the above expression, can be obtained as follows [3]. Under mild regularity conditions,

$$E_\theta\{S(x;\theta)g(x)\} = \frac{\partial E_\theta\{g(x)\}}{\partial \theta} \tag{4.26}$$

for any function $g(x)$. Furthermore, $E_\theta\{S(x;\theta)\} = 0$, and hence $\boldsymbol{\lambda} = \mathrm{cov}\{S(x;\theta), \mathbf{g}(x)\} = \frac{\partial \boldsymbol{\gamma}(\theta)}{\partial \theta} = \dot{\boldsymbol{\gamma}}(\theta)$, where $\boldsymbol{\gamma}(\theta) = E_\theta\{\mathbf{g}(x)\}$. By using (4.25), we find that for any set of functions $\mathbf{g}(x)$, an approximation to the score function is given by:

$$\hat{S}(x;\theta) = \dot{\boldsymbol{\gamma}}(\theta)^H \boldsymbol{\Sigma}^{-1}(\theta)[\mathbf{g}(x) - \boldsymbol{\gamma}(\theta)]. \tag{4.27}$$

According to standard arguments, the approximate expression for the Fisher information is given by:

$$\mathcal{I}(\theta) = \dot{\boldsymbol{\gamma}}(\theta)^H \boldsymbol{\Sigma}^{-1}(\theta)\dot{\boldsymbol{\gamma}}(\theta). \tag{4.28}$$

Returning to the bearing estimation problem in the presence of additive complex isotropic α-stable noise, we consider the case of a single source impinging on the array from direction θ. Then, the likelihood function is given by

$$L(\mathbf{X};\gamma,\mathbf{S},\theta) = \sum_{t=1}^{M}\sum_{i=1}^{r} \log \chi_{\alpha,\gamma}\left(|x_i(t) - a_i(\theta)s(t)|\right). \tag{4.29}$$

The approximation to the maximum likelihood equations based on the data samples at the array sensors can be written as

$$\hat{S}(\mathbf{X};\gamma,\mathbf{S},\theta)=\sum_{t=1}^{M}\sum_{i=1}^{r}\hat{S}_{(t,i)}(x_i(t);\gamma,s(t),\theta)=0. \tag{4.30}$$

By choosing the kernel functions to be $g(\omega,x)=\exp[\jmath\Re\{\omega x\}]$, the approximation of the score functions $\hat{S}_{(t,i)}(x_i(t);\gamma,s(t),\theta)$ will depend on the characteristic functions $\phi_{(t,i)}(\omega;\gamma,s(t),\theta)$. Using (3.7), we can write

$$\phi_{(t,i)}(\omega;\gamma,s(t),\theta)=\exp[\jmath(\omega_{\Re}\Re\{a_i(\theta)s(t)\}+\omega_{\Im}\Im\{a_i(\theta)s(t)\})-\gamma|\omega|^{\alpha}], \tag{4.31}$$

where $\omega=\omega_{\Re}+\jmath\omega_{\Im}$, and $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary part operators. Evaluation in a complex grid $\mathbf{\Omega}=\boldsymbol{\omega}_{\Re}\times\boldsymbol{\omega}_{\Im}$ where $\boldsymbol{\omega}_{\Re}=\boldsymbol{\omega}_{\Im}=[-k\tau,\ldots,-\tau,\tau,\ldots,k\tau]^T$ and application of (4.27) give the following expression for $\hat{S}(\mathbf{X};\theta)$:

$$\hat{S}(\mathbf{X};\theta)=\sum_{t=1}^{M}\sum_{i=1}^{r}[\jmath\mathbf{D}_{\mathbf{\Omega}}(\theta)\boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega},\theta)]^H\mathbf{\Sigma}_{\mathbf{\Omega}}^{-1}(\theta)[\mathbf{g}(\mathbf{\Omega},x_i(t))-\boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega},\theta)]. \tag{4.32}$$

Also, because of the independence of the observed data, and by using (4.28), the approximate Fisher information is given by:

$$\mathcal{I}(\theta)=\sum_{t=1}^{M}\sum_{i=1}^{r}\boldsymbol{\phi}_{(t,i)}^{H}(\mathbf{\Omega},\theta)\,\mathbf{D}_{\mathbf{\Omega}}^{H}(\theta)\,\mathbf{\Sigma}_{\mathbf{\Omega}}^{-1}(\theta)\,\mathbf{D}_{\mathbf{\Omega}}(\theta)\,\boldsymbol{\phi}_{(t,i)}(\mathbf{\Omega},\theta). \tag{4.33}$$

The vectors and matrices in (4.32) and (4.33) are defined as follows:

$$\mathbf{g}(\mathbf{\Omega},x_i(t))=\begin{bmatrix}\exp\left[\jmath\left(-k\tau x_{i,\Re}(t)-k\tau x_{i,\Im}(t)\right)\right]\\ \exp\left[\jmath\left(-k\tau x_{i,\Re}(t)+(-k+1)\tau x_{i,\Im}(t)\right)\right]\\ \vdots\\ \exp\left[\jmath\left(-k\tau x_{i,\Re}(t)+k\tau x_{i,\Im}(t)\right)\right]\\ \exp\left[\jmath\left((-k+1)\tau x_{i,\Re}(t)-k\tau x_{i,\Im}(t)\right)\right]\\ \vdots\\ \exp\left[\jmath\left(k\tau x_{i,\Re}(t)+k\tau x_{i,\Im}(t)\right)\right]\end{bmatrix}_{4k^2\times 1}, \tag{4.34}$$

where $x_{i,\Re}(t) = \Re\{x_i(t)\}$ and $x_{i,\Im}(t) = \Im\{x_i(t)\}$. Also,

$$\phi_{(\mathrm{t,i})}(\mathbf{\Omega},\theta) = \begin{bmatrix} \exp[\jmath(-k\tau\Re\{a_i(\theta)s(t)\} - k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[(-k\tau)^2 + (-k\tau)^2]^{\frac{\alpha}{2}}] \\ \exp[\jmath(-k\tau\Re\{a_i(\theta)s(t)\} + (-k\tau+1)\Im\{a_i(\theta)s(t)\}) - \gamma[(-k\tau)^2 + ((-k+1)\tau)^2]^{\frac{\alpha}{2}}] \\ \vdots \\ \exp[\jmath(-k\tau\Re\{a_i(\theta)s(t)\} + k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[(-k\tau)^2 + (k\tau)^2]^{\frac{\alpha}{2}}] \\ \exp[\jmath((-k+1)\tau\Re\{a_i(\theta)s(t)\} - k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[((-k+1)\tau)^2 + (-k\tau)^2]^{\frac{\alpha}{2}}] \\ \vdots \\ \exp[\jmath(k\tau\Re\{a_i(\theta)s(t)\} + k\tau\Im\{a_i(\theta)s(t)\}) - \gamma[(k\tau)^2 + (k\tau)^2]^{\frac{\alpha}{2}}] \end{bmatrix}_{4k^2\times 1} . \tag{4.35}$$

The matrix $\mathbf{D}_{\mathbf{\Omega}}(\theta)$ is the $4k^2 \times 4k^2$ diagonal matrix

$$\begin{aligned} \mathbf{D}_{\mathbf{\Omega}}(\theta) \;=\; & \mathrm{diag}[-k\tau\Re\{d_i(\theta)s(t)\} - k\tau\Im\{d_i(\theta)s(t)\}, \\ & -k\tau\Re\{d_i(\theta)s(t)\} + (-k+1)\tau\Im\{d_i(\theta)s(t)\}, \ldots, \\ & -k\tau\Re\{d_i(\theta)s(t)\} + k\tau\Im\{d_i(\theta)s(t)\}, \ldots, \\ & k\tau\Re\{d_i(\theta)s(t)\} + k\tau\Im\{d_i(\theta)s(t)\}]. \end{aligned} \tag{4.36}$$

Finally, $\mathbf{\Sigma}_{\mathbf{\Omega}}(\theta)$ is the $4k^2 \times 4k^2$ covariance matrix $\mathrm{cov}\{\mathbf{g}(\mathbf{\Omega}, x_i(t))\}$.

We can make the following remarks concerning the described approximation.

Remark 1: In this section, we considered only the case of a single source impinging on the array. The generalization to multiple sources of a known number is a conceptually straightforward problem, but one that involves a considerable computational load since it requires the solution of a nonlinear system of equations of the form (4.32).

Remark 2: In (4.32) we considered the approximation of the derivatives of the likelihood function only with respect to the parameter of interest θ, i.e., we assumed that the noise dispersion γ and the signal S are known, or can be estimated via some method other than ML (see also Section 4.1 for the Cauchy case).

Remark 3: The kernel class $G = \{\exp[\jmath\Re\{\omega x\}];\ \omega \in \mathcal{C}\}$ is complete and so an approximation, $\hat{h}(x)$, of any function can be made arbitrarily close to $h(x)$. In practice, we use a finite number of functions indexed by $\mathbf{\Omega}$. For the case of a linear array of sensors and a single impinging source from direction θ, the likelihood function is periodic with respect to θ and its domain lies in the interval $[-\pi/2, \pi/2]$. If the domain of interest is the interval $[-F, F]$, an appropriate choice for τ is $\tau = 2\pi/F$. The approximation will depend on the values of the characteristic function at the points of the grid $\mathbf{\Omega}$. This approach is basically equivalent to sampling the probability density function at the nodes of an orthogonal grid

with spacing $2\pi/\tau$. The computational complexity associated with the method is increasing with order $\mathcal{O}(k^2)$. We found that a reasonable choice for k is between 5 and 10 for the general case of the α-stable law ($1 \leq \alpha < 2$).

4.4 A Performance Study

To demonstrate the performance of the proposed method for the direction-of-arrival (DOA) problem, we conducted several simulation experiments where we compared the ML estimator based on the Cauchy noise assumption (MLC) with the ML estimator based on the Gaussian noise assumption (MLG), and with the MUSIC estimator.

In all the experiments the array is linear with five sensors spaced a half-wavelength apart. A single signal impinges to the array from a source located at a direction of 5^o. At first, the signal is assumed to be known at the receiver and the ML method is applied to estimate the direction of arrival. Also, for the unknown signal case the pseudo ML method is applied by using a least-squares estimate for the signal. The noise is assumed to follow the bivariate isotropic stable distribution. We chose p to be $p = 0.5$ when we estimated the noise dispersion γ via expression (4.12). In theory, expression (3.9) holds for any value of p as long as $0 < p < \alpha$. In practice, the variance of the estimator is a function of p. It is shown in [76] that values of p in the range $[1/2, \alpha/2]$ give estimators with the smallest variance.

In every experiment we perform 500 Monte-Carlo runs and compute the mean square error (MSE) of the DOA estimates. The optimization for the MLC and MLG methods is performed by a steepest-descent algorithm with variable stepsize selected by means of Armijo's rule [44].

In Figures 4.1, 4.2, and 4.3 we plot the likelihood functions of the MLC and MLG methods for three trials. The likelihood functions are formed by using $M = 20$ snapshots of data from two incoming source signals located at directions $\theta = [-5^o, 5^o]$. The signals are assumed to be known at the receiver. In Figure 4.1, since the additive noise to the sensors is Gaussian ($\alpha = 2$), the MLG likelihood function is the one based on the correct assumption about the noise distribution. On the other hand, in Figure 4.2, the additive noise is Cauchy distributed ($\alpha = 1$) and therefore, the MLC likelihood function is the one based on the correct assumption about the noise distribution. Finally, in Figure 4.3, the additive noise to the sensors is α-stable with $\alpha = 0.5$ and neither the MLG nor the MLC likelihood functions rely on the correct assumption about the noise distribution. As we

can see from the figures, the MLC likelihood function, based on the Cauchy assumption, attains its maximum value very close to the true directions of arrival in all three different cases of additive stable noise. On the other hand, the MLG likelihood function, based on the Gaussian assumption, cannot localize the two sources accurately when the actual data distribution deviates from the Gaussian case. The robustness of the MLC method can clearly be seen in Figure 4.3, where the Cauchy beamformer is still able to localize the two sources even though the characteristic exponent of the additive noise is $\alpha = 0.5$. In this case, the Gaussian beamformer totally fails. These observations, are quantified in the following subsections through extensive Monte-Carlo runs.

Since the alpha-stable family for $\alpha < 2$ determines processes with infinite variance, we define two alternative signal-to-noise ratios (SNR's). Namely, we define the *Generalized SNR (GSNR)* to be the ratio of the signal power over the noise dispersion γ:

$$GSNR = 10 \log(\frac{1}{\gamma M} \sum_{t=1}^{M} |s(t)|^2). \tag{4.37}$$

Also, for finite sample realizations, we define the *Pseudo-SNR (PSNR)* as:

$$PSNR = 10 \log(\frac{\sum_{t=1}^{M} |s(t)|^2}{\sum_{t=1}^{M} |n(t)|^2}). \tag{4.38}$$

In the following, we present a comparative simulation study for the estimation accuracy and the probability of convergence of the aforementioned algorithms.

4.4.1 Estimation Accuracy

In this example we study the estimation accuracy of MLC, MLG and MUSIC as a function of three parameters, namely the number of snapshots M, the noise dispersion γ, and the noise characteristic exponent α.

Number of snapshots M. In the first experiment we study the influence of the number of snapshots M to the performance of the algorithms. The noise follows the complex isotropic Cauchy distribution with dispersion $\gamma = 1$ (cf. (3.8)). For this experiment, the GSNR is kept almost constant at 22.5 dB as shown on Table 4.1. The PSNR is different in every Monte-Carlo run, so we calculate the average PSNR over the 500 Monte-Carlo runs (cf. Table 4.1). As the number of snapshots M increases, the PSNR decreases because more and more impulsive noise samples are incorporated into the data.

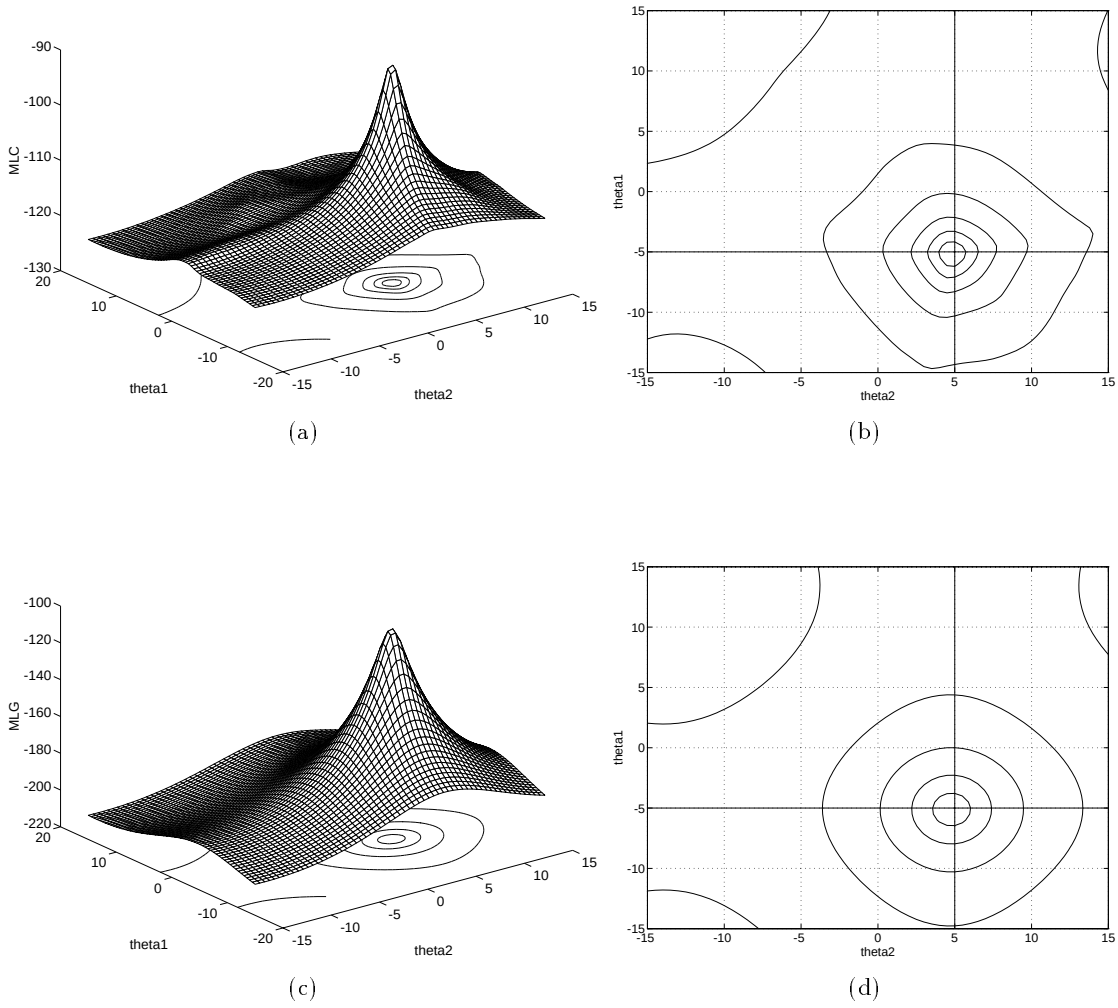


Figure 4.1: MLC (a-b) and MLG (c-d) likelihood functions ($r = 5$, $M = 20$, $\mathbf{\Theta} = [-5^o, 5^o]$). Additive Gaussian noise ($\alpha = 2.0$, $\gamma = 1$).

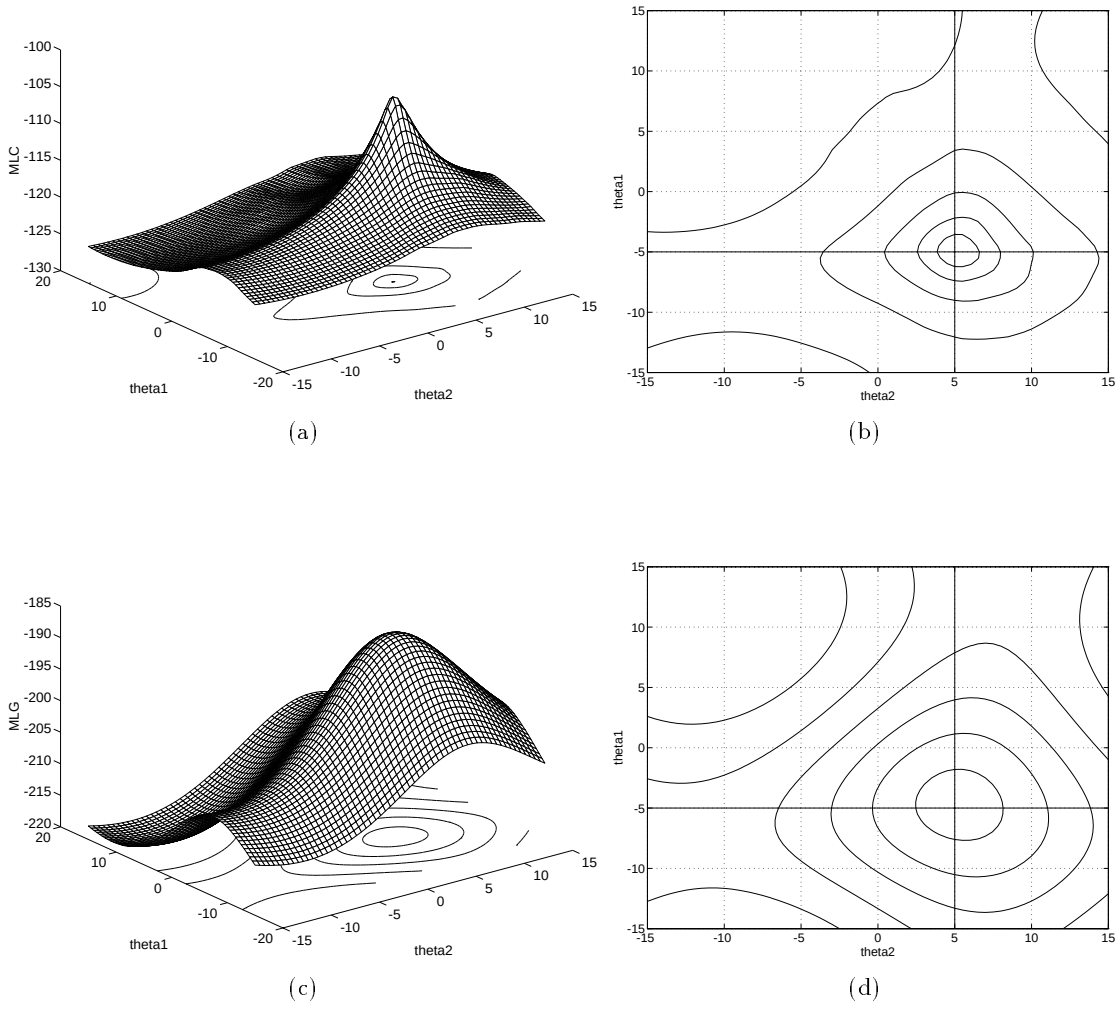


Figure 4.2: MLC (a-b) and MLG (c-d) likelihood functions ($r = 5$, $M = 20$, $\boldsymbol{\Theta} = [-5^o, 5^o]$). Additive Cauchy noise ($\alpha = 1.0$, $\gamma = 1$).

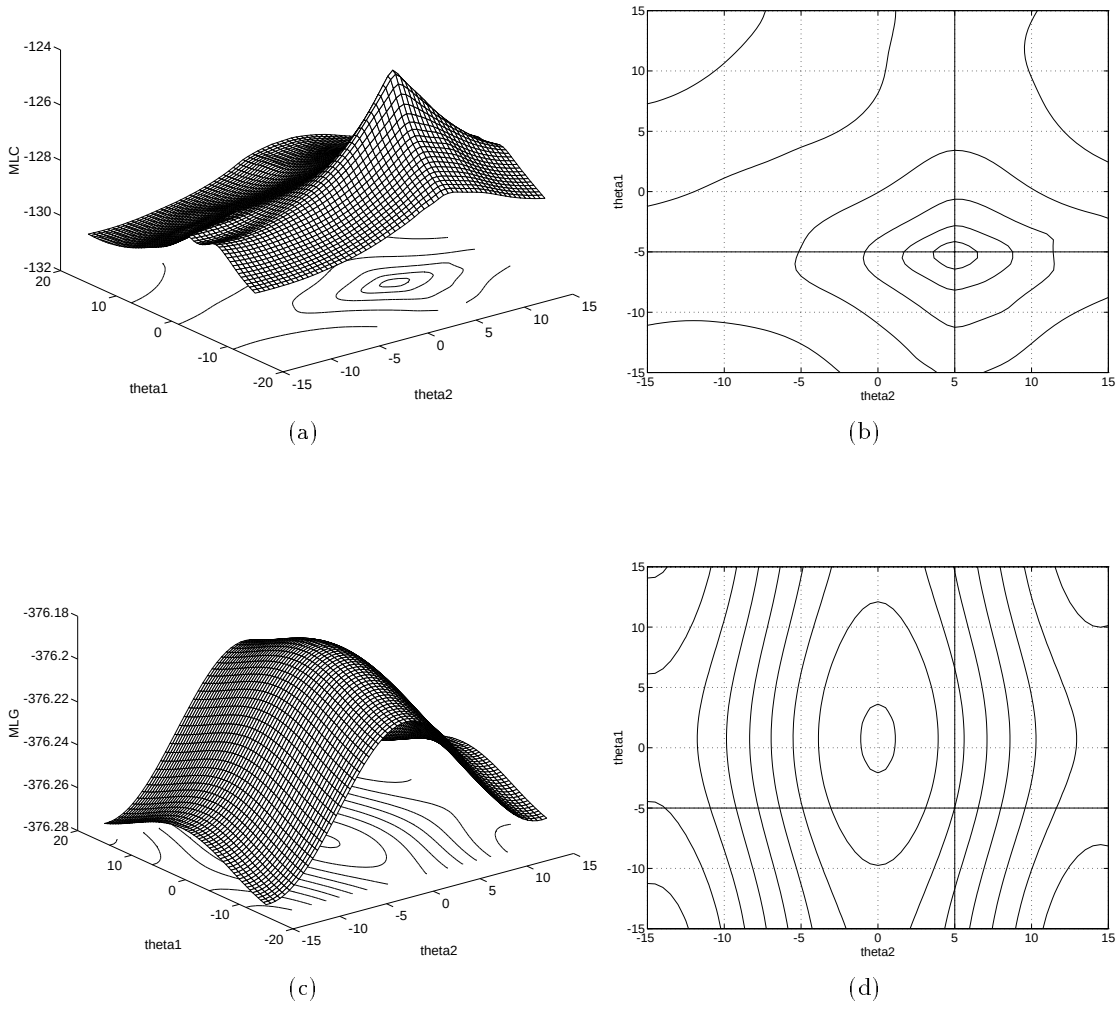


Figure 4.3: MLC (a-b) and MLG (c-d) likelihood functions ($r = 5$, $M = 20$, $\mathbf{\Theta} = [-5^o, 5^o]$). Additive stable noise ($\alpha = 0.5$, $\gamma = 1$).

Table 4.1: GSNR and average PSNR for different values of M.

	Number of snapshots, M				
	$M = 5$	$M = 10$	$M = 20$	$M = 50$	$M = 100$
GSNR [dB]	22.6951	22.7323	22.7416	22.5003	22.6543
PSNR [dB]	7.3712	3.3938	0.5371	-4.3158	-7.149

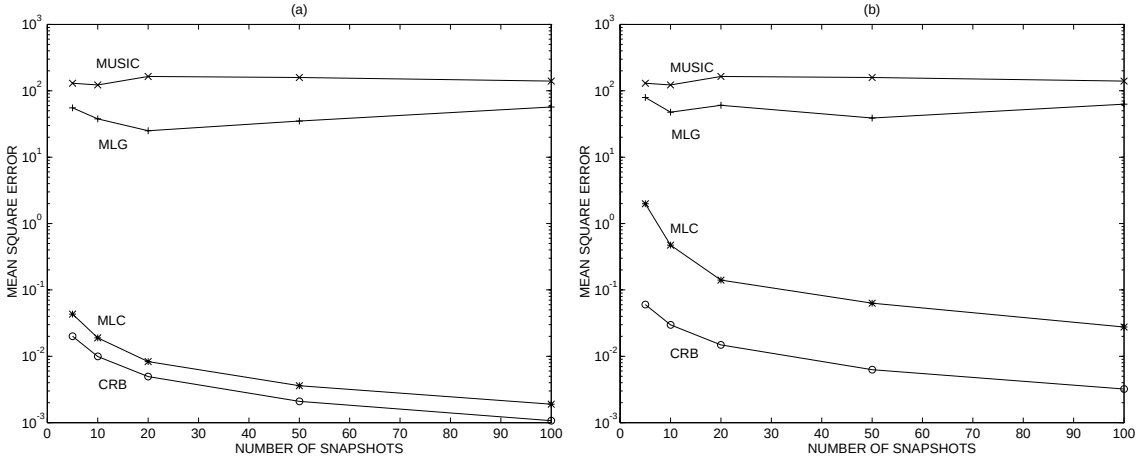


Figure 4.4: MSE of the estimated DOA and CRB as functions of the number of snapshots M. (a) Exact signal knowledge, (b) Least-squares estimate of the signal.

Figure 4.4(a) shows the resulting MSE of the estimated DOA as a function of the number of snapshots when the signal is known. The CRB is also plotted. As expected, the MLC estimator has the best performance since it is the optimal estimator for this type of noise. Also, the complete failure of the MLC and MUSIC processors for this type of impulsive noise is apparent. Figure 4.4(b) shows similar plots for the case of the pseudo ML estimators where we use a LS estimate for the signal. The MLC estimator has again the least MSE. Comparing these curves with the analogous curves obtained assuming exact signal knowledge, we observe a larger MSE for the pseudo ML estimates, as expected.

Noise dispersion γ. In the second experiment we study the influence of the noise dispersion γ, i.e., the influence of the GSNR to the performance of the methods. Here also, the noise follows the bivariate isotropic Cauchy distribution with dispersion γ. The number of snapshots available to the algorithms is $M = 20$. The GSNR and average PSNR for this experiment are shown on Table 4.2.

Table 4.2: GSNR and average PSNR for different values of γ.

	Noise Dispersion, γ						
	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 2$	$\gamma = 4$	$\gamma = 6$	$\gamma = 8$	$\gamma = 10$
GSNR [dB]	25.7519	22.7416	19.7313	16.7210	14.9601	13.7107	12.7416
PSNR [dB]	6.5577	0.5371	-5.4835	-11.5041	-15.0259	-17.5247	-19.4629

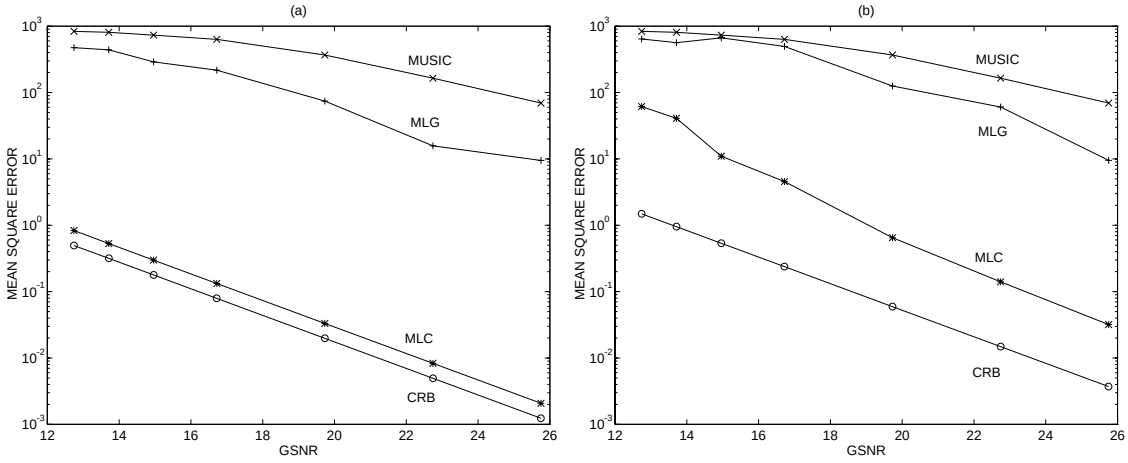


Figure 4.5: MSE of the estimated DOA and CRB as functions of the GSNR. (a) Exact signal knowledge, (b) Least-squares estimate of the signal.

Figure 4.5 shows the resulting MSE of the estimated DOA as a function of the GSNR. Again the MLC estimate has the best performance. As evident in Figure 4.5(b), the performance of the MLC using a LS estimate for the signal degrades more rapidly for large values of the noise dispersion, γ (low GSNR values). The reason is that for large values of the noise dispersion, there is a small probability that assumption B.2 made in Section 4.1 holds. Thus, the MLC estimate suffers from suboptimal signal estimation. It is noted however that the MLC estimate still has the best performance.

Characteristic exponent α. The importance of this experiment rests in its study of the robustness of the algorithms in different noise environments. Of course, by design, the MLG estimator is optimal for additive Gaussian noise ($\alpha = 2$), and the introduced MLC estimator is optimal for additive Cauchy noise ($\alpha = 1$). An important property of any processor is to be able to perform reasonably well in a wide range of noise environments ($1 < \alpha < 2$). Here, we test the performance of the estimators when the characteristic exponent, α, of the noise stable law is changing.

Table 4.3: GSNR and average PSNR for different values of α.

	Noise Characteristic Exponent, α					
	$\alpha = 1.0$	$\alpha = 1.2$	$\alpha = 1.4$	$\alpha = 1.6$	$\alpha = 1.8$	$\alpha = 2.0$
GSNR [dB]	22.7416 ($\gamma = 1$)					
PSNR [dB]	0.5371	6.2171	10.1869	13.1035	15.3269	20.0383

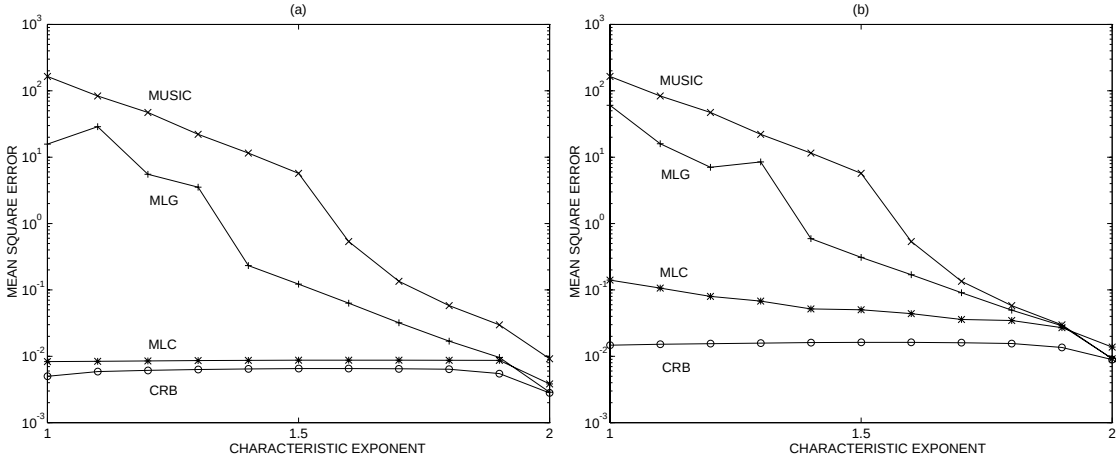


Figure 4.6: MSE of the estimated DOA and CRB as functions of the characteristic exponent α. (a) Exact signal knowledge, (b) Least-squares estimate of the signal.

Figure 4.6 shows the resulting MSE curves as functions of the characteristic exponent α. The number of snapshots available to the algorithms is $M = 20$. The GSNR is 22.7416 dB ($\gamma = 1$) and is shown together with the average PSNR, on Table 4.3. The CRB as a function of α was computed by means of (4.33), where we used the values $k = 10$, and $\tau = 0.157$.

As we can clearly see, the Cauchy beamformer is practically insensitive to the changes of α, and for exact signal knowledge it almost achieves the CRB for the whole range of values α. On the other hand, both the MLG and the MUSIC algorithms exhibit very large mean-square estimation errors for non-Gaussian noise environments. Note that when $\alpha = 2$, i.e., for the Gaussian noise case, the MLG method has the least MSE, as expected.

The experiment demonstrates that for values of α in [1,2], the ML method based on the Cauchy noise assumption exhibits performance very close to the optimum. On the other hand, the optimum ML estimator for the general $S\alpha S$ case with $\alpha \in (1, 2)$

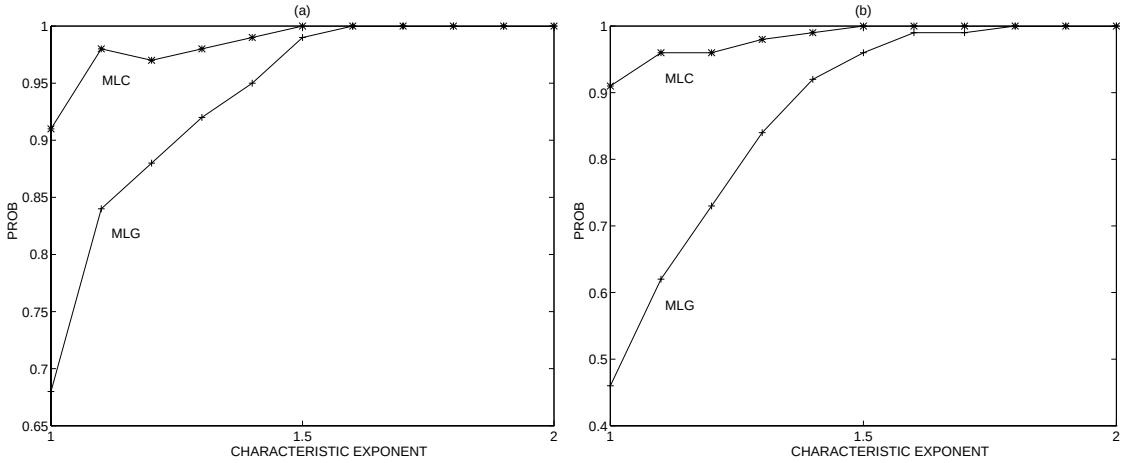


Figure 4.7: Probability of convergence within 1^o of the true DOA as a function of the characteristic exponent α. (a) Exact signal knowledge, (b) Least-squares estimate of the signal.

involves the computationally intensive solution of $\hat{S}(\mathbf{X};\theta) = 0$ with $\hat{S}(\mathbf{X};\theta)$ given by (4.32). These observations, combined with the fact that the ML method based on the Cauchy assumption has computational complexity similar to the ML method based on the Gaussian assumption, justify the importance of the Cauchy beamformer for the DOA estimation problem in practice.

4.4.2 Initialization and Convergence

In this example, we study the convergence of the MLC and MLG algorithms to the true DOA value as a function of the characteristic exponent α of the noise. The initialization of the optimization procedure is done by means of the MUSIC bearing estimates. Again, the number of snapshots available to the algorithms is $M = 20$ and the GSNR and average PSNR are as shown on Table 4.3. In Figure 4.7 we plot the probability that the processor will converge within 1^o from the true DOA as a function of the characteristic exponent of the noise. It is shown that for the known signal case, MLC exhibits higher probability of convergence than MLG in the range $1 \leq \alpha \leq 1.6$. The range of the superior performance of MLC is $1 \leq \alpha \leq 1.8$ when a LS estimate for the signal waveforms is used.

4.5 Concluding Remarks

In this chapter, we have presented a novel approach to the DOA estimation problem in the presence of impulsive interference. The method is based on the maximum likelihood estimation technique where the noise is modeled as a complex isotropic $S\alpha S$ process. The Cauchy beamformer has been shown to give better bearing estimates than the Gaussian beamformer in a wide range of impulsive noise environments, and for very low SNR values. The technique inherits the computational complexity of the ML family of methods but avoids the eigen-decomposition operations of the eigenvector-based methods. The simulations show that for very large values of the noise dispersion, the Cauchy beamformer suffers when the LS estimate for the signal is used. Still, its performance is better than the performance of existing methods.

Chapter 5

Subspace Techniques with Alpha-Stable Distributions

The optimal ML techniques employed in Chapter 4 are often regarded as exceedingly complex due to the high computational load of the multivariate nonlinear optimization problem involved with these techniques. Hence, sub-optimal methods need to be developed for the solution of the DOA problem in the presence of impulsive noise, when reduced computational cost is a crucial design requirement.

In this chapter, we present subspace techniques for the source localization problem, techniques which are based on the geometrical properties of the data model. Considerable research has been done in this area under the framework of Gaussian-distributed signals and/or noise. The better known of the so-called eigendecomposition-based methods are the MUSIC [66], the Minimum Norm [41, 57], and the ESPRIT method [61]. These methods estimate the bearings of the source signals by performing an eigendecomposition on the spatial covariance matrix of the array sensor outputs. Studies concerning the statistical efficiency of the most popular eigendecomposition-based method, namely the MUSIC algorithm, have been done in [38, 72]. The relationship between the MUSIC and ML methods has also been studied in [72]. Since $S\alpha S$ processes do not possess finite pth order moments for $p \geq \alpha$, traditional subspace techniques employing second- and higher-order moments cannot be applied in impulsive noise environments modeled under the stable law. Instead, properties of fractional lower-order moments (FLOM's) and covariations should be used.

In the following sections, we extend the subspace-based techniques for bearing estimation to processes with finite moments of order p ($p < 2$) and to complex isotropic $S\alpha S$ processes. In Section 5.1, we discuss the development of subspace techniques in the presence of α-stable distributed signals and noise. Our analysis is based on the formulation of the covariation matrix of the array sensor outputs. In Section 5.2, we present consistent

estimators for the covariation matrix and we study their asymptotic performance. Simulation experiments are presented in Section 5.3, and conclusions are drawn in Section 5.4.

5.1 The Array Covariation Matrix

In this section, we assume that the q signal waveforms are non-coherent, statistically independent, complex isotropic $S\alpha S$ $(1 < \alpha \leq 2)$ random processes with zero location parameter and covariation matrix $\mathbf{\Gamma}_S = \text{diag}(\gamma_{s_1}, \ldots, \gamma_{s_q})$. In addition, the noise vector $\mathbf{n}(t)$ is a complex isotropic $S\alpha S$ random process with the same characteristic exponent α as the signals. The noise is assumed to be independent of the signals and has covariation matrix $\mathbf{\Gamma}_N = \gamma_n \mathbf{I}$.

Equation (2.2) can be written as

$$\mathbf{x}(t) = \mathbf{w}(t) + \mathbf{n}(t), \tag{5.1}$$

where $\mathbf{w}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)$. By the stability property of stable processes, it follows that $\mathbf{w}(t)$ is also a complex isotropic $S\alpha S$ random vector with components

$$w_i(t) = \mathbf{A}_i(\boldsymbol{\theta})\mathbf{s}(t) = a_i(\theta_1)s_1(t) + \cdots + a_i(\theta_q)s_q(t) \quad i = 1, \ldots, r. \tag{5.2}$$

In addition, it holds that $\mathbf{w}(t)$ is independent of $\mathbf{n}(t)$.

Now, we define the *covariation matrix*, $\mathbf{\Gamma}_X$, of the observation vector process $\mathbf{x}(t)$ as the matrix whose elements are the covariations $[x_i(t), x_j(t)]_\alpha$ of the components of $\mathbf{x}(t)$. We have that

$$\begin{aligned} [x_i(t), x_j(t)]_\alpha &= [w_i(t) + n_i(t), w_j(t) + n_j(t)]_\alpha \\ &= [w_i(t), w_j(t)]_\alpha + [w_i(t), n_j(t)]_\alpha + [n_i(t), w_j(t)]_\alpha + [n_i(t), n_j(t)]_\alpha. \end{aligned} \tag{5.3}$$

By the independence assumption of $\mathbf{w}(t)$ and $\mathbf{n}(t)$ and by property P3 of Section 3.4, we have that

$$[w_i(t), n_j(t)]_\alpha = 0, \tag{5.4}$$

and

$$[n_i(t), w_j(t)]_\alpha = 0. \tag{5.5}$$

Also, by using (5.2) and properties P1 and P2, it follows that

$$\begin{aligned}
[w_i(t), w_j(t)]_\alpha &= [\sum_{k=1}^{q} a_i(\theta_k) s_k(t), w_j(t)]_\alpha \\
&= \sum_{k=1}^{q} a_i(\theta_k) [s_k(t), w_j(t)]_\alpha \\
&= \sum_{k=1}^{q} a_i(\theta_k) [s_k(t), \sum_{l=1}^{q} a_j(\theta_l) s_l(t)]_\alpha \\
&= \sum_{k=1}^{q} a_i(\theta_k) a_j^{<\alpha-1>}(\theta_k) \gamma_{s_k}, \qquad (5.6)
\end{aligned}$$

where $\gamma_{s_k} = [s_k, s_k]_\alpha$. Finally, due to the noise assumption made earlier, it holds that

$$[n_i(t), n_j(t)]_\alpha = \gamma_n \delta_{i,j}, \qquad (5.7)$$

where $\delta_{i,j}$ is the Kronecker delta function. Combining (5.3)-(5.7) we obtain the following expression for the covariations of the sensor measurements:

$$[x_i(t), x_j(t)]_\alpha = \sum_{k=1}^{q} a_i(\theta_k) a_j^{<\alpha-1>}(\theta_k) \gamma_{s_k} + \gamma_n \delta_{i,j} \quad i, j = 1, \ldots, r. \qquad (5.8)$$

In addition, the dispersion and covariation coefficients of the array sensor measurements are given respectively by

$$\gamma_{x_j(t)} = \sum_{k=1}^{q} |a_j(\theta_k)|^\alpha \gamma_{s_k} + \gamma_n \quad j = 1, \ldots, r, \qquad (5.9)$$

and

$$\lambda_{x_i(t), x_j(t)} = \frac{\sum_{k=1}^{q} a_i(\theta_k) a_j^{<\alpha-1>}(\theta_k) \gamma_{s_k} + \gamma_n \delta_{i,j}}{\sum_{k=1}^{q} |a_j(\theta_k)|^\alpha \gamma_{s_k} + \gamma_n} \quad i, j = 1, \ldots, r. \qquad (5.10)$$

In matrix form, (5.8) gives the following expression for the covariation matrix of the observation vector:

$$\mathbf{\Gamma}_X \stackrel{\triangle}{=} [\mathbf{x}(t), \mathbf{x}(t)]_\alpha = \mathbf{A}(\boldsymbol{\theta}) \mathbf{\Gamma}_S \mathbf{A}^{<\alpha-1>}(\boldsymbol{\theta}) + \gamma_n \mathbf{I}, \qquad (5.11)$$

where the (i,j)th element of matrix $\mathbf{A}^{<\alpha-1>}(\boldsymbol{\theta})$ results from the (j,i)th element of $\mathbf{A}(\boldsymbol{\theta})$ according to the operation

$$[\mathbf{A}^{<\alpha-1>}(\boldsymbol{\theta})]_{i,j} = [\mathbf{A}(\boldsymbol{\theta})]_{j,i}^{<\alpha-1>} = |\,[\mathbf{A}(\boldsymbol{\theta})]_{j,i}\,|^{\alpha-2}\,[\mathbf{A}(\boldsymbol{\theta})]_{j,i}^{*} \tag{5.12}$$

Clearly, when $\alpha = 2$, i.e., for Gaussian distributed signals and noise, the expression for the covariation matrix is reduced to the form shown in (2.31).

When the amplitude response of the sensors equals unity, i.e., for steering vectors of the form $\mathbf{a}(\theta_k) = [1, e^{-\jmath\omega\tau_2(\theta_k)}, \ldots, e^{-\jmath\omega\tau_r(\theta_k)}]^T$, it follows that

$$[\mathbf{A}^{<\alpha-1>}(\boldsymbol{\theta})]_{i,j} = |e^{-\jmath\omega\tau_j(\theta_i)}|^{\alpha-2} e^{\jmath\omega\tau_j(\theta_i)} = [\mathbf{A}(\boldsymbol{\theta})]_{j,i}^{*}\,, \tag{5.13}$$

and thus, the covariation matrix can be written as

$$\boldsymbol{\Gamma}_X = \mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}_S\mathbf{A}^H(\boldsymbol{\theta}) + \gamma_n\mathbf{I}. \tag{5.14}$$

Finally, by using (5.9) and (5.10), the dispersion and covariation coefficients of the array sensor measurements can be written as

$$\gamma_{x_j(t)} = \sum_{k=1}^{q} \gamma_{s_k} + \gamma_n \quad j = 1, \ldots, r, \tag{5.15}$$

and

$$\lambda_{x_i(t),x_j(t)} = \frac{\sum_{k=1}^{q} a_i(\theta_k)a_j^*(\theta_k)\gamma_{s_k} + \gamma_n\delta_{i,j}}{\sum_{k=1}^{q}\gamma_{s_k} + \gamma_n} \quad i,j = 1, \ldots, r. \tag{5.16}$$

Hence, in this case, we can apply the subspace techniques described in Section 2.2.2, to the covariation or the covariation coefficient matrices of the observation vector to extract the bearing information. Since several estimators have been proposed for the covariation coefficient of two $S\alpha S$ random variables, in the following we will use the covariation coefficient matrix of the array sensor measurements to estimate the bearings of the sources.

We will refer to the new algorithm resulting from the eigendecomposition of the array covariation coefficient matrix as the *Robust Covariation-Based MUSIC* or *ROC-MUSIC*. In practice, we have to estimate the covariation matrix from a finite number of array sensor measurements. The following section describes one such estimator, based on fractional lower order moments of the stable process, and studies its asymptotic statistical properties.

5.2 Covariation Estimators

An estimator for the covariation coefficient $\lambda_{X,Y}$ defined in (3.24), is called the *fractional lower order (FLOM) estimator* and was proposed in [69]. Expanding the p-support range of the FLOM estimator we define the *modified FLOM (MFLOM)* estimator

$$\hat{\lambda}_{X,Y}(p) = \frac{\sum_{i=1}^{n} X_i Y_i^{<p-1>}}{\sum_{i=1}^{n} |Y_i|^p} \tag{5.17}$$

for independent observations $(X_1, Y_1), \ldots, (X_n, Y_n)$. In the above expression, $1/2 < p < \alpha$ if X and Y are real $S\alpha S$ random variables, and $0 < p < \alpha$ if X and Y are complex isotropic $S\alpha S$ random variables. The estimator given in (5.17) is a moments-based estimator of the *modified covariation coefficient function* defined as

$$\lambda_{X,Y}(p) = \frac{E\{XY^{<p-1>}\}}{E\{|Y|^p\}}, \qquad \begin{array}{ll} 1/2 < p < \alpha & \text{if } X \text{ and } Y \text{ are real} \\ 0 < p < \alpha & \text{if } X \text{ and } Y \text{ are complex} \end{array} \tag{5.18}$$

The modified covariation coefficient function is well defined (finite) for the aforementioned values of the parameter p as follows from the following proposition:

Proposition 5.1 *Let X and Y be jointly $S\alpha S$ random variables. It holds that $E\{|X||Y|^{p-1}\} < \infty$ and $E\{|X|^2|Y|^{2p-2}\} < \infty$ if $1/2 < p < \alpha/2$ ($0 < p < \alpha/2$) for X and Y real (isotropic complex.)*

Proof See Appendix B.

Clearly, when $1 \leq p < \alpha$ the function $\lambda_{X,Y}(p)$ equals the covariation coefficient $\lambda_{X,Y}$ as defined in (3.24)-(3.25). The theoretical performance of the MFLOM estimator is given by the following theorem:

Theorem 5.1 *The estimator $\hat{\lambda}_{X,Y}(p)$ given by (5.17) is consistent and asymptotically normal with mean equal to the true modified covariation coefficient function $\lambda_{X,Y}(p)$, and variance*

$$\begin{aligned} \frac{1}{n}\sigma_{\lambda}^2(p) \;=\; & \frac{1}{n}\left[\left[E\{|X|^2|Y|^{2p-2}\} - |E\{XY^{<p-1>}\}|^2\right]\frac{1}{(E\{|Y|^p\})^2} - \right. \\ & 2\Re\left\{\left[E\{XY^{<p-1>}|Y|^p\} - E\{XY^{<p-1>}\}E\{|Y|^p\}\right]\frac{E\{XY^{<p-1>}\}}{(E\{|Y|^p\})^3}\right\} + \\ & \left.\left[E\{|X|^{2p}\} - (E\{|Y|^p\})^2\right]\left(\frac{|E\{XY^{<p-1>}\}|}{(E\{|Y|^p\})^2}\right)^2\right], \end{aligned} \tag{5.19}$$

where $\Re[z]$ denotes the real part of z and p varies in the range $0 < p < \alpha/2$ for the complex isotropic case. In notational form,

$$\sqrt{n}(\hat{\lambda}_{X,Y}(p) - \lambda_{X,Y}(p)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_{\lambda}(p)), \tag{5.20}$$

where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution, and $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2.

Proof See Appendix C.

For the case of real jointly $S\alpha S$ random variables, the asymptotic variance of the estimator is finite for values of p in the range $1/2 < p < \alpha/2$. The simulation experiments in the following section give some significant insight on the performances of the MFLOM estimator and the proposed ROC-MUSIC algorithm.

5.3 Simulations

We performed three simulation experiments to assess the performance of the MFLOM estimator and to compare the MUSIC and ROC-MUSIC algorithms. The first experiment illustrates the behavior of the MFLOM estimator of the covariation coefficient as a function of p and compares MFLOM with the least-squares and screened ratio estimators. The second and third experiments study the performance of the MUSIC and ROC-MUSIC algorithms in the presence of simulated $S\alpha S$ noise and real radar clutter, respectively. The improved performance of the ROC-MUSIC method in terms of resolution capability, bias and mean square error is apparent in the simulation results.

5.3.1 Performance of the MFLOM Estimator

The purpose of this experiment is to study the influence of the parameter p to the performance of the MFLOM estimator of the covariation coefficient. Two real $S\alpha S$ $(1 < \alpha \leq 2)$ random variables, X and Y, are defined as

$$X = a_{11}U_1 + a_{12}U_2,$$

$$Y = a_{21}U_1 + a_{22}U_2,$$

where U_1, and U_2 are independent, $S\alpha S$ random variables. The model coefficients $\{a_{ij};\ \ i,j=1,2\}$ are given by

$$[a_{ij}] = \begin{bmatrix} -0.75 & 0.25 \\ 0.18 & 0.78 \end{bmatrix}. \tag{5.21}$$

It follows from (5.10) that the true covariation coefficient λ of X with Y is

$$\lambda = \frac{\sum_{j=1}^{2} a_{1j} a_{2j}^{<\alpha-1>}}{\sum_{j=1}^{2} |a_{2j}|^{\alpha}}. \tag{5.22}$$

We generated $n = 5,000$ independent samples of U_1 and U_2, and we calculated the MFLOM estimator by means of the expression

$$\hat{\lambda}_{MFLOM}(p) = \frac{\sum_{i=1}^{n} X_i Y_i^{<p-1>}}{\sum_{i=1}^{n} |Y_i|^{p}} \tag{5.23}$$

for different values of p in the range [0,2]. We run $K = 1,000$ Monte Carlo experiments and compared the performance of the MFLOM estimator to that of the least-squares and the screened ratio estimators. The two later estimators are defined as follows:

Least-squares estimator:

$$\hat{\lambda}_{LS} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} |Y_i|^2} \tag{5.24}$$

Screened ratio estimator [37]:

$$\hat{\lambda}_{SR} = \frac{\sum_{i=1}^{n} X_i Y_i^{-1} \chi_{Y_i}}{\sum_{i=1}^{n} \chi_{Y_i}}, \tag{5.25}$$

where

$$\chi_Y = \begin{cases} 1 & \text{if } c_1 < |Y| < c_2 \\ 0 & \text{otherwise.} \end{cases}$$

In [37], it is shown that the least-squares estimator $\hat{\lambda}_{LS}$ is not consistent for $1 < \alpha < 2$ while the screened ratio estimator $\hat{\lambda}_{SR}$ is strongly consistent.

Figure 5.1 shows the standard deviation of the MFLOM estimator of the modified covariation coefficient as a function of the parameter p and for different values of the characteristic exponent α. As we can see, for the case of non-Gaussian stable signals $(1 < \alpha < 2)$, the values of p in the range $(1/2, \alpha/2)$ result into the smallest standard

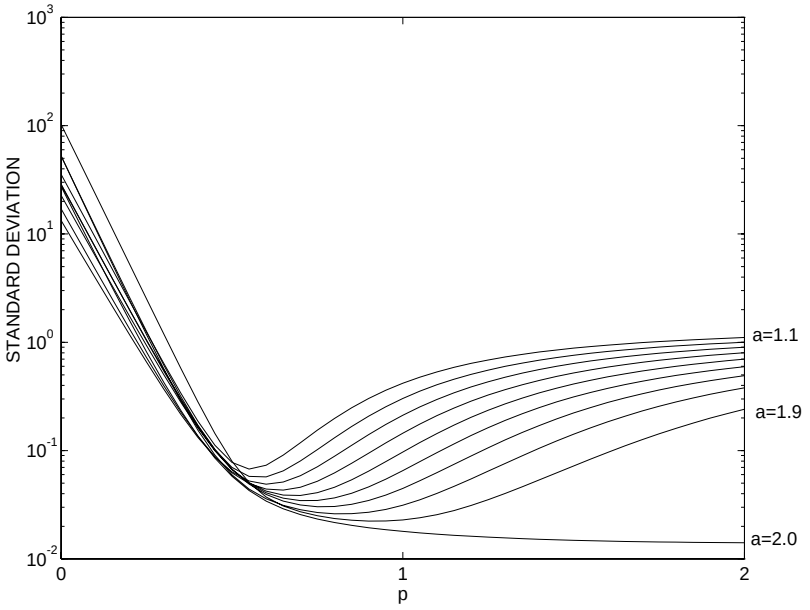


Figure 5.1: Standard deviation of the MFLOM estimates of the modified covariation coefficient as a function of the parameter p.

deviations. For Gaussian signals, the optimal value of p is 2 and the resulting MFLOM estimator is simply the least-squares estimator, as expected.

In Figure 5.2, we plot the MFLOM estimates for $\alpha = 1.5$ and several values of p. We also plot in Figure 5.3 the corresponding running sample variance estimates S_k^2 defined as follows:

$$S_k^2 = \frac{1}{k-1}\sum_{i=1}^{k}(\hat{\lambda}_{MFLOM}^{(i)}(p) - \bar{\hat{\lambda}}_{MFLOM}^{(k)}(p))^2 \quad k = 1, \ldots, K, \tag{5.26}$$

where

$$\bar{\hat{\lambda}}_{MFLOM}^{(k)}(p) = \frac{1}{k}\sum_{i=1}^{k}\hat{\lambda}_{MFLOM}^{(i)}(p). \tag{5.27}$$

If the population of the MFLOM estimates $\{\hat{\lambda}_{MFLOM}^{(i)}(p)\}_{i=1}^{K}$ has finite variance, S_k^2 will converge to a finite value (converging variance test). As we can see, only for the value of p in the range $(1/2, \alpha/2)$, i.e., for $p = 0.7$, is the MFLOM estimator normally distributed with finite variance, which supports the theoretical analysis of Section 5.2. Figure 5.4 shows the screened ratio estimates and the corresponding running sample variance.

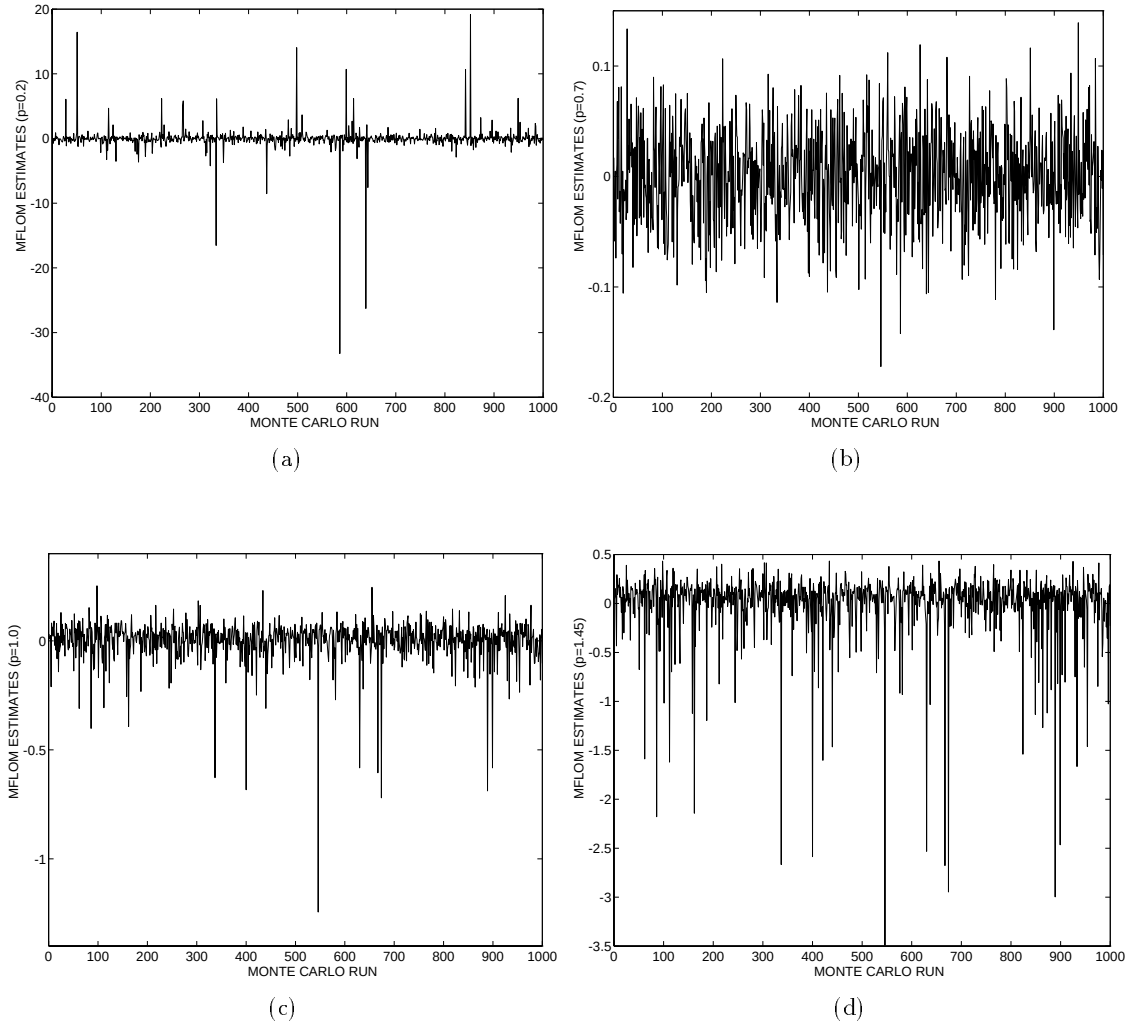


Figure 5.2: MFLOM estimates for $\alpha = 1.5$ and $p = 0.2$ (a), $p = 0.7$ (b), $p = 1.0$ (c), $p = 1.45$ (d).

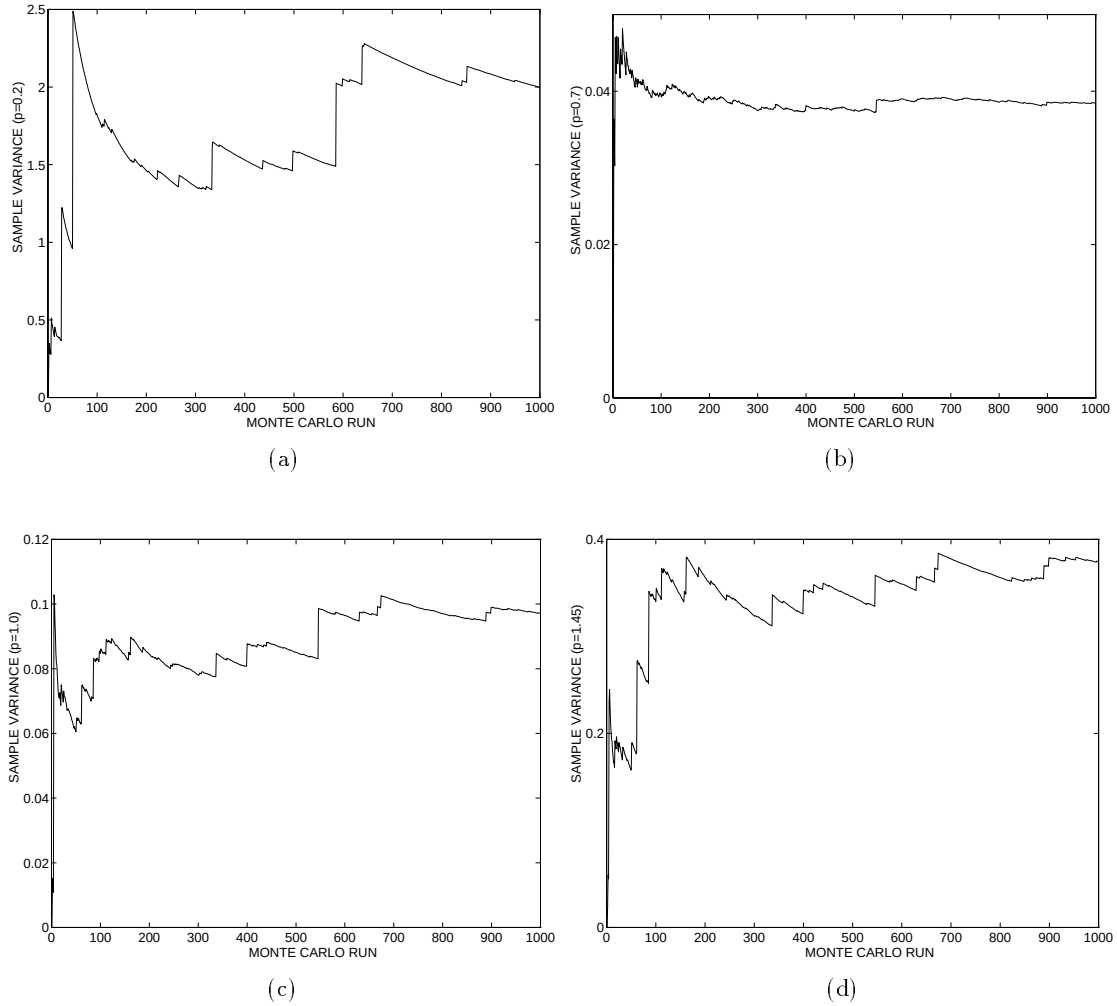
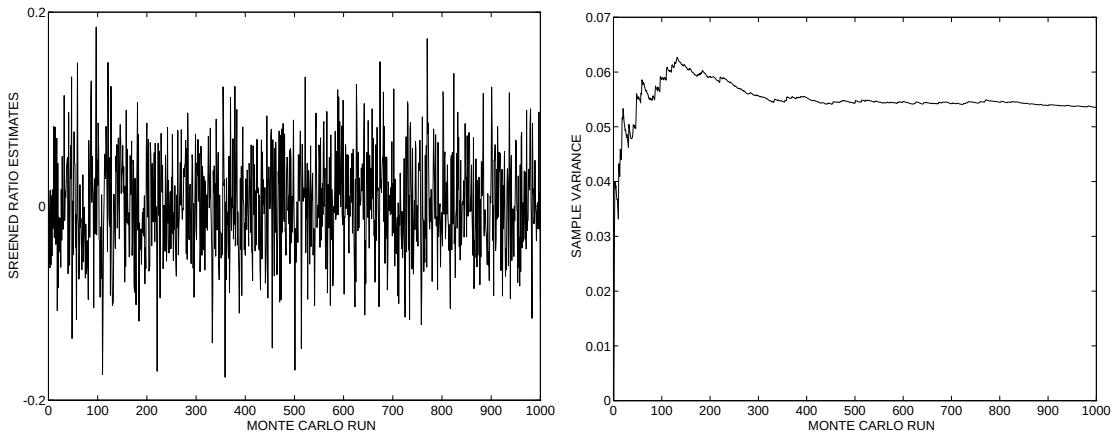


Figure 5.3: MFLOM running variances for $\alpha = 1.5$ and $p = 0.2$ (a), $p = 0.7$ (b), $p = 1.0$ (c), $p = 1.45$ (d).



(a) (b)

Figure 5.4: Screened ratio estimates (a) and running variance (b) for $\alpha = 1.5$.

Table 5.1: Performance of the covariation coefficient estimators.

α	LS	SR	MFLOM	True λ
1.1	-0.4240	-0.4263	-0.4245 $(p = 0.55)$	-0.4252
	(1.1065)	(0.0651)	(0.0636)	
1.2	-0.3371	-0.3396	-0.3380 $(p = 0.58)$	-0.3384
	(1.0010)	(0.0603)	(0.0572)	
1.3	-0.2641	-0.2595	-0.2596 $(p = 0.60)$	-0.2601
	(0.9008)	(0.0574)	(0.0490)	
1.4	-0.1984	-0.1873	-0.1874 $(p = 0.65)$	-0.1900
	(0.8007)	(0.0565)	(0.0432)	
1.5	-0.1386	-0.1250	-0.1258 $(p = 0.68)$	-0.1273
	(0.6986)	(0.0536)	(0.0385)	
1.6	-0.0840	-0.0706	-0.0709 $(p = 0.70)$	-0.0715
	(0.5956)	(0.0522)	(0.0345)	
1.7	-0.0342	-0.0215	-0.0215 $(p = 0.75)$	-0.0222
	(0.4913)	(0.0501)	(0.0303)	
1.8	0.0113	0.0226	0.0225 $(p = 0.80)$	0.0215
	(0.3794)	(0.0472)	(0.0261)	
1.9	0.0537	0.0601	0.0600 $(p = 0.90)$	0.0599
	(0.2407)	(0.0456)	(0.0223)	
2.0	0.0932	0.0917	0.0932 $(p = 2.0)$	0.0936
	(0.0141)	(0.0365)	(0.0141)	

Table 5.1 shows comparative results on the performance of the three estimators. We include the mean of the estimators, the standard deviation in parentheses, and the value of p for which the smallest standard deviation is achieved by the MFLOM estimator. For the screened ratio estimator, we chose $c_1 = 0.1$ and $c_2 = \infty$. Clearly, for $\alpha \neq 2$, the least-squares estimator has large standard deviation compared to the other two estimators. For the Gaussian case ($\alpha = 2$), the least-squares estimator is optimum. The screened ratio estimator exhibits small bias and standard deviation. On the other hand, the MFLOM estimator gives the smallest bias and standard deviation of all three estimators for a choice of the parameter p in the range $(1/2, \alpha/2)$. For such values of p, the MFLOM estimator of the modified covariation coefficient function converges to the true covariation coefficient calculated by (5.22). Hence, we can conjecture that (3.24) holds for $1/2 < p < \alpha$ in the case of real $S\alpha S$ random variables.

5.3.2 Performance Comparison of ROC-MUSIC versus MUSIC for Simulated Data

In this experiment, we compare the performance of the proposed ROC-MUSIC algorithm to MUSIC. The sample covariation coefficient matrix (SCCM), as estimated by (5.17), is not symmetric and hence it has complex eigenvalues in general. The more snapshots are available at the array sensors, the more nearly symmetric SCCM becomes. We come around this problem by performing the eigenvalue decomposition to the sum of the sample covariation coefficient matrix and its Hermitian transpose.

The array is linear with five sensors spaced a half-wavelength apart. Two QAM communication signals independent of each other and of the same power impinge to the array. The number of signals is assumed to be known. The noise is assumed to follow the complex isotropic $S\alpha S$ distribution with dispersion γ. In every experiment we perform 200 Monte-Carlo runs and compute the resolution event probability and the mean-square error (MSE) of the direction-of-arrival estimates averaged for the two sources. The MSE of the DOA estimates was calculated by taking into consideration only the Monte-Carlo runs for which the two algorithms resolved the two sources.

The resolution analysis of the two algorithms was studied by using a popular resolution criterion defined by the following threshold equation [38, 96]:

$$\Lambda(\theta_1, \theta_2) \stackrel{\triangle}{=} P(\theta_m) - \frac{1}{2}\{P(\theta_1) + P(\theta_2)\} > 0, \tag{5.28}$$

Table 5.2: GSNR and average PSNR for different values of α.

	Noise Characteristic Exponent α					
	$\alpha = 1.01$	$\alpha = 1.2$	$\alpha = 1.4$	$\alpha = 1.6$	$\alpha = 1.8$	$\alpha = 2.0$
GSNR [dB]	22.3 ($\gamma = 1$)					
PSNR [dB]	-17.9690	-7.8245	-2.8001	-0.8380	-0.2164	-0.0441

where θ_1 and θ_2 are the angles of arrival of the two signals, $\theta_m = (\theta_1 + \theta_2)/2$ is the mid-range between them, and the *null spectrum* $P(\theta) \stackrel{\triangle}{=} 1/S(\theta)$ is defined as the reciprocal of the spatial spectrum $S(\theta)$ given in (2.41). The two signals are said to be resolvable if inequality (5.28) holds. The inequality implies that the null spectrum magnitude at the mid-angle should lie above the line segment linking the two signal valleys, in order for the two sources to be resolvable. In our experiments, we estimated the null spectrum from a finite number of array sensor measurements and we determined the probability of resolution by averaging the resolution events over the 200 independent Monte Carlo runs.

In the following simulation experiments, we study the resolution capability and estimation accuracy of ROC-MUSIC versus MUSIC as functions of four parameters, namely the noise characteristic exponent α, the number of snapshots M, the noise dispersion γ, and the angular separation of the two sources.

Characteristic exponent α. In this experiment, we evaluate the performance of the two algorithms in a wide range of noise environments, from the more impulsive (α in the neighborhood of 1) to the Gaussian ones ($\alpha = 2$). The angles of arrival for the two signals are $\theta_1 = -5^o$ and $\theta_2 = 5^o$. The number of snapshots available to the algorithms is $M = 1000$. The GSNR is 22.3 dB ($\gamma = 1$) and is shown together with the average PSNR on Table 5.2. The characteristic exponent α of the additive noise is unknown to the ROC-MUSIC algorithm. We use two values of the parameter p in the estimation of the covariation matrix (cf. (5.17)): $p = 0.8$ and $p = 0.4$. Clearly, MUSIC can be thought as a special case of ROC-MUSIC with $p = 2$.

In Figure 5.5, spatial spectral estimates obtained in ten independent trials are shown for the ROC-MUSIC and MUSIC algorithms. Three types of alpha stable noise corresponding to three values of the characteristic exponent $\alpha = 2.0,\ 1.5,\ 1.2$ were used. We can see that the MUSIC method exhibits high-resolution performance only for the case of Gaussian additive noise. On the other hand, the ROC-MUSIC method has considerably

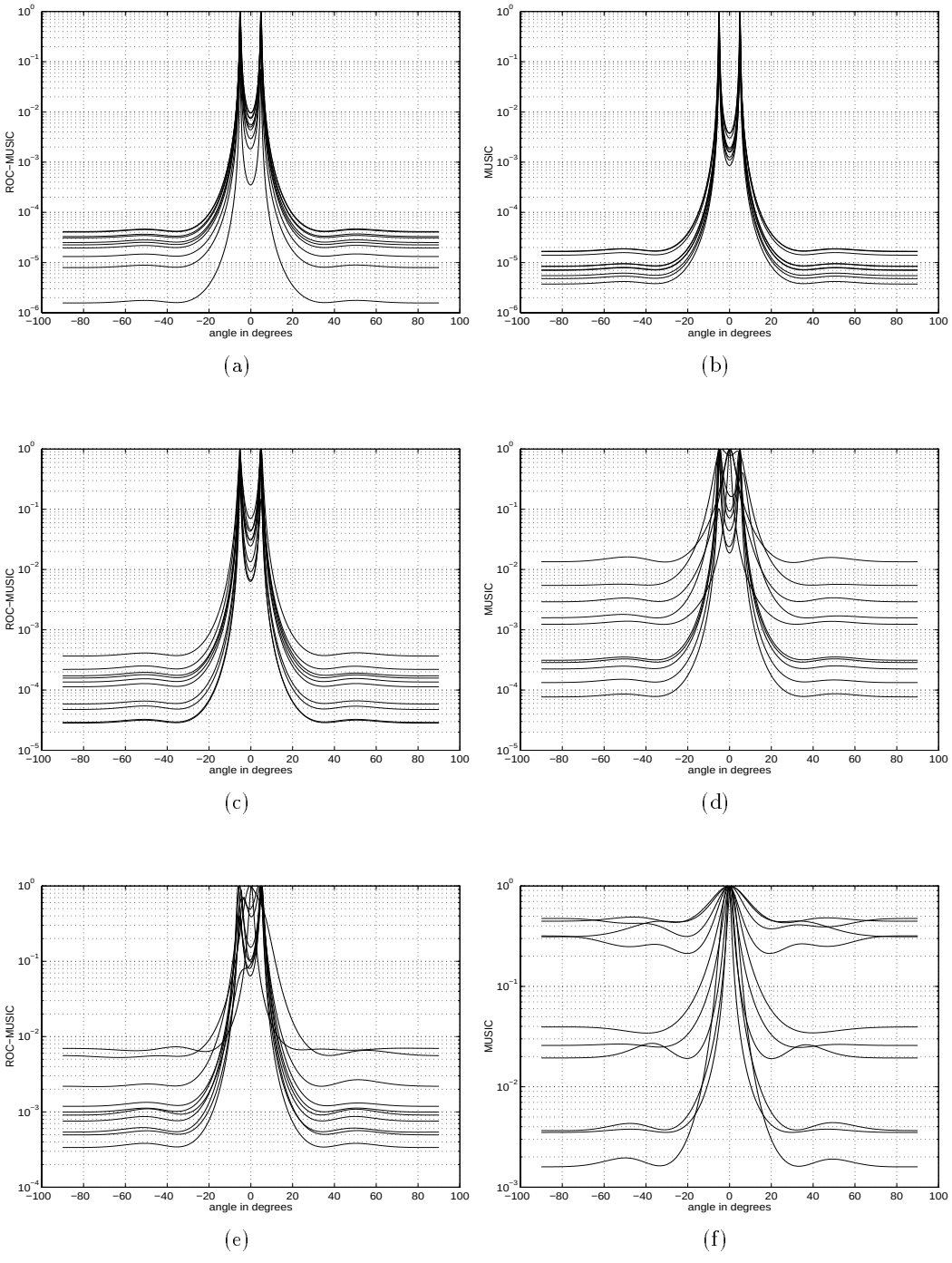


Figure 5.5: Spatial spectral estimates for the ROC-MUSIC (a,c,e) and MUSIC (b,d,f) algorithms ($r = 5$, $M = 1,000$, $\mathbf{\Theta} = [-5^o, 5^o]$). Additive stable noise (a,b): $\alpha = 2.0$, (c,d): $\alpha = 1.5$, (e,f): $\alpha = 1.2$, $\gamma = 1$.

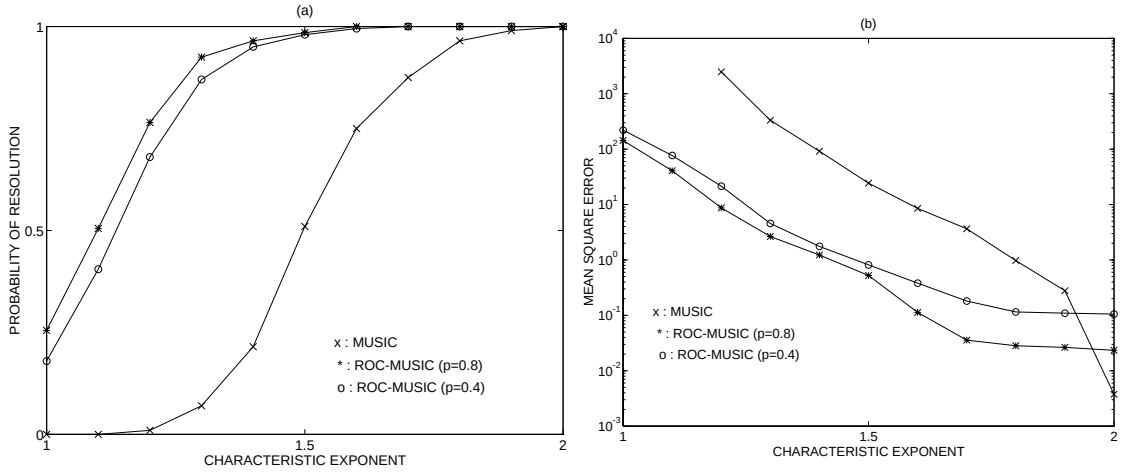


Figure 5.6: Probability of resolution (a) and mean square error (b) as functions of the characteristic exponent α.

less variability from trial to trial than the MUSIC technique, and exhibits better resolution capabilities for non-Gaussian additive noise environments.

Figure 5.6 depicts the improved performance of ROC-MUSIC over that of MUSIC both in terms of resolution probability and MSE, for values of α in the range $(1, 2)$. Note that for $\alpha < 1.2$, MUSIC does not resolve the two sources in any of the 200 Monte Carlo runs. The results suggest that in impulsive noise environments modeled under the stable law, it is more beneficial to use the covariation matrix (lower-order moments) instead of the covariance matrix (second-order moments). Of course, for Gaussian additive noise ($\alpha = 2$) the use of second-order moments ($p = 2$) gives better results. Figure 5.6(b) shows that the choice of $p = 0.8$ in ROC-MUSIC gives better results than the ones obtained when using $p = 0.4$, especially when $\alpha \geq 1.6$ in which case $p < \alpha/2$.

Number of snapshots M. The influence of the number of snapshots M to the performance of the algorithms is shown on Figure 5.7. For this and the following experiments we chose two values for the characteristic exponent of the noise, namely $\alpha = 1.5$ and $\alpha = 1.8$, corresponding to a fairly impulsive noise environment ($\alpha = 1.5$) and an almost Gaussian one ($\alpha = 1.8$). Also, in the implementation of ROC-MUSIC, we used $p = 0.8$ for both cases of the stable noise. For this experiment, the GSNR is kept almost constant at 22.3 dB as shown on Table 5.3. As we can also see from Table 5.3, the PSNR is greater for the case of the less impulsive noise ($\alpha = 1.8$). As illustrated in Figure 5.7, the performance

Table 5.3: GSNR and average PSNR for different values of M.

	Number of snapshots M						
	$M = 50$	$M = 100$	$M = 200$	$M = 400$	$M = 800$	$M = 1,600$	$M = 3,200$
GSNR [dB]	22.4003	22.2257	22.2015	22.3551	22.2832	22.2907	22.3166
PSNR [dB] ($\alpha = 1.5$)	-0.6420	-0.6034	-1.0301	-1.4491	-1.2882	-2.1307	-1.9418
PSNR [dB] ($\alpha = 1.8$)	-0.1551	0.1494	-0.1808	-0.5530	-0.1396	-0.4561	-0.3315

Table 5.4: GSNR and average PSNR for different values of γ.

	Noise Dispersion γ						
	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 2$	$\gamma = 4$	$\gamma = 6$	$\gamma = 8$	$\gamma = 10$
GSNR [dB]	25.3202	22.3099	19.2996	16.2893	14.5284	13.2790	12.3099
PSNR [dB] ($\alpha = 1.5$)	-0.7486	-1.5614	-2.9538	-5.0779	-6.6677	-7.9334	-8.9824
PSNR [dB] ($\alpha = 1.8$)	-0.0730	-0.2164	-0.5031	-1.0462	-1.5426	-1.9969	-2.4147

of ROC-MUSIC is superior to that of MUSIC especially as the number of snapshots M increases and more and more impulsive noise samples are incorporated into the data.

Noise dispersion γ. In this experiment, we study the influence of the noise dispersion γ, i.e., the influence of the GSNR to the performance of the methods. The number of snapshots available to the algorithms is $M = 1,000$. The GSNR and average PSNR for this experiment are shown on Table 5.4. The results are depicted in Figure 5.8. Again, for $\alpha = 1.5$ and GSNR$<$ 14 dB, MUSIC fails to resolve the two sources in all 200 Monte Carlo runs.

Angular separation. Figure 5.9 illustrates the variation of the algorithmic performance with respect to the angle separation of the two incoming signals, for $M = 1,000$, GSNR$=$ 22.3 dB, PSNR($\alpha = 1.5$) $= -1.56$ dB, and PSNR($\alpha = 1.8$) $= -0.22$ dB. As expected, the resolution capability of both algorithms improves with increased angle separation between the two sources. But for a given probability of resolution, the ROC-MUSIC algorithm requires a lower angle separation threshold than MUSIC.

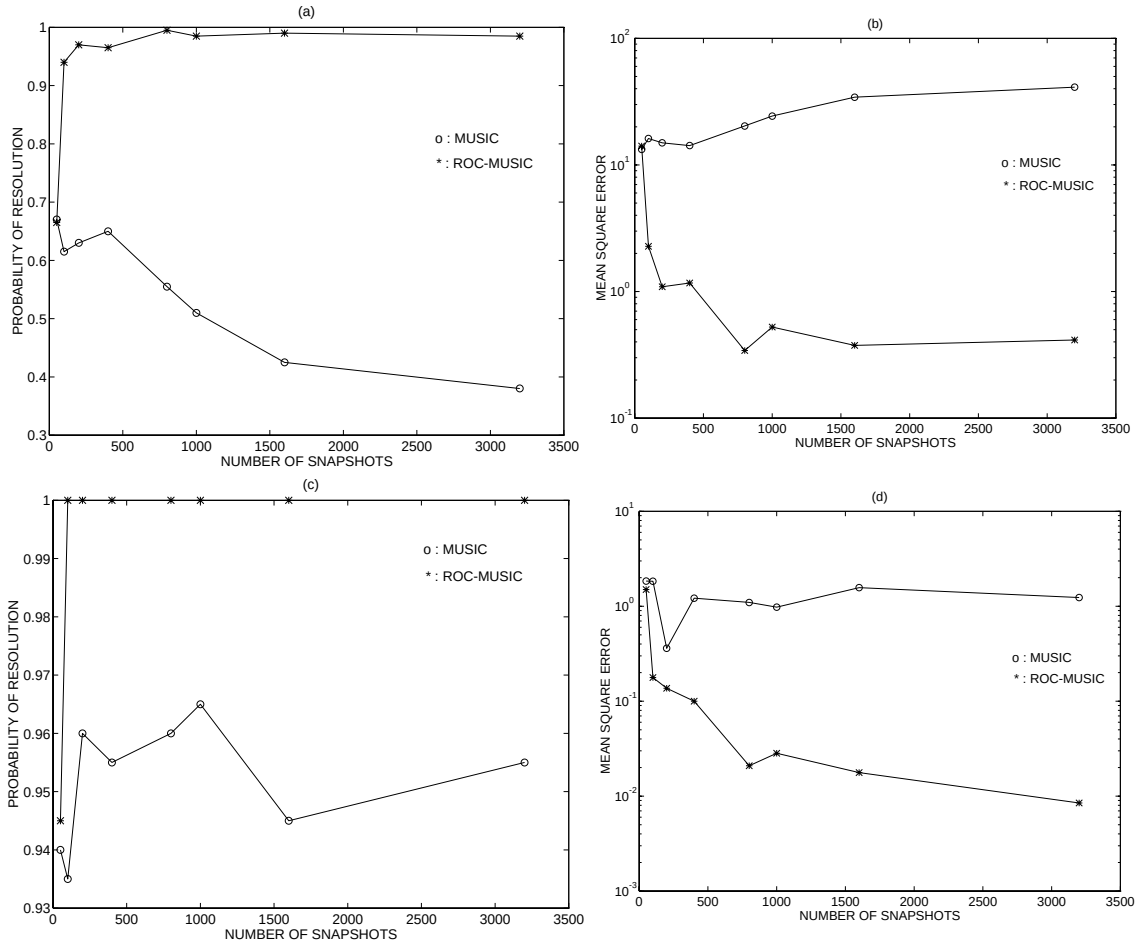


Figure 5.7: Probability of resolution and mean square error as functions of the number of snapshots, (a-b): $\alpha = 1.5$, (c-d): $\alpha = 1.8$.

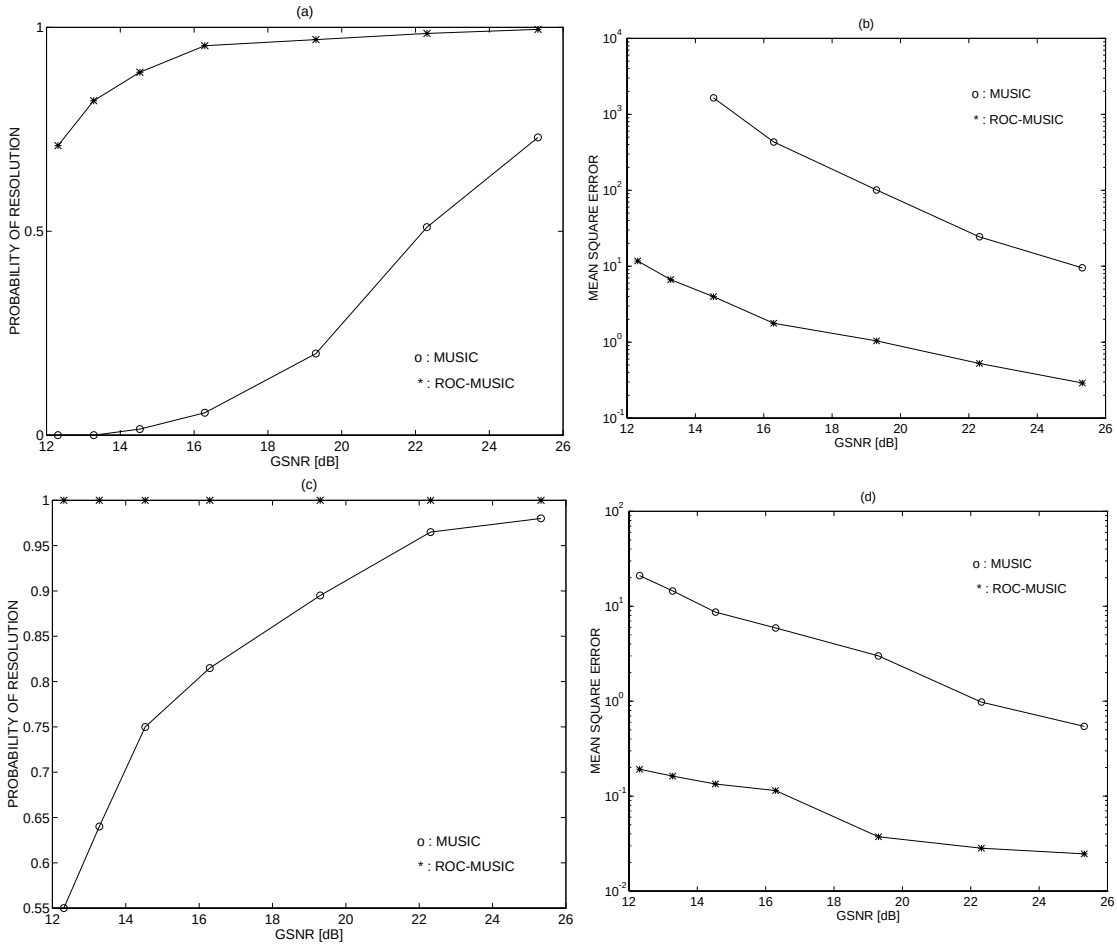


Figure 5.8: Probability of resolution and mean square error as functions of the generalized signal-to-noise ratio (GSNR), (a-b): $\alpha = 1.5$, (c-d): $\alpha = 1.8$.

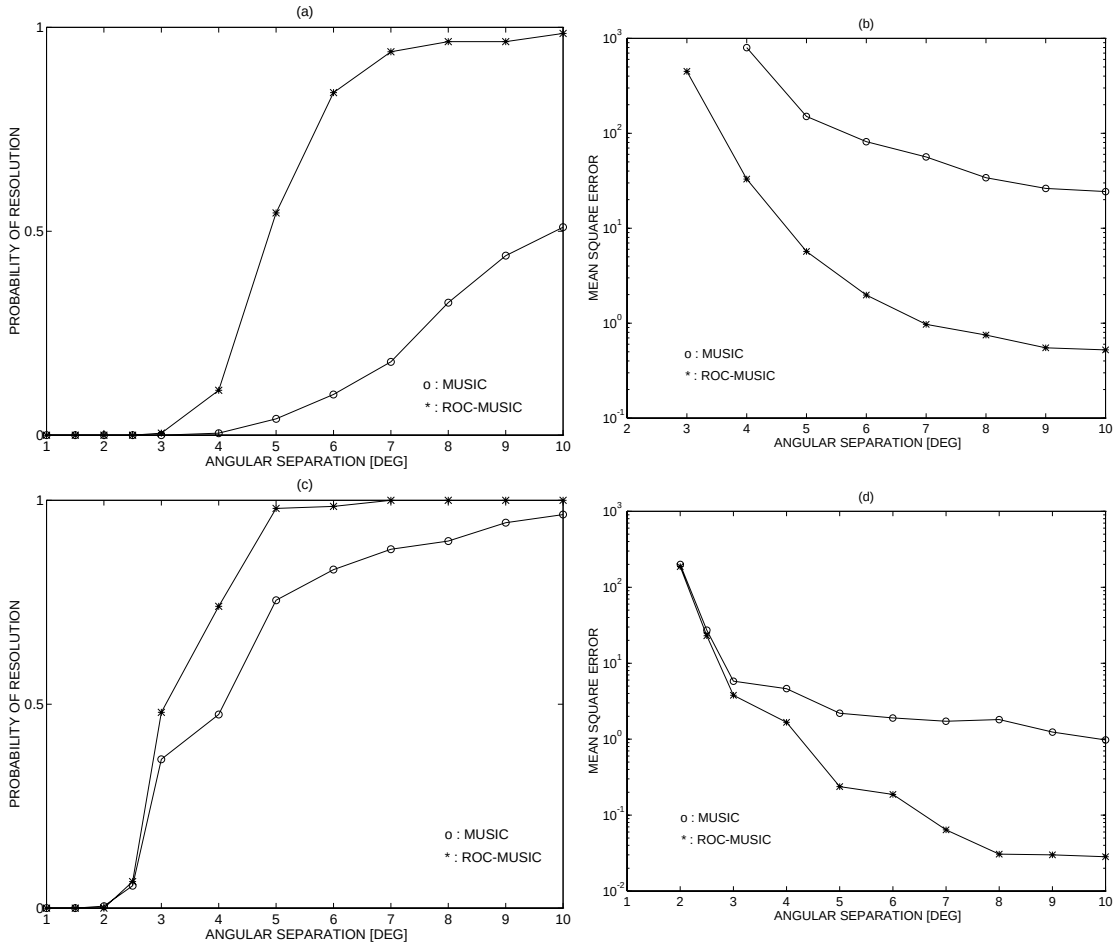


Figure 5.9: Probability of resolution and mean square error as functions of the source angular separation, (a-b): $\alpha = 1.5$, (c-d): $\alpha = 1.8$.

Table 5.5: GSNR and average PSNR for different values of γ (real clutter data).

	Noise Dispersion γ					
	$\gamma = 1$	$\gamma = 2$	$\gamma = 4$	$\gamma = 8$	$\gamma = 16$	$\gamma = 32$
GSNR [dB]	22.3290	19.3187	16.3084	13.2981	10.2878	7.2775
PSNR [dB]	-0.4998	-0.8081	-1.4173	-2.5602	-4.5344	-7.5873

5.3.3 Performance Comparison of ROC-MUSIC versus MUSIC for Real Clutter Data

The proposed algorithm has been tested with real radar sea clutter data provided by the Naval Surface Warfare Center, Carderock Division, Bethesda, Maryland. Clutter is a group of unwanted radar returns due to scattering centers such as precipitation, birds, and ocean waves. The received clutter signal can be represented in terms of its *in-phase* (I) and *quadrature* (Q) components. A typical sample set of the sea clutter data is shown in Figure 5.10. The spiky nature of these radar returns is obvious, and it has been shown, using the algorithms developed in [77], that they can be modeled as $S\alpha S$ processes with $\alpha = 1.85$ and $\gamma = 0.19$.

We studied the performance of ROC-MUSIC versus MUSIC in the presence of radar clutter and the results are shown in Figure 5.11. Two uncorrelated radar waveforms were used and their number was assumed to be known. For the ROC-MUSIC, we used $p = 0.85$ in the expression for the FLOM estimator of the covariation. Figure 5.11(a-b) shows the probability of resolution and the MSE for various values of GSNR (cf. Table 5.5). The number of snapshots was $M = 600$ and the two sources had angles of arrival taking values $\theta_1 = -5^o$ and $\theta_2 = 5^o$. As we can see, for values of GSNR less than 13 dB (PSNR< -2 dB), ROC-MUSIC exhibits a 10 dB gain in MSE performance. Hence, for low GSNR values in which most radars operate, ROC-MUSIC shows a significant performance improvement over MUSIC. Figure 5.11(c-d) depicts the algorithmic behavior as a function of the angle separation when GSNR$= 19.3$ dB (PSNR$= -0.8$ dB).

5.4 Concluding Remarks

We have formulated the covariation matrix of the array outputs for the case of complex isotropic $S\alpha S$ signals and noise. We showed that, for the special case of uncorrelated

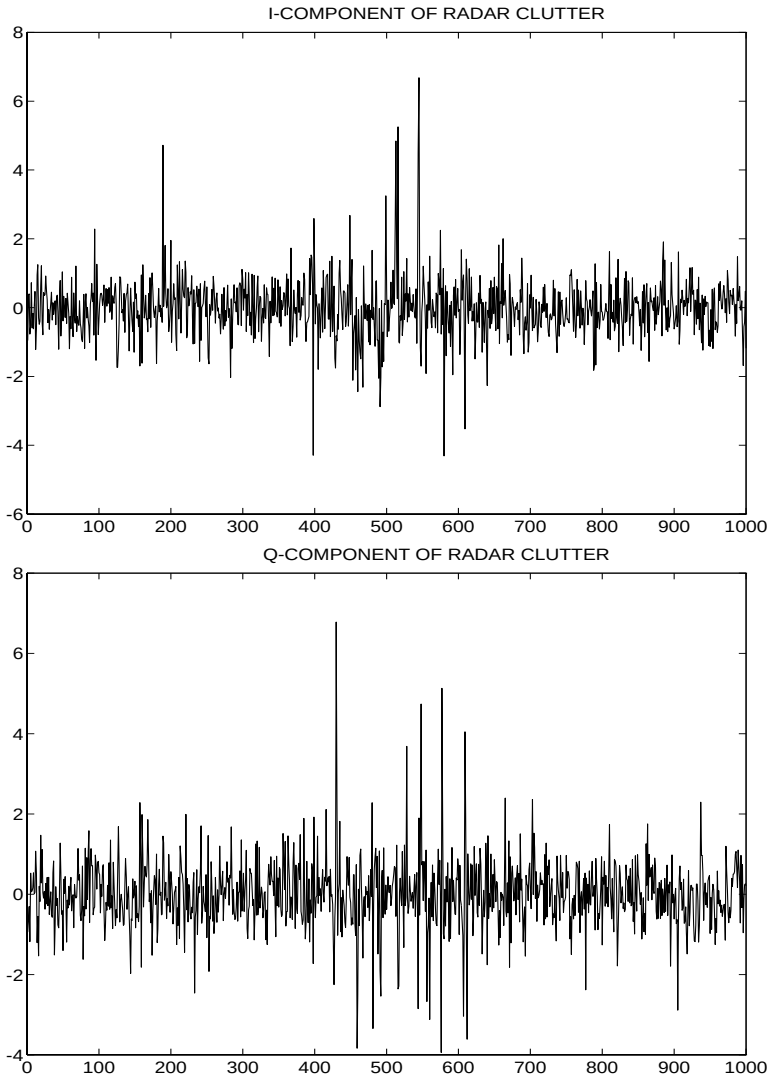


Figure 5.10: In-phase (I) and Quadrature (Q) components of radar clutter.

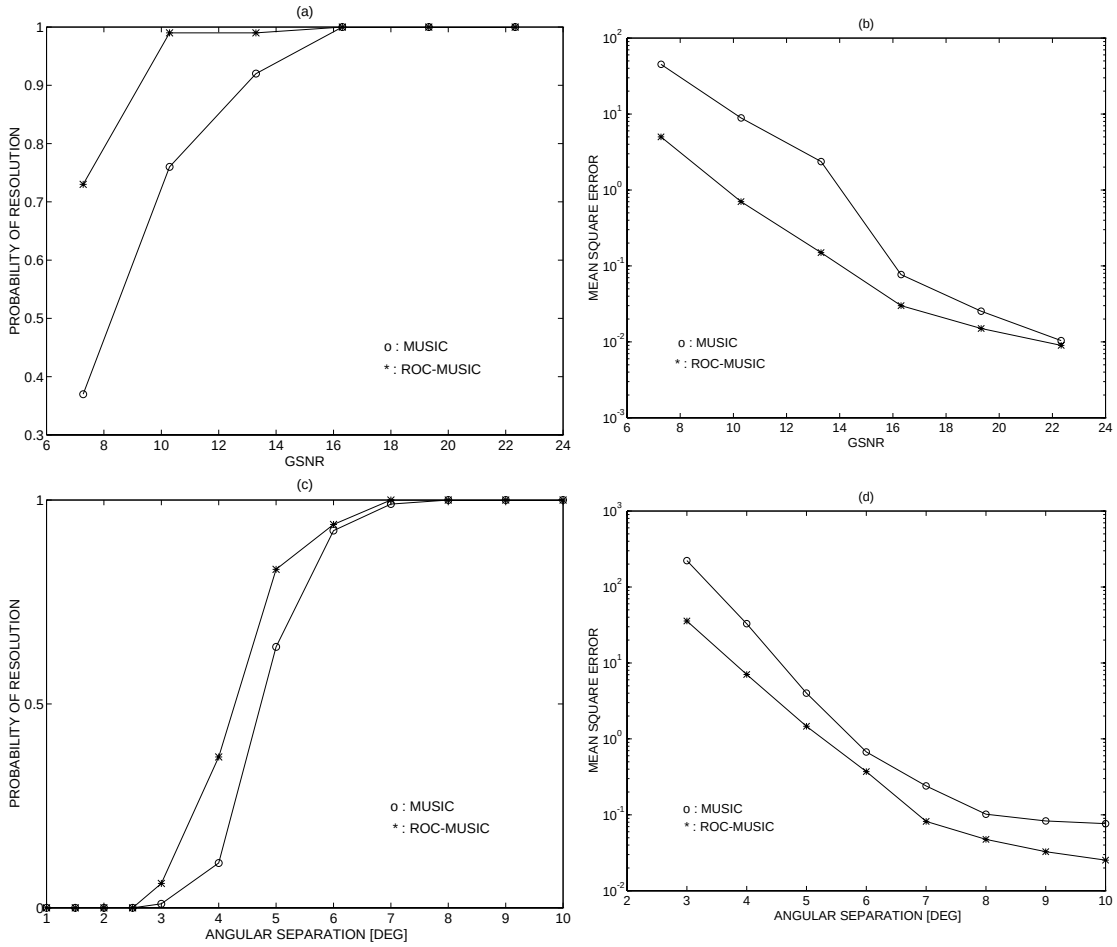


Figure 5.11: Probability of resolution and mean square error as functions of the generalized signal-to-noise ratio (GSNR) (a-b), and of the source angular separation (c-d). Real clutter data experiment.

signals and of an array of sensors with unit amplitude response, the covariation matrix has similar form to the covariance matrix of Gaussian distributed signals. Therefore, subspace-based bearing estimation techniques can be applied to the covariation matrix resulting to improved direction of arrival estimates in the presence of impulsive noise environments. We presented consistent estimators for the marginals of the covariation matrix and we studied their asymptotic performance through both theory and simulations. The improved performance of the proposed ROC-MUSIC algorithm in the presence of a wide range of impulsive noise environments was demonstrated via Monte Carlo experiments.

Chapter 6

Wide-Band Source Localization in the Alpha-Stable Framework

In this chapter, we study the problem of localizing wide-band sources in the presence of noise modeled as a complex isotropic stable process. We consider the frequency-domain representation of the sensor outputs and show that the spectral density of complex stable processes plays a role in array processing problems analogous to that played by the power spectral density of second-order processes.

6.1 Introduction

In Chapter 4, we addressed the solution of the signal parameter estimation problem through the use of sensor array data retrieved in the presence of impulsive interference. We introduced optimal, maximum likelihood-based approaches to the localization problem of *narrow-band* sources in the presence of noise modeled as a complex isotropic Cauchy process. Additionally, in Chapter 5, we developed subspace methods based on fractional lower-order statistics and introduced the ROC-MUSIC algorithm for array processing applications where reduced computational cost is a crucial design parameter. In this chapter, we develop techniques for direction finding of *wide-band* sources in impulsive noise environments. Focused wide-band methods are used so that broad-band sources can be represented by rank-one models at the receiver. Following an approach similar to the one in [40], the covariation matrix of sensor outputs is formed after steering delays are inserted to form a conventional delay-and-sum beamformer. The resulting *steered spectral covariation matrix (SSCM)* focuses wide-band arrivals from its steering direction. By using a different SSCM for every direction of interest, all broad-band sources can be handled by a

rank-one model. With this approach, the full signal bandwidth is exploited and thus, the observation time to achieve high-resolution performance is reduced.

In Section 6.2, we present some necessary preliminaries on the spectral representation of complex stationary α-stable processes. In Section 6.3, we introduce the steered spectral covariation matrix (SSCM) and present wide-band source localization methods in the presence of α-stable distributed signals and noise. Simulation experiments demonstrating the performance of the proposed methods are presented in Section 6.4. Finally, conclusions are drawn in Section 6.5.

6.2 Spectral Representation of Alpha-Stable Processes

Statistical signal processing has traditionally been based on the theory of second-order stationary Gaussian processes, $\{X(t);\ -\infty < t < \infty\}$. According to the spectral representation theorem, this type of processes can be expressed as:

$$X(t) = \int_{-\infty}^{\infty} e^{\jmath t\omega}\, d\xi(\omega); \quad -\infty < t < \infty, \tag{6.1}$$

where the Gaussian process $d\xi(\omega)$ has orthogonal increments and

$$E|d\xi(\omega)|^2 = \phi(\omega)\, d\omega, \tag{6.2}$$

with $\phi(\omega)$ being the power spectral density of $X(t)$. The *covariance function* of $X(t)$ plays an important role in linear prediction and filtering problems and is defined as

$$\text{Covariance}\{X(t), X(s)\} = \int_{-\infty}^{\infty} e^{\jmath(t-s)\omega}\phi(\omega)\, d\omega. \tag{6.3}$$

In this chapter, we consider an important class of stationary processes which have the same spectral representation shown in (6.1), but whose second-order moments are infinite [27]. Namely, we consider strictly stationary complex symmetric alpha-stable ($S\alpha S$) processes ($0 < \alpha < 2$) having the spectral representation (6.1) where $d\xi(\omega)$ is a $S\alpha S$ process with independent increments satisfying

$$\{E|d\xi(\omega)|^p\}^{\alpha/p} = C(p,\alpha)\phi(\omega)\, d\omega \quad \text{for all } \ 0 < p < \alpha, \tag{6.4}$$

where $C(p,\alpha)$ is a constant depending on p and α and $\phi(\omega)$ is a nonnegative function called the *spectral density* of $X(t)$. The spectral density $\phi(\omega)$ describes fully the distribution of the process $X(t)$. The formal definition of a symmetric alpha-stable process $X(t)$ is as follows:

Definition 6.1 *A stochastic process* $\{X(t);\ -\infty < t < \infty\}$ *is called symmetric alpha-stable (*$S\alpha S$*) for* $\alpha \in (0,2]$*, if for every* $n \in \mathcal{N}$ *and any* $a_1, \ldots, a_n \in \mathcal{R}$*,* $t_1, \ldots, t_n \in \mathcal{R}$*, the random variable* $Y = \sum_{i=1}^{n} a_i X(t_i)$ *has a symmetric stable distribution with index* α*.*

The process $\{X(t);\ -\infty < t < \infty\}$ defined by (6.1) is strictly stationary if and only if ξ has isotropic or rotationally invariant increments, i.e., the distribution of the process of increments $e^{\jmath\vartheta}\, d\xi(\omega),\ -\infty < \omega < \infty$, does not depend on the rotation ϑ [47].

The *covariation function* plays for stable processes a role analogous to the role that the covariance plays for Gaussian processes and is given by:

$$[X(t), X(s)]_\alpha \stackrel{\text{def}}{=} \text{Covariation}\{X(t), X(s)\} = \int_{-\infty}^{\infty} e^{\jmath(t-s)\omega} \phi(\omega)\, d\omega. \tag{6.5}$$

As we see in (6.5), the covariation function of an alpha-stable process has an identical representation in terms of the spectral density ϕ as the covariance function of a second-order process. The covariation of complex jointly $S\alpha S$ r.v.'s is not generally symmetric and has the following properties [6]:

Q1 If $X(r)$, $X(s)$ and $X(t)$ are jointly $S\alpha S$, then

$$[aX(r) + bX(s), X(t)]_\alpha = a[X(r), X(t)]_\alpha + b[X(s), X(t)]_\alpha \tag{6.6}$$

for any complex constants a and b.

Q2 If $X(s)$ and $X(t)$ are independent and $X(r)$, $X(s)$ and $X(t)$ are jointly $S\alpha S$, then

$$[aX(r), bX(s) + cX(t)]_\alpha = ab^{<\alpha-1>}[X(r), X(s)]_\alpha + ac^{<\alpha-1>}[X(r), X(t)]_\alpha \tag{6.7}$$

for any complex constants a, b and c.

Q3 If $X(s)$ and $X(t)$ are independent $S\alpha S$, then $[X(s), X(t)]_\alpha = 0$.

6.3 The Wide-Band Source Localization Problem

Consider an array of r sensors that receive signals generated by q *wide-band* sources $s_1, s_2, \ldots, s_q$ with identical bandwidth and locations $\theta_1, \theta_2, \ldots, \theta_q$. The signals are assumed to be noncoherent, ergodic, and stationary complex isotropic $S\alpha S$ random processes. In addition, the noise vector $\mathbf{n}(t) = [n_1(t), n_2(t), \ldots, n_r(t)]^T$ is a complex $S\alpha S$ random process with the same characteristic exponent α as the signals. The signal received at the ith sensor can be expressed as

$$x_i(t) = \sum_{k=1}^{q} s_k(t - \tau_i(\theta_k)) + n_i(t), \tag{6.8}$$

where $\tau_i(\theta_k)$ is the propagation delay between the first and the ith sensor for a waveform coming from direction θ_k. Unlike the narrow band case where the problem is usually formulated in terms of the sampled data, in the wide-band case it is convenient to formulate the problem in the frequency domain. Assuming the availability of M observations in the interval $(-T/2, T/2)$, let $\mathbf{X}(m)$ denote the frequency-domain vector with elements $X_i(m)$ corresponding to the Fourier series coefficients of $x_i(t)$ at frequency $\omega_m = 2\pi m/T$. It holds that $X_i(m)$ can be expressed as

$$\begin{aligned} X_i(m) &= \sum_{t=1}^{M} x_i(t) e^{-\jmath \omega_m t} \\ &= \sum_{t=1}^{M} \left[\sum_{k=1}^{q} s_k(t - \tau_i(\theta_k)) + n_i(t) \right] e^{-\jmath \omega_m t} \\ &= \sum_{k=1}^{q} a_i(\theta_k) S_k(m) + N_i(m), \end{aligned} \tag{6.9}$$

with $S_k(m)$ and $N_i(m)$ corresponding to the Fourier series coefficients of $s_k(t)$ and $n_i(t)$ respectively at frequency $\omega_m = 2\pi m/T$, and where $a_i(\theta_k) = e^{-\tau_i(\theta_k)\omega_m}$. Under the assumption that the sensor outputs are approximately band-limited to $\omega_l \leq \omega \leq \omega_h$, the frequency-domain sensor output vector $\mathbf{X}(m)$ can be expressed in matrix form:

$$\mathbf{X}(m) = \mathbf{A}(\omega_m, \boldsymbol{\theta})\mathbf{S}(m) + \mathbf{N}(m); \quad \omega_l \leq \omega_m \leq \omega_h, \tag{6.10}$$

where

- $\mathbf{A}(\omega_m, \boldsymbol{\theta}) = [\mathbf{a}(\omega_m, \theta_1), \ldots, \mathbf{a}(\omega_m, \theta_q)]$ is the $r \times q$ matrix of the array steering vectors at frequency ω_m, with $\mathbf{a}(\omega_m, \theta_k) = [1, e^{-\jmath\omega_m \tau_2(\theta_k)}, \ldots, e^{-\jmath\omega_m \tau_r(\theta_k)}]^T$ being the frequency steering vector of the array toward direction θ_k and,

- $\mathbf{S}(m) = [S_1(m), \ldots, S_q(m)]^T$ and $\mathbf{N}(m) = [N_1(m), \ldots, N_r(m)]^T$ are the frequency-domain signal and noise vectors, respectively.

Note that, by the stability property of the alpha-stable processes, the frequency-domain elements $X_i(m)$ are also $S\alpha S$ processes with the same characteristic exponent as the time domain samples $x_i(t)$. Also, assuming that the observation time T is large, the Fourier coefficients $\mathbf{X}(m)$ and $\mathbf{X}(n)$ are independent for $m \neq n$ [47].

6.3.1 The Spectral Covariation Matrix

Denote by $\mathbf{K}_s(\omega_m)$ and $\mathbf{K}_n(\omega_m)$ the source and noise spectral covariation matrices, respectively. Under the assumptions made above about the statistical independence of the signals and the noise, both matrices are diagonal with $\mathbf{K}_s(\omega_m) = \text{diag}(\gamma_{s_1}(\omega_m), \ldots, \gamma_{s_q}(\omega_m))$ and $\mathbf{K}_n(\omega_m) = \gamma_n(\omega_m)\mathbf{I}$, where $\gamma_{s_k}(\omega_m) = [S_k(m), S_k(m)]_\alpha$, $k = 1, \ldots, q$, and $\gamma_n(\omega_m) = [N_i(m), N_i(m)]_\alpha$, $i = 1, \ldots, r$.

Equation (6.10) can be written as

$$\mathbf{X}(m) = \mathbf{W}(m) + \mathbf{N}(m); \quad \omega_l \leq \omega_m \leq \omega_h, \tag{6.11}$$

where $\mathbf{W}(m) = \mathbf{A}(\omega_m, \boldsymbol{\theta})\mathbf{S}(m)$. By the stability property, it follows that $\mathbf{W}(m)$ is also a complex isotropic $S\alpha S$ random vector with components

$$W_i(m) = \mathbf{A}_i(\theta_m, \boldsymbol{\theta})\mathbf{S}(m) = a_i(\omega_m, \theta_1)S_1(m) + \cdots + a_i(\omega_m, \theta_q)S_q(m); \quad i = 1, \ldots, r. \tag{6.12}$$

Also, it holds that $\mathbf{W}(m)$ is independent of $\mathbf{N}(m)$.

Now, we define the *spectral covariation matrix*, $\mathbf{K}(\omega_m)$, of the frequency-domain vector process $\mathbf{X}(m)$ as the matrix whose elements are the covariations $[X_i(m), X_j(m)]_\alpha$ of the components of $\mathbf{X}(m)$. We have that

$$\begin{aligned} [\mathbf{K}(\omega_m)]_{ij} &= [X_i(m), X_j(m)]_\alpha = [W_i(m) + N_i(m), W_j(m) + N_j(m)]_\alpha \\ &= [W_i(m), W_j(m)]_\alpha + [W_i(m), N_j(m)]_\alpha + \\ &\quad [N_i(m), W_j(m)]_\alpha + [N_i(m), N_j(m)]_\alpha. \end{aligned} \tag{6.13}$$

By the independence assumption of $\mathbf{W}(m)$ and $\mathbf{N}(m)$ and by property Q3 we have that

$$[W_i(m), N_j(m)]_\alpha = 0, \tag{6.14}$$

and

$$[N_i(m), W_j(m)]_\alpha = 0. \tag{6.15}$$

Also, by using (6.12) and properties Q1 and Q2 it follows that

$$\begin{aligned}
[W_i(m), W_j(m)]_\alpha &= [\sum_{k=1}^{q} a_i(\omega_m, \theta_k) S_k(m), W_j(m)]_\alpha \\
&= \sum_{k=1}^{q} a_i(\omega_m, \theta_k) [S_k(m), W_j(m)]_\alpha \\
&= \sum_{k=1}^{q} a_i(\omega_m, \theta_k) [S_k(m), \sum_{l=1}^{q} a_j(\omega_m, \theta_l) S_l(m)]_\alpha \\
&= \sum_{k=1}^{q} a_i(\omega_m, \theta_k) a_j^{<\alpha-1>}(\omega_m, \theta_k) \gamma_{s_k}(\omega_m),
\end{aligned} \tag{6.16}$$

where $\gamma_{s_k}(\omega_m) = [S_k(m), S_k(m)]_\alpha$. Finally, due to the noise assumption made earlier, it holds that

$$[N_i(m), N_j(m)]_\alpha = \gamma_n(\omega_m) \delta_{i,j}, \tag{6.17}$$

where $\delta_{i,j}$ is the Kronecker delta function. Combining (6.13)-(6.17) we obtain the following expression for the covariations of the sensor measurements:

$$[X_i(m), X_j(m)]_\alpha = \sum_{k=1}^{q} a_i(\omega_m, \theta_k) a_j^{<\alpha-1>}(\omega_m, \theta_k) \gamma_{s_k}(\omega_m) + \gamma_n(\omega_m) \delta_{i,j} \quad i, j = 1, \ldots, r. \tag{6.18}$$

In matrix form, (6.18) gives the following expression for the spectral covariation matrix of the observation vector:

$$\mathbf{K}_X(\omega_m) \stackrel{\triangle}{=} [\mathbf{X}(m), \mathbf{X}(m)]_\alpha = \mathbf{A}(\omega_m, \boldsymbol{\theta}) \mathbf{K}_S(\omega_m) \mathbf{A}^{<\alpha-1>}(\omega_m, \boldsymbol{\theta}) + \gamma_n(\omega_m) \mathbf{I}, \tag{6.19}$$

where the (i,j)th element of matrix $\mathbf{A}^{<\alpha-1>}(\omega_m, \boldsymbol{\theta})$ results from the (j,i)th element of $\mathbf{A}(\omega_m, \boldsymbol{\theta})$ according to the operation

$$[\mathbf{A}^{<\alpha-1>}(\omega_m, \boldsymbol{\theta})]_{i,j} = [\mathbf{A}(\omega_m, \boldsymbol{\theta})]_{j,i}^{<\alpha-1>} = |\, [\mathbf{A}(\omega_m, \boldsymbol{\theta})]_{j,i} \,|^{\alpha-2} \, [\mathbf{A}(\omega_m, \boldsymbol{\theta})]_{j,i}^{*} \tag{6.20}$$

Since the frequency steering vectors are of the form $\mathbf{a}(\omega_m, \theta_k) = [1, e^{-\jmath\omega_m \tau_2(\theta_k)}, \ldots, e^{-\jmath\omega_m \tau_r(\theta_k)}]^T$, it follows that

$$[\mathbf{A}^{<\alpha-1>}(\omega_m, \boldsymbol{\theta})]_{i,j} = |e^{-\jmath\omega_m \tau_j(\theta_i)}|^{\alpha-2} e^{\jmath\omega_m \tau_j(\theta_i)} = [\mathbf{A}(\omega_m, \boldsymbol{\theta})]^*_{j,i}\,, \tag{6.21}$$

and thus the spectral covariation matrix can be written as

$$\mathbf{K}_X(\omega_m) = \mathbf{A}(\omega_m, \boldsymbol{\theta})\mathbf{K}_S(\omega_m)\mathbf{A}^H(\omega_m, \boldsymbol{\theta}) + \gamma_n(\omega_m)\mathbf{I}. \tag{6.22}$$

Clearly, when $\alpha = 2$, i.e., for Gaussian distributed signals and noise, the expression for the spectral covariation matrix is identical to the well-known expression for the power spectral density matrix:

$$\mathbf{P}_X(\omega_m) = \mathbf{A}(\omega_m, \boldsymbol{\theta})\mathbf{P}_S(\omega_m)\mathbf{A}^H(\omega_m, \boldsymbol{\theta}) + \sigma^2(\omega_m)\mathbf{I}, \tag{6.23}$$

where $\mathbf{P}_S(\omega_m)$ is the source power spectral density matrix.

In practice, in order to estimate $\mathbf{K}_X(\omega_m)$, we divide the T second observation into N nonoverlapping segments of ΔT seconds each and apply the discrete Fourier transform to obtain uncorrelated frequency-domain vectors $\mathbf{X}_n(m)$ for each segment. Then, the spectral covariation matrix $\mathbf{K}_X(\omega_m)$ is estimated as

$$\hat{\mathbf{K}}_X(\omega_m) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{X}_n(m)\mathbf{X}_n^{<p-1>}(m), \tag{6.24}$$

for some $0 < p < \alpha$.

Extending the narrow-band source localization techniques of Chapter 5 to the wide-band case, the series of spectral covariation matrices over the receiver band can be used to obtain narrow-band spatial estimation results in each frequency. Then, the bearings of the wide-band sources can be estimated by combining the narrow-band location estimates. Hence, we define $P_{IRM}(\theta)$ as the incoherent wide-band ROC-MUSIC estimator obtained by summing the narrow-band ROC-MUSIC results over the band of interest:

$$P_{IRM}(\theta) = \sum_{\omega_m=\omega_l}^{\omega_h} \frac{1}{\sum_{i=q+1}^{r} |\mathbf{a}^H(\omega_m, \theta)\mathbf{v}_i(\omega_m)|^2}, \tag{6.25}$$

where $\{\mathbf{v}_i(\omega_m);\ \ i = q+1, \ldots, r\}$ are the eigenvectors corresponding to the $r-q$ smallest eigenvalues of the sample spectral covariation matrix $\hat{\mathbf{K}}_X(\omega_m)$ estimated in (6.24).

The performance of the incoherent ROC-MUSIC in localizing wide-band sources is limited because it requires long observation times to obtain statistically stable estimates of the frequency-dependent spectral covariation matrices [40, 83]. In many practical situations, the observation time available at the receiver may be limited due to nonstationary propagation environments and fast moving signal sources. Such a scenario causes a severe degradation to the performance of the incoherent wide-band methods. To reduce the observation time required to achieve high-resolution broad-band source localization, we present in the following section a focused wide-band array processing method which extends results obtained by Krolik [40] to alpha-stable signals and noise.

6.3.2 The Steered Spectral Covariation Matrix

According to the steered spectral covariation methods, steering delays are inserted into the sensor output measurements to ensure that sources from a particular direction have the same rank-one representation at all temporal frequencies. Under the assumption that the sensor outputs are approximately band-limited to $\omega_l \leq \omega \leq \omega_h$, and that the observation time $T >> 2\pi/(\omega_h - \omega_l) + \max(\tau_i(\theta))$, the *steered sensor output vector* $\mathbf{y}(t,\theta)$ at the direction of interest θ is defined as [40]

$$\mathbf{y}(t,\theta) = \sum_{m=l}^{h} \mathbf{T}_m(\theta)\mathbf{X}(m)e^{\jmath\omega_m t}, \tag{6.26}$$

where

$$\mathbf{T}_m(\theta) = \mathrm{diag}\{e^{\jmath\omega_m \tau_0(\theta)}, \ldots, e^{\jmath\omega_m \tau_{r-1}(\theta)}\}. \tag{6.27}$$

By substituting $\mathbf{T}_m(\theta)$ into (6.26), we can express the ith component $y_i(t,\theta)$ of the steered time-domain sensor output vector $\mathbf{y}(t,\theta)$ as follows

$$y_i(t,\theta) = \sum_{m=l}^{h} e^{\jmath\omega_m \tau_i(\theta)} X_i(m) e^{\jmath\omega_m t} = x_i(t + \tau_i(\theta)). \tag{6.28}$$

Hence, the steered output vector can be expressed as

$$\mathbf{y}(t,\theta) = [x_0(t + \tau_0(\theta)), \ldots, x_{r-1}(t + \tau_{r-1}(\theta))]^T \tag{6.29}$$

Now, we define the *Steered Spectral Covariation Matrix (SSCM)*, $\mathbf{C}(\theta)$, of the steered sensor output process $\mathbf{y}(t,\theta)$ as the matrix whose elements are the covariations

$[y_i(t,\theta), y_j(t,\theta)]_\alpha$ of the components of $\mathbf{y}(t,\theta)$. Assuming that the observation time T is large, the Fourier coefficients $\mathbf{X}(m)$ and $\mathbf{X}(n)$ are independent for $m \neq n$ [47]. Then, by following similar steps as in the derivation of the spectral covariation matrix in Section 6.3.1, $\mathbf{C}(\theta)$ can be expressed as

$$\mathbf{C}(\theta) = \sum_{m=l}^{h} \mathbf{T}_m(\theta)\mathbf{K}(\omega_m)\mathbf{T}_m^H(\theta), \tag{6.30}$$

where $\mathbf{K}(\omega_m) = [\mathbf{X}(m), \mathbf{X}(m)]_\alpha$ is the covariation matrix of the Fourier coefficient vector at frequency ω_m. The structure of $\mathbf{C}(\theta)$ is apparent when we consider the case of a line array with sensors equally spaced at distance d apart. In this case, $\tau_i(\theta) = i\tau(\theta)$, where $\tau(\theta) = (d/c)\sin(\theta)$ and c is the propagation speed. Then, according to the model of (6.8) the ijth element of $\mathbf{C}(\theta)$ is given by

$$\begin{aligned}
[\mathbf{C}(\theta)]_{ij} &= [y_i(t,\theta), y_j(t,\theta)]_\alpha = [x_i(t+\tau_i(\theta)), x_j(t+\tau_j(\theta))]_\alpha \\
&= [\sum_{k=1}^{q} s_k(t+\tau_i(\theta)-\tau_i(\theta_k)) + n_i(t+\tau_i(\theta)), \\
&\quad \sum_{k'=1}^{q} s_{k'}(t+\tau_j(\theta)-\tau_j(\theta_{k'})) + n_j(t+\tau_j(\theta))]_\alpha \\
&= \sum_{k=1}^{q} \zeta_k(h(\tau(\theta)-\tau(\theta_k))) + \gamma_n\delta_{i,j},
\end{aligned} \tag{6.31}$$

where $h = i - j$, $\zeta_k(h(\tau(\theta)-\tau(\theta_k))) = [s_k(t+i(\tau(\theta)-\tau(\theta_k))), s_k(t+j(\tau(\theta)-\tau(\theta_k)))]_\alpha$ is the covariation function of the kth source signal and $\gamma_n = [n_i(t), n_i(t)]_\alpha$. Equation (6.31) is valid because we assumed that the signals and noise are stationary and uncorrelated among them stable processes. For steering direction $\theta = \theta_l$ we have that

$$[\mathbf{C}(\theta_l)]_{ij} = \sum_{k=1,k\neq l}^{q} \zeta_k(h(\tau(\theta_l)-\tau(\theta_k))) + \gamma_{s_l} + \gamma_n\delta_{i,j}, \tag{6.32}$$

where $\gamma_{s_l} = [s_l(t), s_l(t)]_\alpha$ is the covariation of the lth source. Hence, the spectral covariation matrix $\mathbf{C}(\theta)$ steered at the direction of a source contains a constant component equal to the source covariation regardless of the source's spectral signature. Also, the off-steering direction summation terms in (6.32) decrease with increasing h and increasing separation $|\tau(\theta_l) - \tau(\theta_k)|$ from the steering direction. It follows that the spatial spectral signature of a

source in the steering direction $\theta = \theta_l$ can be estimated by measuring the constant component γ_{s_l}. The resulting spatial spectral estimate is called the *Steered Spectral Covariation (SSC)* method given by

$$P_{SSC}(\theta) = [\mathbf{1}^H \mathbf{C}(\theta)^{-1} \mathbf{1}]^{-1}, \tag{6.33}$$

where $\mathbf{1}$ is an $r \times 1$ vector of ones, and $\mathbf{C}(\theta)$ is given by (6.30). In practice, the SSCM $\mathbf{C}(\theta)$ is computed for each steering direction, θ, of interest. Hence, SSCM-based methods are more computationally intensive than incoherent subspace methods. Their advantage lies in the fact that they take advantage of the full time-bandwidth product of the observations, thus resulting in more stable statistical estimates.

6.4 Simulation Results

In this section, the performance of the proposed steered spectral covariation (SSC) method for localizing wide-band sources is assessed through simulation experiments. We use a uniformly spaced linear array consisting of $r = 10$ omnidirectional sensors which are spaced a half-wavelength apart at the normalized frequency of π radians. The two source signals are modeled as mutually uncorrelated communication signals with bandpass spectral spectra with $\omega_l = 0.125\pi$ and $\omega_h = \pi$. The source directions are -20^o and 30^o. The additive noise process is modeled as a stable process with flat spectral density over the same bandpass range as the signal. A total of 8 narrow-band spectral covariation matrices $\mathbf{K}(\omega_m)$, equally spaced from ω_l to ω_h, were used when applying the IRM and SSC methods (cf. (6.25) and (6.33), respectively). We also implement the incoherent MUSIC (IM) and steered minimum variance (STMV) [40] methods which are based on a second-order statistics formulation.

We study the estimation accuracy of the four methods as a function of the observation time (number of snapshots at each frequency, N) and the characteristic exponent α of the signals and noise. The corresponding PSNR and GSNR values as functions of the experiment parameters are shown in Tables 6.1 and 6.2.

In Figure 6.1 we plot spatial spectral estimates obtained in ten independent trials for the four methods. Ten independent trials per method are shown with $N = 625$ snapshots of data at each frequency ω_m; $m = 1, \ldots, 8$. The characteristic exponent of the additive stable noise is $\alpha = 1.5$. The larger variability of the incoherent methods is obvious from this figure.

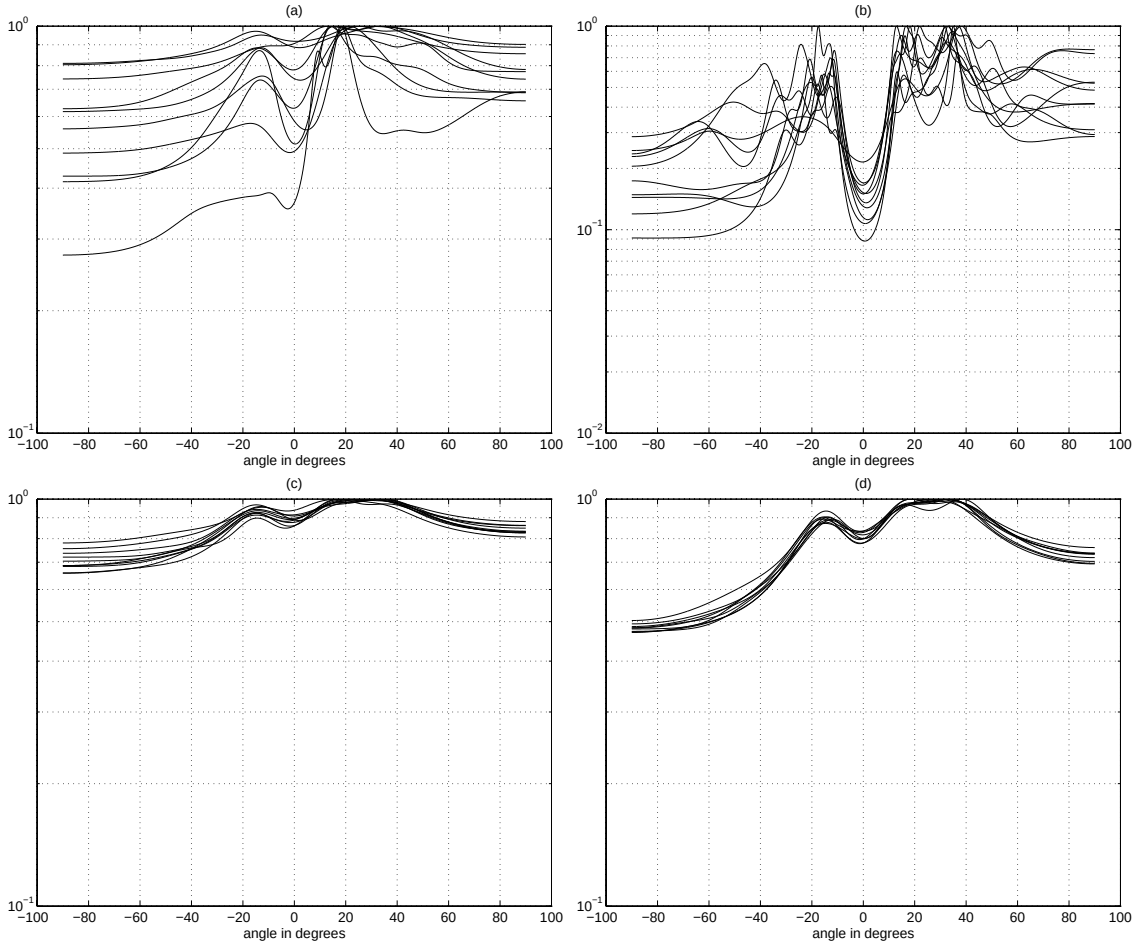


Figure 6.1: Broad-band spatial spectral estimates for the incoherent MUSIC (IM) (a), incoherent ROC-MUSIC (IRM) (b), steered minimum variance (STMV) (c), and steered spectral covariation (SSC) (d). Ten independent trials per method with $N = 625$ snapshots of data at each frequency ω_m.

Table 6.1: GSNR and average PSNR for different values of N ($\alpha = 1.5$).

	Number of snapshots at frequency ω_m, N			
	$N = 625$	$N = 1,250$	$N = 2,500$	$N = 5,000$
GSNR [dB]	22.5643	22.8876	22.6512	22.4403
PSNR [dB]	4.9623	4.2564	3.5671	3.2155

Table 6.2: GSNR and average PSNR for different values of α ($N = 625$).

	Noise Characteristic Exponent, α						
	$\alpha = 1.0$	$\alpha = 1.2$	$\alpha = 1.4$	$\alpha = 1.6$	$\alpha = 1.8$	$\alpha = 1.9$	$\alpha = 2.0$
GSNR [dB]	22.8767						
PSNR [dB]	3.3682	4.6223	4.8075	5.1944	6.6755	9.0353	11.0233

Figure 6.2(a) shows the resulting MSE of the estimated DOA as a function of the number of snapshots, N, at each frequency ω_m for the different methods. Figure 6.2(b) shows the resulting MSE curves as functions of the characteristic exponent α. The number of snapshots at each frequency available to the receiver is $N = 625$. The GSNR is 22.8767 dB and is shown together with the average PSNR, on Table 6.2. Clearly, for non-Gaussian additive noise, SSC exhibits less mean-square estimation error than the other three methods.

6.5 Concluding Remarks

Until recently, statistical signal processing with alpha-stable distributions has not been popular due to the fact that the linear space of a stable process is not a Hilbert space, as in the case of Gaussian processes, but either a Banach ($1 < \alpha < 2$) or a metric space ($0 < \alpha \leq 1$) both of which are more unyielding in their structure. In this chapter, we presented new approaches to the wide-band DOA estimation problem in the presence of impulsive interference. We defined the steered spectral covariation matrix of an array of sensors, and applied steered subspace-based bearing estimation techniques resulting to improved bearing estimates in the presence of impulsive additive noise.

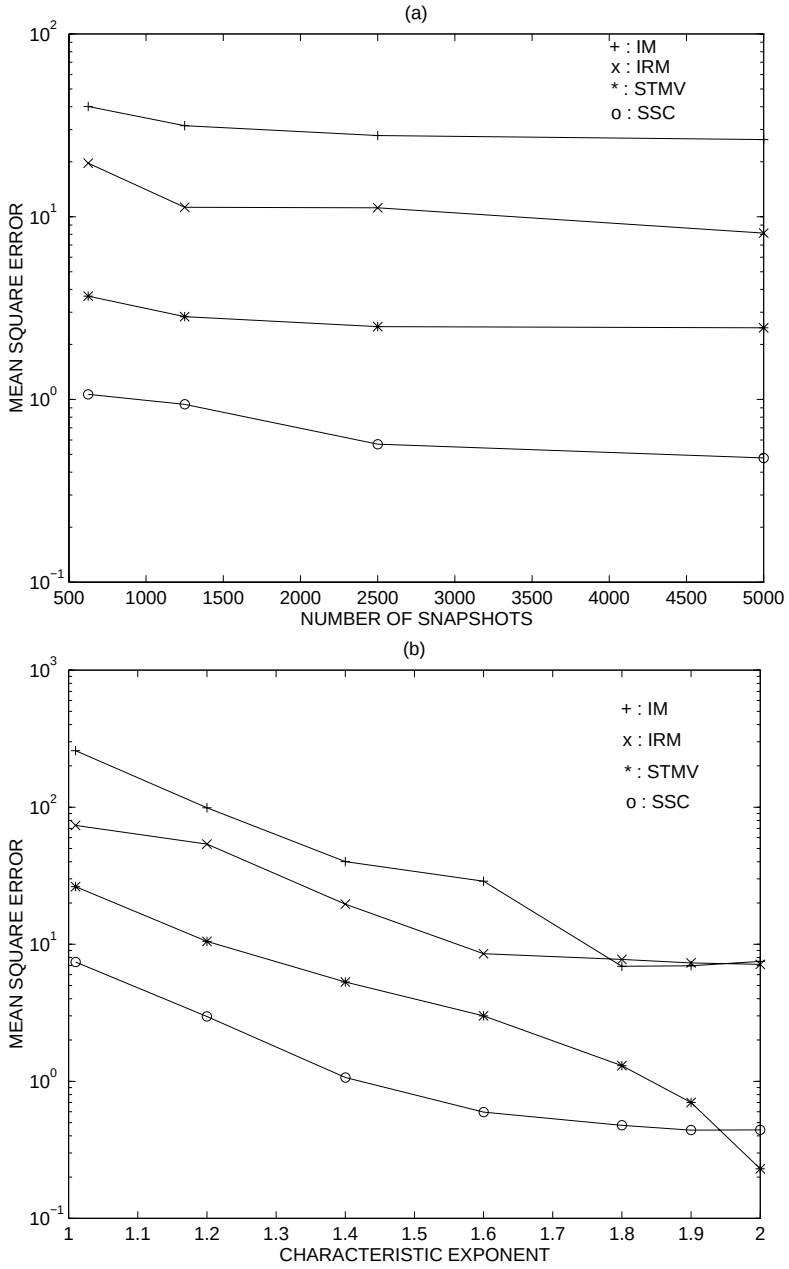


Figure 6.2: MSE curves of the estimated DOA as functions of the number of snapshots N (a), and the characteristic exponent α (b).

Chapter 7

Future Work

Traditionally, classical array signal processing has dealt with two major aspects of spatial processing, namely localization of a signal of interest or of an interfering signal and adaptation of the spatial response of an array of sensors to achieve null steering in a given direction. The achieved spatial focusing in the direction of interest and rejection of jamming signals makes array signal processing a necessary component of any wireless communication system. Furthermore, antenna arrays have been used in radar and sonar applications to detect targets and to estimate important parameters such as target range, velocity, and location.

In today's era of multimedia technologies, statistical array signal processing is used for fast blind equalization of multipath channels without the use of a training sequence, as well as for blind signal copy, i.e., for blind estimation of the transmitted waveform. Also, phased antenna arrays play an increasingly crucial role in remote sensing and imaging with applications in diverse fields such as radar, sonar, geophysical and celestial exploration and mapping, diagnostic medicine, and climate monitoring, to name a few.

Hence, today more than ever, successful research in statistical signal processing is dependent upon consideration of the interrelations of the different components of a signal processing system. Phased array imaging is an example of such an application. On one hand, it requires very high resolution, adaptive beamforming, and blind calibration capabilities in the array processing unit. On the other hand, it requires sophisticated spectral estimation and signal enhancement methods in the image processing unit and in the object detection and identification units.

In this dissertation, we have demonstrated the importance of extending the statistical array signal processing methodology to the so-called alpha-stable framework. We developed array processing methods for a larger class of random processes which include the

Gaussian processes as special elements. Our proposed methods can be applied in environments which, while sharing many characteristics, also differ from Gaussian environments in significant ways. In particular, we developed techniques such as the Cauchy Beamformer and the ROC-MUSIC algorithm, which perform very robustly in detecting and localizing sources buried a wide range of impulsive noise environments.

Part of our future research interests lies in the wireless distributed multimedia communication networks where the spatial selectivity of an array of sensors can be used in order to operate in close-frequency bands and suppress undesirable noise and interference in favor of the signal of interest. Our emphasis will be in the development of methods for direction finding, null- and beam-steering, and waveform recovery in mobile networks. The objective will be to permit successful communication and signal direction-of-arrival tracking in the presence of a large number of sidelobe interferers in a possibly hostile environment.

Several major design considerations are involved in the aforementioned scenario in which moving antenna arrays play an important role. First, signal processing design must accommodate arrays with a time-varying manifold. The calibration of such arrays will have to be done blindly since in many situations it is either impossible or impractical to deploy calibrating sources. Secondly, the environment can be highly non-stationary and unpredictable, characterized by unfavorable ever-changing propagation conditions due to the presence of impulsive interferers and jammers. Therefore, the design of robust beamformers that perform well in a wide range of interference and noise environments is of great importance. Thirdly, the proposed solutions must be robust in the face of hardware problems. Since it is costly and sometimes impossible to replace faulty hardware, the proposed methods should be able to work by suitable processing of the remaining available measurements.

7.1 Detection of Sources in Impulsive Interference Environments

The detection of the number of sources impinging on a sensor array is a critical and difficult issue. Wax *et al.* [86, 88, 89] approached the problem based on the application of information theoretic criteria for model selection introduced by Akaike (AIC) and Schwartz and Rissanen (MDL). Unlike conventional hypothesis-testing-based methods that require subjective threshold settings, the information theoretic approaches give the number of

signals by minimizing the AIC or the MDL criteria. We propose to reformulate the information theoretic criteria for the case of signals in the presence of additive $S\alpha S$ noise. We expect the new methods to be able to detect the number of signals in severe impulsive interference environments.

7.2 Moving Arrays and Array Calibration in Severe Noise

The array processing methods mentioned in the previous chapters assume a static calibrated array whose manifold is precisely known. This is not a realistic assumption in a scenario where the antenna arrays are mounted on moving platforms. Relatively little work has been done in the area of time-varying arrays. Notably, this type of problem has been investigated for underwater sonar systems employing a spatially-referenced towed array consisting of a towed line subarray and a fixed position line array of passive equispaced hydrophones [17]. Further work by Zeira and Friedlander showed that the use of Doppler information in antenna arrays, mounted on platforms moving in an arbitrary but known way, makes them more robust to ambiguity errors than static arrays of comparable dimensions [95].

Sensor location uncertainty poses the even more difficult task of array calibration. This has to be done blindly and requires the calibration of sensors with imprecisely known locations, gains, and phases [60]. Rockah and Schultheiss have studied this problem using the Cramér-Rao bound (CRB) for the bearing estimates in the presence of sensor location uncertainty. Friedlander approached the calibration problem in a "blind" way, i.e., without knowledge of the direction of the calibrating source and with minimal knowledge about the array manifold [22].

In our future research we will extend our methods to handle sensor location uncertainty for the case of moving platforms in impulsive noise environments. Also, we will develop iterative ML methods for simultaneously estimating source signal locations, waveforms, and sensor locations.

Appendix A

Derivation of CRB for Complex Isotropic Cauchy Noise

In this appendix, we derive the CRB for the most general case of multiple signal sources in the presence of complex isotropic Cauchy noise of unknown dispersion. As a first step, we derive a useful proposition about the magnitude r and phase ϕ of the noise:

Proposition A.1 *The amplitude r and phase ϕ of the noise following the complex isotropic Cauchy distribution are independent and the magnitude r satisfies:*

$$E\left\{\frac{r^2}{(\gamma^2+r^2)^2}\right\}=\frac{2}{15\gamma^2}. \tag{A.1}$$

Proof The complex noise samples $n_i(t)=|n_i(t)|e^{\jmath\phi_i(t)};\quad i=1,\ldots,p;\quad t=1,\ldots,M$, follow the bivariate isotropic Cauchy distribution with dispersion γ. In other words, the real and imaginary parts of the noise $n=n_{\Re}+\jmath n_{\Im}$ are jointly Cauchy with probability density $f_1(n_{\Re},n_{\Im})$ given by:

$$f_1(n_{\Re},n_{\Im})=\frac{1}{2\pi}\frac{\gamma}{(n_{\Re}^2+n_{\Im}^2+\gamma^2)^{3/2}}\qquad -\infty<n_{\Re},n_{\Im}<\infty. \tag{A.2}$$

As we can see, the probability density $f_1(n_{\Re},n_{\Im})$ can be expressed as a function of the noise magnitude $r=|n|=\sqrt{n_{\Re}^2+n_{\Im}^2}$,

$$f_1(n_{\Re},n_{\Im})=\chi_1(r)=\frac{1}{2\pi}\frac{\gamma}{(r^2+\gamma^2)^{3/2}}\qquad r\geq 0. \tag{A.3}$$

The joint density of the noise magnitude r and phase $\phi=\arctan(n_{\Im}/n_{\Re})$ is given by [53]:

$$f(r,\phi)=r f_1(r\cos\phi,r\sin\phi)\qquad r\geq 0,\ \ \phi\in[0,2\pi]. \tag{A.4}$$

From (A.3), it follows that

$$f(r,\phi) = \frac{1}{2\pi}\frac{\gamma r}{(r^2+\gamma^2)^{3/2}} \qquad r \geq 0, \quad \phi \in [0, 2\pi]. \tag{A.5}$$

Since the density $f(r,\phi)$ is independent of ϕ, the noise phase ϕ is uniformly distributed in $[0, 2\pi]$ and is independent of the magnitude r. Hence,

$$f(r) = \frac{\gamma r}{(r^2+\gamma^2)^{3/2}} \qquad r \geq 0, \tag{A.6}$$

and

$$f(\phi) = \frac{1}{2\pi} \qquad \phi \in [0, 2\pi]. \tag{A.7}$$

By using (A.6), we get

$$\begin{aligned} E\left\{\frac{r^2}{(\gamma^2+r^2)^2}\right\} &= \int_0^{+\infty} \frac{r^2}{(\gamma^2+r^2)^2}\frac{\gamma r}{(\gamma^2+r^2)^{3/2}}dr \\ &= \gamma\int_0^{+\infty}\frac{r^3}{(\gamma^2+r^2)^{7/2}}dr. \end{aligned} \tag{A.8}$$

A general form of the above integral can be found in [25]

$$\int_0^{+\infty}\frac{r^{n+1}}{(\alpha r^2+2\beta r+c)^{n+3/2}}dr = \frac{n!}{(2n+1)!!\sqrt{\alpha}(\sqrt{\alpha c}+\beta)^{n+1}}, \tag{A.9}$$

where $\alpha \geq 0$, $c > 0$, and $\beta > -\sqrt{\alpha c}$. Setting $n = 2$, $\alpha = 1$, $c = \gamma^2$, and $\beta = 0$, we get

$$\int_0^{+\infty}\frac{r^3}{(\gamma^2+r^2)^{7/2}}dr = \frac{2}{15\gamma^3}, \tag{A.10}$$

and (A.1) follows. ∎

Now, we proceed by proving Theorem 4.1. Let $\boldsymbol{\eta}$ represent the $(2Mq+q+1)$-dimensional vector of unknown parameters in the stochastic model:

$$\begin{aligned} \boldsymbol{\eta} &= [\gamma, s_{1,\Re}(1), \ldots, s_{q,\Re}(1), s_{1,\Im}(1), \ldots, s_{q,\Im}(1), \ldots, \\ &\quad s_{1,\Re}(M), \ldots, s_{q,\Re}(M), s_{1,\Im}(M), \ldots, s_{q,\Im}(M), \boldsymbol{\theta}^T]^T \\ &= [\gamma, \mathbf{s}_{\Re}^T(1), \mathbf{s}_{\Im}^T(1), \ldots, \mathbf{s}_{\Re}^T(M), \mathbf{s}_{\Im}^T(M), \boldsymbol{\theta}^T]^T, \end{aligned} \tag{A.11}$$

where $\mathbf{s}_{\Re}(t) = [s_{1,\Re}(t), \ldots, s_{q,\Re}(t)]^T$, $\mathbf{s}_{\Im}(t) = [s_{1,\Im}(t), \ldots, s_{q,\Im}(t)]^T$; $t = 1, \ldots, M$, and $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_q]^T$ is the vector of the unknown directions of arrival. The Fisher Information Matrix $\mathbf{J}(\boldsymbol{\eta})$ is defined as

$$\mathbf{J}(\boldsymbol{\eta}) = E\left\{ \left(\frac{\partial L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})}{\partial \boldsymbol{\eta}} \right) \left(\frac{\partial L(\mathbf{X}; \gamma, \mathbf{S}, \boldsymbol{\theta})}{\partial \boldsymbol{\eta}} \right)^T \right\}. \tag{A.12}$$

First, we calculate the derivatives of the log likelihood function given in (4.3) with respect to the components of $\boldsymbol{\eta}$. We have

$$\frac{\partial L}{\partial \theta_k} = 3 \sum_{t=1}^{M} \sum_{i=1}^{r} \frac{\Re\{s_k^*(t) d_i^*(\theta_k) n_i(t)\}}{\gamma^2 + |n_i(t)|^2}; \quad k = 1, \ldots, q, \tag{A.13}$$

where $d_i(\theta_k) = \frac{\partial a_i(\theta_k)}{\partial \theta_k}$. In addition,

$$\frac{\partial L}{\partial \gamma} = \frac{Mr}{\gamma} - 3\gamma \sum_{t=1}^{M} \sum_{i=1}^{r} \frac{1}{\gamma^2 + |n_i(t)|^2}, \tag{A.14}$$

$$\frac{\partial L}{\partial s_{k,\Re}(t)} = 3 \sum_{i=1}^{r} \frac{\Re\{a_i^*(\theta_k) n_i(t)\}}{\gamma^2 + |n_i(t)|^2}; \quad k = 1, \ldots, q; \;\; t = 1, \ldots, M, \tag{A.15}$$

and

$$\frac{\partial L}{\partial s_{k,\Im}(t)} = 3 \sum_{i=1}^{r} \frac{\Im\{a_i^*(\theta_k) n_i(t)\}}{\gamma^2 + |n_i(t)|^2}; \quad k = 1, \ldots, q; \;\; t = 1, \ldots, M. \tag{A.16}$$

In the following derivations, we will extensively use assumption A.4 of Section 2.1 which states that the noise samples are spatially and temporally independent, and Proposition A.1 which states that the noise phase is uniformly distributed in $[0, 2\pi]$, in order to get simplifications in the expressions. First, we have

$$\begin{aligned} E\left(\frac{\partial L}{\partial \theta_k} \frac{\partial L}{\partial \theta_l} \right) &= 9E\left\{ \sum_{t=1}^{M} \sum_{t'=1}^{M} \sum_{i=1}^{r} \sum_{j=1}^{r} \frac{\Re\{s_k^*(t) d_i^*(\theta_k) n_i(t)\}}{\gamma^2 + |n_i(t)|^2} \frac{\Re\{s_l^*(t') d_j^*(\theta_l) n_j(t')\}}{\gamma^2 + |n_j(t')|^2} \right\} \\ &= 9E\left\{ \sum_{t=1}^{M} \sum_{i=1}^{r} \frac{\Re\{s_k^*(t) d_i^*(\theta_k) n_i(t)\}}{\gamma^2 + |n_i(t)|^2} \frac{\Re\{s_l^*(t) d_i^*(\theta_l) n_i(t)\}}{\gamma^2 + |n_i(t)|^2} \right\}. \end{aligned} \tag{A.17}$$

Setting $n_i(t) = |n_i(t)|e^{j\phi_i(t)}$, $s_k(t) = |s_k(t)|e^{j\beta_k(t)}$, and $a_i^*(\theta_k) = e^{j\omega_i(\theta_k)}$, (A.17) can be written as

$$\begin{aligned}
E\left(\frac{\partial L}{\partial \theta_k}\frac{\partial L}{\partial \theta_l}\right) &= 9E\left\{\sum_{t=1}^{M}\sum_{i=1}^{r}|s_k(t)d_i(\theta_k)|\frac{|n_i(t)|}{\gamma^2+|n_i(t)|^2}\right. \\
&\quad \cos(-\beta_k(t)+\omega_i(\theta_k)+\frac{\pi}{2}+\phi_i(t)) \\
&\quad \left.|s_l(t)d_i(\theta_l)|\frac{|n_i(t)|}{\gamma^2+|n_i(t)|^2}\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}+\phi_i(t))\right\} \\
&= 9\sum_{t=1}^{M}\sum_{i=1}^{r}|s_k(t)d_i(\theta_k)||s_l(t)d_i(\theta_l)|E\left\{\frac{|n_i(t)|^2}{\left(\gamma^2+|n_i(t)|^2\right)^2}\right\} \\
&\quad \frac{1}{2}\cos\left(\beta_k(t)-\beta_l(t)-\omega_i(\theta_k)+\omega_i(\theta_l)\right) \\
&= \frac{3}{5\gamma^2}\sum_{t=1}^{M}\sum_{i=1}^{r}\Re\left\{s_k(t)s_l^*(t)d_i(\theta_k)d_i^*(\theta_l)\right\}; \quad k,l=1,\ldots,q. \qquad \text{(A.18)}
\end{aligned}$$

Equation (A.18) can be written compactly as

$$E\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = \frac{3}{5\gamma^2}\sum_{t=1}^{M}\Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\mathbf{D}\mathbf{S}(t)\right\}. \qquad \text{(A.19)}$$

Now,

$$\begin{aligned}
E\left(\frac{\partial L}{\partial \gamma}\right)^2 &= \frac{M^2r^2}{\gamma^2} - 6Mr\sum_{t=1}^{M}\sum_{i=1}^{r}E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\} \\
&\quad +9\gamma^2\sum_{t=1}^{M}\sum_{i=1}^{r}E\left\{\frac{1}{\left(\gamma^2+|n_i(t)|^2\right)^2}\right\} \\
&\quad +9\gamma^2\underbrace{\sum_{t=1}^{M}\sum_{t'=1}^{M}\sum_{i=1}^{r}\sum_{j=1}^{r}}_{t'\neq t \text{ or } j\neq i}E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\}E\left\{\frac{1}{\gamma^2+|n_j(t')|^2}\right\} \\
&= \frac{M^2r^2}{\gamma^2} - 6M^2r^2E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\} + 9\gamma^2MrE\left\{\frac{1}{\left(\gamma^2+|n_i(t)|^2\right)^2}\right\} \\
&\quad +9\gamma^2Mr(Mr-1)\left(E\left\{\frac{1}{\gamma^2+|n_i(t)|^2}\right\}\right)^2 \\
&= \frac{M^2r^2}{\gamma^2} - 6M^2r^2\frac{1}{3\gamma^2} + 9\gamma^2Mr\frac{1}{5\gamma^4} + 9\gamma^2Mr(Mr-1)\frac{1}{9\gamma^4} \qquad \text{(A.20)}
\end{aligned}$$

Therefore,

$$E\left(\frac{\partial L}{\partial \gamma}\right)^2 = \frac{4}{5}\frac{Mr}{\gamma^2}. \tag{A.21}$$

In addition,

$$\begin{aligned}
E\left(\frac{\partial L}{\partial s_{k,\Re}(t)}\frac{\partial L}{\partial s_{l,\Re}(t')}\right) &= 9E\left\{\sum_{i=1}^{r}\sum_{j=1}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_j^*(\theta_l)n_j(t')\}}{\gamma^2+|n_j(t')|^2}\right\} \\
&= 9E\left\{\sum_{i=1}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_i^*(\theta_l)n_i(t')\}}{\gamma^2+|n_i(t')|^2}\right\} \\
&\quad +9E\left\{\sum_{i=1}^{r}\sum_{j=1,j\neq i}^{r}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_i^*(\theta_l)n_j(t')\}}{\gamma^2+|n_j(t')|^2}\right\} \\
&= 9\sum_{i=1}^{r}E\left\{\frac{|n_i(t)|^2}{(\gamma^2+|n_i(t)|^2)^2}\right\} \\
&\quad E\left\{\cos(\omega_i(\theta_k)+\phi_i(t))\cos(\omega_i(\theta_l)+\phi_i(t))\right\}\delta_{t,t'} \\
&\quad +9\sum_{i=1}^{r}\sum_{j=1,j\neq i}^{r}E\left\{\frac{|n_i(t)|}{\gamma^2+|n_i(t)|^2}\right\}E\left\{\cos(\omega_i(\theta_k)+\phi_i(t))\right\} \\
&\quad E\left\{\frac{|n_j(t')|}{\gamma^2+|n_j(t')|^2}\right\}E\left\{\cos(\omega_j(\theta_l+\phi_j(t'))\right\} \\
&= 9\frac{2}{15\gamma^2}\frac{1}{2}\sum_{i=1}^{r}\cos(\omega_i(\theta_k)-\omega_i(\theta_l))\delta_{t,t'} \\
&= \frac{3}{5\gamma^2}\sum_{i=1}^{r}\Re\left\{a_i(\theta_k)a_i^*(\theta_l)\right\}\delta_{t,t'}.
\end{aligned} \tag{A.22}$$

Equation (A.22) can be written compactly as

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t')}\right)^T = \frac{3}{5\gamma^2}\Re\left\{\mathbf{A}^H\mathbf{A}\right\}\delta_{t,t'}. \tag{A.23}$$

Similarly,

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t')}\right)^T = -\frac{3}{5\gamma^2}\Im\left\{\mathbf{A}^H\mathbf{A}\right\}\delta_{t,t'}, \tag{A.24}$$

and

$$E\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t)}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t')}\right)^T = \frac{3}{5\gamma^2}\Re\left\{\mathbf{A}^H\mathbf{A}\right\}\delta_{t,t'}. \tag{A.25}$$

Also,

$$
\begin{aligned}
E\left(\frac{\partial L}{\partial \theta_l}\frac{\partial L}{\partial s_{k,\Re}(t)}\right) &= 9E\left\{\sum_{t'=1}^{M}\sum_{i=1}^{r}\sum_{j=1}^{r}\frac{\Re\{s_l^*(t')d_i^*(\theta_l)n_i(t')\}}{\gamma^2+|n_i(t')|^2}\frac{\Re\{a_j^*(\theta_k)n_j(t)\}}{\gamma^2+|n_j(t)|^2}\right\} \\
&= 9E\left\{\sum_{i=1}^{r}\frac{\Re\{s_l^*(t)d_i^*(\theta_l)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\frac{\Re\{a_i^*(\theta_k)n_i(t)\}}{\gamma^2+|n_i(t)|^2}\right\} \\
&= 9\sum_{i=1}^{r}|s_l(t)||d_i(\theta_l)|E\left\{\frac{|n_i(t)|^2}{(\gamma^2+|n_i(t)|^2)^2}\right\} \\
&\quad E\left\{\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}+\phi_i(t))\cos(\omega_i(\theta_k)+\phi_i(t))\right\} \\
&= 9\sum_{i=1}^{r}|s_l(t)||d_i(\theta_l)|\frac{2}{15\gamma^2} \\
&\quad \frac{1}{2}E\left\{\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}+\omega_i(\theta_k)+2\phi_i(t))+\right. \\
&\quad \left.\cos(-\beta_l(t)+\omega_i(\theta_l)+\frac{\pi}{2}-\omega_i(\theta_k))\right\} \\
&= \frac{3}{5\gamma^2}\sum_{i=1}^{r}\Re\left\{s_l^*(t)d_i^*(\theta_l)a_i(\theta_k)\right\}, \qquad \text{(A.26)}
\end{aligned}
$$

which in compact form can be written as

$$
E\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = \frac{3}{5\gamma^2}\Re\left\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\right\}. \qquad \text{(A.27)}
$$

Similarly,

$$
E\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t)}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = \frac{3}{5\gamma^2}\Im\left\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\right\}. \qquad \text{(A.28)}
$$

Finally,

$$
E\left(\frac{\partial L}{\partial \gamma}\right)\left(\frac{\partial L}{\partial \boldsymbol{\theta}}\right)^T = 0, \qquad \text{(A.29)}
$$

$$
E\left(\frac{\partial L}{\partial \gamma}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Re}(t)}\right)^T = 0, \qquad \text{(A.30)}
$$

and

$$
E\left(\frac{\partial L}{\partial \gamma}\right)\left(\frac{\partial L}{\partial \mathbf{s}_{\Im}(t)}\right)^T = 0. \qquad \text{(A.31)}
$$

The Fisher Information Matrix $\mathbf{J}(\boldsymbol{\eta})$ can now be written as

$$\mathbf{J}(\boldsymbol{\eta}) = \begin{bmatrix} \Gamma & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \boldsymbol{\Sigma}_{\Re} & -\boldsymbol{\Sigma}_{\Im} & \cdots & 0 & 0 & \boldsymbol{\Delta}_{\Re}(1) \\ 0 & \boldsymbol{\Sigma}_{\Im} & \boldsymbol{\Sigma}_{\Re} & \cdots & 0 & 0 & \boldsymbol{\Delta}_{\Im}(1) \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \boldsymbol{\Sigma}_{\Re} & -\boldsymbol{\Sigma}_{\Im} & \boldsymbol{\Delta}_{\Re}(M) \\ 0 & 0 & 0 & \cdots & \boldsymbol{\Sigma}_{\Im} & \boldsymbol{\Sigma}_{\Re} & \boldsymbol{\Delta}_{\Im}(M) \\ 0 & \boldsymbol{\Delta}_{\Re}^T(1) & \boldsymbol{\Delta}_{\Im}^T(1) & \cdots & \boldsymbol{\Delta}_{\Re}^T(M) & \boldsymbol{\Delta}_{\Im}^T(M) & \boldsymbol{\Theta} \end{bmatrix}, \tag{A.32}$$

where

$$\Gamma = \frac{4}{5}\frac{Mr}{\gamma^2}, \tag{A.33}$$

$$\boldsymbol{\Sigma}_{\Re} = \Re\{\boldsymbol{\Sigma}\} = \frac{3}{5\gamma^2}\Re\left\{\mathbf{A}^H\mathbf{A}\right\}, \tag{A.34}$$

$$\boldsymbol{\Sigma}_{\Im} = \Im\{\boldsymbol{\Sigma}\} = \frac{3}{5\gamma^2}\Im\left\{\mathbf{A}^H\mathbf{A}\right\}, \tag{A.35}$$

$$\boldsymbol{\Delta}_{\Re}(t) = \Re\{\boldsymbol{\Delta}\} = \frac{3}{5\gamma^2}\Re\left\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\right\}, \tag{A.36}$$

$$\boldsymbol{\Delta}_{\Im}(t) = \Im\{\boldsymbol{\Delta}\} = \frac{3}{5\gamma^2}\Im\left\{\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\right\}, \tag{A.37}$$

and

$$\boldsymbol{\Theta} = \frac{3}{5\gamma^2}\sum_{t=1}^{M}\Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\mathbf{D}\mathbf{S}(t)\right\}. \tag{A.38}$$

Clearly,

$$CRB(\gamma) = \frac{1}{\Gamma} = \frac{5}{4}\frac{\gamma^2}{Mr}. \tag{A.39}$$

Setting $\boldsymbol{\Xi} \stackrel{\text{def}}{=} \boldsymbol{\Sigma}^{-1}$ and using a standard result on the inverse of a partitioned matrix [71], the CRB for the direction of arrival $\boldsymbol{\theta}$ can be expressed as

$$CRB^{-1}(\boldsymbol{\theta}) = \boldsymbol{\Theta} - \left[\boldsymbol{\Delta}_{\Re}^T(1), \boldsymbol{\Delta}_{\Im}^T(1), \ldots, \boldsymbol{\Delta}_{\Re}^T(M), \boldsymbol{\Delta}_{\Im}^T(M)\right]$$

$$
\begin{bmatrix}
\boldsymbol{\Xi}_{\Re} & -\boldsymbol{\Xi}_{\Im} & \cdots & 0 & 0 \\
\boldsymbol{\Xi}_{\Im} & \boldsymbol{\Xi}_{\Re} & \cdots & 0 & 0 \\
\vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & \boldsymbol{\Xi}_{\Re} & -\boldsymbol{\Xi}_{\Im} \\
0 & 0 & \cdots & \boldsymbol{\Xi}_{\Im} & \boldsymbol{\Xi}_{\Re}
\end{bmatrix}
\begin{bmatrix}
\boldsymbol{\Delta}_{\Re}(1) \\
\boldsymbol{\Delta}_{\Im}(1) \\
\vdots \\
\boldsymbol{\Delta}_{\Re}(M) \\
\boldsymbol{\Delta}_{\Im}(M)
\end{bmatrix}
$$

$$
\begin{aligned}
&= \boldsymbol{\Theta} - \sum_{t=1}^{M} \Re\left\{\boldsymbol{\Delta}^H(t)\boldsymbol{\Xi}\boldsymbol{\Delta}(t)\right\} \\
&= \frac{3}{5\gamma^2}\sum_{t=1}^{M} \Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\mathbf{D}\mathbf{S}(t) - \mathbf{S}^H(t)\mathbf{D}^H\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{D}\mathbf{S}(t)\right\} \\
&= \frac{3}{5\gamma^2}\sum_{t=1}^{M} \Re\left\{\mathbf{S}^H(t)\mathbf{D}^H\left[I - \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\right]\mathbf{D}\mathbf{S}(t)\right\} \qquad \text{(A.40)}
\end{aligned}
$$

and the proof is completed. ∎

Appendix B

Fractional Lower Order Moments of Products of $S\alpha S$ Random Variables

It is known that if X is a $S\alpha S$ random variable and $p > 0$, then $E\{|X|^p\} < \infty$ if and only if $p < \alpha$ [6]. Also, if $X_1, \ldots, X_n$ are n-fold dependent $S\alpha S$ random variables and $p_1, \ldots, p_n$ are positive numbers, then

$$E\{|X_1|^{p_1} \cdots |X_n|^{p_n}\} < \infty \text{ if and only if } p_1 + \cdots + p_n < \alpha. \tag{B.1}$$

Recently, it was proven in [45] that for a real $S\alpha S$ random variable X it holds

$$E\{|X|^p\} < \infty \quad \text{for } -1 < p < \alpha. \tag{B.2}$$

Similarly, for an isotropic complex $S\alpha S$ random variable X it holds

$$E\{|X|^p\} < \infty \quad \text{for } -2 < p < \alpha. \tag{B.3}$$

We now consider the problem of determining a range of values of the parameter p for which $E\{|X||Y|^{p-1}\} < \infty$ and $E\{|X|^2|Y|^{2p-2}\} < \infty$.

Proof of Proposition 5.1 First, consider $E\{|X|^2|Y|^{2p-2}\}$. It follows from (B.1) that $E\{|X|^2|Y|^{2p-2}\} < \infty$ for $p < \alpha/2$, when X and Y are jointly $S\alpha S$ random variables. If μ is the measure induced on R^2 by X and Y, then $E\{|X|^2|Y|^{2p-2}\}$ can be written as

$$E\{|X|^2|Y|^{2p-2}\} = \int_{R^2} |x|^2|y|^{2p-2} d\mu(x,y) =: I_1. \tag{B.4}$$

But $I_1 < \infty$ if and only if [94]

$$I_1^{'} := \int_{R^2} |y|^{2p-2} d\mu(x,y) < \infty \tag{B.5}$$

By using (B.2) and (B.3) we get that $I_1^{'} < \infty$ if $p > 1/2$ ($p > 0$) when X and Y are real (complex isotropic) $S\alpha S$ random variables. Hence, it follows that $E\{|X|^2|Y|^{2p-2}\} < \infty$ if and only if $1/2 < p < \alpha/2$ ($0 < p < \alpha/2$) when X and Y are real (complex isotropic). Finally, since $E\{|X|^2|Y|^{2p-2}\} < \infty$ implies $E\{|X||Y|^{p-1}\} < \infty$, it follows that for the aforementioned values of p, $E\{|X||Y|^{p-1}\}$ is also finite. ∎

Appendix C

Asymptotic Performance of the MFLOM Estimator

The fractional lower-order (MFLOM) estimator for the modified covariation coefficient function $\lambda_{X,Y}(p)$ given in (5.17) can be written as the quotient of two statistics:

$$\hat{\lambda}_{X,Y}(p) = \frac{T_{1n}}{T_{2n}} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i Y_i^{<p-1>}}{\frac{1}{n}\sum_{i=1}^{n} |Y_i|^p}. \tag{C.1}$$

First, we prove the following Lemma for the case of complex isotropic $S\alpha S$ random variables.

Lemma C.1 *Given the two-dimensional statistic (T_{1n}, T_{2n}) described in (C.1) with $0 < p < \alpha/2$, the asymptotic distribution (a.d.) of $\sqrt{n}(T_{1n} - \theta_1, T_{2n} - \theta_2)$ is bivariate normal with mean zero and covariance matrix*

$$\mathbf{S} = \frac{1}{n}\begin{bmatrix} E\{|X|^2|Y|^{2p-2}\} - |E\{XY^{<p-1>}\}|^2 & E\{XY^{<p-1>}|Y|^p\} - E\{XY^{<p-1>}\}E\{|Y|^p\} \\ E\{X^*Y^{*<p-1>}|Y|^p\} - E\{X^*Y^{*<p-1>}\}E\{|Y|^p\} & E\{|Y|^{2p}\} - (E\{|Y|^p\})^2 \end{bmatrix}, \tag{C.2}$$

where

$$\theta_1 = E\{XY^{<p-1>}\}, \tag{C.3}$$

and

$$\theta_2 = E\{|Y|^p\}. \tag{C.4}$$

Proof Clearly,

$$E\{T_{1n}\} = \frac{1}{n}\sum_{i=1}^{n} E\{X_i Y_i^{<p-1>}\} = E\{XY^{<p-1>}\}, \tag{C.5}$$

and

$$E\{T_{2n}\} = \frac{1}{n}\sum_{i=1}^{n} E\{|Y_i|^p\} = E\{|Y|^p\}. \tag{C.6}$$

In addition,

$$\begin{aligned} E\{T_{1n}T_{1n}^*\} &= \frac{1}{n^2}E\{\sum_{i=1}^{n}\sum_{j=1}^{n} X_i Y_i^{<p-1>} X_j^* Y_j^{*<p-1>}\} \\ &= \frac{1}{n^2}\sum_{i=1}^{n} E\{|X_i Y_i^{<p-1>}|^2\} + \frac{1}{n^2}\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}}_{i\neq j} E\{X_i Y_i^{<p-1>}\}E\{X_j^* Y_j^{*<p-1>}\} \\ &= \frac{1}{n}E\{|X|^2|Y|^{2p-2}\} + \frac{n-1}{n}|E\{XY^{<p-1>}\}|^2, \end{aligned} \tag{C.7}$$

and

$$\begin{aligned} E\{T_{2n}T_{2n}^*\} &= \frac{1}{n^2}E\{\sum_{i=1}^{n}\sum_{j=1}^{n} |Y_i|^p |Y_j|^p\} \\ &= \frac{1}{n^2}\sum_{i=1}^{n} E\{|Y_i|^{2p}\} + \frac{1}{n^2}\underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}}_{i\neq j} E\{|Y_i|^p\}E\{|Y_j|^p\} \\ &= \frac{1}{n}E\{|Y|^{2p}\} + \frac{n-1}{n}(E\{|Y|^p\})^2, \end{aligned} \tag{C.8}$$

where it follows from Proposition 5.1 that it must be $0 < p < \alpha/2$ so that $E\{|X|^2|Y|^{2p-2}\} < \infty$ and $E\{|Y|^{2p}\} < \infty$. By using (C.5)-(C.8) we have that

$$\begin{aligned} \mathrm{var}(T_{1n}) &= E\{T_{1n}T_{1n}^*\} - |E\{T_{1n}\}|^2 \\ &= \frac{1}{n}\sigma_{T_{1n}}^2 = \frac{1}{n}\left[E\{|X|^2|Y|^{2p-2}\} - |E\{XY^{<p-1>}\}|^2\right] < \infty, \quad 0 < p < \frac{\alpha}{2}, \end{aligned} \tag{C.9}$$

and

$$\begin{aligned} \mathrm{var}(T_{2n}) &= E\{T_{2n}T_{2n}^*\} - |E\{T_{2n}\}|^2 \\ &= \frac{1}{n}\sigma_{T_{2n}}^2 = \frac{1}{n}\left[E\{|Y|^{2p}\} - (E\{|Y|^p\})^2\right] < \infty, \quad 0 < p < \frac{\alpha}{2}. \end{aligned} \tag{C.10}$$

Hence, by means of the central limit theorem, we conclude that the statistics T_{1n} and T_{2n} are asymptotically normal:

$$\sqrt{n}(T_{1n} - \theta_1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_{T_{1n}}^2), \tag{C.11}$$

and

$$\sqrt{n}(T_{2n} - \theta_2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_{T_{2n}}). \tag{C.12}$$

Now, for every complex numbers a and b define the statistic $T_n = aT_{1n} + bT_{2n}$. Then, it can be easily shown that the asymptotic distribution of $\sqrt{n}(T_n - aE\{XY^{<p-1>}\} - bE\{|Y|^p\})$ is normal with variance

$$\begin{aligned}
&|a|^2 \left[E\{|X|^2|Y|^{2p-2}\} - |E\{XY^{<p-1>}\}|^2\right] \\
&-2\Re\left\{ab^* \left[E\{XY^{<p-1>}|Y|^p\} - E\{XY^{<p-1>}\}E\{|Y|^p\}\right]\right\} + \\
&|b|^2 \left[E\{|Y|^{2p}\} - (E\{|Y|^p\})^2\right] < \infty.
\end{aligned} \tag{C.13}$$

It follows that the two statistics T_{1n} and T_{2n} are asymptotically jointly normal. Finally, the covariance of T_{1n} and T_{2n} can be shown to be

$$\mathrm{cov}(T_{1n}, T_{2n}) = \mathrm{cov}^*(T_{2n}, T_{1n}) = \frac{1}{n}[E\{XY^{<p-1>}|Y|^p\} - E\{XY^{<p-1>}\}E\{|Y|^p\}] < \infty, \tag{C.14}$$

and the proof of Lemma C.1 is complete. ∎

Now, let g be the totally differentiable function of the two statistics T_{1n} and $T_{2n} \neq 0$ with

$$g(T_{1n}, T_{2n}) = \frac{T_{1n}}{T_{2n}}. \tag{C.15}$$

Then, from convergence theory [56, p. 321] and Lemma C.1, it follows that the asymptotic distribution of

$$\sqrt{n}[g(T_{1n}, T_{2n}) - g(\theta_1, \theta_2)]$$

is normal with mean zero and variance

$$\frac{1}{n}\sigma^2_\lambda(p) = \sum_{i=1}^{2}\sum_{j=1}^{2} \mathbf{S}_{i,j} \frac{\partial g}{\partial \theta_i} \left(\frac{\partial g}{\partial \theta_j}\right)^*. \tag{C.16}$$

Combining (C.2)-(C.4), (C.9)-(C.10), and (C.14)-(C.16) we get that

$$\sqrt{n}\left[\frac{\frac{1}{n}\sum_{i=1}^{n} X_i Y_i^{<p-1>}}{\frac{1}{n}\sum_{i=1}^{n}|Y_i|^p} - \frac{E\{XY^{<p-1>}\}}{E\{|Y|^p\}}\right] \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2_\lambda(p)), \tag{C.17}$$

with $\sigma^2_\lambda(p)$ as shown in (5.19). ∎

Bibliography

[1] S. P. Applebaum and D. J. Chapman. Adaptive arrays with main beam constraints. *IEEE Trans. Antennas Prop.*, 24:650–662, 1976.

[2] J. F. Bohme. Estimation of spectral parameters of correlated signals in wavefields. *Signal Processing*, 10:329–337, 1986.

[3] R. Brant. Approximate likelihood and probability calculations based on transforms. *Ann. Statist.*, 12:989–1005, 1984.

[4] L. E. Brennan and I. S. Reed. Theory of adaptive radar. *IEEE Trans. Aerosp. Electron. Syst.*, 9:237–252, 1973.

[5] J. P. Burg. The relationship between maximum entropy spectra and maximum likelihood spectra. *Geophysics*, 37:375–376, 1972.

[6] S. Cambanis. Complex symmetric stable variables and processes. In P. Sen, editor, *Contributions to Statistics: Essays in Honor of Norman L. Johnson*, pages 63–79. North-Holland, New York, 1983.

[7] S. Cambanis, C. D. Hardin, and A. Weron. Ergodic properties of stationary stable processes. *Stochastic Processes and their Applications*, 24:1–18, 1987.

[8] S. Cambanis, C. D. Hardin, and A. Weron. Innovations and Wold decompositions of stable sequences. *Probab. Th. Rel. Fields*, 79:1–27, 1988.

[9] S. Cambanis and A. G. Miamee. On prediction of harmonizable stable processes. *Sankhyā: The Indian Journal of Statistics*, 51:269–294, 1989.

[10] S. Cambanis and G. Miller. Linear problems in pth order and stable processes. *SIAM J. Appl. Math.*, 41:43–69, 1981.

[11] S. Cambanis, G. Samorodnitsky, and M. S. Taqqu, editors. *Stable Processes and Related Topics*. Birkhauser, Boston, 1991.

[12] S. Cambanis and A. R. Soltani. Prediction of stable processes: Spectral and moving average representations. *Z. Wahrsch verw. Gebiete*, 66:593–612, 1984.

[13] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE*, 57:1408–1418, 1969.

[14] J. M. Chambers, C. L. Mallows, and B. W. Stuck. A method of simulating stable random variables. *Journal of the American Statistical Association*, 71:340–344, 1976.

[15] M. C. Doğan and J. M. Mendel. Applications of cumulants to array processing, Part I: Aperture extension and array calibration. *IEEE Trans. Signal Processing*, 43:1200–1216, 1995.

[16] M. C. Doğan and J. M. Mendel. Applications of cumulants to array processing, Part II: Non-Gaussian noise suppression. *IEEE Trans. Signal Processing*, 43:1663–1676, 1995.

[17] G. S. Edelson and D. W. Tufts. On the ability to estimate narrow-band signal parameters using towed arrays. *IEEE J. Oceanic Eng.*, 17:48–61, 1992.

[18] M. Feder and E. Weinstein. Parameter estimation of superimposed signals using the EM algorithm. *IEEE Trans. Acoust., Speech, and Signal Process.*, 36:477–489, 1988.

[19] P. Forster and C. L. Nikias. Bearing estimation in the bispectrum domain. *IEEE Trans. Signal Processing*, 39:1994–2006, 1991.

[20] P. Forster and G. Vezzosi. Application of spheroidal sequences to array processing. In *Proc. ICASSP'87*, pages 2268–2271, Dallas, TX, 1987.

[21] B. Friedlander. A sensitivity analysis of the MUSIC algorithm. *IEEE Trans. Acoust., Speech, and Signal Process.*, 38:1740–1751, 1990.

[22] B. Friedlander. Some issues in blind array calibration. In *Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1052–1056, Pacific Grove, CA, November 1993.

[23] J. A. Gadzow. A high resolution direction-of-arrival algorithm for narrow-band coherent and incoherent sources. *IEEE Trans. Acoust., Speech, and Signal Process.*, 36:965–979, 1988.

[24] G. Gong and F. J. Samaniego. Pseudo maximum likelihood estimation:Theory and applications. *Ann. Statist.*, 9:861–869, 1981.

[25] I. S. Gradshteyn and I. M. Ryzhik. *Tables of Integrals, Series, and Products.* Academic Press, New York, 1980.

[26] D. L. Hall. *Mathematical Techniques in Multisensor Data Fusion.* Artech House, Boston, MA, 1992.

[27] C. D. Hardin. On the spectral representation of symmetric stable processes. *J. Multivariate Analysis*, 12:385–401, 1982.

[28] S. Haykin, editor. *Advances in Spectrum Analysis and Array Processing*, volume 2. Prentice Hall, Englewood Cliffs, 1991.

[29] Y. Hosoya. Discrete-time stable processes and their certain properties. *The Annals of Probability*, 6:94–105, 1978.

[30] H. Hung and M. Kaveh. Focusing matrices for coherent signal-subspace processing. *IEEE Trans. Acoust., Speech, and Signal Process.*, 36:1272–1281, 1988.

[31] A. G. Jaffer. Maximum likelihood direction finding of stochastic sources: A separable solution. In *Proc. ICASSP 1988*, pages 2893–2896, New York, NY, 1995.

[32] A. Janicki and A. Weron. *Simulation and Chaotic Behavior of α-Stable Stochastic Processes*. Dekker, New York, 1993.

[33] A. Janicki and A. Weron. Can one see α-stable variables and processes. *Statistical Science*, 9:109–126, 1994.

[34] D. Johnson and D. Dudgeon. *Array Signal Processing:Concepts and Techniques*. Prentice Hall, Englewood Cliffs, 1993.

[35] M. Kanter. Linear sample spaces and stable processes. *Adv. Funct. Anal.*, 9:441–459, 1972.

[36] M. Kanter. Stable densities under change of scale and total variation inequalities. *The Annals of Probability*, 3:697–707, 1975.

[37] M. Kanter and W. L. Steiger. Regression and autoregression with infinite variance. *Adv. Appl. Prob.*, 6:768–783, 1974.

[38] M. Kaveh and A. J. Barabell. The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise. *IEEE Trans. Acoust., Speech, and Signal Process.*, 34:331–341, 1986.

[39] D. Kletter and H. Messer. The role of third-order spectrum in maximum likelihood time delay estimation of a random multitone signal in noise. In *Proc. ICASSP'89*, pages 2310–2313, Glasgow, Scotland, May 1989.

[40] J. Krolik. Focused wide-band array processing for spatial spectral estimation. In S Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, volume 1, pages 221–261. Prentice Hall, Englewood Cliffs, 1991.

[41] R. Kumaresan and D. W. Tufts. Estimating the angles of arrival of multiple plane waves. *IEEE Trans. Aerosp. Electron. Syst.*, 19:134–139, 1983.

[42] E. L. Lehmann. *Theory of Point Estimation*. Wiley, New York, 1983.

[43] R. LePage. Conditional moments for coordinates of stable vectors. In S. Cambanis and A. Weron, editors, *Probability Theory on Vector Spaces IV*, pages 148–163. Springer's LNM, Lancut, 1987.

[44] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison Wesley, Menlo Park, second edition, 1984.

[45] X. Ma and C. L. Nikias. On blind channel identification for impulsive signal environments: A method-of-moments approach. Technical Report USC-SIPI, University of Southern California, September 1994.

[46] M. Marques and S. Cambanis. Admissible and singular translates of stable processes. In S Cambanis and A Weron, editors, *Probability Theory on Vector Spaces IV*, pages 239–257. Springer's LNM, Lancut, 1987.

[47] E. Masry and S. Cambanis. Spectral density estimation for stationary stable processes. *Stochastic Processes and their Applications*, 18:1–31, 1984.

[48] G. Miller. Properties of certain symmetric stable distributions. *J. Multivariate Analysis*, 8:346–360, 1978.

[49] M. I. Miller and D. R. Fuhrmann. Maximum-likelihood narrow-band direction finding and the EM algorithm. *IEEE Trans. Acoust., Speech, and Signal Process.*, 38:1560–1577, 1990.

[50] U. Nickel. Algebraic formulation of Kumaresan-Tufts superresolution method, showing relation to ME and MUSIC methods. *IEE-F Proc. (London)*, 135:7–10, 1988.

[51] C. L. Nikias and A. Petropulu. *Higher-Order spectra analysis: A nonlinear signal processing framework*. Prentice Hall, Englewood Cliffs, 1993.

[52] C. L. Nikias and M. Shao. *Signal Processing with Alpha-Stable Distributions and Applications*. John Wiley and Sons, New York, 1995.

[53] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, second edition, 1987.

[54] S. Pei, C. C. Yeh, and S. C. Chiu. Modified spatial smoothing for coherent jammer suppression without signal cancellation. *IEEE Trans. Acoust., Speech, and Signal Process.*, 36:412–414, 1988.

[55] U. Pillai and B. Kwon. Forward/backward spatial smoothing schemes for coherent signal identification. *IEEE Trans. Acoust., Speech, and Signal Process.*, 37:8–15, 1989.

[56] C. R. Rao, editor. *Linear Statistical Inference and its Applications*. John Wiley and Sons, New York, 1965.

[57] S. S. Reddi. Multiple source location–A digital approach. *IEEE Trans. Aerosp. Electron. Syst.*, 15:95–105, 1979.

[58] V. Reddy, A. Paulraj, and T. Kailath. Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing. *IEEE Trans. Acoust., Speech, and Signal Process.*, 35:927–936, 1987.

[59] I. S. Reed, J. D. Mallett, and L. E. Brennan. Rapid convergence rate in adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.*, 10:853–863, 1974.

[60] Y. Rockah and P. M. Schultheiss. Array shape calibration using sources in unknown locations–Part I: Far-field sources, and Part II: Near-field sources and estimation implementation. *IEEE Trans. Acoust., Speech, and Signal Process.*, 35:286–299 and 724–735, 1987.

[61] R. Roy, A. Paulraj, and T. Kailath. ESPRIT–A subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Trans. Acoust., Speech, and Signal Process.*, 34:1340–1342, 1986.

[62] M. Rutkowski. Optimal linear filtering and smoothing for a discrete-time stable linear model. *J. Multivariate Analysis*, 50:68–92, 1994.

[63] H. Sakai. Estimation of frequencies of sinusoids in colored noise. In *Proc. ICASSP'86*, pages 177–180, Tokyo, Japan, 1986.

[64] G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, New York, 1994.

[65] M. Schilder. Some structure theorems for the symmetric stable laws. *The Annals of Mathematical Statistics*, 41:412–421, 1970.

[66] R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Prop.*, 34:276–280, 1986.

[67] T. J. Shan, M. Wax, and T. Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Trans. Acoust., Speech, and Signal Process.*, 33:806–811, 1985.

[68] M. Shao and C. L. Nikias. On symmetric stable models for impulsive noise. Technical Report USC-SIPI-231, University of Southern California, Feb. 1993.

[69] M. Shao and C. L. Nikias. Signal processing with fractional lower order moments: Stable processes and their applications. *Proc. IEEE*, 81:986–1010, 1993.

[70] D. Starer and A. Nehorai. Newton algorithms for conditional and uncoditional maximum likelihood estimation of the parameters of exponential signals in noise. *IEEE Trans. Signal Processing*, 40:1528–1534, 1992.

[71] P. Stoica and A. Nehorai. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Trans. Acoust., Speech, and Signal Process.*, 37:720–741, 1989.

[72] P. Stoica and A. Nehorai. Statistical efficiency study of direction estimation methods part I: Analysis of MUSIC and preliminary study of MLM. In S Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, volume 1, pages 263–305. Prentice Hall, Englewood Cliffs, 1991.

[73] A. L. Swindlehurst and T. Kailath. A performance analysis of subspace-based methods in the presence of model errors, Part I: The MUSIC algorithm. *IEEE Trans. Signal Processing*, 40:1758–1774, 1992.

[74] P. Tsakalides and C. L. Nikias. Testing of a new direction finding method in impulsive interference environments. Technical Report USC-SIPI-257, University of Southern California, March 1994.

[75] P. Tsakalides and C. L. Nikias. Maximum likelihood localization of sources in noise modeled as a stable process. *IEEE Trans. Signal Processing*, November 1995.

[76] P. Tsakalides and C. L. Nikias. The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments. Technical Report USC-SIPI-278, University of Southern California, Jan. 1995.

[77] G. A. Tsihrintzis and C. L. Nikias. Fast estimation of the parameters of alpha-stable impulsive interference. *IEEE Trans. Signal Processing.* submitted on September, 1994, pp. 27.

[78] G. A. Tsihrintzis and C. L. Nikias. On the detection of impulsive stochastic transients. *Signal Processing*, 41:1–16, 1994.

[79] G. A. Tsihrintzis and C. L. Nikias. Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. *IEEE Trans. Comm.*, 43:904–914, Mar. 1995.

[80] B. D. Van Veen. An analysis of several partially adaptive beamformer designs. *IEEE Trans. Acoust., Speech, and Signal Process.*, 37:192–203, 1989.

[81] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, pages 4–24, 1988.

[82] M. Viberg, B. Ottersten, and T. Kailath. Detection and estimation in sensor arrays using weighted subspace fitting. *IEEE Trans. Acoust., Speech, and Signal Process.*, 39:2436–2449, 1991.

[83] R. S. Walker. Bearing accuracy and resolution bounds of high-resolution beamformers. In *Proc. ICASSP 1985*, pages 1784–1787, Tampa, FL, 1985.

[84] H. Wang and M. Kaveh. Estimation of angles of arrival for wide-band sources. In *Proc. ICASSP 1984*, pages 7.5.1–7.5.4, San Diego, CA, 1984.

[85] H. Wang and M. Kaveh. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. *IEEE Trans. Acoust., Speech, and Signal Process.*, 33:823–831, 1985.

[86] M. Wax. Detection and localization of multiple sources via the stochastic signals model. *IEEE Trans. Acoust., Speech, and Signal Process.*, 39:2450–2456, 1991.

[87] M. Wax and T. Kailath. Optimal localization of multiple sources by passive arrays. *IEEE Trans. Acoust., Speech, and Signal Process.*, 31:1210–1218, 1983.

[88] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Acoust., Speech, and Signal Process.*, 33:387–392, 1985.

[89] M. Wax and I. Ziskind. Detection of the number of coherent signals by the mdl principle. *IEEE Trans. Acoust., Speech, and Signal Process.*, 37:1190–1196, 1989.

[90] A. Weron. Stable processes and measures; A survey. In D. Szynal and A. Weron, editors, *Probability Theory on Vector Spaces III*, pages 306–364. Springer's LNM, Lublin, 1983.

[91] B. Widrow, K. M. Duvall, R. P. Gooch, and W. C. Newman. Signal cancellation phenomena in adaptive antennas: Causes and cures. *IEEE Trans. Antennas Prop.*, 30:469–478, 1982.

[92] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode. Adaptive antenna systems. *Proc. IEEE*, 55:2143–2159, 1967.

[93] R. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul. An improved spatial smoothing technique for bearing estimation in a multipath environment. *IEEE Trans. Acoust., Speech, and Signal Process.*, 36:425–432, 1988.

[94] W. Wu and S. Cambanis. Conditional variance of symmetric stable variables. In S. Cambanis, G. Samorodnitsky, and M. S. Taqqu, editors, *Stable Processes and Related Topics*, pages 85–99. Birkhauser, Boston, 1991.

[95] A. Zeira and B. Friedlander. Direction estimation using time-varying arrays. In *Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1116–1120, Pacific Grove, CA, November 1993.

[96] Q. T. Zhang. Probability of resolution of the MUSIC algorithm. *IEEE Trans. Signal Processing*, 43:978–987, 1995.

[97] I. Ziskind and M. Wax. Maximum likelihood localization of multiple sources by alternating projection. *IEEE Trans. Acoust., Speech, and Signal Process.*, 36:1553–1560, 1988.

[98] V. M. Zolotarev. Integral transformations of distributions and estimates of parameters of multidimensional spherically symmetric stable laws. In J. Gani and V. K. Rohatgi, editors, *Contributions to Probability*, pages 283–305. Academic Press, New York, 1981.