Robust Sensor Array Processing for Non-stationary Signals

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

von
Waqas Sharif, M.Sc.
geboren am 13.08.1981 in Gujrat, Pakistan

Referent:	Prof. Dr.-Ing. Abdelhak M. Zoubir
Korreferent:	Prof. Dr. Wilhelm Stannat
Tag der Einreichung:	02.07.2012
Tag der mündlichen Prüfung:	05.11.2012

D 17
Darmstadt, 2013

Acknowledgments

I would like to thank all people who have helped me during my PhD studies.

First of all, I would like to express my special appreciation and sincere gratitude to my advisor Prof. Dr.-Ing. Abdelhak M. Zoubir. You have been a tremendous mentor for me and I could not have imagined having a better supervision. I would like to thank you for your encouragement and support throughout the course of my PhD studies. Your advice on both research as well as on my career has always been invaluable.

I am also grateful to Prof. Dr. Wilhelm Stannat for his supervision, guidance and precious time. I really benefited a lot from all my interactions with you.

I also acknowledge my gratitude to Prof. Dr.-Ing. Thomas Weiland, Prof. Dr.-Ing. Ralf Steinmetz and Prof. Dr.-Ing. Ulrich Konigorski who acted as the chair and examiners in the PhD committee.

I would like to thank all my colleagues at the Signal Processing Group at TU Darmstadt. I treasure my memories of those joyous days. Many thanks to Michael Muma, Ahmed Mostafa, Fiky Suratman, Dr. Zhihua Lu, Dr. Raquel Fandos, Dr. Christian Debes, Dr. Weaam Alkhaldi, Dr. Yacine Chakhchoukh, Sara Al-Sayed, Mouhammad Alhumaidi, Nevine Demitri, Gökhan Gül, Dr. Philipp Heidenreich, Sahar Khawatmi and Gebremichael Teame as well as Renate Koschella and Hauke Fath.

I want to thank Heike Hoffmann, Carina Schuster, Markus Lazanowski, Florain van der Loo and Christian Schmitt at the Graduate School of Computational Engineering for their kind support.

I also wish to thank Dr. Faisal Shafait for providing me with valuable advice in all sorts of matters. I am also thankful to Syed Zahid Ali, Abdul Majid Malik, Iftikhar Ahmed and Adnan Shah for the happy time we spent together during the years in Darmstadt.

Last but not the least, I wish to thank my family for their unconditional love and support throughout my life.

Darmstadt, 05.11.2012

Kurzfassung

Nichtstationäre Signale kommen in vielen Praktischen Anwendungen vor, so z.B. im Radar, im Sonar und in der Mobilkommunikation. Eine wichtige Aufgabe in diesen Anwendungen is die Schätzung der Einfallsrichtung (direction-of-arrival, DOA) der Signale, was zu der Lokalisierung dieser über verrauschte übertragungskanäle gesendeten Signale notwendig ist. Die existierenden Methoden für nichtstationäre Signale basieren auf der Annahme einer Gaußschen Rauschverteilung. In der Praxis ist das Rauschen allerdings häufig impulsiv, was zu einem signifikanten Abfall der Leistungsfähigkeit der traditionellen Methoden führt. Deshalb wird in dieser Arbeit das Problem der Einfallsrichtungsschätzung von nichtstationären Signalen in Gegenwart von impulsivem Rauschen behandelt.

Der entwickelte Algorithmus zur robusten Einfallsrichtungsschätzung basiert auf einer robusten Instantanfrequenzschätzung (instantaneous frequency, IFreq) der individuellen Quellen im Empfangssignal. Dann wird die Raum-Zeit-Frequenz-Verteilungsmatrix (STFD) geschätzt als Mittelwert jedes IFreq Segments, um die DOA Schätzung für ein Signal zu erhalten. Die IFreq Schätzung in dieser Arbeit basiert auf morphologischer Signalverarbeitung des räumlich gemittelten robust berechneten Auto-Raum-Zeit-Frequenz (TFD) Bildes. Diese Arbeit liefert auch robuste Methoden STFD Matritzen zu berechnen, die viel benutzt werden in der Sensorgruppensignalverarbeitung zur blinden Quellentrennung (BSS) und DOA Schätzung.

Die vorgeschlagenen Methoden zur robusten Schätzung von STFD Matrizen können in drei Klassen eingeteilt werden: Vorverarbeitungs basierte, Robuste Positionsschätzung basierte und robuste nichtiterative Techniken. Die Vorverarbeitung basierten STFDs haben einen geringen Berechnungsaufwand, sind adaptiv, einfach zu implementieren und hocheffektiv in stark impulsiven Rauschumgebungen. Die Methoden der zweiten Klasse benutzen Methoden der robusten Schätzung aus der statistischen Literatur. Diese Arbeit liefert STFD Schätzmethoden basierend auf dem M-Schätzer, dem hochrobusten S-Schätzer und dem MM-Schätzer. Die robuste Positionsschätzungsmethoden liefern eine verbesserte Impulsunterdrückungsfähigkeit im Vergleich zu den Vorverarbeitungsbasierten Methoden. Sie benötigen jedoch im Vergleich einen höheren Berechnungsaufwand. Eine Kompromiss zwischen Robustheit und Berechnungsaufwand stellen die nichtiterativen Methoden dar. Für alle Methoden wurden Simulationen durchgeführt um ihre Leistungsfähigkeit zu evaluieren. Die vorgeschlagenen Methoden wurden verglichen durch den erreichten gewurzelten quadatischen mittleren Fehler (RMSE) für DOA Schätzung unter variierenden Signal zu Rausch Leistungen (SNR) und variierenden Anteilen an impulsivem Rauschen. Die vorgeschlagenen robusten

Methoden erreichen einen niedrigeren RMSE für DOA Schätzung im Vergleich zu den traditionellen nichtrobusten Methoden.

Die Dissertation liefert auch eine Analysemethode für STFD Matritzen. Die Analyse ist begründet auf der sogenannten influence function. Diese ist ein Mass für qualitative Robustheit eines Schätzers. Die influence function beschreibt die zusätzliche Bias aufgrund einer infinitesimalen Kontaminierung an einem beliebigen Punkt. Für die vor kurzem vorgeschlagenen robusten Methoden wurden analytische Ausdrücke für die influence function hergeleitet. Ausserdem wurde zur Analyse des Falles der endlichen Anzahl an Samples die empirische influence function ausgewertet. Die Ergebnisse zeigen, dass die traditionellen nichtrobusten STFD Schätzer eine unbegrenzte Bias haben, was ihre nichtrobustheit analytisch bestätigt. Die robusten Methoden auf der anderen Seite haben eine begrenzte influence function was wiederum ihre Robustheit bestätigt.

Abstract

Non-stationary signals arise in many practical applications such as radar, sonar and mobile communication. An important task in these applications is to estimate the direction-of-arrival (DOA) of the signals in order to locate the desired signals transmitted via inherently noisy wireless channels. The existing methods for non-stationary DOA estimation are based on the assumption that the noise is Gaussian. However, in practice, the noise is often non-Gaussian and impulsive which leads to a significant performance loss in traditional methods. Therefore, in this thesis, the problem of DOA estimation of non-stationary signals in the presence of impulsive noise is considered.

The developed algorithm for robust DOA estimation is based on the robust instantaneous frequency (IFreq) estimation of the individual sources present in the mixture. Then robustly estimated spatial time-frequency distribution matrices (STFD) are averaged across each IFreq segment to obtain the DOA estimate of that particular signal. For IFreq estimation, the presented approach is based on morphological image processing of the spatially averaged, robustly computed auto time-frequency distribution (TFD) image. This thesis also provides robust methods to compute the STFD matrices which are widely used in array signal processing for blind source separation (BSS) and DOA estimation.

The proposed methods for the robust estimation of STFD matrices can be categorized into three classes: pre-processing based, robust position based and robust non-iterative techniques. The pre-processing based STFDs are computationally lightweight, adaptive, easily implementable and are effective in highly impulsive noise environments. For the second class, methods are based on the robust estimation techniques from the statistical literature. This dissertation provides STFD estimation techniques based on M-estimator, highly robust S-estimator and MM-estimator. The robust position based STFD estimation methods provide improved impulsive rejection capability and at the same time enhanced efficiency when compared to the pre-processing based methods. The robust position based methods, however, require expansive computations in comparison to the pre-processing based techniques. To provide a compromise in terms of robustness, efficiency and computational burden, non-iterative robust methods are proposed. For all the proposed methods, simulations have been conducted to exhibit their efficacy. The proposed methods are compared in terms of their achieved root-mean-square-error (RMSE) for DOA estimation under varying signal-to-noise-ratio (SNR) and under varying amount of impulsive contamination. The proposed robust methods provide an improved (lower) RMSE for DOA estimation as compared to the classical non-robust methods.

This dissertation also presents robustness analysis for STFD matrices. The analysis is based on the influence function. The influence function is also a measure of qualitative robustness of an estimator. It gives us the additional bias due to an infinitesimal contamination at a certain point. For recently proposed robust techniques, analytical expressions for influence functions have been provided. Moreover, finite sample counterpart of influence function, which is called the empirical influence function, has also been evaluated. The results show that the classical non-robust STFD estimators yield unbounded influence functions that confirms the non-robustness of classical approaches. On the other hand the proposed robust methods result in bounded influence functions and thus confirm the robustness of the presented robust estimators.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Non-stationary signals are encountered in many practical applications, such as, radar, sonar, seismology, mobile communications, vibrational motion analysis and biomedicine [1–4]. These signals exhibit time-varying spectra. The examples of such signals are the linear or higher-order frequency modulated (FM) signals, speech signals and the signals which arise due to the Doppler shift introduced by the relative motion of the source with respect to the receiver. Furthermore, in the aforementioned applications, usually a number of sensors are employed to sample the impinging signals spatially. These spatially collected observations of the sensors are then used to extract certain information about the arriving signals. The extracted information depends upon the particular application. In radar or sonar, e.g., one may wish to find the direction and the position of the target. In mobile communications, it is required to find the direction of a mobile user to enhance the quality of service to that user and to suppress the interference arriving from all other directions. Therefore, the direction-of-arrival (DOA) estimation is one of the fundamental tasks in sensor array applications [5, 6].

For analysis of non-stationary signals, time-frequency distributions (TFDs) have evolved as a natural and powerful mean [1, 7, 8]. TFDs map a time-only signal into time and frequency domain. Especially, in the context of array processing, the concept of spatial time-frequency distribution (STFD) matrices introduced in [9] has received significant attention over the last decade. STFDs are powerful means to exploit the spatiality of the array and the non-stationarity of the signals. STFDs are widely used in DOA estimation and blind source separation (BSS) of non-stationary signals [10–16]. In comparison to the conventional approaches for DOA estimation, the STFD based methods utilize the time-frequency (TF) structure of non-stationary sources which leads to an enhanced effective signal-to-noise-ratio (SNR) and therefore an improved estimation accuracy when compared to the conventional approaches for DOA estimation [10, 12–14].

1.1 Motivation

The existing STFD based methods for sensor array processing of non-stationary signals assume that the noise is Gaussian [17]. However, in reality this assumption does not hold and the noise is often non-Gaussian due to the existence of many interference sources which are either natural or man-made. Many measurement campaigns have indeed confirmed the presence of impulsive noise in many practical applications [17–26]. The sources of impulsiveness are e.g., natural phenomena such as lightning strikes, geo-magnetic storms etc [19, 27]. The impulsive noise is also caused by man-made interference sources such as jamming, sparking due to car ignitions and neighboring interference sources in wireless communications [17, 20, 22, 23, 25, 26, 28, 29]. Depending on a certain application, a variety of natural and man made impulsive noise sources have been identified, see [30] for a recent overview.

Investigations of TFD under non-Gaussian and impulsive noise have been performed in the literature in [31–34]. Based on the findings of [31–34], one can draw the conclusions that the performance of TFDs is severely degraded in the presence of impulsive noise. Since STFDs are obtained by computing TFDs across the sensor arrays, they also suffer in the presence of impulsive noise [35, 36]. In the literature, different robust methods for estimating the TFDs have been proposed [31–34]. However, to date their efficacy to DOA estimation in sensor arrays has not be established. The objective of this thesis is to fill the gap and to investigate the performance of existing classical DOA estimation methods for non-stationary sources in the presence of impulsive and to propose robust DOA estimation methods which can mitigate the effect of impulsive noise. This brings us to the contributions of the thesis.

This thesis proposes robust methods for STFD matrix estimation. The proposed methods has been classified into the pre-processing based, the robust position based- and the non-iterative robust techniques. A comparison of the proposed robust STFD matrix estimator to the classical STFD matrix estimator has been performed to assess the achieved performance for the DOA estimation of non-stationary sources. An algorithm to robustly estimate the instantaneous frequencies (IFreqs) of multi-component signals has been discussed. Furthermore, the robustness of the proposed estimators is assessed analytically based on the influence functions. The influence function is a widely used tool to assess the qualitative robustness of an estimator [37, 38]. This thesis provides analytical expressions for the influence function of different kinds of robust STFD estimators. It also presents finite sample counterpart of the influence function which is the empirical influence function. The point-wise contributions of this thesis are formulated in the following Section.

1.2 Contributions

The summary of original contributions is as follows:

1. An algorithm is provided to estimate the DOA of non-stationary signals robustly. The presented approach uses

   (a) robust instantaneous frequency estimation and

   (b) robust spatial time-frequency distribution matrix estimation.

   Simulations have been conducted to show the efficacy of the proposed approach. The proposed approach does not require to know the number of sources and provides lower RMSE for estimated DOAs when compared to the classical non-stationary DOA estimation and to the conventional robust estimation.

2. For robust IFreq estimation, a variation of morphological image processing technique [14] has been proposed. The proposed technique is based on TF image which is obtained by spatially averaging the robustly computed auto TFDs of the sensor signals.

3. Robust estimation of STFD matrices is proposed. The proposed approaches can be classified into

   (a) pre-processing,

   (b) robust position based and

   (c) robust non-iterative techniques for STFD estimation.

4. This thesis also presents an analytical approach for robustness analysis of S/TFDs. Analytical expressions for the influence functions of recently proposed STFD matrix estimators have been derived. Moreover, definition of finite sample counterpart of influence function for STFDs is also provided.

1.3 Publications

The following publications have been produced during the period of PhD candidacy.

Internationally Refereed Journal Articles

- W. Sharif, M. Muma and A. M. Zoubir, "Robustness analysis of spatial time-frequency distributions based on the influence function", IEEE Transactions on Signal Processing, May 2012 (accepted for publication).
- W. Sharif, Y. Chakhchoukh and A.M. Zoubir, "Robust Spatial Time-Frequency Distribution Matrix Estimation with Application to Direction-of-Arrival Estimation", Signal Processing 91(11):2630-2638, November 2011, DOI: 10.1016/j.sigpro.2011.05.022

Internationally Refereed Conference Papers

- W. Sharif, Y. Chakhchoukh and A. M. Zoubir, "Direction-of-arrival estimation of FM sources based on robust spatial time-frequency distribution matrices," In Proc. of IEEE Workshop on Statistical Signal Processing (SSP), Nice, France, June 2011.
- W. Sharif, P. Heidenreich and A. M. Zoubir, "Robust Direction-of-Arrival Estimation for FM Sources in the Presence of Impulsive Noise", In Proc. of the 35th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) Dallas TX, USA March 2010.

1.4 Thesis overview

The remaining part of this dissertation is divided into the following five chapters. The organization is also provided in Figure 1.1 and a chapter-wise summary for each chapter is provided.

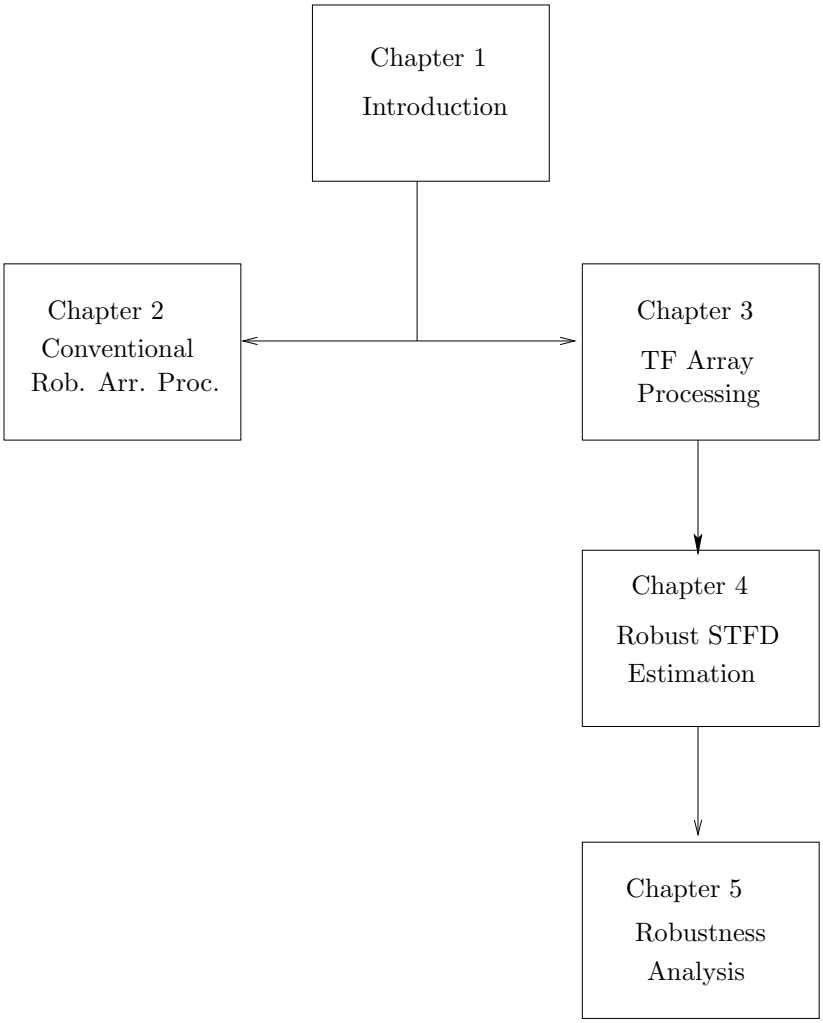

Figure 1.1. The structure of the thesis

Chapter 2

Chapter 2 provides problem formulations and presents state-of-the-art methods for DOA estimation under Gaussian and non-Gaussian noise for conventional array processing. In particular, it provides an insight into robust subspace based methods for DOA estimation. Furthermore, robust estimators of covariance matrix from statistical literature are presented.

Chapter 3

Chapter 3 covers the time-frequency based array processing and an algorithm for robustly estimating the DOA of non-stationary sources is outlined in this Chapter. It presents methods for instantaneous frequency estimation and proposed variations to counter the effect of impulsive noise in practical scenarios. Simulations are performed to depict the efficacy of robust TF array processing than robust conventional array processing.

Chapter 4

This Chapter describes several methods for estimating the STFD matrices robustly. These methods are classified into three classes which include pre-processing, robust position based and robust non-iterative techniques. Simulations are performed to compare different kinds of robust estimators in terms of achieved RMSE for DOA estimation under varying SNR and varying impulsive contamination conditions.

Chapter 5

Chapter 5 provides a framework for robustness analysis of selected robust techniques from Chapter 4. The analysis is based on the influence function. The analytical as well as empirical influence functions are described. Different estimators are compared on the basis of boundedness and continuity of their influence functions.

Chapter 6

Chapter 6 presents the concluding remarks and an outlook on the directions for the future work of this thesis.

Chapter 2
Conventional Robust Array Processing

This Chapter provides an introduction to conventional robust array processing for DOA estimation. Section 2.1 briefly introduces fundamental problems in sensor array processing while Section 2.2 provides problem formulations for the DOA estimation of non-stationary signals. Existing conventional state-of-the-art methods for the DOA estimation are described in Section 2.3 while the remaining part of this Chapter is dedicated to the state-of-the-art robust DOA estimation methods in conventional array processing. The presented techniques can be classified into two types: (i) the pre-processing based and (ii) the robust methods for covariance matrix estimation from the statistical literature [37, 38]. Simulation results are presented in Section 2.6 and Section 2.7 concludes the Chapter with a short summary.

2.1 Introduction

In array processing, a number of sensors are employed to observe a certain phenomenon. The corresponding observations of the sensors are then used to infer certain information about the impinging signals [39]. The tasks in array processing are e.g., to estimate the number of sources "*source enumeration*", to separate individual sources from the mixture "*source separation*" and to estimate the angles describing the directions from which signals are impinging on the array "*direction-of-arrivals* (DOAs)". The focus of this thesis is on the DOA estimation of non-stationary signals.

DOA estimation is one of the fundamental tasks in array processing and has received a considerable amount of attention over the last decades [6, 39]. DOA estimation enables to localize sources in many practical applications such as radar, sonars, wireless communications and in seismology etc. In conventional array signal processing, array processing methods do not utilize the signal's structure on the contrary to the non-stationary array processing where time-frequency (TF) structure of signals is utilized to enhance the estimation accuracy of the estimates.

Conventional DOA estimation has been a well investigated problem over the last several decades and a number of state-of-the-art methods exist [6, 39, 40]. In the following, the signal model in array processing is introduced and a brief recap of commonly used conventional approaches is provided.

2.2 Problem formulation

As shown in Figure 2.1, an array of sensors receives signals emitted from sources present in different directions. The signals denoted by s_k are considered to be narrow-band and impinge on an array of sensors with m elements. The received signal $\boldsymbol{x}(t)$ of the array is then modeled as:

$$\boldsymbol{x}(t) = \sum_{k=1}^{K} \boldsymbol{a}(\theta_k) s_k(t) + \boldsymbol{n}(t),\ t = \{1, 2, \ldots, N\} \tag{2.1}$$

where

- $\boldsymbol{x}(t) \in \mathbb{C}^m$ is the array observation vector at time t and
- N is the total number of observations, i.e., $t \in \{1, 2, \ldots, N\}$,
- $\boldsymbol{a}(\theta_k) \in \mathbb{C}^M$ denotes the array response (*steering vector*) of the signal arriving from the direction θ_k. For an array in uniform geometry (*uniform linear array* (ULA)), the steering vector is given by the following

  $$\boldsymbol{a}(\theta_k) = \left[1\ \mathrm{e}^{-j\frac{2\pi}{\lambda}d\sin(\theta_k)} \ldots \mathrm{e}^{-j(M-1)\frac{2\pi}{\lambda}d\sin(\theta_k)}\right]^T \tag{2.2}$$

  where m is the total number of sensors in the array, λ is the source signal wavelength and d is the distance between the sensors.

- $s_k(t)$ denotes the source symbol at time t and in the present context sources are considered to be frequency modulated (FM) signals, of the form:

  $$s_k(t) = A_k(t)\mathrm{e}^{-j\phi_k(t)}, \quad k = 1, 2, \ldots, K, \tag{2.3}$$

  where $\phi_k(t)$ and $A_k(t)$ denote the instantaneous phase and the amplitude of the k^{th} signal, respectively. The non-stationary sources are well defined by their IFreqs, given by the time derivative of the instantaneous phase

  $$f_k(t) = \frac{1}{2\pi}\frac{\mathrm{d}\phi_k(t)}{\mathrm{d}t} \tag{2.4}$$

  The IFreq of the signals based on Weierstrass's theorem is modeled as a polynomial phase signal model given by [41]

  $$f_k(t) = \sum_{p=1}^{P} \alpha_p^k t^{p-1} \tag{2.5}$$

  where α_p denotes the parameters of the FM signal and P is the order of the polynomial. Linear and quadratic FM signals which exists in many practical applications are described by polynomials of the order of $P = 2$ and $P = 3$, respectively.

- $\boldsymbol{n}(t) \sim \mathcal{N}_{\mathbb{C}}^{M}(\mathbf{0}, \sigma^2\boldsymbol{I})$) models the noise at the sensors and consists of thermal and surrounding noise where $\mathcal{N}_{\mathbb{C}}^{M}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes circular complex Gaussian with mean $\boldsymbol{\mu}$ and covariance of $\sigma^2\boldsymbol{I}$. In classical methods, the noise is assumed to be independently and identically distributed (i.i.d.) complex circular Gaussian. In practice, however, the noise is impulsive and modeled as a mixture of complex Gaussian distributions, given by:

$$\boldsymbol{n}(t) \sim (1-\epsilon)\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2\boldsymbol{I}) + \epsilon\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \kappa\sigma^2\boldsymbol{I}) \tag{2.6}$$

  where ϵ denotes the percentage of contamination and usually $\epsilon < 0.5$, κ denotes the factor of impulsiveness with $\kappa \gg 1$.

The objectives in array processing for non-stationary DOA estimation are two-fold:

1. To estimate the instantaneous frequencies of the sources in the mixture, i.e., $\{f_k(t)\}\ \forall k \in \{1, K\},\ t \in \{1, N\}$ and then use these instantaneous frequencies,
2. To estimate the DOAs $\{\theta_k\}_{k=1}^{K}$ provided N observations of above model in Eq. (2.1),

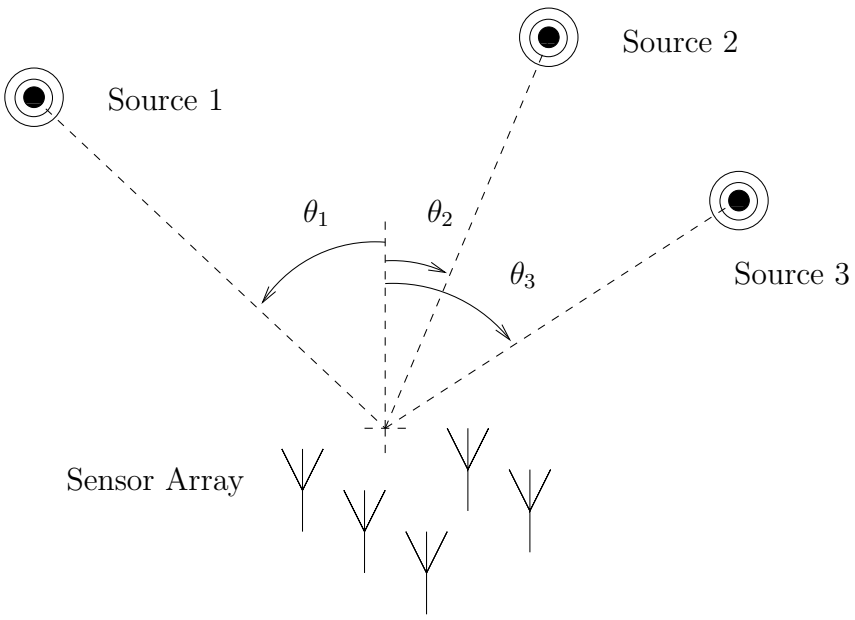

Figure 2.1. An array of sensors observing signals from different directions

2.3 Conventional DOA estimation methods

This Section briefly describes conventional methods for DOA estimation. These methods are the maximum-likelihood DOA estimation [42,43] and the subspace based DOA estimation method, namely multiple signal classification (MUSIC) [44].

2.3.1 Maximum-likelihood DOA estimation

In maximum likelihood DOA estimation, the observations are modeled as multivariate Gaussian denoted as $\boldsymbol{x}(t) \sim \mathcal{N}_{\mathbb{C}}^{M}\left(\boldsymbol{A}\boldsymbol{s}(t), \sigma^2\boldsymbol{I}\right)$, where $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_K)]$ is the matrix of steering vectors for each of the K sources and $\boldsymbol{s}(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T$ is the vector containing the symbols for all the sources at time t. Let $\hat{\boldsymbol{\theta}} = [\theta_1, \theta_2, \ldots, \theta_K]^T$ denote the maximum-likelihood estimates which are obtained by solving the following optimization problem, given below:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta},\boldsymbol{s}} \prod_{t=1}^{N} \frac{1}{\sqrt[M]{2\pi}\sqrt{|\boldsymbol{R}|}} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}(t)-\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t))^H \boldsymbol{R}^{-1}(\boldsymbol{x}(t)-\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t))} \tag{2.7}$$

where $\boldsymbol{R}$ denotes the covariance matrix. Taking the natural-log and solving for $\boldsymbol{\theta}$ and $\boldsymbol{s}$, we obtain the following non-linear least squares formulation, given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta},\boldsymbol{s}} \sum_{t=1}^{N} |\boldsymbol{x}(t) - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t)|^2 \tag{2.8}$$

The minimization of above cost function requires search over a multi-dimensional and multi-modal cost function with local minima's. This is computationally very expensive. However, this method yields the best performance and achieves the minimum of variance which is given by the Cramér-Rao bound (CRB) [40,45]. Alternatively, computationally efficient ML type algorithms have also been proposed in the literature, e.g., in-case of known signals the efficient decoupled ML algorithm [46] and RELAX algorithm [47] can be used. The next Section introduces a subspace based method for DOA estimation which reduces the computational complexity of ML type of DOA estimators.

2.3.2 Multiple signal classification (MUSIC)

Subspace methods are based on the eigenvalue decomposition of the covariance matrix. Among subspace methods, MUSIC is the most widely used for DOA estimation of

sources. It utilizes the structure of the covariance matrix which on the basis of signal model in Eq. (2.1) bears the following structure:

$$\boldsymbol{R} = \mathrm{E}\left[\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right] = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^H + \sigma^2\boldsymbol{I} \tag{2.9}$$

where $\boldsymbol{R}$ denotes the spatial covariance matrix, $\mathrm{E}[\cdot]$ is the statistical expectation and $\boldsymbol{R}_{ss}$ is the source covariance matrix. Eigenvalue decomposition is then performed on $\boldsymbol{R}$ to obtain the signal and noise subspaces. Let $\boldsymbol{R} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$ denote the eigenvalue decomposition of $\boldsymbol{R}$. $\boldsymbol{U}$ and $\boldsymbol{\Lambda}$ consist of eigenvectors and eigenvalues, respectively, given by

$$\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m], \quad \boldsymbol{\Lambda} = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_m] \tag{2.10}$$

Based on the model in Eq. (2.1), the eigenvalues exhibit the following structure.

$$\lambda_1 > \lambda_2 > \cdots \lambda_K > \lambda_{K+1} = \lambda_{K+2} \cdots = \lambda_m = \sigma^2$$

The K largest eigenvalues belong to the signal and the corresponding eigenvectors $\boldsymbol{E}_s = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_K]$ constitute the signal subspace, while the smallest $m - K$ eigenvalues belong to the noise and the corresponding eigenvectors forms the noise subspace $\boldsymbol{E}_n = [\boldsymbol{u}_{K+1}, \ldots, \boldsymbol{u}_m]$. The MUSIC estimator utilizes the orthogonality between the signal and the noise subspace. The DOA estimates for $\boldsymbol{\theta}$ based on MUSIC are obtained by finding the K peaks in the MUSIC pseudo-spectrum given by the following

$$P_{MUSIC}(\theta) = \frac{1}{\|\hat{\boldsymbol{E}}_n\boldsymbol{a}(\theta)\|^2} \tag{2.11}$$

In practice, however, $\boldsymbol{R}$ is not known and must be estimated. The maximum-likelihood estimate under Gaussian distribution is the sample covariance matrix, given by:

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

2.3.3 Beamformer

In this approach, output signals from all the sensors are coherently combined using a spatial matched filter $\boldsymbol{a}(\theta)$, matched to the angle θ. The estimates are then obtained by finding the peaks in the beamformer spectrum, given by

$$P_{BF}(\theta) = \frac{1}{MN}\sum_{t=1}^{N}\left|\boldsymbol{a}^H(\theta)\boldsymbol{x}(t)\right|^2 \tag{2.13}$$

This methods is computationally simple and optimal in the case of single source DOA estimation (and is equivalent to ML). For this reason, it is of particular importance

in TF array processing for DOA estimation, where single source DOA estimation is performed.

In the presence of impulsive noise, the methods described above lead to erroneous estimates and, therefore, robust methods are needed. In the following Section some of the existing robust estimators for DOA estimation in conventional array processing are briefly reviewed.

2.4 Conventional robust DOA estimation methods

Impulsive noise leads to the performance degradation of standard DOA estimates and the remedy is to look for robust methods. In literature some examples of robust approaches can be found in e.g., [48–50]. The existing robust methods for conventional DOA estimation can be classified into two main types, the robust maximum-likelihood and the robust subspace based methods.

The robust maximum-likelihood algorithms are based on the modeling of the noise as heavy tailed distributions. The heavy-tailed noise can be modeled as a Gaussian mixture, Cauchy distribution and in the form of Student-t distribution [51]. Two examples of robust ML algorithms for DOA estimation are mentioned below.

2.4.1 Maximum-likelihood type of estimators

Tsakalides et al. [49] modeled the heavy-tailed noise with α stable distributions and use the Cauchy distribution which is a special form of the α stable distribution with $\alpha = 1$. The Cauchy distribution is given by:

$$f(\delta, \gamma; n_p) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (n_p - \delta)^2}, \tag{2.14}$$

where γ and δ are the dispersion and the location parameters, respectively and n_p denotes the noise at the p^{th} sensor element. Based on this noise model, the corresponding likelihood function for estimating the parameters $\{\theta_k\}_{k=1}^{K}$ is given by the following

$$f(\boldsymbol{x}; \hat{\boldsymbol{\theta}}) = \prod_{t=1}^{N} \prod_{p=1}^{m} \frac{1}{\pi} \frac{\gamma}{\left(\gamma^2 + \left| x_p(t) - \sum_{k=1}^{K} \boldsymbol{a}(\theta_k) s_k(t) \right|^2 \right)} \tag{2.15}$$

where $\boldsymbol{x} = [x_1, \ldots, x_m]^T$ and again $\boldsymbol{a}(\theta_k)$ is array steering vector for the source arriving from the direction θ_k. The estimates for the DOA estimates are found by maximizing log-likelihood with respect to the parameters θ_k and s_k with $k \in \{1, \ldots, K\}$. This requires to solve a multi-dimensional optimization of the above likelihood function to estimate the DOA. Similarly, Sadler et al. modeled the impulsive noise as a Gaussian mixture [48], which can be written as follows:

$$n_p(t) = \sum_{i=1}^{P} \epsilon_i \mathcal{N}_{\mathbb{C}}(0, \sigma_p^2), \tag{2.16}$$

where again n_p is the noise at the p^{th} sensor of the array, ϵ_i denotes the fraction (or the probability) of occurrence of i^{th} component of the Gaussian mixture and with $\sum_{i=1}^{P} \epsilon_i = 1$. Again $\mathcal{N}_{\mathbb{C}}$ denote the circular complex Gaussian noise with a mean of 0 and a variance of σ_p^2. The authors used the space alternating generalized expectation maximization (SAGE) algorithm [52] to solve the corresponding log-likelihood function and to obtain the DOAs. The ML type of algorithms obtain the optimal performance in terms of estimation accuracy given that the distributional assumption hold, however, such algorithms require strict computational requirements which prohibit their usage in the practical applications. The robust subspace methods have been investigated in the literature over the last decade. In the next Section, some of the algorithms to obtain robust DOA estimates based on subspace methods are described.

2.5 Robust subspace based DOA estimation methods

Since subspace based DOA estimation methods depend on the covariance matrix [44] and therefore corresponding robust counterparts require a robust estimate of the covariance matrix. The classical MUSIC uses the sample covariance matrix which is maximum-likelihood estimate of covariance matrix under Gaussian noise. In the presence of impulsive noise, this non-robust estimate can be replaced by its robust counterpart. This is also depicted in Figure 2.5 which shows that given the heavy-tailed noisy observations, a robust estimate of covariance followed by the computation of MUSIC spectrum yields the estimated DOAs $\{\hat{\theta}_k\}_{k=1}^{K}$.

In the following, robust methods for estimating the covariance matrix are discussed. Several methods to estimate the covariance matrix robustly have been proposed in the

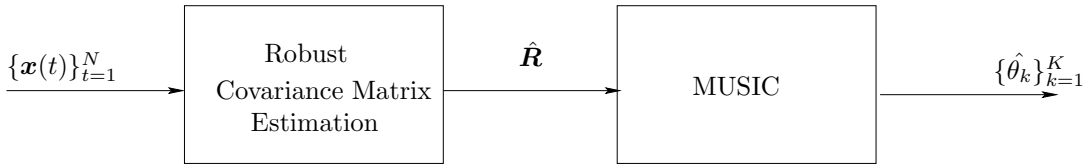

Figure 2.2. Robust subspace based estimation under impulsive noise

literature [50, 53–56]. The existing methods are classified into the methods based on the pre-processing and the methods based on robust estimation approaches. A few examples from the pre-processing based robust estimation are provided in the next Section.

2.5.1 Spatial sign function

Visuri et al. [50] proposed to suppress the impulsive interference by applying spatial sign function, given by

$$\boldsymbol{x}_{ssf}(t) = \begin{cases} \dfrac{\boldsymbol{x}(t)}{\|\boldsymbol{x}(t)\|} & \text{if } \boldsymbol{x}(t) \neq \boldsymbol{0} \\ \boldsymbol{0} & \text{otherwise,} \end{cases} \tag{2.17}$$

where $\boldsymbol{x}_{ssf}(t)$ denotes the spatial sign function. Then, the covariance matrix can be estimated by the pre-processed observations $\boldsymbol{x}_{ssf}(t)$ instead of $\boldsymbol{x}(t)$ and is given by

$$\hat{\boldsymbol{R}}_{ssf} = \sum_{t=1}^{N} \boldsymbol{x}_{ssf}(t)\boldsymbol{x}_{ssf}^{H}(t). \tag{2.18}$$

The advantage of using spatial sign function is that it is easy to compute and is parameter free, i.e., does not require any tuning parameters.

2.5.2 Trimming based on hypothesis testing

Lim et al. proposed an approach to trim the samples iteratively from given observations until the hypothesis for the normal data is accepted [55]. Therein, Shapiro-Wilk's test is used to assess the normality of data. Let $\boldsymbol{Z}$ be a matrix of size $m \times N$ consisting of zeros for all the trimmed points and ones elsewhere and $\hat{\boldsymbol{x}}_{pre}(t)$ denote the trimmed observation. Then the robust covariance matrix based on the adaptive trimming is given by the following

$$\hat{\boldsymbol{R}}_{pre} = \frac{1}{\boldsymbol{Z}\boldsymbol{Z}^{H}} \odot \sum_{t=1}^{N} \hat{\boldsymbol{x}}_{pre}(t)\hat{\boldsymbol{x}}_{pre}^{H}(t), \tag{2.19}$$

where $\odot$ denotes the element-wise multiplication. This algorithm is computationally more expensive if compared to the spatial sign function approach because it requires order statistics and evaluation of Shapiro Wilks normality test statistics. It also needs tunning parameters in the form of as to how many number of observations have to be trimmed in each of the iterations.

2.5.3 Modulus transformation

This methods uses a parameter free transformation proposed in [57] to suppress the impulsive noise. The transformation is called the modulus transformation and bears the following form:

$$x_{mod} = \begin{cases} \text{sign}(x) \log_{10} |x+1| & \text{if } x \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.20}$$

Here, $\text{sign}(\cdot)$ is the sign function and $|\cdot|$ denotes the modulus operation. Since in the context of sensor array processing, the signal is complex-valued. The transformation is applied individually on the real and the imaginary part. A modified version for the complex data in the context of array signal processing was then used by Lim et al. in [56] for robustly estimating the covariance matrix. The authors in [56] termed this modulus operation as the modified modulus transformation. It can be written as follows:

$$x_{mod} = \begin{cases} \text{sign}(\bar{x}) \log_{10} |\bar{x}+1| + j\text{sign}(\bar{x}) \log_{10} |\tilde{x}+1| & \text{if } |x| \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.21}$$

Herein, $\bar{x}$ and $\tilde{x}$ denotes the real and the imaginary part of x, respectively. Let $\boldsymbol{x}_{mod}(t)$ denote the observation obtained after applying the modulus transformation, then the corresponding robust covariance matrix is obtained by:

$$\hat{\boldsymbol{R}}_{mod} = \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{x}_{mod}(t) \boldsymbol{x}_{mod}{}^{H}(t) \tag{2.22}$$

Apart from these methods, there are several other transformations and procedures to suppress the impulsive noise and can be found in [53, 54, 58]. In an alternate approach to pre-processing, robust statistics provides numerous methods for robust estimation of covariance matrix [37, 38]. One such example is the M-estimator which has been employed in the context of sensor array processing [59, 60]. The following Section revisits some of the robust covariance estimation techniques from statistical literature.

2.5.4 M-estimation of covariance matrix

The M-estimator is a generalization of maximum-likelihood estimation [37]. It provides robustification of ML estimators where the log-likelihood function is replaced by a ρ-function. The estimation of covariance matrix in the framework of M-estimation is defined as the minimization of the following cost functions [37, 38]

$$\hat{\boldsymbol{R}} = \arg\min_{\boldsymbol{R},\boldsymbol{\mu}} \sum_{t=1}^{N} \rho\left(d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R})\right), \tag{2.23}$$

where $\boldsymbol{\mu}$ denotes the mean and $\boldsymbol{R}$ denotes the covariance matrix, $\rho(\cdot)$ is the ρ-function and it provides weighting to suppress the influence of outliers in the observations. $d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R})$ denotes the Mahalanobis distance and is given by the following

$$d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R}) = \sqrt{(\boldsymbol{x}(t) - \boldsymbol{\mu})^H \boldsymbol{R}^{-1} (\boldsymbol{x}(t) - \boldsymbol{\mu})} \tag{2.24}$$

Several choices for ρ are available and the final choice depend on the application, robustness requirement and the efficiency. The ρ-function tunes the impulsive rejection capability and efficiency of the estimator. Commonly used ρ-functions are the Huber's and the Tukey's ρ-functions denoted by ρ_H and ρ_T, given below:

$$\rho_H(d) = \begin{cases} \dfrac{d^2}{2} & \text{if } |d| \leq c_H \\ c_H \text{sign}(d) - \dfrac{c_H^2}{2} & \text{if } |d| > c_H \end{cases} \tag{2.25}$$

$$\rho_T(d) = \begin{cases} 1 - \left[1 - \dfrac{d^2}{c_T^2}\right]^3 & \text{if } |d| \leq c_T \\ 0 & \text{if } |d| > c_T \end{cases} \tag{2.26}$$

$$\tag{2.27}$$

where c_H and c_T are the parameters for Huber's and Tukey's ρ-functions, respectively. c_H and c_T are the thresholds to weigh the outliers differently. The corresponding system of equations to compute the robust estimate of the mean and the covariance matrix are:

$$\sum_{t=1}^{N} W_1(d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R}))(\boldsymbol{x}(t) - \boldsymbol{\mu}) \;=\; \boldsymbol{0}, \tag{2.28}$$

$$\sum_{t=1}^{N} W_2(d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R}))(\boldsymbol{x}(t) - \boldsymbol{\mu})(\boldsymbol{x}(t) - \boldsymbol{\mu})^H \;=\; \boldsymbol{R} \tag{2.29}$$

where W_1 and W_2 are the weights and depend upon the used ρ-function. Weights in case of Huber's and Tukey's ρ-functions denoted by w_H and w_T, respectively are given by:

$$w_H(d) = \min\left[1, \frac{c_H}{|d|}\right] \tag{2.30}$$

$$w_T(d) = \begin{cases} \left[1 - \frac{|d|^2}{c_T^2}\right]^2, & \text{if } |d| \leq c_T \\ 0, & \text{otherwise} \end{cases} \tag{2.31}$$

This system of equations in (Eqs. 2.28 and 2.29) is then solved by iterative re-weighted least squares (IRLS) algorithm. The robustness of an estimator is measured by tools such as breakdown point (BP), asymptotic maximum bias, influence function (IF) and efficiency. The influence function defines the robustness of an estimator to an infinitesimal contamination. The efficiency is the performance of an estimator under the nominal distribution. M-estimator of covariance offers good efficiency and bounded influence [38,61]. However, the M-estimator has a low BP which is defined as the measure of maximum amount of fraction of outliers beyond which the estimator becomes unstable. For Huber's M-estimator, the BP is $\leq \frac{1}{m+1}$ [38]. This means that the BP of M-estimator decreases with increasing dimensionality of the observations. The next Section presents a robust estimator which provides a higher BP than M-estimator.

2.5.5 S-estimation of covariance matrix

The traditional least-square solution can also be interpreted as a minimization of standard deviations of the residuals. Since the standard deviation (scale) is non-robust, thus a robust estimate can be obtained by minimizing a robust scale of the residuals. This is the concept of an S-estimation. The robust estimate of multivariate covariance based on S-estimator can be found in [62,63]. Let $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{R}}$ denote the S-estimator of multivariate location and covariance, respectively. Then the S-estimator of the mean and the covariance are defined as follows:

$$\begin{aligned} & \min \det\left(\hat{\boldsymbol{R}}\right) \quad s.t. \\ \frac{1}{N}\sum_{t=1}^{N} \rho\left(d(\boldsymbol{x}(t), \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{R}})\right) \quad &= \quad b_0 \end{aligned} \tag{2.32}$$

where $\det(\cdot)$ denotes the determinant and the ρ is the ρ-function while b_0 is a constant chosen based on the distribution of the observations $\boldsymbol{x}(t),\ t \in \{1, N\}$ [62]. The choice of b_0 also tunes the BP and the efficiency of S-estimator. A maximally possible achievable breakdown point of 0.5 is reached by S-estimator [64]. The asymptotic properties of S-estimator have been derived in [64,65]. A numerical algorithm to solve the S-estimator for real observations is presented in [66]. However, in sensor array processing, the observations are complex-valued which are then converted into the real-valued observations by stacking the real and imaginary part of the observation. The S-estimator is an affine equivariant and has high breakdown point which makes it a suitable candidate for robust subspace based DOA estimates. The next Section provides another robust

estimator of covariance matrix which provides a high breakdown point and also a high efficiency.

2.5.6 MM-estimation of covariance matrix

The MM-estimator of covariance was first introduced by Tatsuoka et al. [67]. It is highly robust and an efficient estimator. MM-estimator of covariance is estimated using a two-step approach. An initial estimate of covariance is obtained using a very robust estimator such as, an S-estimator of covariance. Then, the final estimate of covariance matrix is computed by using a more efficient M-estimator. Given an S-estimator of covariance in Eq. (2.32) and let $\hat{\sigma}_N := |\hat{\boldsymbol{R}}_S|^{2m}$ where m is the number of sensors, then MM-estimator is the solution to the minimization over all $(\boldsymbol{\mu}, \boldsymbol{R}) \in \mathbb{R}^{2m} \times PDS(2m \times 2m)$ and for which $|\boldsymbol{R}| = 1$ the following [68]

$$\frac{1}{N}\sum_{t=1}^{N} \rho_1 \left(\frac{d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R})}{\hat{\sigma_N}} \right) \tag{2.33}$$

where ρ_1 is a ρ-function to obtain an efficient estimator. In above a $PDS(2m \times 2m)$ denotes a positive definite symmetric matrix of dimensions $2m \times 2m$. Let $\boldsymbol{R}_{\mathrm{MM}}$ denote the corresponding $m \times m$ covariance matrix estimate of sensor array observations. It is obtained from real-valued $\boldsymbol{R}$, given by:

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_A & \boldsymbol{R}_B \\ \boldsymbol{R}_C & \boldsymbol{R}_D \end{pmatrix} \tag{2.34}$$

where $\boldsymbol{R}_A, \boldsymbol{R}_B, \boldsymbol{R}_C$ and $\boldsymbol{R}_D$ are $m \times m$ real covariance matrices. Then $\boldsymbol{R}_{MM}$ is obtained by:

$$\boldsymbol{R}_{MM} = \boldsymbol{R}_A + \boldsymbol{R}_D + j(\boldsymbol{R}_C - \boldsymbol{R}_B) \tag{2.35}$$

A further robust estimator of covariance which trades off breakdown point and efficiency is described next.

2.5.7 Minimum Covariance Determinant estimation of covariance matrix

The minimum covariance determinant (MCD) estimator was first introduced by Rousseeuw et al in [69]. The MCD estimator of covariance is defined as the covariance of selected h such observations which minimize the determinant of the covariance matrix. A fast algorithm to obtain MCD estimate of covariance has been proposed

in [69]. Let $\mathcal{S}_h$ denote the selected observations for which a minimum of determinant is achieved, then the corresponding MCD-estimator of covariance matrix can be written as follows

$$\hat{\boldsymbol{R}}_{rob} = \frac{1}{h} \sum_{i \in \mathcal{S}_h} \boldsymbol{x}(i) \boldsymbol{x}^H(i) \tag{2.36}$$

with $h < N$. A standard choice of h which yields a good trade-off between efficiency and robustness is $h = 0.75N$.

2.5.8 Weighted covariance matrix estimator

In weighted covariance matrix, each observation is weighted based on its distance from the mean. For distance, Mahalanobis distance is used which is given by:

$$d(\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{R}) = \sqrt{(\boldsymbol{x}(t) - \boldsymbol{\mu})^H \boldsymbol{R}^{-1} (\boldsymbol{x}(t) - \boldsymbol{\mu})} \tag{2.37}$$

In the presence of outliers, the estimates for the mean and the covariance are not robust. This results in erroneous distances where outliers do not necessarily get larger distances. It is called the masking effect. Alternatively, robust distances which are obtained by replacing the non-robust estimates of the mean $\boldsymbol{\mu}$ and the covariance $\boldsymbol{R}$ by their robust counterpart, can be used. For robust estimate, MCD estimate of the mean and the covariance matrix defined in Section 2.5.7 can be used. Let $\boldsymbol{\eta}$ and $\boldsymbol{C}$ denote the robust MCD estimate of the mean and the covariance, respectively. Then the corresponding robust distance (RD) is

$$\mathrm{RD}(\boldsymbol{x}(t)) = \sqrt{(\boldsymbol{x}(t) - \boldsymbol{\eta})^H \boldsymbol{C}^{-1} (\boldsymbol{x}(t) - \boldsymbol{\eta})} \tag{2.38}$$

These distances are then used to weigh the observations for the computation of the covariance matrix. The weights can be chosen based on Huber or Tukey's ρ-function and for the case of Huber's ρ-function are given by

$$w(d(\boldsymbol{x}(t), \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{R}})) := \min \left[1, \frac{c_H}{|d(\boldsymbol{x}(t), \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{R}})|} \right] \tag{2.39}$$

where c_H describes threshold for the Huber's ρ-function. The following Section provides simulation results which demonstrate the efficacy of above described robust techniques in the presence of impulsive noise.

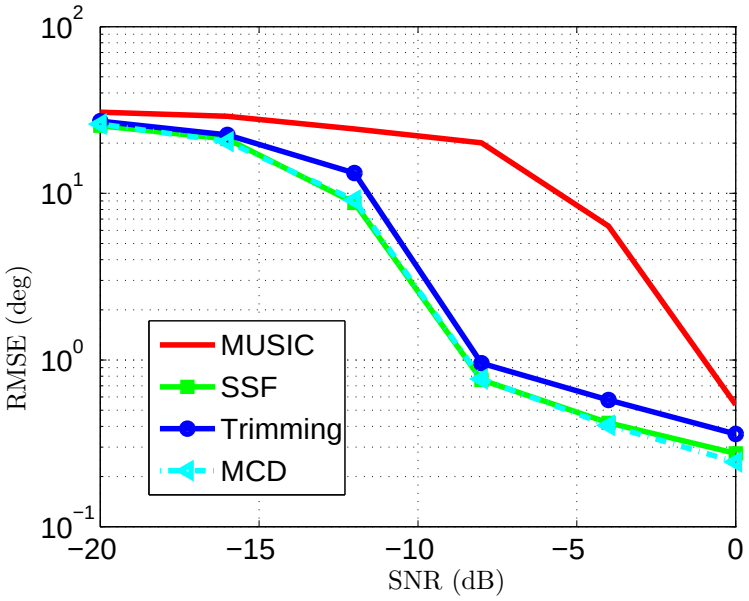

Figure 2.3. The results of RMSE obtained for the conventional, SSF, trimming and MCD based MUSIC for DOA estimation for $\epsilon = 0.1$ and $\kappa = 50$.

2.6 Simulation results

For simulations, three FM (two linear and one hyperbolic) sources impinging on a uniform linear array (ULA) with $m = 8$ sensors are considered. The sources arrive from the directions of $\boldsymbol{\theta} = [-3^\circ\ 5^\circ\ 25^\circ]$ from the broadside. A total of $N = 128$ observations of the array are collected. The noise is impulsive and is modeled by two-term Gaussian mixture model given in Eq. (2.6).

In the following simulations, contamination factor $\epsilon = 0.1$ and impulsiveness factor of $\kappa = 50$ are used. Figures 2.3 depicts the RMSE obtained for the DOA estimation for the pre-processing based and for a chosen MCD estimator. It is evident from the figure that the conventional approach yields DOAs with a mean error of more than 10°. The pre-processed robust approaches yield lower RMSE of DOA when compared to the non-robust covariance approach.

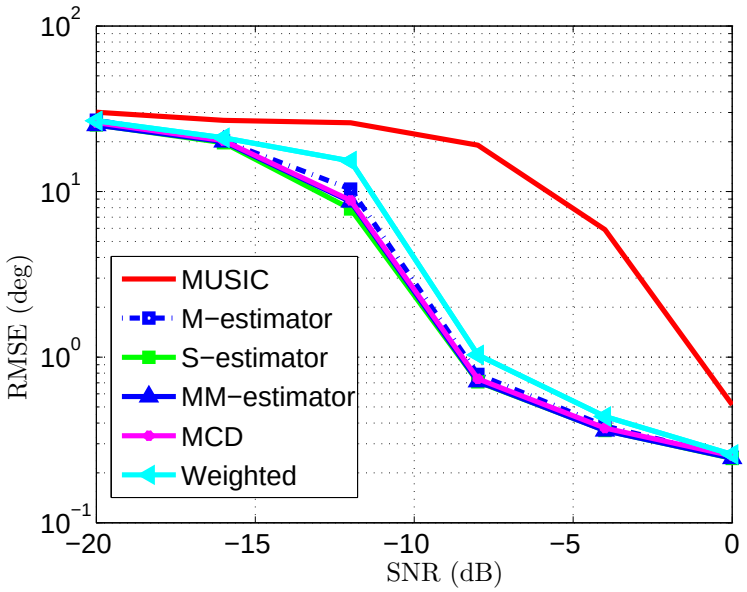

Figure 2.4. The results of RMSE obtained for the conventional, M-estimator, S-estimator, MM-estimator, MCD and weighted covariance based MUSIC for DOA estimation for $\epsilon = 0.1$ and $\kappa = 50$.

In the second simulation, robust covariance matrix estimators from the robust statistics are considered. The results are shown in Figure 2.4 which depict an improvement in the obtained RMSE. The obtained RMSE is comparable to the one obtained by the pre-processing techniques, however, robust approaches from the statistical literature provide better theoretical properties and are more efficient in the case of Gaussian noise. This aspect is discussed in later parts of this thesis.

2.7 Summary

This chapter introduced the problem of DOA estimation for non-stationary sources. The state-of-art conventional methods for the robust DOA estimation were presented. The estimators from the robust statistical literature such as M-estimator, S-estimator and the MM-estimator are shown to produce good results in the presence of impulsive noise. The following chapter introduces DOA estimation of sources by exploiting the non-stationarity of the sources. It also highlights remedies to obtain robust DOAs in the presence of impulsive noise.

Chapter 3

Robust Array Processing For Non-stationary Signals

This Chapter provides an overview to classical and robust DOA estimation approaches for non-stationary sources. Section 3.1 introduces classical non-stationary DOA estimation. Advantages of exploiting the non-stationarity of the signals are highlighted in Section 3.2. The robust DOA estimation algorithm is outlined in Section 3.3. Section 3.4 is dedicated to instantaneous frequency estimation of non-stationary sources. Simulation results are presented in Section 3.5 while Section 3.6 concludes the Chapter with a brief summary.

3.1 Introduction

Non-stationary sources such as FM signals arise frequently in many practical applications. These applications include e.g., radar, sonar, mobile communications, biomedical signals, in vibrational analysis and transient analysis [12]. In radar chirp (linear FM) signals are commonly used [70,71]. Marine animals such as dolphins and whales sounds can be well described by hyperbolic FM signals [2]. Non-stationary signals may also arise due to the Doppler shift introduced by the relative motion between the receiver and the transmitter. In vibrational motion analysis, sinusoidal FM signals are generated due to a rotational motion. Moreover, FM signals are also present in wireless communication and broadcasting. Therein, it is desired to estimate the direction in order to localize the sources in the spatial domain. However, non-stationary sources are well defined by their instantaneous frequencies where *instantaneous frequency is defined as the time-derivative of the instantaneous phase*. The energy of non-stationary signals is concentrated around their instantaneous frequencies in the time-frequency (TF) domain and this is utilized by using time-frequency distributions (TFDs) [1,8,72]. TFDs have been developed and extensively used. TFDs maps a time-only signal into the time and the frequency plane [8]. The following Section provides a brief introduction to TFDs and list some of the commonly used TFDs.

3.1.1 Time-frequency distributions

TFD for a non-stationary signal $y(t)$ is denoted by $D_{yy}(t,f)$ and for the case of Cohen's class of TFDs is defined as follows [8]

$$D_{yy}(t,f) = \frac{1}{4\pi^2}\int\int\int \phi(\theta,\tau)y(t'+\tau/2)y^*(t'-\tau/2)e^{-j\theta t - j2\pi f\tau + j\theta t'}\mathrm{d}t'\mathrm{d}\tau\mathrm{d}\theta \quad (3.1)$$

where $\phi(\theta,\tau)$ denotes the kernel function [1,72]. Kernel function determines the properties of a TFD, such as the cross-terms suppression and the resolution in TF plane. Here, y^* is the complex-conjugate of y. Examples of widely used TFDs include, e.g., the Wigner-Ville distribution (WVD), the Modified B-distribution (MBD) and the Margenau-Hill distribution (MHD) etc. The kernel functions for some of important TFDs are given in Table 3.1.

Name	Kernel Function
Wigner-Ville (WV)	1
Pseudo Wigner-Ville (PWV)	$h(\tau)$
Margenau-Hill (MH)	$\cos(\theta\tau/2)$
Page distribution (PD)	$e^{j\theta\lvert\tau\rvert/2}$
Born-Jordan (BJ)	$(1/2\alpha\tau)\,\mathrm{rect}\,(t/2\alpha\tau)$
Choi-Williams (CW)	$(\sqrt{\pi\sigma}/\tau)\exp\left(-\pi^2\sigma t^2/\tau^2\right)$
B-distribution (BD)	$\lvert\tau\rvert^{\beta}\cosh^{-2\beta}t$
Modified B-distribution (MBD)	$\left(\cosh^{-2\beta}t\right)/\left(\int_{-\infty}^{\infty}\cosh^{-2\beta}\zeta\mathrm{d}\zeta\right)$

Table 3.1. Examples of commonly used TFDs and their kernel function.

For DOA estimation of non-stationary sources, TFDs have received a significant attention over last two decades [9,12,14,73–75]. The following Section will describe the classical approach for TFD based DOA estimation of non-stationary signals.

3.2 Classical DOA estimation for non-stationary signals

In the context of sensor array processing, computation of TFDs across the sensor array yields the spatial time-frequency distribution (STFD) matrix. STFD matrix is an analogue to the covariance matrix which is used in conventional array processing [39]

for DOA estimation. STFD matrix consists of auto- and cross- TFDs of sensor signals as its diagonal and off-diagonal elements, respectively. It is given by the following:

$$[\boldsymbol{D_{xx}}(t,f)]_{pq} = D_{x_p x_q}(t,f,\phi), \quad p,q \in \{1,m\}, \quad t,f \in \mathbb{R} \tag{3.2}$$

$$\boldsymbol{D_{xx}}(t,f) = \begin{bmatrix} D_{x_1x_1}(t,f) & D_{x_1x_2}(t,f) & \dots & D_{x_1x_m}(t,f) \\ D_{x_2x_1}(t,f) & D_{x_2x_2}(t,f) & \dots & D_{x_2x_m}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ D_{x_mx_1}(t,f) & D_{x_mx_2}(t,f) & \dots & D_{x_mx_m}(t,f) \end{bmatrix} \tag{3.3}$$

where $[\cdot]_{pq}$ denotes the $(p,q)^{th}$ element of the STFD matrix and $D_{x_p x_q}(t,f)$ with $p \neq q$ denote the cross-TFD of the signals corresponding to the p^{th} and the q^{th} sensor element of the array, x_q^* is the complex-conjugate of x_q. For the purpose of DOA estimation of non-stationary sources, a commonly used TFD is the WVD as defined in Table 3.1. The $(p,q)^{th}$ element of STFD matrix based on the WVD is given by the following

$$\hat{D}_{x_p x_q}(t,f) = \sum_{l=-\infty}^{\infty} x_p(t+l)x_q^*(t-l)e^{-j4\pi fl} \tag{3.4}$$

In practice, only a finite number of samples are available and therefore a finite window of length L is used. The resulting $D_{x_p x_q}$ is then called the pseudo-WVD(PWVD) given by:

$$\hat{D}_{x_p x_q}(t,f) = \sum_{l=-L/2}^{L/2} x_p(t+l)x_q^*(t-l)e^{-j4\pi fl} \tag{3.5}$$

Given the signal model of Eq. (2.1), the structure of the STFD matrix is also analogous to the structure of the covariance matrix.

$$\boldsymbol{R} = \mathrm{E}\left[\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right] = \boldsymbol{A}\boldsymbol{R_{ss}}\boldsymbol{A}^H + \sigma^2\boldsymbol{I} \tag{3.6}$$

$$\mathrm{E}\left[\boldsymbol{D_{xx}}(t,f)\right] = \boldsymbol{A}\boldsymbol{D_{ss}}(t,f)\boldsymbol{A}^H + \sigma^2\boldsymbol{I} \tag{3.7}$$

Here, $\boldsymbol{D_{xx}}$ denotes the STFD matrix of the array signal $\boldsymbol{x}$, while $\boldsymbol{D_{ss}}$ is the matrix containing the auto- and cross-source TFDs of the source signals, given by

$$\boldsymbol{D_{ss}}(t,f) = \begin{bmatrix} D_{s_1s_1}(t,f) & D_{s_1s_2}(t,f) & \dots & D_{s_1s_K}(t,f) \\ D_{s_2s_1}(t,f) & D_{s_2s_2}(t,f) & \dots & D_{s_2s_K}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ D_{s_Ks_1}(t,f) & D_{s_Ks_2}(t,f) & \dots & D_{s_Ks_K}(t,f) \end{bmatrix} \tag{3.8}$$

and $\boldsymbol{s}(t) = [s_1(t), s_2(t), \dots, s_K(t)]^T$ is the source vector at time t. Since the covariance matrix and the STFD matrix bear similar structures, therefore, subspace based DOA estimation can also be applied on STFD matrices. But, what is the advantage of using the STFD matrix instead? These are formulated as follows:

1. Due to non-stationarity of the signals the covariance matrix changes over time and, therefore, STFD is a relevant statistic for non-stationary signals [1].

2. The non-stationary signals might have distinct instantaneous frequencies and such signals can be discriminated based on their instantaneous frequencies. This allows that the STFD matrices can be computed for a subset of the sources. This relaxes the requirement of having less number of sources than sensors ($K < M$) required for conventional DOA estimation. Hence, the STFD based DOA estimation also provide us with an additional degree of freedom to treat the under-determined DOA estimation [1, 11, 73].

3. The instantaneous frequency of sources can be used for signal excision in the TF plane. This leads to better noise rejection and enhanced SNR as compared to the conventional methods for DOA estimation [1, 11, 76, 77].

However, an additional requirement for the STFD based DOA estimation method is the knowledge of the instantaneous frequency of the sources. The estimation of instantaneous frequency is discussed in Section 3.4. In the next Section, the algorithm for DOA estimation exploiting the non-stationarity of signals is described.

3.2.1 Algorithm for non-stationary DOA estimation

DOA estimation by exploiting the non-stationarity of signals has been investigated over the last two decades. In the literature, a few examples of previous works on non-stationary DOA estimation include, e.g. [11–14, 73, 74]. DOA estimation algorithm for non-stationary signal is outlined in Figure 3.1.

It is evident from Figure 3.1 that given the observations $\{\boldsymbol{x}(t)\}_{t=1}^{N}$, the instantaneous frequency is estimated using the spatially averaged auto-TFD image. The IFreq estimation is detailed in Section 3.3. STFD matrices are computed only for the TF points which are obtained during the IFreq estimation step. Let $\boldsymbol{\mathcal{M}}_k,\ k \in \{1, P\}$ denote the obtained STFD matrices computed across IFreq of the sources. DOA estimation then requires then estimation of the noise and the signal subspace in case of MUSIC [44]. For the P STFD matrices defined above, the subspaces can be estimated either by joint block diagonalization (JBD) as proposed in [74] or by STFD matrix averaging as proposed in [73]. These two methods are explained in the next Section.

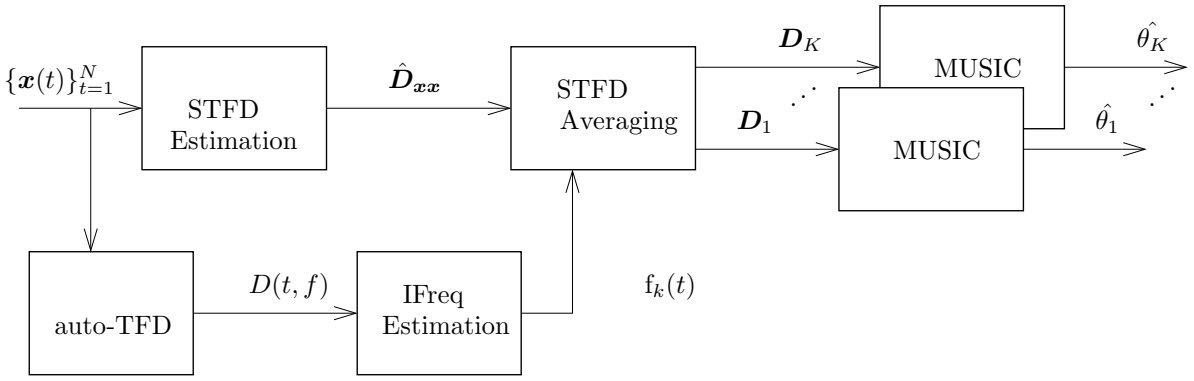

Figure 3.1. Direction-of-arrival estimation of non-stationary signals based on time-frequency distributions.

3.2.2 Time-frequency MUSIC

Let $\boldsymbol{\mathcal{M}}$ denote the set of STFD matrices which are computed along the IFreqs of at least $K_0 < K$ non-stationary signals. To estimate the joint signal and the noise subspaces, Belouchrani et al. [74] proposed to find a unitary transformation $\boldsymbol{U}$ which jointly block diagonalizes the set of STFD matrices $\boldsymbol{\mathcal{M}}$. This is achieved by finding $\boldsymbol{U}$ which maximizes the following cost function [78]:

$$\hat{\boldsymbol{U}} = \arg\max_{\boldsymbol{U}} \sum_{k\in\boldsymbol{\mathcal{M}}} \sum_{i,j=1}^{K_0} \boldsymbol{u}_i^H \boldsymbol{D}_i \boldsymbol{u}_j \tag{3.9}$$

$$\hat{\boldsymbol{U}} = [\underbrace{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{K_0}}_{\hat{\boldsymbol{E}}_s}, \underbrace{\boldsymbol{u}_{K0+1}, \ldots, \boldsymbol{u}_m}_{\hat{\boldsymbol{E}}_n}] \tag{3.10}$$

where $\hat{\boldsymbol{E}}_s$ and $\hat{\boldsymbol{E}}_n$ denote the signal and the noise subspaces, respectively. The DOAs are then obtained finding the largest K_0 peaks in the MUSIC spectrum [44], given by

$$P_{MUSIC}(\theta) = \frac{1}{|\hat{\boldsymbol{E}}_n \boldsymbol{a}(\theta)|^2} \tag{3.11}$$

An efficient Jacobi-like technique can be used for joint block-diagonalization of STFD matrices [78, 79]. This algorithm requires more computational effort than an alternate approach where STFD matrices are averaged across the IFreq of the signals and is explained in the next Section.

3.2.3 Spatial time-frequency distribution matrices averaging

Amin et al. [73] proposed an alternate approach to obtain the joint signal and noise subspaces by averaging the STFD matrices over IFreqs belonging to a subset of $K_0 < K$

sources. The averaging of the STFD matrices over the IFreqs is comparatively simpler than the TF MUSIC. In this work STFD averaging is used for DOA estimation and the corresponding algorithm is depicted in Figure 3.1. The averaging of the STFD matrices results in the magnification of signal eigenvalues as compared to the noise eigenvalues [11, 76, 77]. This leads to an improvement in the estimation accuracy. This approach also allows single source DOA estimation treatment by averaging the STFD matrices along the IFreq belonging to a single source.

In the following Section, the problem of DOA estimation for non-stationary sources in the presence of impulsive noise is investigated. Though, conventional DOA estimation under non-Gaussian noise for conventional array processing has been treated in the literature and numerous methods can be found in [48–50, 53, 55, 56]. However, robust methods for DOA estimation of non-stationary under impulsive noise do not exist. The proposed algorithm to obtain robust DOA estimates for non-stationary signals is discussed in the next Section.

3.3 Robust DOA estimation for non-stationary signals

The DOA estimation of non-stationary sources depends on the TFDs which are not robust in the presence of impulsive noise [34, 80, 81]. It is also evident from Eq. (3.5) that the outliers can have an unbounded influence on the final estimate. The resulting estimates of STFD matrix are unreliable which lead to the performance breakdown of classical DOA estimation methods for non-stationary signals.

Under such scenario, this work proposes to use robust IFreq and STFD matrices estimation instead of classical estimates. The proposed approach for DOA estimation of non-stationary sources in the presence of impulsive noise is depicted in Figure 3.2. For better illustration, the steps of the algorithm to estimate DOA of non-stationary sources robustly are also outlined in Table 3.2.

Robust estimation of STFD matrix is treated in Chapter 4 while IFreq estimation based on robust TFDs is discussed in the next Section.

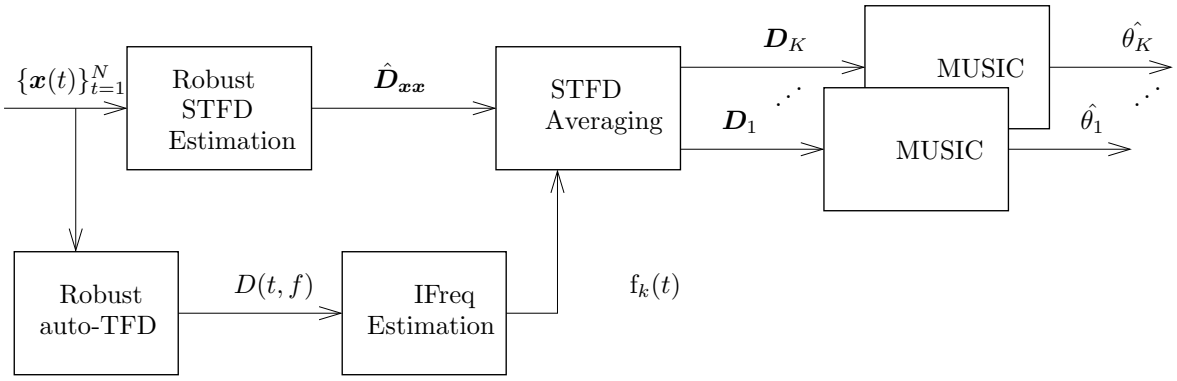

Figure 3.2. Robust direction-of-arrival estimation of non-stationary signals based on the time-frequency distributions.

3.4 Instantaneous frequency estimation

Non-stationary signals are well described by their IFreqs. In the context of radar, sonar and mobile communication, the relative motion between the receiver and the target leads to the frequency shift due to Doppler phenomenon. Therefore, the estimation of IFreq allows us to infer speed, direction and trajectory of the moving target. Moreover, in telecommunications, the interference signals often belong to a class of non-stationary signals and the estimation of IFreq allows to design methods to suppress the interference. In vibrational analysis, the IFreq gives valuable information about the underlying motion. In underwater navigation, marine animals can be classified based on their IFreqs. Hence, the problem of IFreq estimation bears a significant importance in many practical applications and some of the existing methods to estimate the IFreq can be found in [33, 82–88].

The basic introduction to the problem of IFreq estimation is provided in [82,83,89] while IFreq estimation method based on the discrete evolutionary transform (DET) exists in [90], based on time-varying auto-regressive (TVAR) modeling in [91], peak search in TFDs is presented in [33, 34, 81, 84, 85, 92]. However, IFreq estimation methods which utilize the TFDs are the focus of this work.

IFreq estimation methods based on TFDs can be classified into parametric or non-parametric approaches. The examples of parametric approaches are the parameter estimation based on Hough transform proposed in [75,93–96]. In [97], Radon transform has been used to estimate the IFreq of the signals. In non-parametric methods, IFreq estimation based on the morphological image processing has been proposed in [14, 87]. This work robustify the method proposed in [14] while the major advantage of such a

Table 3.2. Robust DOA estimation of FM sources.

1. Compute the spatially averaged robust TFD for IFreq estimation of the impinging sources. The objective of this step is to obtain non-overlapping IFreq segments. This robust IFreq estimation part is discussed in Section 3.4.
2. For each IFreq segment, compute an averaged STFD matrix by using robust STFD estimation techniques described in more detail in Chapter 4.
3. Estimate the DOA of each source by using MUSIC algorithm for each averaged STFD.
4. The obtained broken segments belonging to a single source yield DOAs which are close w.r.t. estimation variance. Herein, some clustering algorithms can be used to combine the results from the previous step.

technique is that it does not require any preliminary information about the number of signals and/or the types of FM sources.

Although, robust techniques for IFreq estimation exist, however, robust IFreq estimation in particular for DOA estimation has not been addressed in literature. The organization of the remaining part of this Chapter is as follows: problem formulation for IFreq estimation is described in Section 3.4.1. Section 3.4.2 describe two state-of-the-art approaches while proposed improvements to counter the effect of impulsive noise in practical scenarios are listed in Section 3.4.4. Section 3.4.5 describes performance measures to assess the accuracy of estimated IFreq. Moreover, the demonstrate the advantages of using robust non-stationary DOA estimation in comparison to the conventional robust techniques is simulation examples are provided in Section 3.5. The chapter is concluded with a short summary in Section 3.6.

3.4.1 Problem formulation

Consider again, the received signal at the p^{th} sensor of the array based on the signal model in Section 2.2 is given by

$$x_p(t) = \sum_{k=1}^{K} a_p(\theta_k)\mathrm{e}^{-j\phi_k(t)} + n_p(t), \quad f_k(t) = \frac{\partial\phi_k(t)}{2\pi\partial t} \tag{3.12}$$

where ϕ_k denotes the phase of the k^{th} signal and the IFreq of the signal is the time derivative of the instantaneous phase given above. The IFreq of the signal $x_p(t)$ can be modeled as a polynomial of a certain order P, given by:

$$f_k(t) = \sum_{i=0}^{P} \alpha_i^k t^i \tag{3.13}$$

where α_i denotes the parameters describing the IFreq of the k^{th} signal. The objective is to estimate the parameters $\alpha_i^k, \quad \forall k \in \{1, K\}, \quad i \in \{1, P\}$ or equivalently the IFreqs $f_k(t)\ \forall k \in \{1, K\}, t \in \{1, N\}$ given the N observations of the array $x_p(t),\ \forall\ t \in \{1, N\},\ p \in \{1, m\}$ in the presence of impulsive noise. In the following Section, two state-of-the approaches for the IFreq estimation are described.

3.4.2 State-of-the art: parametric approaches

The parametric approaches require prior knowledge about the number of sources in the mixture and the polynomial order of the IFreq of each of the individual source. The parametric approaches provide an improved estimation quality when compared to the non-parametric approaches. In parametric IFreq estimation, the signal phase is modeled as a polynomial phase signal (PPS) of a certain order P as in Eq. (3.13) and several techniques exist to estimate the parameters of the polynomial. Among those a widely used technique called the Wigner-Hough transform (WHT) is based on the Hough transform [98]. It was proposed by Barbarossa et al. in [94, 95]. In this approach, the signals $x_p(t)$ are firstly transformed into TF domain by using auto-WVD. The resulting TFD is then mapped to the parameters domain by using the Hough transform [98]. The peaks in the parameter's domain then correspond to the estimates of the IFreq parameters of the signals. This is more suited for the linear FM signals as the WVD produces the best concentration along the IFreq of signals for linear FM signals [1]. For non-linear FM signals such as hyperbolic, sinusoidal FM signals an approach based on pattern recognition of the TF images was proposed in [99].

In the context of array processing, the application of WHT requires that (i) WVDs are computed for each sensor signal (ii) WVDs of all the sensors are averaged to obtain a single TF image (iii) the parameters are estimated by projecting the obtained image to parameters domain by Hough transform. The averaged TFD across the sensor array is given by

$$\begin{aligned} D_{\text{WVD}}^{av}(t,f) &= \frac{1}{m}\sum_{p=1}^{m} \Re\left(\sum_{l} x_p(t+l)x_p^*(t-l)e^{-j4\pi fl}\right), \\ & \quad t, f \in \mathbb{R} \end{aligned} \tag{3.14}$$

where $\Re(\cdot)$ denotes the real of the complex-valued WVD $\sum_l x_p(t+l)x_p^*(t-l)e^{-j4\pi fl}$. The application of Hough transform maps given TFD image into the parameters of the IFreq, i.e., the $\{\alpha_i^k\},\ i \in \{1, P\},\ k \in \{1, K\}$. The estimates for the parameters then correspond to the K peaks in the cost function for the WHT given by:

$$P_{\mathrm{WHT}}(\alpha_i^k) = \sum_{t=1}^{N} D_{\mathrm{WVD}}^{av} e^{-j4\pi f(\alpha_i^k)t}, \quad i \in \{1, P_k\}, \quad k \in \{1, K\} \tag{3.15}$$

A variation of WHT which uses a pseudo-Wigner Ville distribution (PWVD) instead of the WVD was proposed in [75]. The resulting approach is termed as pseudo Wigner-Hough transform (PWHT) and provides a significant reduction in terms of computational effort, however at a reduced output SNR as compared to WHT [75]. PWHT based cost function is given by:

$$P_{\mathrm{PWHT}}(\alpha_i^k) = \sum_{t=1}^{N/2} \sum_{l=-L/2}^{L/2} x(t+l)x^*(t-l)e^{-j4\pi(f+\sum_{k=1}^{K}\alpha_i^k)t} \tag{3.16}$$

Table 3.3 lists a a few other examples of existing parametric approaches for IFreq estimation which can be employed in the context of array processing. The drawback

Table 3.3. Parametric approaches for IFreq estimation

1. Hough transform based on the local polynomial phase transform (LPP) [96]
2. Product higher order ambiguity function (PHAF) [99]
3. Non-linear least square (NLS) approach as proposed in [100]
4. Parameter estimation based on the Radon transform [97]

of parametric techniques is the requirement of having the prior information about the number of sources and types of their IFreqs, which in practice is not available. Alternate is to use non-parametric approaches. The following Section reviews a non-parametric technique for IFreq estimation of non-stationary sources.

3.4.3 State-of-the-art: non-parametric approaches

In non-parametric approaches, this work considers an approach which is based on the morphological image processing of the TF images and was proposed in [13, 14, 87].

The steps of a morphological image processing based IFreq estimation are as follows: (i) the sensor signals are transformed in to the TF domain by using a cross-term free and a high resolution TFD, (ii) the TFDs are averaged across sensors to obtain an averaged TF image and (iii) morphological image processing operations are performed to extract non-overlapping IFreq segments.

This approach requires a TFD which has a high resolution in TF plane and is free of cross-terms. The examples of such TFDs are the spectrogram, modified B-distribution (MBD) and a recently proposed complex-lag distribution [101]. The averaging to obtain TF image leads to enhancement of the signal component over the noise. The morphological image processing operations are as follows:

1. Thresholding
2. Erosion and dilation
3. Thinning
4. Removing junction points and removing spurs

In the following, each step of the morphological image processing technique is detailed.

Thresholding: The objective of thresholding is to obtain a binary image which contains ones for all those points which are segmented as source TF points while the points labeled as zeros belong to noise TF point. This is achieved by selecting a threshold and all the TF points for which the value exceeds the threshold are labeled as the source TF points and all remaining TF points with value lower than that of the threshold are discarded as noise TF points.

This work suggests to obtain the threshold based on a histogram based iterative approach proposed in [102]. This technique is summarized in Table 3.4. Alternatively, Gaussian fitting [103] can also be used to find the threshold which then segments the TF image into the signal and noise TF points.

Morphological opening and closing: Morphological opening and closing operations are performed to obtain smoothed objects. These operations use dilation and erosion. Dilation is defined as follows:

$$C \oplus E = \bigcup_{b \in C} C_b \tag{3.17}$$

Table 3.4. Iterative thresholding for retrieving the source TF points

1. Initialization: Compute the histogram of the TF image and divide the histogram into two equal regions. The initial threshold is set to be the mean value of the image.
2. Given the initial/previous threshold, re-label the pixels of the image as the signal and the noise pixels. Then, compute new mean for each of the class and let the corresponding means be m_s and m_n.
3. Compute the new threshold as the average of means of the two classes as $t_{new} = \dfrac{m_s + m_n}{2}$
4. Repeat steps 2 and 3 until the means m_s and m_n do not change over two successive iterations.

where E denotes the structuring element and C is the image area equal in size to E, $\bigcup$ denotes the union operation. The dilation is an increasing operation and results in objects to grow. Similarly, erosion is the opposite of the dilation and leads to the contraction of objects, given by:

$$C \ominus E = \bigcap_{b \in C} C_b \tag{3.18}$$

The morphological *opening* consists of erosion followed by dilation $(A \ominus E) \oplus E$ while *closing* is obtained by performing dilation before erosion $(A \oplus E) \ominus E$. The opening operation produces smoothed edges by removing small objects while the closing operation fills small objects. In this step of algorithm, structuring element to be used for obtaining smoothed objects needs to be specified. For this purpose, it is remarked from simulations that a diamond shaped structuring element provides good results.

Thinning: Morphological thinning is then performed to extract medial axis as estimates of IFreq. Let I denote the image and E be the structuring element, then thinning operation is defined as follows:

$$\text{thin}(I, E) = I \cap \overline{I \otimes E} \tag{3.19}$$

where $\cap$ denote the intersection, $\otimes$ is the hit-and-miss transform and $\bar{\cdot}$ denotes the logical complement operation (The interested readers are referred to [104] for further details and examples for obtaining the medial axis). This is the final estimate unless there are overlapping components. In case of overlapping IFreq segments, an additional

step has to be performed to remove the overlaps. For better understanding, algorithm to obtain the IFreq estimates given the observations of a sensor array is summarized in Table 3.5. In the presence of impulsive noise the above mentioned techniques suffers a

Table 3.5. Morphological image processing based instantaneous frequency estimation

1. Compute spatially averaged TF image based on a TFD which is free of cross-terms.
2. Perform the image thresholding to convert to a binary image consisting of the TF points belong to the source as one and all others as zeros.
3. Perform dilation and erosion on the thresholded image to obtain smoothed objects.
4. Perform thinning to extract the medial axis of the smoothed objects from the above step.
5. Remove the junction points and obtain the non-overlapping IFreq segments and remove small spurious components.

performance loss for IFreq estimators. This work propose to use robust TFDs instead of classical TFD estimates for morphological image processing based IFreq extraction. In the following Section, robust versions of TFDs that can be used for morphological image processing technique are described.

3.4.4 Robust TFDs for IFreq estimation

The robust IFreq estimation has been previously considered in the literature and existing robust methods for IFreq estimation can be found in [80, 92, 105, 106]. In [80], an IFreq estimator is proposed which is based on peak search in robust spectrogram. The spectrogram is computed based on L-estimation. Also, in [105], Barakat et al. proposed an IFreq estimator based on L-estimation of WVD. The spectrogram offers poor resolution and the WVD contains cross-terms. Therefore, this work proposes to use robust MBD. The generalized TFD in discrete form can also be written as follows [34]:

$$D_{x_p x_p}(t,f) = \sum_{l=-L/2}^{L/2} \mathcal{G}(t,l) e^{-j4\pi fl} \tag{3.20}$$

where $\mathcal{G}(t,l)$ denotes the time convolution between the TFD kernel function $\varphi(t,l)$ and the instantaneous auto-correlation function defined as $R_{x_px_p}(t,l) := x_p(t+l)x_p^*(t-l)$. The kernel function for MBD is:

$$\varphi(t,l) = \left[\frac{|l|}{cosh(t)}\right]^{\beta} \tag{3.21}$$

where parameter β tunes the cross-term suppression. The range of β is between $0 < \beta < 1$ and a best trade-off between resolution and cross-term resolution is obtained for $\beta = 0.1$ [107]. Then, the robust version of MBD based can be defined as a solution to the minimization of the following cost function [1]

$$D_{x_px_p}(t,f) = \arg\min_{D} \sum_{l=-L/2}^{L/2} |\mathcal{G}(t,l)e^{-j4\pi fl} - D|^{\alpha} \tag{3.22}$$

where α is the factor determining the robustness of the resulting MBD. For $\alpha = 1$, resulting MBD is called the median based MBD which is robust. The solution to the above minimization problem leads to the following:

$$D_{xx}(t,f) = \frac{1}{a_h(l)} \sum_{l=-L/2}^{L/2} d(t,f,l)\mathcal{G}(t,l)e^{-j4fl\pi} \tag{3.23}$$

where $d(t,f,l)$ and $a_h(t,f,l)$ are given by

$$d(t,f,l) = \frac{1}{|\mathcal{G}(t,l)e^{-j4\pi fl} - D_{xx}(t,f)|} \tag{3.24}$$

and $a_h(t,f,l) = 1/\sum_l d(t,f,l)$, respectively. Eq. (3.23) can be solved iteratively and iterations are run until the relative change between two successive iterations is very small. The iterative approach is computationally expensive, therefore, robust MBD using vector median based approach is also proposed. To explain the vector median approach, let

$$\begin{aligned} \bar{\boldsymbol{v}} &:= \Re\big(\mathcal{G}(t,l)e^{-j4\pi fl}\big), \quad l \in \{-L/2,\ L/2\} \\ \tilde{\boldsymbol{v}} &:= \Im\big(\mathcal{G}(t,l)e^{-j4\pi fl}\big) \end{aligned} \tag{3.25}$$

denote the vectors containing the real and the imaginary parts of the complex-valued $\mathcal{G}(t,l),\ l \in \{-L/2,\ L/2\}$. For robust IFreq extraction, robust vector median based approach can be written as follows:

$$D^{med}_{x_px_p}(t,f) = \text{median}\,(\bar{\boldsymbol{v}}) + j\text{median}\,(\tilde{\boldsymbol{v}}) \tag{3.26}$$

Though median based estimators provide robustness and high breakdown point, but lack in terms of efficiency (which is defined as the performance under nominal model).

The alternate efficient approaches are the weighted TFDs wherein Eq. (3.23) can be interpreted as mean of observations $\mathcal{G}(t,l)e^{-j4\pi fl},\ l \in \{-L/2,\ L/2\}$. The weights for the observations are chosen to suppresses the influence of outliers in the final estimate. The weighted MBD can be written as follows:

$$D_{x_p x_p}(t,f) = \sum_l w(l)\mathcal{G}(t,l)e^{-j4\pi fl} \tag{3.27}$$

There are numerous ways to choose weights and will be discussed in more detail in Chapter 4. A simple procedure for weights can be e.g.,

$$w(l) := \frac{1}{|\mathcal{G}(t,l)e^{-j4\pi fl} - \hat{\eta}|}. \tag{3.28}$$

Here, weights are defined as inverse of the distance of the observation $\mathcal{G}(t,l)e^{-j4\pi fl}$ from the estimate of mean $\hat{\eta}$ of the sequence $\mathcal{G}(t,l)e^{-j4\pi fl},\ l \in \{-L/2,\ L/2\}$. The performance of the resulting IFreq estimators is highly dependent on the type and the properties of TFDs. The desirable properties are cross-term reduction and the resolution of the TFDs in the TF plane. Following Section describes two metrics to assess the performance of an IFreq estimator.

3.4.5 Performance measures

The performance of a parametric IFreq estimators can be measured by Cramér Rao bound (CRB). CRB analysis under Gaussian noise for parametric IFreq has been performed in [89, 100, 108]. However, in the case of non-parametric approaches, mean square error (MSE) can also be used as a measure to express the accuracy for a non-parametric IFreq estimator. It is defined as

$$\mathrm{MSE}_f = \frac{\sum_{k,t,j}(f_k(t) - \hat{f}_k^j(t))^2}{NTK} \tag{3.29}$$

where $\hat{\cdot}$ is denoting the estimated quantities, $\hat{f}_k^j(t)$ being the estimate of IFreq parameter for the k^{th} signal at j^{th} trial and T is the number of Monte-Carlo trials.

For the morphological image processing based algorithm, the performance of the IFreq estimator hugely depends on the readability and resolution of the TFD. This can be measured by the concentration of energy in the TF plane and as defined in [1, 96],

$$B = 10\log\left(\frac{\mathrm{average}(D(t,f),\ t,f \in \mathcal{S})}{\mathrm{average}(D(t,f),\ t,f \notin \mathcal{S})}\right) \tag{3.30}$$

where $\mathcal{S}$ denote the set of TF points which belong to the signal. This ratio measures relative concentration of the signal's energy in the TF domain.

3.5 Simulation results

To give an example of robust DOA estimation of non-stationary sources employing robust IFreq estimation, two linear FM sources in the presence of impulsive noise are considered. Median-based MBDs as defined in Eq. (3.26) and Eq. (3.23) followed by morphological image processing are used to extract the IFreq estimates. Figure 3.3 shows the non-robust MBD of the two sources without noise, and in impulsive noise, as well as the robust MBD and the extracted IFreq using the described technique. From Figure 3.3, it is evident that the application of morphological operations on the

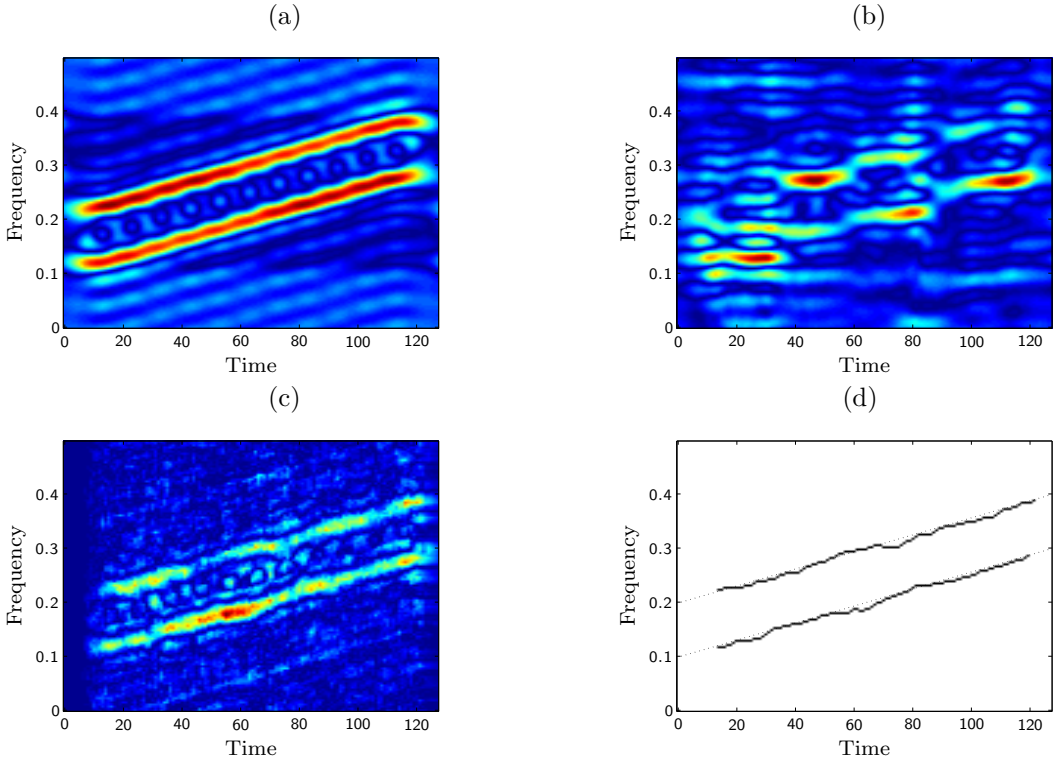

Figure 3.3. (a) MBD of the noise-free signal (b) Non-robust MBD of the signal at SNR = −5 dB (c) Median-based robust MBD (d) IFreq extraction with median-based MBD.

non-robust TF image most likely lead to an incorrect IFreq estimate. On the other hand, the robust median-based approach is able to provide an accurate estimate of the TF signature.

A comparison based on MSE of obtained IFreq estimates and relative concentration ratio has also been performed. Figures 3.4 shows the results for MSE of IFreq estimates based on morphological image processing with automatic threshold selection while Figure 3.5 depicts the concentration ratios. It is evident that the vector median based

approach provides good concentration and MSE for the IFreq estimation. To evaluate

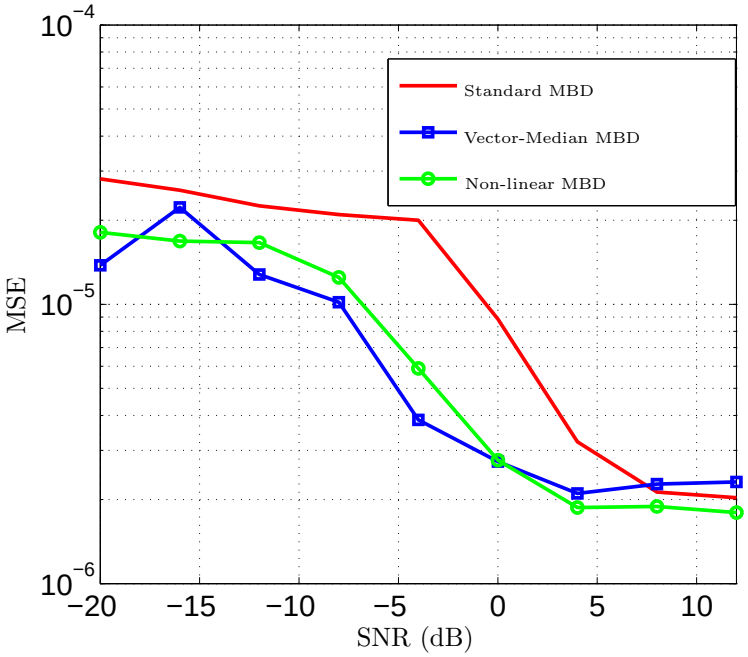

Figure 3.4. MSE of Standard MBD, vector-median and iterative MBD based IFreq estimation in impulsive noise.

the performance of the proposed method in comparison to standard robust methods for DOA estimation in impulsive noise, 500 Monto-Carlo trials are conducted. In the setup, $M = 8$ sensors in an ULA geometry and two linear FM signals, as shown in Figure 3.3, from DOAs $[-3^\circ\ 2^\circ]$ have been simulated. All sources have the same power and $N = 128$ snapshots are used for estimation. The noise $\boldsymbol{n}(t)$ in Eq. (2.2) is modeled as an ϵ-contaminated mixture, given by

$$\boldsymbol{n}(t) \sim (1-\epsilon)\mathcal{N}_C(\mathbf{0}, \sigma^2\mathbf{I}) + \epsilon\mathcal{N}_C(\mathbf{0}, \kappa\sigma^2\mathbf{I}) \tag{3.31}$$

In the simulations, $\epsilon = 0.2$ and $\kappa = 20$ was used. For IF extraction, the robust MBD using the median-based procedures of Eq. (3.26) and the iterative procedures of Eq. (3.23) with a window length of $L = 21$ was used. For the computation of STFD matrices, the iterative robust WVD (described in Chapter 4) with a window length of $L = 29$ was used. Figures 3.6 and 3.7 show the RMSE of DOA estimates versus SNR for both sources.

It can be observed in both figures that all TF methods converge to a smaller RMSE when compared with the non-TF methods, since effectively single-source DOA estimation is carried out. In terms of standard robust methods, the Visuri method provides

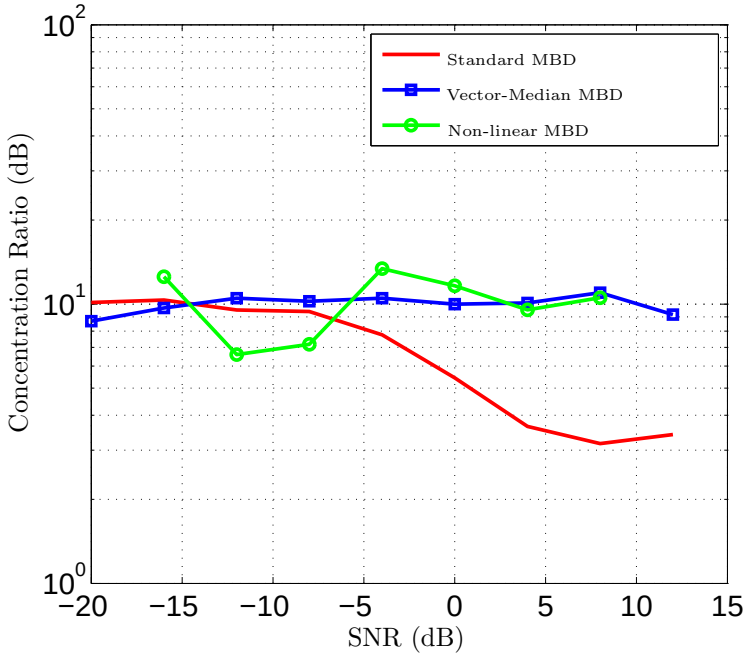

Figure 3.5. Concentration ratio for standard, vector-median and iterative MBD.

the lowest RMSE while the method without any impulsive remedy requires a high SNR to achieve reasonable results. It can also be seen that the median-based techniques is able to provide a cleaner image which eases the IFreq extraction when compared with the iterative based methods which is computationally more expensive. In the case of known IFreq, the robust TF averaging provides the best RMSE.

3.6 Summary

In this chapter, robust DOA estimation of non-stationary sources was discussed. The robust IFreq estimation algorithm utilizes the robust vector-median MBD in combination with the morphological image processing technique. STFD based approaches with both known and unknown IFreq yield lower RMSE in comparison to the robust procedures from the conventional array processing.

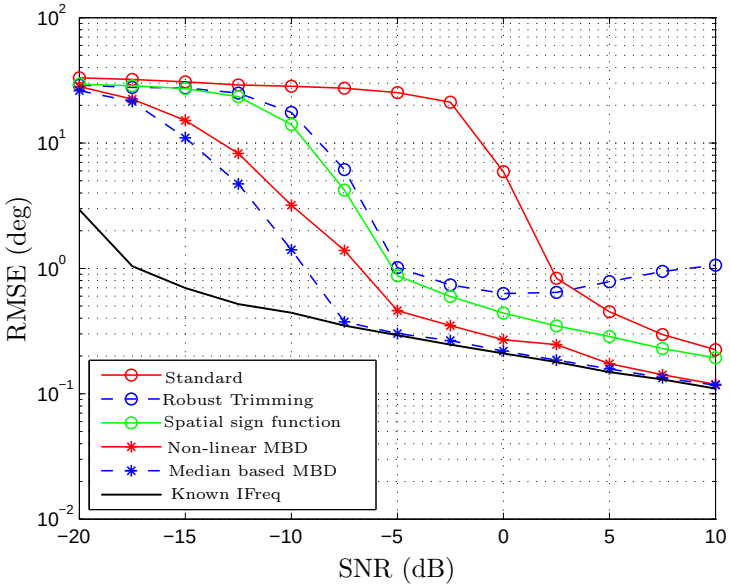

Figure 3.6. RMSE of DOA estimate for the source from $-3°$.

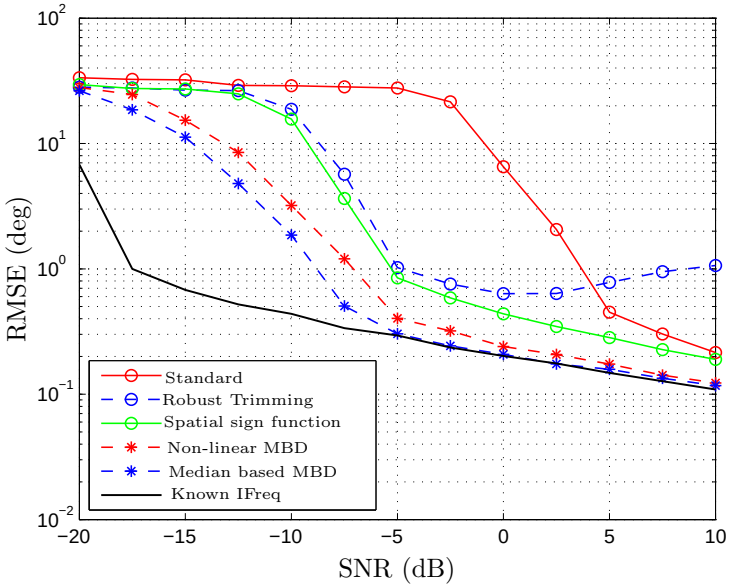

Figure 3.7. RMSE of DOA estimate for the source from $2°$.

Chapter 4

Robust Spatial Time-Frequency Distribution Matrix Estimation

[1] This Chapter considers robust estimation of the spatial time-frequency distribution (STFD) matrices. It is organized as follows: Section 4.1 briefly introduces the problem of robust STFD matrices estimation in sensor array applications. Robust DOA estimation based on STFD matrices is discussed in Section 4.2. Section 4.3 lists preprocessing techniques which can be applied for estimating the STFD matrices robustly in the context of DOA estimation. The robust position based STFD matrix estimation is described in Section 4.4 while simpler non-iterative robust methods are listed in Section 4.5. The simulations and an objective comparison of proposed techniques for DOA estimation are presented in Section 4.6 and a summary is provided in Section 4.7.

4.1 Introduction

To exploit the time-varying spectra of non-stationary signals which arise in many practical applications, spatial time-frequency distribution (STFD) matrices constitutes a common and powerful tool. STFD matrices were firstly introduced in [110,111]. STFD matrices enhance the spatial resolution and estimation accuracy in sensor array processing applications such as DOA and blind source separation (BSS) by capturing well the different sources' signatures present in the time-frequency (TF) plane [14, 76, 112, 113]. DOA estimation based on STFDs has been treated in [14, 76, 110–113] while blind source separation (BSS) based on STFDs can be found in [9, 16, 75, 114]. However, main focus of this Chapter is to develop robust STFD methods with an application to DOA estimation of non-stationary signals in the presence of impulsive noise.

Existing techniques for STFD based DOA estimation consider only the sensors' thermal noise, which is assumed to be Gaussian. As described earlier, in practical radio environments, the noise is often non-Gaussian and impulsive due to many natural and

[1] Parts of this Chapter are based on the article "Robust spatial time-frequency distribution matrices with application to DOA estimation" in Elsevier Signal Processing by W. Sharif, Y. Chakhchoukh and A. M. Zoubir (2011) [109] and the paper "Direction-of-arrival estimation of FM sources based on robust spatial time-frequency distribution matrices" in IEEE Workshop on Statistical Signal Processing by W. Sharif, Y. Chakhchoukh and A. M. Zoubir (2011) [36].

man-made phenomena [21, 25]. Hence, the estimation is performed in a contaminated environment where the observations contain outliers and deviate from the Gaussian assumption.

In [31, 105, 115], it has been shown that standard time-frequency distributions (TFDs) are severely degraded in the presence of impulsive noise. Since STFD matrices are obtained by computing TFDs across the sensor array, they also suffer from a lack of robustness as well. Consequently, the performance of DOA estimation based on standard STFD matrices degrades significantly. The goal is to propose DOA estimation methods based on robust STFD estimates.

Examples of commonly used existing robust TFD estimation procedures are the fractional-lower-order moment based TFDs in [115] and median based TFDs proposed in [105]. However, their applicability for STFD based DOA estimation has not been investigated so far. Furthermore, these basic techniques do not offer a good theoretical trade-off between efficiency and robustness [37, 38]. These two aspects are treated in this Chapter.

This Chapter presents newly developed robust methods for STFD estimation and performs a comparison of these techniques for DOA estimation of non-stationary signals. The proposed methods follow performant statistical robust procedures [37, 38]and are classified into pre-processing, robust position based and non-iterative robust STFD estimation.

4.1.1 Problem statement

Given N observations $\{\mathbf{x}(t)\}_{t=1}^{N}$ of signal model in Eq. (2.1) in the presence of impulsive noise, the objective is to estimate the DOAs $\theta_k,\ k \in \{1, K\}$ of non-stationary signals described by their IFreqs $f_k(t),\ k \in \{1, K\}$ which are assumed to be known.

4.2 Robust estimation of STFD matrices

For DOA estimation, commonly used STFD matrix is based on pseudo-WVD (PWVD). The $(p, q)^{th}$ element of PWVD based STFD matrix in discrete-time form is [8],

$$D_{x_p x_q}(t, f) = \sum_{l=-L/2}^{L/2} x_p(t + l)x_q^*(t - l)e^{-j4\pi fl} \tag{4.1}$$

It is clear from Eq. (4.1) that the presence of an outlier in the observations $x_p(t)$ or $x_q(t)$ can have an unbounded effect on the estimate of the $(p,q)^{th}$ element of the STFD matrix and, therefore, in the presence of even a small fraction of outlying observations in $\mathbf{x}(t)$, will lead to wrong estimates of DOAs. In order to obtain robust DOA estimates, this work proposes two different approaches based on robust estimation of the STFD matrices. The first approach is based on suppression of outlying observations before standard STFD estimation, i.e., pre-processing while in the second approach STFD matrix is robustly estimated. Figure 4.1 illustrates these two schemes for the problem of direction finding. The following Sections describe different pre-processing and robust STFD estimation procedures and provide simulation results which demonstrate the efficacy of the proposed robust techniques for DOA estimation.

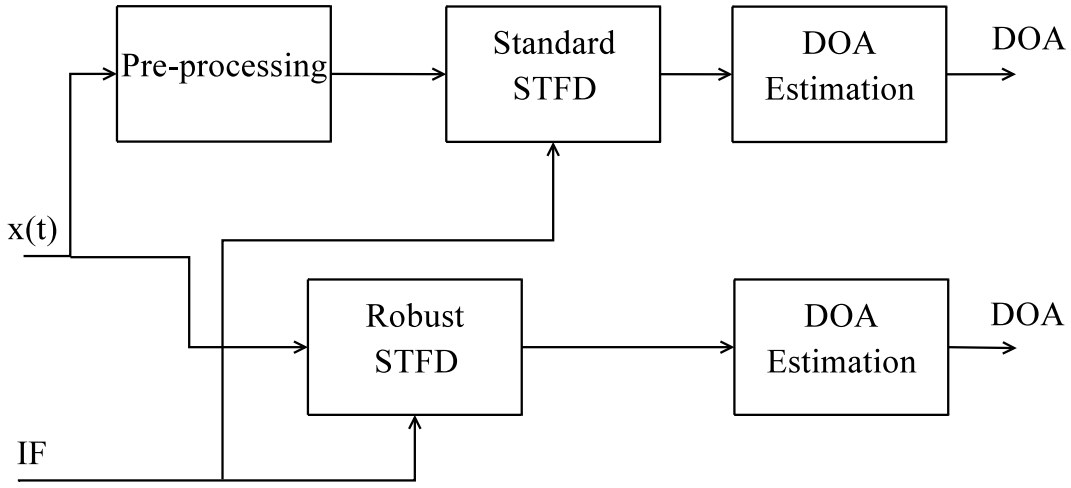

Figure 4.1. Proposed schemes for STFD based DOA estimation in impulsive noise

4.3 Pre-processing based robust estimation

The objective of a pre-processing techniques is to suppress impulsive noise from the observations. Several pre-processing techniques exist which are employed to suppress impulsive interference in many practical applications. The examples of pre-processing techniques which have been particularly applied in conventional robust direction finding can be found in [55,56,116]. The pre-processing techniques can be classified into transformation based and the diagnostic approaches. Transformation based pre-processing techniques use transformations which resist impulsiveness while diagnostic based ap-

proaches require detection and trimming of impulsive contaminated samples. For robust STFD matrices following pre-processing techniques can be used:

4.3.1 Spatial sign function

This transformation based pre-processing technique was first proposed by Visuri et al. in [116], and has been used for direction finding in impulsive noise for conventional array processing. After applying the spatial sign function, the transformed observations are given by

$$\mathbf{x}_{ssf}(t) = \begin{cases} \dfrac{\mathbf{x}(t)}{\|\mathbf{x}(t)\|} & \text{if } \mathbf{x}(t) \neq \mathbf{0} \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{4.2}$$

where $\|\cdot\|$ denotes the Frobenious norm and $\|\mathbf{x}\| = \sqrt{(\mathbf{x}^H\mathbf{x})}$. Then, STFD estimation can be applied on $\mathbf{x}_{ssf}(t)$ instead of $\mathbf{x}(t)$.

4.3.2 Modified modulus transformation

This technique uses modulus transformations to suppress impulsiveness and yields transformed data which resembles more to a normal distributed rather than to a heavy-tailed distribution [55,57]. For a real input observation y, the transformed data is given by

$$y_{mod} = \begin{cases} \text{sign}(y)((|y|+1)^{\lambda} - 1) & \text{if } \lambda \neq 0 \\ \text{sign}(y)(\log_{10}|y| + 1) & \text{if } \lambda = 0 \end{cases} \tag{4.3}$$

where 'sign' denotes the sign function and λ is a parameter. In the context of array processing, observations are complex, therefor, MT is performed separately on the real and the imaginary parts of the observations [55]. Hence, transformed sensor data in the case of $\lambda = 0$ is given by:

$$x_{p,mod}(t) = \begin{cases} \text{sign}(\bar{x}_p)[\log_{10}|\bar{x}_i| + 1] + \cdots & \\ \quad j\text{sign}(\tilde{x}_p)[\log_{10}|\tilde{x}_p| + 1] & \text{if } x_p(t) \neq 0 \\ 0 & \text{if } x_p(t) = 0 \end{cases} \tag{4.4}$$

where $x_p(t)$ denotes the p^{th} sensor data at time t, $\bar{x}_p$ and $\tilde{x}_i$ denote the real and imaginary parts of $x_p(t)$, respectively. This technique can also be interpreted as a compression where input signal is compressed using a log-like function. This compression leads to the attenuation of the influence of large outliers while preserving the phase of the input signal. Moreover, based on the results of [57], it can be concluded that the

MT transformed signal is no longer heavy-tailed rather it approximates more a Gaussian distribution. This means that the standard STFD estimation can be performed, given by

$$\boldsymbol{D}_{mod}(t,f) = \sum_{l=-L/2}^{L/2} \boldsymbol{x}_{mod}(t+l)\boldsymbol{x}_{mod}{}^{H}(t-l)e^{-j4\pi fl} \tag{4.5}$$

where L denotes the window length and $\boldsymbol{x}_{mod}(t) = [x_{1,mod}(t), x_{2,mod}(t), \ldots, x_{m,mod}(t)]^T$. Note that the modulus transformation is an adaptive technique as it neither requires any tuning parameter nor makes any assumption about the data.

4.3.3 3-σ rejection pre-processing

This diagnostic based pre-processing technique detects and trims the contaminated samples. The detection of a contaminated sample is performed through comparison between the magnitude of the observation and 3σ, where σ denotes the scale. A sample with magnitude greater than 3σ is considered impulsive and is discarded. Since, in practice σ is not known and therefore normalized median absolute deviation (MADN) [37] is used as a robust estimate of scale. MADN is defined to be the normalized median of absolute deviations from the median. The normalization is performed in order to obtain a consistent estimate of standard deviation under Gaussian data. Given the observations $\mathbf{y} = \{y_1, y_2, \ldots, y_n\}$ of a random variable, the MADN is given by,

$$\text{MADN}\{\mathbf{y}\} = \frac{\text{Med}\,|\mathbf{y} - \text{Med}(\mathbf{y})|}{0.6745} \tag{4.6}$$

Let $\bar{\mathbf{x}}_p$ and $\tilde{\mathbf{x}}_p$ denote the vectors of the real and the imaginary parts of the p^{th} sensor observations $\mathbf{x}_p \in \mathbb{C}^N$ respectively, then the 3-σ rejection pre-processing is performed as described in Table 4.1.

4.3.4 Normalization technique

This is another transformation based technique in which the influence of outliers is suppressed by applying normalization. Herein, each sensor observation is individually normalized by its magnitude. The observations after applying the normalization are given by

$$y_p(t) = \frac{x_p(t)}{|x_p(t)|} \tag{4.7}$$

where $x_p(t)$ denotes the p^{th} sensor signal at time t. This technique differs from the one in Section 4.3.1 where an array observation $\mathbf{x}(t) \in \mathbb{C}^m$ at time t is normalized by its norm.

Table 4.1. 3-σ rejection pre-processing

1. Compute robust estimate of scale for $\bar{\mathbf{x}}_p$ and $\tilde{\mathbf{x}}_p$, e.g., MADN. Let $\bar{\sigma}_p$ and $\tilde{\sigma}_p$ denote the robust scale estimates of $\bar{\mathbf{x}}_p$ and $\tilde{\mathbf{x}}_p$ respectively.
2. Record the indices's of $\bar{\mathbf{x}}_p$ and $\tilde{\mathbf{x}}_p$ for which $|\bar{\mathbf{x}}_p| > 3\bar{\sigma}_p$ and $|\tilde{\mathbf{x}}_p| > 3\tilde{\sigma}_p$. Let $\bar{\mathbf{v}}_p$ and $\tilde{\mathbf{v}}_p$ denote the indices.
3. Clip the observations in $\mathbf{x}_p$ to 0 that are indexed by $\mathbf{v}_p = \bar{\mathbf{v}}_p \cup \tilde{\mathbf{v}}_p$.
4. Repeat the above steps for all $p \in \{1, m\}$.
5. Apply classical STFD matrix estimation on the new observations.

Table 4.2. Summary of pre-processing techniques for robust STFD matrix estimation with application to non-stationary DOA estimation

Technique	Parameter/properties
Spatial sign function	scaling, parameter free, adaptive
Modulus transformation	compression, λ,
3-σ rejection	parameter free
Normalization	compression, parameter free

Once having the normalized observations, standard techniques for STFD estimation can be applied to get the direction estimates. Table 4.3.4 presents a summary of pre-processing techniques which can be used for STFD estimation with application to DOA estimation of non-stationary signals.

4.4 Robust position based STFD estimation

This section describes robust STFD matrix estimation based on robust estimation of position. Let us consider only the $(p, q)^{th}$ element of the STFD matrix at a given (t, f) TF point. In the case of a PWVD given by

$$\hat{D}_{x_p x_q}(t, f) = \sum_{l=-L/2}^{L/2} x_p(t+l) x_q^*(t-l) e^{-j4\pi fl} \tag{4.8}$$

where $x_p(t+l)$ denotes the p^{th} sensor signal at time $t+l$. Let us denote the summand $x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}$ by $\mathcal{G}_{pq}(t,f,l)$. Furthermore, remark that $\{\mathcal{G}_{pq}(t,f,l),\ l \in \{-L/2, L/2\}\} \in \mathbb{C}^{L+1}$ where $L+1$ is the window length.

4.4.1 Median-based STFD estimation

Robust position based TFD estimation described in [105] is based on the reformulation of TFD estimation as an estimation of position of a complex vector $\{\mathcal{G}_{pq}(t,f,l)\ l \in [-L/2, L/2]\}$ of length $L+1$. The $(p,q)^{th}$ element of STFD matrix is defined as solution to the minimization of following cost function:

$$\hat{D}_{x_p x_q}(t,f) = \arg\min_{D_{pq}} \mathcal{Q}_{pq}(t,f,l,D_{pq}) \tag{4.9}$$

where

$$\mathcal{Q}_{pq} = \sum_{l=-L/2}^{L/2} \rho\left(e_{pq}(t,f,l,D_{pq})\right) \tag{4.10}$$

and $e_{pq}(t,f,l) = \mathcal{G}_{pq}(t,f,l) - D_{pq}$ is the error function depending on (t,f,l) and where $D_{pq} \in \mathbb{C}$ denotes the quantity to be estimated. The standard TFD is obtained by taking $\rho(e_{pq}(t,f,l,D_{pq})) = |e_{pq}(t,f,l,D_{pq})|^2$. The resulting TFD is the mean of the vector $\{\mathcal{G}_{pq}(t,f,l),\ l \in [-L/2,\ldots,L/2]\}$. The estimators based on $|e_{pq}(t,f,l)|^2$ are highly sensitive to outliers and result in biased estimates even in the presence of small fraction of outliers. To obtain a robust estimate, impulsive resistant cost functions must be utilized.

Barkat et al. [105] and Sahmoudi et al. [117] proposed to use the L_1-norm based cost function, i.e., $\rho(e(\cdot)) = |e(\cdot)|$. The estimator corresponding to minimizing an L_1-norm is the median which is highly robust against outliers. Minimization of L_1-norm based cost function requires solving for $\partial\mathcal{Q}_{pq}/\partial D_{pq}^* = 0$ which results in non-linear equations and D_{pq} is obtained iteratively [1]. The iterative procedure requires complex operations since $D_{pq} \in \mathbb{C}$. In [1], the authors also proposed to estimate the real and the imaginary parts of D_{pq} separately. However, separate estimation requires independence of the real and the imaginary parts of $\mathcal{G}_{pq}(t,f,l)$ which does not necessarily exist. To address these issues and to provide more efficient robust estimators for STFD, the following Section proposes to use multi-dimensional position based M-estimation.

4.4.2 M-estimation based STFD

This estimator takes into account the existing correlation between the real and the imaginary parts of $\mathcal{G}_{pq}(t,f,l)$ and treats the complex observations in the sequence

$\{\mathcal{G}_{pq}(t,f,l),\ l \in \{-L/2,\ldots,L/2\}\}$ as vector of stacked real and the imaginary parts, i.e. $\boldsymbol{\mathcal{G}}_{pq}(t,f,l) = [\bar{\mathcal{G}}_{pq}(t,f,l),\ \tilde{\mathcal{G}}_{pq}(t,f,l)]^T$, where $[\cdot]^T$ is the transpose. Let $\hat{\mathbf{d}}_{x_px_q}(t,f) = [\Re(\hat{D}_{x_px_q}(t,f)),\ \Im(\hat{D}_{x_px_q}(t,f))]^T$ denote the vector containing the real and imaginary parts of $\hat{D}_{x_px_q}(t,f)$. Then, STFD based on M-estimation can be written as follows

$$\hat{\mathbf{d}}_{x_px_q}(t,f) = \arg\min_{\mathbf{d}_{pq}} \sum_{l} \rho(d(t,f,l,\mathbf{d}_{pq})) \tag{4.11}$$

where $\mathbf{d}_{pq} = [\Re(D_{pq}),\ \Im(D_{pq})]^T$ denotes the quantity to be estimated and $d(t,f,l,\mathbf{d}_{pq})$ is 2-dimensional Mahalanobis distance given by

$$d(t,f,l,\mathbf{d}_{pq}) = \ (\boldsymbol{\mathcal{G}}_{pq}(t,f,l) - \mathbf{d}_{pq})^T\, \boldsymbol{\Sigma}^{-1}\, (\boldsymbol{\mathcal{G}}_{pq}(t,f,l) - \mathbf{d}_{pq}) \tag{4.12}$$

with $\boldsymbol{\Sigma} \in \mathbb{R}^{2\times 2}$ being the covariance matrix of $\boldsymbol{\mathcal{G}}_{pq}(t,f,l),\ l \in \{-L/2,\ldots,L/2\}$. The $\rho(\cdot)$ function can be chosen in different ways. The most common ones are Huber's ρ-function and Tukey's bi-weight function [38]. Here, we choose Huber's ρ-function given by:

$$\rho_H(d) = \begin{cases} d^2/2 & \text{if } |d| \le c_H \\ c_H|d| - c_H^2/2 & \text{if } |d| > c_H \end{cases} \tag{4.13}$$

where $c_H \in \mathbb{R}_+$ is an appropriately chosen constant and acts as a threshold. It ensures that larger weights are applied to the normal observations as compared to the outliers. It also provides a trade-off between robustness and efficiency [38]. It can either be chosen empirically or can be computed, e.g. based on the distribution of the Mahalanobis' distances. The majority of Mahalanobis' distances follow a χ^2_p distribution, where p denotes the degree of freedom and is equal to the dimension of the observation [118]. with $p = 2$, the threshold c can be set to $\sqrt{\chi^2_{p,0.975}}$, where $\chi^2_{p,0.975}$ denotes the 97.5% quantile of χ^2_p distribution.

4.4.3 S-estimation based STFD

Consider again the classical cost function given by

$$\mathcal{Q}_{pq} = \sum_{l=-L/2}^{L/2} d(t,f,l,\mathbf{d}_{pq})^2 \tag{4.14}$$

where applying STFD estimation means finding a $\mathbf{d}_{pq}$ that minimizes the classical estimate of variance of the sequence $\{d(t,f,l,\mathbf{d}_{pq}), l = -L/2,\ldots,L/2\}$. It can also be interpreted that the classical STFD is obtained by minimizing the classical scale estimate which is non-robust and is highly influenced in the presence of outliers. One way

to achieve robust STFD is by performing minimization on a robust scale instead of the classical scale. The chosen robust scale can be e.g. an M-estimate of scale [38]. Therefore, the objective of S-estimation based STFD is to find $\mathbf{d}_{pq}$ and $\mathbf{\Sigma}$ which minimize the robust M-scale estimate $\hat{\sigma}_M$ of distances $\{d(t,f,l,\mathbf{d}_{pq}(t,f)),\ l \in \{-L/2, L/2\}\}$. This is given by

$$\begin{aligned} &\min\ \hat{\sigma}_M \\ &\quad s.t.\quad \left(\tfrac{1}{L+1}\right)\sum_{l=-L/2}^{L/2}\rho\left(\tfrac{d(t,f,l,\mathbf{d}_{pq})}{\hat{\sigma}_M}\right) = b_0 \end{aligned} \tag{4.15}$$

where $d(t,f,l,\mathbf{d}_{pq}(t,f))$ is defined in Eq. (4.12) and b_0 is an appropriately chosen constant which ensures a certain efficiency under the Gaussian distribution and a certain breakdown point (BP). The BP is defined as the maximum percentage of outliers beyond which the estimator does not give meaningful estimate. Further details on choosing b_0 can be found in [38]. Again, $\rho(\cdot)$ can be either Huber's or Tukey's bi-weight ρ-function. Here, Tukey's bi-weight function is used, given by

$$\rho_T(d) = \begin{cases} 1-[1-(d/c_T)^2]^3 & \text{if } |d| \le c_T \\ 1 & \text{if } |d| > c_T \end{cases} \tag{4.16}$$

where c_T is an appropriately chosen constant that depends on a certain efficiency and a breakdown point. S-estimator provides maximally achievable breakdown point but at this BP it lacks efficiency. For better efficiency and a higher breakdown point, an MM-estimator of STFD can be used which is described in the following Section.

4.4.4 MM-estimation based STFD

MM-estimator of STFD is obtained in two steps. Firstly, an initial estimate is obtained based on a very robust estimator, such as an S-estimator. Then, given the S-estimate of STFD, the MM-estimate of STFD is obtained by computing an estimator which is more efficient such as an M-estimator, given by

$$D_{x_p x_q}(t,f) = \min_{\boldsymbol{d}_{pq},\mathbf{\Sigma}} \sum_{l=-L/2}^{L/2} \rho_1\left(\frac{d(t,f,l,\mathbf{d}_{pq})}{\hat{\sigma_M}}\right) \tag{4.17}$$

where $\rho_1(\cdot)$ is a ρ-function and is tuned to obtain an efficient estimate. MM-estimator inherits the breakdown properties of an S-estimator and an efficiency of an M-estimator [68]. However, MM-estimator is computationally more demanding than all of the above estimators. This estimator is mentioned here for the sake of completeness and is not considered in the simulation section. Table 4.4.4 provides a brief summary of robust position based STFD estimators.

Table 4.3. Summary of robust position based STFD matrix estimation with application to non-stationary DOA estimation

Methods	Remarks
M-estimation	General framework for STFD estimation
Median-based STFD	A special case of M-estimation
S-estimation based STFD	A highly robust estimator
MM-estimation based STFD	Highly robust and efficient estimator

4.5 Non-iterative robust STFD estimation

The robust position based methods are iterative and require a significant amount of computational effort for the estimation of STFDs. This Section deals with simpler non-iterative robust STFD procedures. It describes an existing and presents newly proposed robust non-iterative STFD estimation techniques. The proposed techniques are (i) STFD estimation based on robust covariance (ii) STFD estimation based on weighted mean and (iii) STFD estimation based on robust distances.

4.5.1 Fractional-lower-order-moment based STFD

Fractional-lower-order-moment based estimate of $(p,q)^{th}$ element of STFD matrix at a given TF point (t,f) is given by [115]:

$$\hat{D}_{x_p x_q}(t,f) = \sum_l x_p^{<\alpha>}(t+l) x_q^{-<\alpha>}(t-l) e^{-j4\pi fl} \tag{4.18}$$

where $x_p^{<\alpha>} = |x_p|^{\alpha+1}/x_p^*$ and $0 < \alpha < 1$. From Eq. 4.18, it is evident that the FLOM operator influences only the magnitude of x while phase remains unchanged. It is required to tune the value of α to get robust estimates. Also it should be noted that for $\alpha = 1$, it yields the normalized TFD.

4.5.2 Robust covariance based STFD estimation

Robust estimation of covariance is a well investigated problem in the statistical literature [37, 38]. For robust STFD estimation, the idea is to define the $(p,q)^{th}$ element of STFD matrix $\hat{D}_{x_p x_q}(t,f)$ as a function of covariance between pairwise sequences of

the real and the imaginary parts of x_p and x_q. Consider again the standard estimate of $(p,q)^{th}$ element of STFD matrix, given by

$$\hat{D}_{x_p x_q}(t,f) = \sum_{l=-L/2}^{L/2} x_p(t+l)x_q^*(t-l)e^{-j4\pi fl} \tag{4.19}$$

Let us define:
$\mathbf{x}_p = \{x_p(t+l),\ l \in \{-L/2,\ldots,L/2\}\}$,
$\mathbf{y}_q = \{x_q^*(t-l)e^{-j4\pi fl},\ l \in \{-L/2,\ldots,L/2\}\}$.
where $\mathbf{x}_p, \mathbf{y}_q \in \mathbb{C}^{L+1}$. Let $\bar{\mathbf{x}}_p$, $\bar{\mathbf{y}}_q$ and $\tilde{\mathbf{x}}_p$, $\tilde{\mathbf{y}}_q$ denote the real and imaginary parts of $\mathbf{x}_p$ and $\mathbf{y}_q$, respectively. Then (4.19) can also be written as

$$\begin{aligned}\hat{D}_{x_p x_q}(t,f) &= \sum_{l=-L/2}^{L/2} \bar{\mathbf{x}}_p(l)\bar{\mathbf{y}}_q(l) - \sum_{l=-L/2}^{L/2} \tilde{\mathbf{x}}_p(l)\tilde{\mathbf{y}}_q(l) \\ &+ j\sum_{l=-L/2}^{L/2} \bar{\mathbf{x}}_p(l)\tilde{\mathbf{y}}_q(l) - j\sum_{l=-L/2}^{L/2} \tilde{\mathbf{x}}_p(l)\bar{\mathbf{y}}_q(l)\end{aligned} \tag{4.20}$$

where $\mathbf{x}_p(l)$ denotes the l^{th} element of the vector $\mathbf{x}_p$. $\bar{\mathbf{x}}_p$, $\tilde{\mathbf{x}}_p$, $\bar{\mathbf{y}}_q$ and $\tilde{\mathbf{y}}_q$ can also be interpreted as realizations of random variables a_p, b_p, c_q and d_q, respectively. Then each of the summations in (4.20) is the covariance of the corresponding two random variables. Hence (4.20) can also be written as follows

$$\hat{D}_{x_p x_q}(t,f) = \mathrm{Cov}(a_p,c_q) - \mathrm{Cov}(b_p,d_q) + j(\mathrm{Cov}(a_p,d_q) + \mathrm{Cov}(b_p,c_q)) \tag{4.21}$$

where $\mathrm{Cov}(\cdot)$ denotes the covariance operator. A robust estimate of the STFD matrix can be obtained by replacing these non-robust covariance operators by their robust counterparts. On this concept, two simple methods to compute robust STFDs are proposed. The first approach is based on a robust estimate of the scale. For example, the robustified first term $\mathrm{Cov}(a_p,c_q)$ in Eq. (4.21) can be written as [38]:

$$\begin{aligned}\mathrm{RCov}(a_p,c_q) = \frac{\hat{\sigma}_r(a_p)\hat{\sigma}_r(c_q)}{4}\Bigg(&\hat{\sigma}_r\left(\frac{a_p}{\hat{\sigma}_r(a_p)} + \frac{c_q}{\hat{\sigma}_r(c_q)}\right)^2 - \\ &\hat{\sigma}_r\left(\frac{a_p}{\hat{\sigma}_r(a_p)} - \frac{c_q}{\hat{\sigma}_r(c_q)}\right)^2\Bigg)\end{aligned} \tag{4.22}$$

where $\hat{\sigma}_r(\cdot)$ denotes the robust estimator of the scale. Herein, MADN is used as robust estimate of scale. The second robust approach uses such a function $\psi(\cdot)$ which suppresses outliers. For this kind of robust covariance, the first term in Eq. (4.21), for example, can be written as [38]:

$$\mathrm{RCov}(a_p,c_q) = \hat{\sigma}_r(a_p)\hat{\sigma}_r(c_q)\sum_l \left[\psi\left(\frac{a_p(l)-\hat{\mu}_r(a_p)}{\hat{\sigma}_r(a_p)}\right)\psi\left(\frac{c_q(l)-\hat{\mu}_r(c_q)}{\hat{\sigma}_r(c_q)}\right)\right] \tag{4.23}$$

where ψ is a bounded monotone or re-descending $\psi(\cdot)$ function, $\hat{\sigma}_r$ and $\hat{\mu}_r$ are robust estimates of scale and location, respectively. This approach uses $\psi(\cdot) = \text{sgn}(\cdot)$, median and MADN as robust estimates of the location and the scale, respectively. For better clarity, the proposed algorithm to compute the STFD matrix based on robust covariance is summarized in Table 4.4.

Table 4.4. Robust covariance based STFD estimation

For a given (t, f) point and for a certain (p, q):

1. Compute robust estimates of the scale and the location for each of the vectors a_p, b_p, c_q and d_q. MADN and median can be used as robust estimates of the scale and the location, respectively.
2. Compute the robust estimates of each of the covariance in Eq. (4.21) by using either Eq. (4.22) or Eq. (4.23).
3. Compute final estimate of $(p, q)^{th}$ element of STFD matrix by using Eq. (4.21).
4. Repeat Steps (i) to (iii) for all $(t, f) \in \mathcal{S}$ and all $(p, q) \in \{1, m\}$ where $\mathcal{S}$ is the set of all the TF points belonging to a single source.

4.5.3 Weighted STFD based on robust distances

This approach employs detection and weighting of the outliers for robust estimation. The outliers are detected based on their distances from the bulk of the data. The distances are measured based on m-dimensional Mahalanobis' distance (MD), given by

$$\widehat{\text{MD}}(\boldsymbol{x}(t)) = \sqrt{(\boldsymbol{x}(t) - \hat{\boldsymbol{\mu}})^H \hat{\boldsymbol{R}}^{-1} (\boldsymbol{x}(t) - \hat{\boldsymbol{\mu}})} \tag{4.24}$$

where $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{\mu}}$ denote the sample covariance matrix and the sample mean of the observations $\{\boldsymbol{x}(t)\}_{t=1}^{N}$, respectively. As the sample mean and the sample covariance matrix are non-robust, the outliers may not have large distances. This effect is known as the *masking effect* on the classical Mahalanobis distance. Alternate is to use robust distance instead, i.e., distance measures which are based on robust estimators of both the mean and the covariance. A suitable choice for robust mean and robust covariance is the minimum covariance determinant (MCD) estimator which has good statistical

and outlier rejection capabilities [119]. Given N observations, the MCD estimator of mean and covariance selects $h < N$ such observations which minimize the determinant of the covariance matrix [119]. The details about MCD estimator can be found in Chapter 2. Here, a fast algorithm for MCD called FAST-MCD is applied [69]. Let $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{C}}$ be the MCD estimate of the mean and the covariance, respectively. Then the robust distance (RD) is given by

$$\widehat{\mathrm{RD}}(\boldsymbol{x}(t)) = \sqrt{(\boldsymbol{x}(t) - \hat{\boldsymbol{\eta}})^H \hat{\boldsymbol{C}}^{-1} (\boldsymbol{x}(t) - \hat{\boldsymbol{\eta}})} \tag{4.25}$$

The outlying observations can be easily identified now as they have larger distances than the distances obtained for the normal observations. Figure 4.2 plots the Mahalanobis

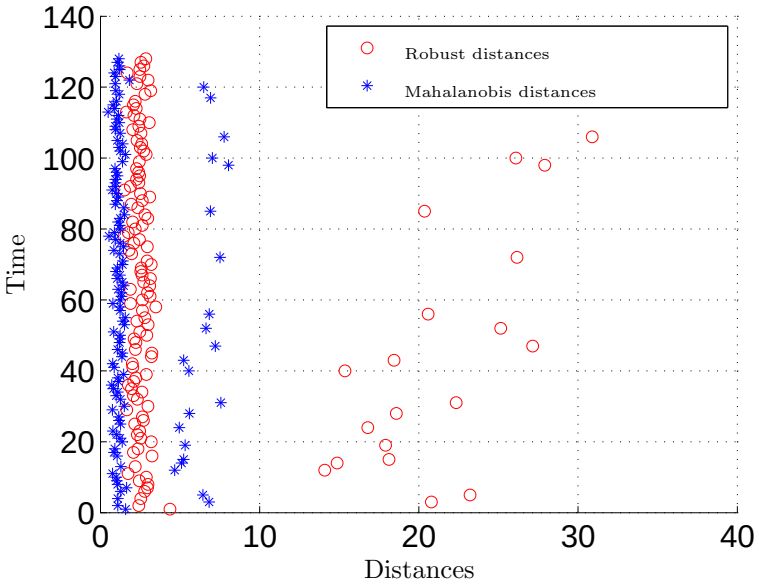

Figure 4.2. Robust distance and Mahalanobis distances

and robust distances where it is evident that with robust distances, it is much easier to discriminate the normal observations from the outliers. The computed robust distances are then used to assign weights to the observations. To obtain a simple and adaptive procedure which is free of tuning parameters, it is proposed to use the inverse of robust distances as weights to the observations during computation of STFD estimate, thus yielding

$$\hat{\boldsymbol{D}}_{\mathrm{MCD}}(t,f) = \sum_{l=-L/2}^{L/2-1} \frac{\boldsymbol{x}(t+l)}{\widehat{\mathrm{RD}}(\boldsymbol{x}(t+l))} \frac{\boldsymbol{x}^H(t-l)}{\widehat{\mathrm{RD}}(\boldsymbol{x}(t-l))} e^{-j4\pi fl} \tag{4.26}$$

4.5.4 Weighted mean based STFD

As described in Section 3.2, an averaged STFD matrix computed over TF points belonging to a single source is required for DOA estimation. The $(p,q)^{th}$ element of such an averaged STFD matrix can also be written as:

$$\hat{\mathbf{d}}_{x_p x_q} = \frac{1}{\#\mathcal{S}} \sum_{(t_n,f_n)\in\mathcal{S}} \sum_{l=-L/2}^{L/2} \boldsymbol{\mathcal{G}}_{pq}(t_n, f_n, l) \tag{4.27}$$

where $\mathcal{S}$ is the set containing all (t,f) points which belong to a single source, $\hat{\boldsymbol{d}}_{x_p x_q} = [\Re(\hat{D}_{x_p x_q}),\ \Im(\hat{D}_{x_p x_q})]^T$ and $\boldsymbol{\mathcal{G}}_{pq}(t_n, f_n, l)$ is as defined in Section 4.4.2. It is remarked from simulations that the elements $\boldsymbol{\mathcal{G}}_{pq}(t_n, f_n, \cdots)$ for different values of $(t_n,\ f_n) \in \mathcal{S}$ are identically distributed when constant amplitude FM signals are considered. Therefore, instead of computing the inner summation in Eq. (4.27) for different $(t_n,\ f_n)$ separately and then averaging, it is proposed to compute the double summation in (4.27) simultaneously instead. Let $(t_1,f_1),\ldots,(t_Z,f_Z) \in \mathcal{S}$ denote the TF points belonging to a single source and Z be the total number of TF points. Furthermore, let

$$\begin{aligned}\mathbf{G}_{pq} = [&\boldsymbol{\mathcal{G}}_{pq}(t_1,f_1,-L/2),\ldots,\boldsymbol{\mathcal{G}}_{pq}(t_1,f_1,L/2),\\ &\boldsymbol{\mathcal{G}}_{pq}(t_2,f_2,-L/2),\ldots,\boldsymbol{\mathcal{G}}_{pq}(t_2,f_2,L/2),\ldots,\\ &\boldsymbol{\mathcal{G}}_{pq}(t_Z,f_Z,-L/2),\ldots,\boldsymbol{\mathcal{G}}_{pq}(t_Z,f_Z,L/2)]^T\end{aligned} \tag{4.28}$$

with $\mathbf{G}_{pq} \in \mathbb{R}^{Z(L+1)\times 2}$ denoting the matrix of all the summands in Eq. (4.27). Then, Eq.(4.27) can also be written in the following form

$$\hat{\boldsymbol{d}}_{x_p x_q} = \frac{1}{Z} \sum_{z=1}^{Z\times(L+1)} \mathbf{G}_{pq}(z) \tag{4.29}$$

The summation in Eq. (4.29) is not robust in the presence of outliers. In order to achieve robustness, weighted mean can be used. The weights are computed based on the outlyingness measures of the elements of $\mathbf{G}_{pq}$. The outlying measures can be determined by using the Mahalanobis distance measure given by

$$d(z) = (\mathbf{G}_{pq}(z) - \mathrm{med}(\mathbf{G}_{pq}))^T {\boldsymbol{\Sigma}_G}^{-1} (\mathbf{G}_{pq}(z) - \mathrm{med}(\mathbf{G}_{pq})) \tag{4.30}$$

where $\boldsymbol{\Sigma}_G \in \mathbb{R}^{2\times 2}$ is the robust covariance estimate of $\mathbf{G}_{pq}$. The weights in case of Tukey's bi-weight function are:

$$w(d) = \begin{cases} \rho_T(d)/d^2 & \text{if } d \neq 0 \\ 6/c_T^2 & \text{if } d = 0 \end{cases} \tag{4.31}$$

where $c_T > 0$ is an appropriately chosen constant. Let $w_z,\ z \in \{1, Z\times(L+1)\}$ denote the weight computed for each of the vector in Eq. (4.28). Then weighted STFD

estimate $\hat{\mathbf{d}}_{x_p x_q}$ is given by:

$$\hat{\boldsymbol{d}}_{x_p x_q} = \frac{\sum_{z=1}^{Z\times(L+1)} w_z \mathbf{G}_{pq}(z)}{\sum_{z=1}^{Z\times(L+1)} w_z} \tag{4.32}$$

This approach is computationally simple and fast as compared to the iterative robust position based techniques. The results in the simulation section show the effectiveness of this approach. The algorithm to compute weighted mean based STFD estimate is summarized in Table 4.5. Since this approach uses a single iteration and for this

Table 4.5. Weighted mean STFD

1. For a given (p, q), compute the matrix $\mathbf{G}_{pq}$ as in Eq. (4.28)
2. Estimate median, robust covariance $\boldsymbol{\Sigma}_{\boldsymbol{G}}$ of $\mathbf{G}_{pq}$ and Mahalanobis distance in Eq. (4.30) for each of the element in $\mathbf{G}_{pq}$.
3. Compute weights for each element of $\mathbf{G}_{pq}$ based on Tukey's bi-weights as defined in Eq. (4.31).
4. Finally compute weighted mean $\hat{\mathbf{d}}_{x_p x_q}$ as in Eq. (4.32).
5. Repeat above steps for all $p, q \in \{1, m\}$.

reason can also be called as one step re-weighted STFD estimators. The following Section provides simulation examples and discussions on the results obtained by using above described robust estimation approaches.

4.6 Simulation results

The simulation setup consists of a uniform linear array (ULA) of 8 sensors and two FM sources impinging from angles $\theta = [-5^\circ\ 4^\circ]$ from broadside. One source is a linear FM signal and the other one is a hyperbolic FM signal. A total of N equal to 128 snapshots are used. Impulsive noise is generated using an ϵ-contaminated Gaussian mixture, given by

$$n_p(t) \sim (1-\epsilon)\mathcal{N}_C(0, \sigma^2) + \epsilon\mathcal{N}_C(0, \kappa\sigma^2) \tag{4.33}$$

A window length $L+1$ of 37 is used for the computation of the STFD matrix. The robust position based methods require fixing some tuning parameters. For the S-

estimator, the required parameters $b_0,\ k$ depend on the desired efficiency under Gaussian noise and the BP of the estimator. Therefore b_0 is chosen to tune efficiency to 95% and the BP of 30%. The parameter c_H for M-estimation is chosen to be equal to 2. The robust covariance based techniques do not require any tuning parameter. The weighted mean based approach requires adjusting the parameter c_T which is equal to 5 in this simulation. For FLOM STFDs, $\alpha = 0.3$ is used.

In the first simulations scenario, $\kappa = 75$ and $\epsilon = 0.2$ are used for the model in Eq. (4.33). The snapshots are generated with a varying signal-to-noise ratio (SNR). STFD matrices are computed using the standard and the proposed techniques, namely pre-processing, robust position and non-iterative robust techniques. The resulting averaged STFD matrices are then used to estimate the DOAs using the MUSIC algorithm [44]. The root mean square error (RMSE) of the estimated DOA for the first source is plotted against the SNR for the pre-processing, the robust position based and the non-iterative robust STFD estimators. Table 4.6 summarizes the robust STFD estimators which are assessed in these simulations.

Table 4.6. Robust methods for STFD matrices estimation.

Pre-proc.	Position based	Non-iterative
Spatial Sign	Median	FLOM
3-σ	M-estimation	Rob. Cov.
Normalization	S-estimation	Weighted mean
Modulus transformation	-	-

Figure 4.3 depicts the RMSE of pre-processing techniques against SNR. The spatial sign function yields biased estimates when the SNR increases. The normalization and the 3-σ methods yield similar results with an increasing SNR. The pre-processing techniques provide a lower RMSE than the one achieved by the standard STFD estimation technique. The results show that the computationally simple pre-processing techniques are effective in suppressing the impulsive interference.

Figure 4.4 shows comparison of different robust position based estimation techniques. The proposed M-estimation and S-estimation based techniques provide lower RMSE if compared with the already existing iterative median based technique. The robust position based techniques are computationally expensive.

Figure 4.5 shows the RMSE of non-iterative robust techniques under impulsive noise. The proposed techniques are the robust covariance and weighted mean based approaches. The robust covariance based STFD and weighted mean approach provide

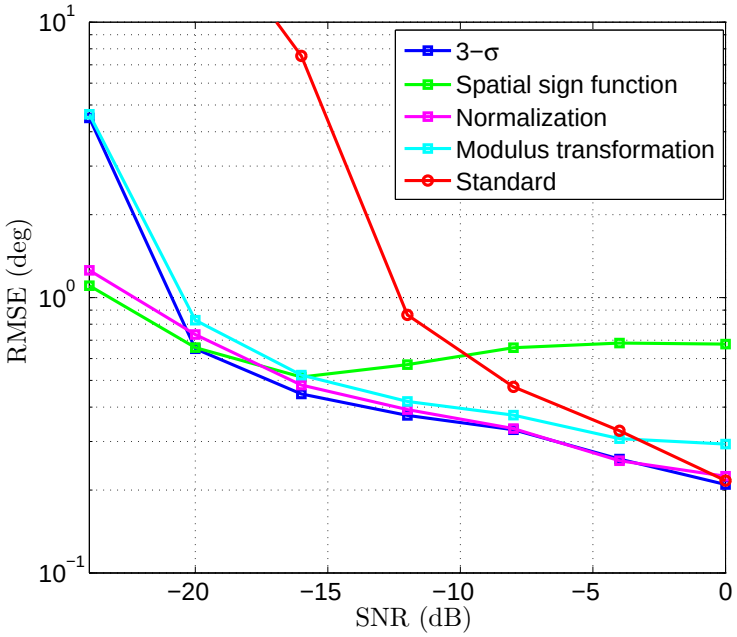

Figure 4.3. Pre processing based techniques under impulsive noise

better RMSE characteristics if compared with the FLOM based approach. The simple weighted mean approach provides the lowest RMSE of all the non-iterative robust techniques. These techniques are computationally simpler than the robust position based estimation techniques.

In the second simulation, the proposed estimators are compared in scenario when the noise is perfectly Gaussian, i.e. there are no outliers. Figures 4.6, 4.7 and 4.8 show the results of RMSE vs. SNR for the pre-processing, the robust position based and the non-iterative robust estimation techniques, respectively.

In the case of the pre-processing technique, the pre-processed data gives a higher RMSE when compared with the standard estimation technique. The spatial sign covariance function provides a higher bias as the SNR increases. The 3-σ rejection rule yields DOA estimates comparable to the estimates obtained by standard techniques. In the case of robust position based estimates, all three techniques provide better estimates when compared with the estimates based on standard techniques. This is due to the fact that the noise of the STFD becomes a mixture of Gaussian and Laplacian even though the noise in the observations is Gaussian [31]. This means that robust methods are suitable even if there is no contamination in the signal. This leads to a performance gain for M-estimator and S-estimator compared to the standard and median based estimators. The median based technique lacks efficiency in the Gaussian case. For the

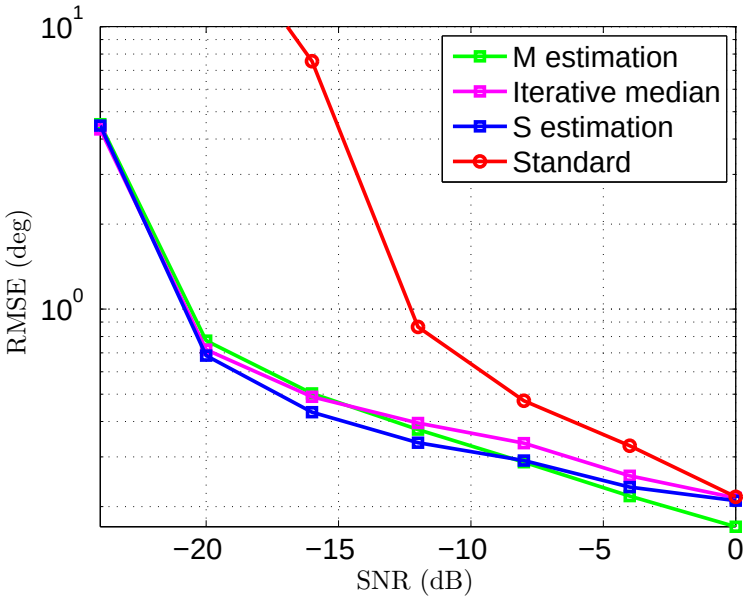

Figure 4.4. Robust position based estimation techniques under impulsive noise

non-iterative robust estimation, the resulting RMSE is lower than the one achieved by the standard techniques for low to moderate SNR values. The proposed weighted mean approach provides the lowest RMSE among the non-iterative techniques. Again CRB under Gaussian noise derived in [45] is also included.

In the third simulation, the sensitivity of different estimators is compared with respect to the percentage of outliers. Figure 4.9 shows the resulting RMSE vs. ϵ. It can be seen that the standard method departs from the true estimates even for a very small fraction of outliers. The FLOM based method starts worsening afterwards. For the robust position based estimation, the S-estimation based STFD and median based STFDs provide a reasonable RMSE for a contamination of up to 40%. The M-estimation based STFD yields correct estimates of DOA for only 30% of contamination. For the non-iterative robust estimation, the robust covariance based estimation techniques could handle 30% of outliers. The weighted-mean based STFD is able to provide good estimates of DOA even up to 40% of outliers. The 3-σ rejection pre-processing departed from the true estimates with 30% of outliers. Hence, most of the proposed robust estimators of STFD are able to handle up to 30% of contamination.

In the fourth simulation, the sensitivity of the estimators with respect to the impulsiveness factor κ is considered. From Figure 4.10, it is evident that all the proposed robust estimators provide a good impulsive rejection over a large range of impulsiveness.

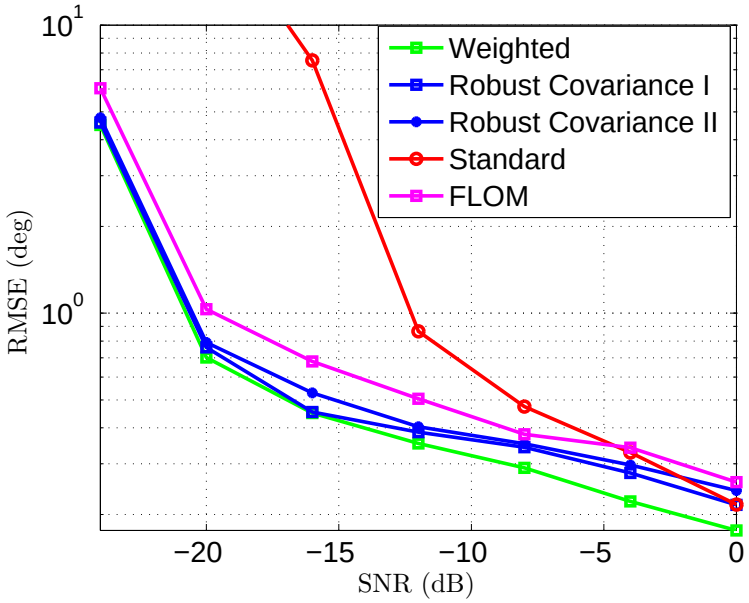

Figure 4.5. Non-iterative robust STFD estimation techniques under impulsive noise

From the above mentioned results, it is clear that the different proposed robust approaches provide comparable results at lower SNR values. Hence, pre-processing is a suitable and constitutes an effective technique for impulsive noise suppression in low SNR regimes. While in the case of low to moderate SNR values, the proposed weighted mean algorithm is a better alternative than pre-processing techniques.

4.7 Summary

This chapter addressed the problem of STFD estimation for frequency modulated (FM) sources in the presence of impulsive noise. The focus was on two different methodologies, namely the pre-processing and robust methods. Different pre-processing procedures which are computationally simple and effective in suppressing the impulsive interference were proposed. In the robust position based estimators, this work extended the existing position based estimators by applying robust statistical techniques, such as, M-estimation and S-estimation. The proposed position based techniques provide an opportunity to improve the performance of DOA estimation under Gaussian and non-Gaussian noise alike. The position based estimators are computationally expensive in general. Therefore, simpler and non-iterative robust estimation techniques were proposed. The proposed non-iterative STFD estimation techniques outperform the existing non-iterative robust techniques. The sensitivity of the proposed estimators

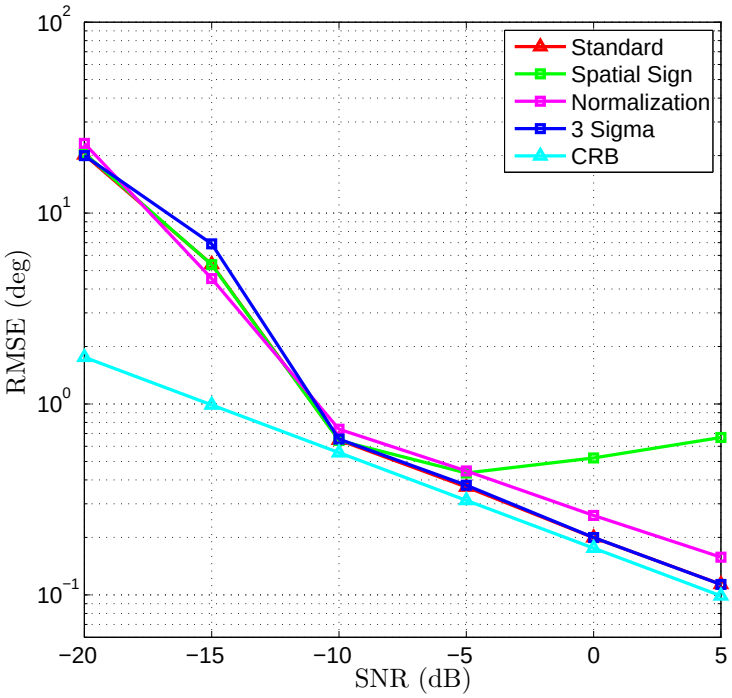

Figure 4.6. Robust pre-processing STFD based DOA estimation techniques under Gaussian noise

against the percentage of contamination and against varying impulsiveness was also considered. The proposed estimators yield reasonable results up to a contamination of 30% and all the estimators show good impulsive rejection capability.

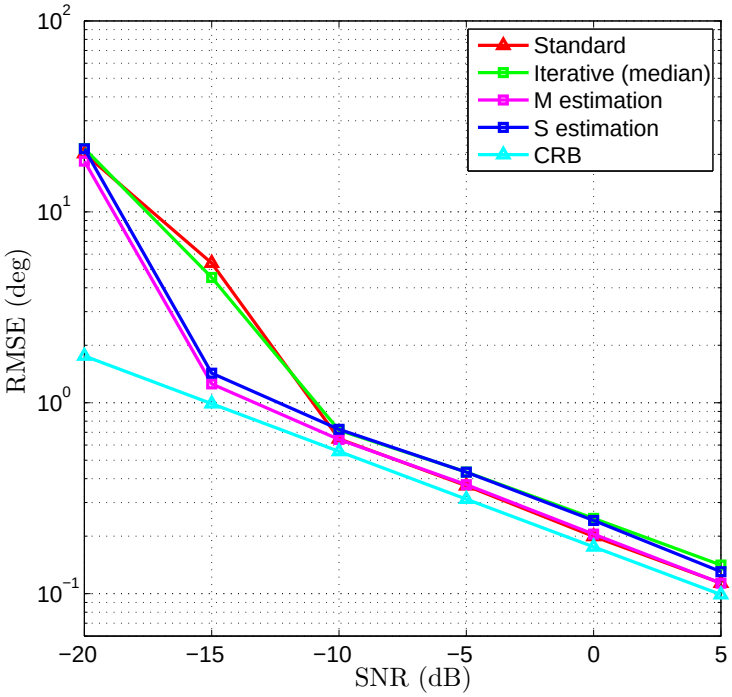

Figure 4.7. Robust position based STFD estimation techniques under Gaussian noise

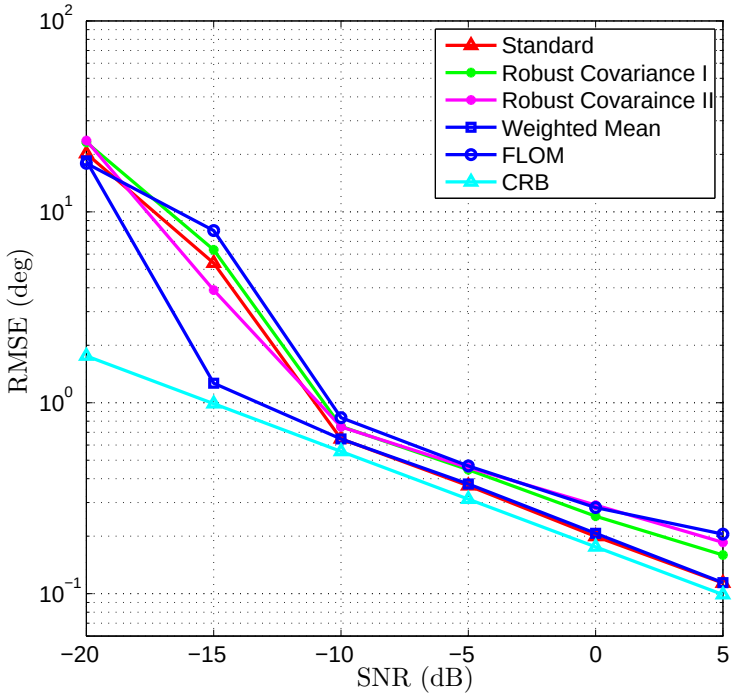

Figure 4.8. Non-iterative robust STFD estimation techniques under Gaussian noise

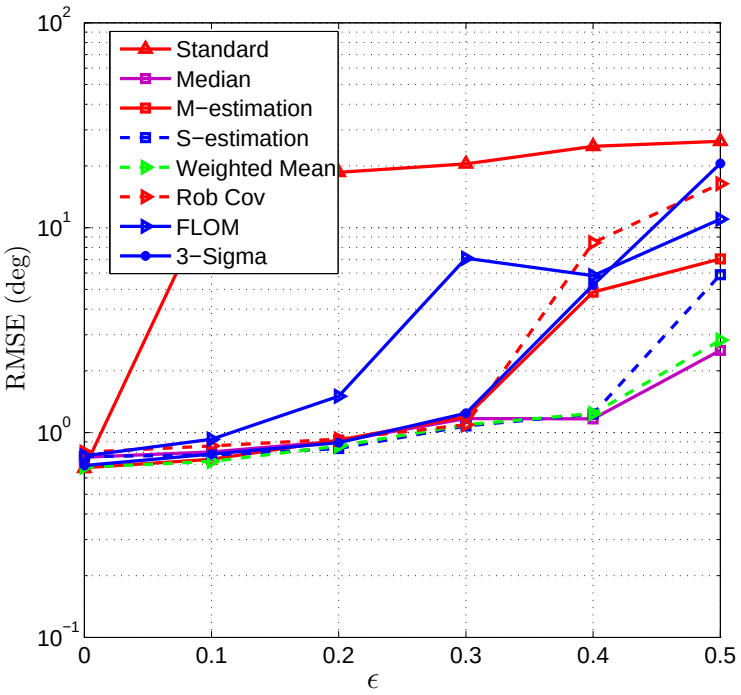

Figure 4.9. Sensitivity with respect to percentage of outliers at an SNR of $-10dB$ and with varying percentage of outliers and $\kappa = [5\ 10\ 20\ 40\ 50\ 60\ 70]$.

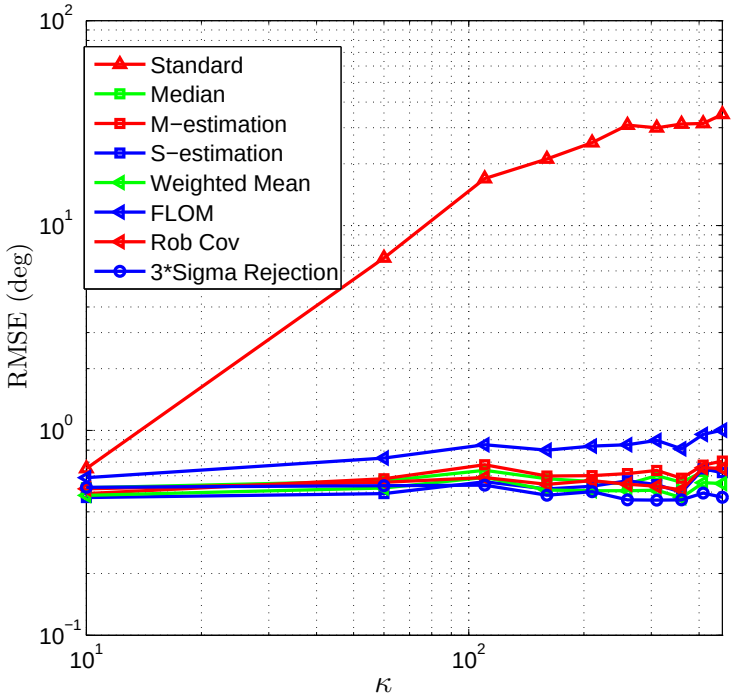

Figure 4.10. Sensitivity with respect to impulsiveness with an SNR of $-5dB$ and $\epsilon = 0.2$.

Chapter 5

Robustness Analysis

[1] This Chapter provides a framework for robustness analysis of spatial time-frequency distributions (STFDs). Section 5.1 provides an introduction and problem formulation for the robustness analysis. Section 5.2 introduces influence function with the help of an easy example and provides definition for the influence function of STFDs. Section 5.3 is dedicated to the robustness analysis of STFD matrix estimators. Simulation results are presented in Section 5.4 while Section 5.5 concludes this Chapter with a short summary.

5.1 Introduction

Robustness analysis is the measure of resistance of an estimator to the deviations from the assumed model. Robust statistics provide three different tools to study the robustness of an estimator. These tools are the breakdown point, the maximum bias curve and the influence function. Breakdown point signifies the maximum percentage of contamination for which the estimator remains bounded [38]. Asymptotic maximum bias is a measure of asymptotic bias of an estimator at a given percentage of contamination while the influence function measures the robustness of an estimator to infinitesimal deviations from the assumed model. It describes the bias impact on the estimator, which is caused by an infinitesimal contamination at an arbitrary point, standardized by the fraction of contamination [121, 122].

The infinitesimal contamination is well motivated in array signal processing applications, due to the transient nature of the impulsiveness. Usually, the energy of impulsive noise is much higher than that of the background (thermal) noise and it is only concentrated in short periods of time. For array covariance matrices, the influence function has been derived in [123] and robustness analysis of the scatter matrix based MVDR beamformer was performed in [124]. For STFD matrices, the claim that a certain STFD estimators is non-robust or robust is merely based on the simulations. To date, a formal robustness analysis of STFD estimators does not exist. Since

[1] This Chapter is based on the article "Robustness analysis of spatial time-frequency distributions on the basis of the influence function" in IEEE Transactions on Signal Processing by W. Sharif, M. Muma and A. M. Zoubir (2013) [120].

STFD matrices have been extensively used in array processing for non-stationary signals [9–11,14–16,35,36,75–77,109,114], therefore, robustness analysis of STFD matrices is highly desired.

The objective of this Chapter is to fill this gap and analyze robustness by means of the influence function [37, 121, 122], which is a useful tool to measure the sensitivity of an estimator to deviations from the distributional assumptions. The influence function not only assures that the STFD estimator perform optimally under the exact theoretical assumptions under which it was derived, but also it ensures that the estimator will work well in a more general case that the assumptions are only approximately fulfilled, which is often the case when dealing with real world problems. The influence function allows to compare the robustness of different STFD estimators and helps to provide answers to the questions, such as: how robust is an estimator? What is the effect of an impulsive disturbance on the bias of the estimator? What is the loss in efficiency at a nominal model when using a robust estimator and more importantly which estimator should one use in a certain scenario? This Chapter answers the aforementioned questions and develops a generally applicable framework of robustness analysis for STFD matrix estimators based on the influence function. Particularly, this Chapter provides influence function and the finite sample version of the influence function which is also called sensitivity curve or the empirical influence function for STFD matrix estimators. It also provides analytical derivations of influence functions for the standard and some recently proposed estimators which have been claimed to be robust. The robustness of recently proposed estimators is proved by the boundedness of influence function. This Chapter also provides array processing examples for the asymptotic and empirical influence functions in case of a uniform linear array observing an FM source.

5.1.1 Objective:

Given the signal model as described in Section 2.2, the goal is to perform the robustness analysis of STFD matrices on the basis of influence function, i.e, to provide a definition of influence function for the STFD matrices, to derive analytical expressions for the influence functions of different kinds of STFD matrix estimators and to provide the definition and computation of finite sample counterpart of influence function. Section 5.2 provides a definition of the influence function and explains the concept of influence function with the help of a toy example of multivariate mean.

5.2 Robustness analysis: the influence function

The asymptotic influence function is basically defined as the first derivative of the functional version of an estimator $\hat{D}$ at a nominal distribution F

$$\mathrm{IF}(z;\hat{D},F)=\lim_{\epsilon\to 0}\frac{\hat{D}_\infty(F_\epsilon)-\hat{D}_\infty(F)}{\epsilon}=\left[\frac{\partial\hat{D}_\infty(F_\epsilon)}{\partial\epsilon}\right]_{\epsilon\downarrow 0}.$$

Here, $\hat{D}_\infty(F)$ and $\hat{D}_\infty(F_\epsilon)$ are the asymptotic values of the estimator when the data is distributed following, respectively, F and the contaminated distribution F_ϵ given by

$$F_\epsilon=(1-\epsilon)F+\epsilon\delta_z, \tag{5.1}$$

with δ_z being the point-mass probability on z and ϵ is the fraction of contamination. To illustrate the concept of the influence function, consider the example of estimating the location $\boldsymbol{\mu}=(\mu_1,\mu_2)^{\mathrm{T}}$ of a real-valued bi-variate Gaussian random variable $\boldsymbol{x}\in\mathbb{R}^2$ with covariance matrix $\boldsymbol{\Sigma}=\sigma\boldsymbol{I}$. Figure 5.1 (top) shows the first coordinate of the asymptotic influence function for the sample mean which is the (optimal) maximum likelihood estimator for this case. It can be seen that the influence function for the sample mean increases linearly with the magnitude of the contamination, which confirms its non-robustness, since outliers have an unbounded effect on the bias of the estimator. The bottom plot depicts the influence function of a robust multivariate location M-estimator [38], which uses Huber's ρ-function (soft limiter) with clipping parameter $c_H=3$. Let $d(\boldsymbol{x})$ be the squared distance given by $d(\boldsymbol{x})=(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$. The influence function of the robust estimator increases linearly with the square root of $d(\boldsymbol{x})$ until the clipping point of $c_H=3$. For a standard normal distribution, $d(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ and therefore, for values $z_1\leq c_H$ and $z_2=0$, the influence functions of both estimators are identical. The soft limiting operation of the robust estimator can best be seen at $\sqrt{d(\boldsymbol{x})}>c_H$ and $z_2=0$: Here, the first coordinate of the influence function takes the constant value of $c_H=3$. Boundedness of the influence function ensures that a small fraction of contamination, e.g., due to impulsive noise, can only have a limited effect on the bias of the estimate. Boundedness and continuity of the influence function together ensure qualitative robustness of the estimator [121]. Furthermore, efficiency of the estimators can be computed by deriving the asymptotic variances analytically from the influence functions. The definition for the influence function of STFD matrix estimator is provided in the following Section.

5.2.1 The influence function of STFD matrix estimators

For the $(p,q)^{th}$ element of the STFD matrix, the influence function is defined as follows:

$$\begin{aligned} \mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), F) &= \left. \frac{\mathrm{E}_{F_\epsilon}\left[\hat{D}_{x_p x_q}(t,f)\right] - \mathrm{E}_F\left[\hat{D}_{x_p x_q}(t,f)\right]}{\epsilon} \right|_{\epsilon \downarrow 0} \\ &= \left. \frac{\partial \mathrm{E}_{\mathcal{F}_\epsilon}\left[\hat{D}_{x_p x_q}(t,f)\right]}{\partial \epsilon} \right|_{\epsilon \downarrow 0} \end{aligned} \tag{5.2}$$

where $\hat{D}_{x_p x_q}(t,f)$ denotes the estimator of the $(p,q)^{th}$ element of the STFD matrix, z_p and z_q are the contaminations in x_p and x_q, respectively, $\mathrm{E}_F[\cdot]$ is the expectation operator under the distribution F. For the signal model defined in Eq. (2.2), F is complex circular Gaussian $\mathcal{N}_{\mathbb{C}}(\mu(t), \sigma^2)$ with variance σ^2 and time varying mean $\mu(t)$ and F_ϵ is the contaminated distribution defined in Eq. (5.1). The following Section introduces the empirical influence function for STFD matrices which is the finite sample counterpart of the influence function in Eq. (5.2).

5.2.2 The empirical influence function of STFD matrix estimators

The empirical influence function for a STFD matrix estimator is obtained by replacing the statistical expectations in Eq. (5.2) by their Monte-Carlo estimates. Let $\bar{D}_{x_p x_q}(t,l)$ be an estimate of $\mathrm{E}_F\left[\hat{D}_{x_p x_q}(t,f)\right]$ obtained by averaging Monte-Carlo realizations of $\hat{D}_{x_p x_q}(t,f)$ based on the signal model defined in Eq. (2.2). Further, let $\bar{D}_{x_p x_q}(t,l,z_p,z_q)$ be the Monte-Carlo estimate of $\mathrm{E}_{\mathcal{F}_\epsilon}\left[\hat{D}_{x_p x_q}(t,f)\right]$ which is obtained by adding single point outliers z_p and z_q to $x_p(t+l)$ and $x_q(t-l)$ with $l \in \{-L/2, \ldots, L/2\}$, respectively. The corresponding empirical influence function is then given by:

$$\mathrm{EIF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), F) = \frac{\bar{D}_{x_p x_q}(t,l,z_p,z_q) - \bar{D}_{x_p x_q}(t,l)}{1/L + 1} \tag{5.3}$$

where L is the window length. The influence functions for different estimators of the STFD matrix is derived next.

5.3 Derivation of the influence function for STFD matrix estimators

As described in Chapter 4, robust STFD estimators are classified into three different types. In the following Sections, one example from each robust class of STFD estimators is considered for robustness analysis. The influence function for the remaining estimators can be derived in an analogous manner.

5.3.1 Derivation of the influence function of the PWVD based STFD matrix estimator

The $(p,q)^{th}$ element of STFD matrix estimator (defined in Section 3.2) based on PWVD is:

$$\hat{D}_{x_p x_q}(t,f) = \sum_{l=-L/2}^{L/2} x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}, \tag{5.4}$$

where L denotes the window length. Since the PWVD is the most commonly used STFD matrix estimator, in the sequel, it is referred as the standard STFD estimator. To derive the influence function of the standard STFD matrix estimator based on the PWVD, it is required to evaluate the expectation of $\hat{D}_{x_p x_q}(t,f)$ in Eq. (5.4) w.r.t. F_ϵ, which can be written as

$$\mathrm{E}_{\mathcal{F}_\epsilon}\left[\hat{D}_{x_p x_q}(t,f)\right] = \sum_{l=-L/2}^{L/2} \mathrm{E}_{\mathcal{F}_\epsilon}\left[x_p(t+l)x_q^*(t-l)\right]e^{-j4\pi fl}. \tag{5.5}$$

Considering the sensor signal model of Eq. (2.2) and the definition of the influence function in Eq. (5.2), it gives

$$\begin{aligned}
\mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), F) \;=\; & -2\sum_{l=-L/2}^{L/2}\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)s_k^*(t-l)a_q^*(\theta_k)e^{-j4\pi fl} \\
& + z_q^* \sum_{l=-L/2}^{L/2}\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl} \\
& + z_p \sum_{l=-L/2}^{L/2}\sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t-l)e^{-j4\pi fl}.
\end{aligned} \tag{5.6}$$

The derivation is given in Appendix 5.6. The influence function for the auto-sensor TFD is obtained by following the same steps, which leads to

$$\begin{aligned}\mathrm{IF}(z_p;\hat{D}_{x_px_p},F) &= -2\sigma^2 - 2\sum_{l=-L/2}^{L/2}\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)s_k^*(t-l)a_p^*(\theta_k)e^{-j4\pi fl} \\ &\quad + z_p^*\sum_{l=-L/2}^{L/2}\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl} \\ &\quad + z_p\sum_{l=-L/2}^{L/2}\sum_{k=1}^{K} a_p^*(\theta_k)s_k^*(t-l)e^{-j4\pi fl}. \end{aligned} \tag{5.7}$$

The first term in Eq. (5.6) is only dependent on the signal parameters while in Eq. (5.7) the first part depends on the signal and the Gaussian noise parameters. The second and the third terms in Eqs. (5.6) and (5.7) are linearly proportional to the magnitude of the contaminations z_q and z_p, respectively. An example for $p = q = 1$ is shown in Figure 5.2. The influence function is unbounded which means that outliers have an unbounded effect on the bias of $\hat{D}_{x_px_q}(t,f)$. This analytically confirms the non-robustness of standard STFD matrix estimators.

5.3.2 Derivation of the influence function for the 3-σ rejection Pre-processing based STFD matrix estimator

The 3-σ rejection pre-processing based robust estimator belongs to the class of pre-processing based robust estimators. The $(p,q)^{th}$ element of the STFD matrix estimator based on 3-σ rejection pre-processing can also be described as:

$$\begin{aligned}\hat{D}_{x_px_q}(t,f) = \sum_{l=-L/2}^{L/2}\Big(&\psi\left(\bar{x}_p(t+l)\right)\mathbb{1}_{(\tilde{x}_p(t+l))}\psi(\bar{x}_q(t-l))\,\mathbb{1}_{(\tilde{x}_q(t-l))} \\ &+ \quad \psi(\tilde{x}_p(t+l))\,\mathbb{1}_{(\bar{x}_p(t+l))}\psi(\tilde{x}_q(t-l))\,\mathbb{1}_{(\bar{x}_q(t-l))} \\ &+ j\left(\psi(\tilde{x}_p(t+l))\,\mathbb{1}_{(\bar{x}_p(t+l))}\psi(\bar{x}_q(t-l))\,\mathbb{1}_{(\tilde{x}_q(t-l))}\right. \\ &\left.- \psi\left(\bar{x}_p(t+l)\right)\mathbb{1}_{(\tilde{x}_p(t+l))}\psi(\tilde{x}_q(t-l))\,\mathbb{1}_{(\bar{x}_q(t-l))}\right)\Big)e^{-j4\pi fl}. \end{aligned} \tag{5.8}$$

Here, $\bar{x}_p$, $\bar{x}_q$ and $\tilde{x}_p$, $\tilde{x}_q$ denote the real and imaginary parts of x_p and x_q, respectively. The function $\psi(\cdot)$ performs the 3-σ trimming operation

$$\psi\left(x\right) = \begin{cases} x & \text{if } |x| < 3\hat{\sigma}_x(t) \\ 0 & \text{otherwise} \end{cases} \tag{5.9}$$

where $x \in \mathbb{R}$ denotes the observation of a univariate random variable with corresponding robust scale estimate $\hat{\sigma}_x(t)$ [37]. $\mathbb{1}(\cdot)$ is the indicator function which is defined as

$$\mathbb{1}(x) = \begin{cases} 1 & \text{if } |x| < 3\hat{\sigma}_x(t) \\ 0 & \text{otherwise.} \end{cases} \tag{5.10}$$

When the real part is trimmed by the trimming operation defined in Eq. (5.9), the imaginary part has to be limited and this is obtained by using the indicator function in Eq. (5.10). For better readability, let us define the following abbreviations:

$$\begin{aligned} \mathcal{I}_1(t,l) &:= \psi(\bar{x}_p(t+l))\,\mathbb{1}_{(\tilde{x}_p(t+l))}\psi(\bar{x}_q(t-l))\,\mathbb{1}_{(\tilde{x}_q(t-l))} \\ \mathcal{I}_2(t,l) &:= \psi(\tilde{x}_p(t+l))\,\mathbb{1}_{(\bar{x}_p(t+l))}\psi(\tilde{x}_q(t-l))\,\mathbb{1}_{(\bar{x}_q(t-l))} \\ \mathcal{I}_3(t,l) &:= \psi(\tilde{x}_p(t+l))\,\mathbb{1}_{(\bar{x}_p(t+l))}\psi(\bar{x}_q(t-l))\,\mathbb{1}_{(\tilde{x}_q(t-l))} \\ \mathcal{I}_4(t,l) &:= \psi(\bar{x}_p(t+l))\,\mathbb{1}_{(\tilde{x}_p(t+l))}\psi(\tilde{x}_q(t-l))\,\mathbb{1}_{(\bar{x}_q(t-l))}. \end{aligned} \tag{5.11}$$

In this way, the influence function of the 3-σ rejection pre-processing based estimator can be written as function of the influence functions for the four terms defined in Eq. (5.11). It can be written as:

$$\begin{aligned} \mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), F) = \sum_{l=-L/2}^{L/2} \big(&\mathrm{IF}(z_p, z_q; \mathcal{I}_1(t,l), F) + \mathrm{IF}(z_p, z_q; \mathcal{I}_2(t,l), F) \\ &+ \mathrm{IF}(z_p, z_q; \mathcal{I}_3(t,l), F) + \mathrm{IF}(z_p, z_q; \mathcal{I}_4(t,l), F)\big) e^{-j4\pi fl} \end{aligned} \tag{5.12}$$

The components of the influence function in Eq. (5.12) are derived in Appendix 5.7 (see Eq. (5.31)). Similarly, the influence function for the auto-sensor element $\hat{D}_{x_p x_p}(t,f)$ of the STFD matrix is given by:

$$\begin{aligned} \mathrm{IF}(z_p; \hat{D}_{x_p x_p}(t,f), F) = \sum_{l=-L/2,\ l\neq 0}^{L/2} \Bigg(\big(&\mathrm{IF}(z_p; \mathcal{I}_1(t,l), F) + \mathrm{IF}(z_p; \mathcal{I}_2(t,l), F) + \mathrm{IF}(z_p; \mathcal{I}_3(t,l), F) \\ &+ \mathrm{IF}(z_p; \mathcal{I}_4(t,l), F)\big) e^{-j4\pi fl} \Bigg) + \mathrm{IF}(z_p; \hat{D}_{x_p x_p}(t,f)\big|_{l=0}, F) \end{aligned} \tag{5.13}$$

From Eq. (5.31), it is evident that for a given signal component and noise variance σ^2, the first term in (5.31) is constant while the other four terms suppress the influence of the outlier and become zero if the outlier increases beyond the 3σ threshold. An example of this type of influence function is shown in Figure 5.3.

5.3.3 Derivation of the influence function for the M-estimation based STFD matrix estimator

M-estimator of STFD is an example of the robust position-based STFD estimators and the $(p,q)^{th}$ element of the STFD matrix based on the M-estimators is solution to the

minimization of the following cost function:

$$\hat{D}_{x_p x_q}(t,f) = \arg\min_{D} \sum_{l=-L/2}^{L/2} \rho\left(d(t,l)\right) \tag{5.14}$$

Here, $d(t,l)$ is the distance, which is given by

$$d(t,l) = \frac{x_p(t+l)x_q^*(t-l)e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)}, \tag{5.15}$$

where $\hat{\sigma}_M(t)$ is the M-estimate of scale [37] of the numerator of Eq. (5.15) for $l = -L/2, \ldots, L/2$, and ρ is a function which ensures robustness if it is chosen in such a way that it leaves the 'good' data untouched and bounds the influence of an outlier. A monotone ρ, ensures convexity of the estimation problem, e.g. Huber's ρ-function [37] is defined as

$$\rho\left(d(t,l)\right) = \begin{cases} |d(t,l)|^2/2 & \text{if } |d(t,l)| < c_H, \\ c_H(|d(t,l)| - c_H/2) & \text{if } |d(t,l)| \geq c_H; \end{cases} \tag{5.16}$$

where c_H is the parameter which determines the threshold to weigh the normal and the outlying observations differently. For details on how to compute $\hat{\sigma}_M(t)$ and how to choose c_H, the reader is referred to [109]. The solution to the minimization of the cost function in Eq. (5.14) can be computed by a gradient descent based approach:

$$\frac{\partial \hat{D}_{pq}(t,f)}{\partial D} = \sum_{l=-L/2}^{L/2} \Psi\left(d(t,l)\right) = 0 \tag{5.17}$$

In the case of Huber's ρ-function given in Eq. (5.16), we obtain the following $\Psi(\cdot)$ function

$$\Psi\left(d(t,l)\right) = \begin{cases} d(t,l) & \text{if } |d(t,l)| < c_H, \\ c_H \operatorname{sign}\left(d(t,l)\right) & \text{if } |d(t,l)| \geq c_H. \end{cases} \tag{5.18}$$

For the complex-valued distance $d(t,l) \in \mathbb{C}$ the $\operatorname{sign}(d(t,l))$ is defined as

$$\operatorname{sign}(d(t,l)) = \frac{d(t,l)}{|d(t,l)|}. \tag{5.19}$$

The solution to Eq. (5.14) when using a monotone ρ, can be computed e.g. by finding the root of Eq. (5.17) with iteratively re-weighted least squares. The derivation of the influence function is given in the Appendix 5.8. The resulting expression for a general

Ψ is:

$$\mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), \mathrm{F}) = \Bigg(2 \sum_{l=-L/2}^{L/2} \mathrm{E}_F \left[\Psi \left(d(t,l) \right) \right] - \sum_{l=-L/2}^{L/2} \mathrm{E}_{\mathrm{F}(x_p)} \left[\Psi \left(d(t,l,z_q) \right) \right] - \sum_{l=-L/2}^{L/2} \left(\mathrm{E}_{\mathrm{F}(x_q)} \left[\Psi \left(d(t,l,z_p) \right) \right] - \mathrm{E}_F \left[\Psi' \left(d(t,l) \right) \frac{\partial d(t,l)}{\partial \hat{\sigma}} \right] \mathrm{IF}(z_p, z_q; \hat{\sigma}_M, \mathrm{F}) \right) \Bigg) \times \left(\sum_{l=-L/2}^{L/2} \mathrm{E}_F \left[\Psi' \left(d(t,l) \right) \frac{\partial d(t,l)}{\partial D} \right] \right)^{-1} \tag{5.20}$$

Here, $\mathrm{E}_{\mathrm{F}(x_p)}[\cdot]$ and $\mathrm{E}_{\mathrm{F}(x_q)}[\cdot]$ denote the expectations w.r.t. the distribution of x_p and x_q. $d(t,l,z_q)$ and $d(t,l,z_p)$ are the distances obtained by setting $x_q = z_q$ and $x_p = z_p$, respectively, in Eq. (5.15). The computation of statistical expectations in Eq. (5.20) for the example of Huber's ρ function is given in the Appendix 5.8 and lead to the final expression given in Eq. (5.46). It can be seen for that all terms in Eq. (5.46) that include z_p or z_q, downweighting of the outliers is performed by means of the indicator functions in the expectations in Eq. (5.45) for $|d(t,l,z)| > c_H$, which ensures a bounded influence function for $z_{p,q} \to \infty$. Figure 5.4 depicts the influence function of an auto-sensor element of the STFD matrix estimator. It can be seen that the influence function is bounded and the outliers influence decreases with increasing magnitude.

5.3.4 Derivation of the influence function for the one-step reweighted STFD matrix estimator

This estimator is a member of non-iterative class of robust STFD estimators and it weights each observation according to its distance from the center of the data. Herein, the $(p,q)^{th}$ element of the STFD matrix estimator is given by

$$\hat{D}_{x_p x_q}(t,f) = \sum_{l=-L/2}^{L/2} w(\hat{d}(l)) x_p(t+l) x_q^*(t-l) e^{-j4\pi fl}. \tag{5.21}$$

Here, $\hat{d}(l)$ is an estimate of the distance, given by

$$d(t,l) = \frac{x_p(t+l)x_q^*(t-l)e^{-j4\pi fl} - \hat{\mu}_W(t)}{\hat{\sigma}_W(t)}, \tag{5.22}$$

where $\hat{\sigma}_W(t)$ is a robust scale estimate, e.g. the normalized median absolute deviation [122] of the numerator of Eq. (5.22), and $\hat{\eta}(t)$ is a robust location estimate, e.g, the

median of $x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}$ for $l \in \{-L/2, \ldots, L/2\}$. A common choice for $w(\cdot)$ is Tukey's bi-weight function for which the weights are defined as [38]:

$$w(d(l)) = \begin{cases} (1 - |d(l)|^2/c_T^2)^2 & \text{if } |d(l)| \leq c_T \\ 0 & \text{otherwise} \end{cases} \tag{5.23}$$

where c_T denotes the cut-off threshold. The influence function of one-step re-weighted estimator of the STFD matrix is defined derived in the Appendix 5.9 and is given by the following expression:

$$\begin{aligned}
\text{IF}\left(z_p, z_q; \hat{D}_{x_p x_q}(t,f), \text{F}\right) =& \sum_{l=-L/2}^{L/2} \Bigg(\frac{1}{\hat{\sigma}_W^2(t)} \Bigg(\left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right)^2 \mathcal{E}_{11} \\
&+ \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right)^2 \mathcal{E}_{12} + \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l)^{-j4\pi fl} \right)^2 \mathcal{E}_{13} + \mathcal{E}_{14} \Bigg) \text{IF}\left(z_p, z_q; \hat{\sigma}_W(t), \text{F}\right) \\
&+ \frac{1}{\hat{\sigma}_W^2(t)} \Bigg(\left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right) \mathcal{E}_{21} \\
&+ \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right) \mathcal{E}_{22} + \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l)^{-j4\pi fl} \right) \mathcal{E}_{23} + \mathcal{E}_{24} \Bigg) \text{IF}\left(z_p, z_q; \hat{\mu}_W(t), \text{F}\right) \\
&- 2 \Bigg(\left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right) \mathcal{E}_{31} + \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right) \mathcal{E}_{32} \\
&+ \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l)^{-j4\pi fl} \right) \mathcal{E}_{33} + \mathcal{E}_{34} \Bigg) + z_q^* e^{-j4\pi fl} \mathcal{E}_{41} + \mathcal{E}_{42} + z_p e^{-j4\pi fl} \mathcal{E}_{51} + \mathcal{E}_{52} \Bigg)
\end{aligned} \tag{5.24}$$

Here, $\mathcal{E}$'s are the statistical expectations as defined in Eq. (5.56). Figure 5.5 shows the influence function of the one-step re-weighted estimator for $p = q = 1$. It can be seen for that all terms in Eq. (5.24) that include z_p or z_q, zero weight is given to the outliers by means of the indicator functions in the expectations in Eq. (5.56) for $|d(t,l,z)| > c_T$, which ensures a bounded and re-descending influence function for $z_{p,q} \to \infty$.

5.4 Simulation results

In the simulations, in the first example, we first use the array signal model defined in Eq. (2.2) for a sensor array of $m = 2$ sensors in a uniform linear array (ULA) geometry observing a single linear FM source ($K = 1$) from the broadside angle of $\theta_1 = -5°$. The instantaneous frequency f_1 of the linear FM source varies from the normalized frequency value of 0.1 to 0.3. An SNR of $-5dB$ is used for plotting the influence function for different STFD estimators. In the case of two sensors, the single

point outlier is $\boldsymbol{z} = [z_1, z_2]^T$, where $z_1, z_2 \in \mathbb{C}$ are complex-valued. In the following, the influence function for the $p = 1,\ q = 1$ auto-sensor TFD is considered and the range of z_1 is $\bar{z}_1, \tilde{z}_1 \in [-50, 50]$. The window length $L = 25$ is used for the computation of the PWVD based STFD matrix. We divide our results into two parts. In the first part, the analytical influence functions are plotted for the above discussed estimators, while in the second part we depict their empirical influence functions. We also give results for a sensor array of $m = 4$ sensors. For $m = 4$, the single point outlier is $\boldsymbol{z} = [z_1, z_2, z_3, z_4]^T$, where $z_1, z_2, z_3, z_4 \in \mathbb{C}$ are complex-valued. All other parameters are as in the previous simulations. Computing the IFs for a $m > 2$ is analogous to $m = 2$, however the influence functions can not be plotted in a 3-D space.

5.4.1 The influence functions for STFD matrix estimators

Figure 5.2 shows the influence function of the standard STFD matrix estimator. The influence function increases linearly with increasing magnitude of contamination meaning that the influence of outliers is unbounded. In Figure 5.3, the influence function for the 3-σ rejection pre-processing based robust STFD matrix estimator is plotted. It is clearly visible that the 3-σ rejection pre-processing operation yields a bounded and constant influence for contaminations beyond the 3-σ threshold, which confirms its robustness. It can also be noticed that the influence function is monotone, thus giving limited but non-zero weights to extreme outliers. Similarly, the influence function of the M-estimation based STFD matrix estimator is plotted in Figure 5.4. The influence of the contamination decreases with an increasing magnitude of the outlier. The largest peak of the influence function for the M-estimator of the STFD matrix is obtained for the case when the contamination $\bar{z}_1 = \tilde{z}_1 = 0$. This means that the uncontaminated data in the observations is given the largest influence on the estimate. For visual clarity, this peak is not plotted in Figure 5.4.

Figure 5.5 shows the influence function for the one-step re-weighted estimator. It is clear from the Figure that again the influence is bounded and re-descends to zero for large contamination. As with the M-estimator, largest influence is given to non-contaminated data.

5.4.2 The empirical influence function for the STFD matrix estimators

This Section contains simulation results for the finite sample counterpart of the influence function, i.e., the empirical influence function (EIF) as defined in Eq. (5.3).

The statistical expectations are estimated by the sample means of the STFDs which are obtained by averaging the STFD estimates from 1000 Monte-Carlo runs. All other simulation parameters are chosen as detailed in the beginning of this Section.

Figure 5.6 shows that the empirical influence for the standard STFD matrix estimator increases with an increase in the magnitude of the contamination and which confirms the finding based on the analytical influence function (see Figure 5.2). Figure 5.7 depicts the EIF for the 3-σ rejection pre-processing based STFD matrix estimator, clearly the influence function does not increase as the magnitude of the outlier increases. The EIF of the M-estimation based STFD matrix estimator is plotted in Figure 5.8. Again, the influence function is bounded. Similarly, the one-step weighted STFD matrix also leads to a bounded EIF and the contamination with large magnitude have nearly zero influence on the final estimate, see Figure 5.9.

5.4.3 The influence functions for STFD matrix estimators for m=4 sensors

Figure 5.10 plots the L_2 norm of influence functions as a function of the outlier magnitude κ. For $m = 4$ sensors, again, the influence function of the standard estimator increases linearly, while it is bounded for all robust estimators. Again, the weighted and M-estimators give highest influence to clean data.

5.5 Summary

This Chapter has provided a framework for a robustness analysis of STFDs. A definition for the influence function of the STFD matrix estimator was given and the analytical expressions for the influence functions for four different types of STFD matrix estimators were provided. Although herein, mainly the robust STFDs were treated, this analysis can also be applied to any type of quadratic TFDs. The influence functions for the robust estimators are bounded and continuous which confirms their qualitative robustness. In addition to the asymptotic analysis, definition for the finite sample counterpart of the influence function also called the EIF or the sensitivity curve was provided . The simulation results for the EIF confirm the analytical results and show the insensitivity to small departures in the distributional assumptions for the robust techniques. The robust location and the one-step re-weighted STFD matrix estimators have a smoother EIF and better bound the influence of outliers in the finite sample case than the 3-σ rejection pre-processing based STFD matrix estimators.

5.6 Appendix I: Derivation of influence function for the PWVD based STFD estimator

To compute the influence function, it is necessary to first evaluate the expectation $\mathrm{E}_{\mathcal{F}_\epsilon}\left[x_p(t+l)x_q^*(t-l)\right]$, which is given by

$$\mathrm{E}_{\mathcal{F}_\epsilon}\left[x_p(t+l)x_q^*(t-l)\right] = \int x_p(t+l)x_q^*(t-l)\mathrm{d}F_\epsilon(x_p(t+l), x_q^*(t-l)), \tag{5.25}$$

where $F_\epsilon(x_p(t+l), x_q^*(t-l))$ denotes the joint contaminated distribution of $x_p(t+l)$ and $x_q^*(t-l)$. For better readability, in the sequel, we abbreviate $F_\epsilon(x_p(t+l), x_q^*(t-l))$ with $F_\epsilon(x_p, x_q)$. Since $x_p(t+l)$ and $x_q(t-l)$ are independently and identically distributed, the joint distribution of $x_p(t+l)x^*(t-l)$ is

$$\begin{aligned} F_\epsilon(x_p, x_q) &= F_\epsilon(x_p)F_\epsilon(x_q) \\ &= (1-\epsilon)^2F(x_p)F(x_q) + (1-\epsilon)F(x_p)\epsilon\delta_{z_q} + \epsilon\delta_{z_p}(1-\epsilon)F(x_q) + \epsilon^2\delta_{z_p}\delta_{z_q}. \end{aligned} \tag{5.26}$$

Here $F(x_p)$ and $F(x_q)$ in Eq. (5.27) are the nominal distributions of $x_p(t+l)$ and $x_q(t-l)$, respectively. Taking the derivative of (5.25) w.r.t. ϵ and with $\epsilon \downarrow 0$ yields:

$$\begin{aligned} \left.\frac{\partial \mathrm{E}_{\mathcal{F}_\epsilon}\left[x_p(t+l)x_q^*(t-l)\right]}{\partial\epsilon}\right|_{\epsilon\downarrow 0} &= \mathrm{E}_{F(x_p)}\left[x_p(t+l)\right]z_q^* - 2\mathrm{E}_{F(x_p)}\left[x_p(t+l)\right]\mathrm{E}_{F(x_q)}\left[x_q^*(t-l)\right] \\ &\quad + \mathrm{E}_{F(x_q)}\left[x_q^*(t-l)\right]z_p \end{aligned} \tag{5.27}$$

Based on the signal model in Eq. (2.2), $F(x_p)$ and $F(x_q)$ are given by

$$F(x_p) = \mathcal{N}_{\mathbb{C}}(\mu_p(t+l), \sigma^2), \quad F(x_q) = \mathcal{N}_{\mathbb{C}}(\mu_q(t-l), \sigma^2),$$

respectively. Here, $\mathcal{N}_{\mathbb{C}}(\mu(t), \sigma^2)$ is a circular complex Gaussian distribution with time-varying mean $\mu(t)$ and variance σ^2. The means at the p^{th} and the q^{th} sensors are given by:

$$\mu_p(t+l) = \sum_{k=1}^{K} a_p(\theta_k)s_k(t+l), \quad \mu_q(t-l) = \sum_{k=1}^{K} a_q(\theta_k)s_k(t-l) \tag{5.28}$$

With Eq. (5.28), we obtain

$$\mathrm{E}_{F(x_p)}\left[x_p(t+l)\right] = \sum_{k=1}^{K} a_p(\theta_k)s_k(t+l), \quad \mathrm{E}_{F(x_q)}\left[x_q^*(t-l)\right] = \sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t-l)$$

and combining Eq. (5.5) and (5.27) leads to the influence function in Eq. (5.6).

5.7 Appendix II: Derivation of the influence function for the 3-σ rejection pre-processing based robust STFD matrix estimator

In the following, we derive the influence function expression for the term $\mathcal{I}_1$ in Eq. (5.11). The expressions for $\mathcal{I}_i,\ i \in \{2,3,4\}$ can be derived in a similar way. Starting from Eq. (5.11), in order to derive the influence function, we first need to compute the expectation of the $\mathcal{I}_1$ w.r.t. the contaminated distribution F_ϵ. Since $\bar{x}_p$, $\tilde{x}_p$, $\bar{x}_q$ and $\tilde{x}_q$ are independently distributed, the joint contaminated distribution $F_\epsilon(\bar{x}_p, \tilde{x}_p, \bar{x}_q, \tilde{x}_q)$ is given by:

$$F_\epsilon(\bar{x}_p, \tilde{x}_p, \bar{x}_q, \tilde{x}_q) = F_\epsilon(\bar{x}_p)F_\epsilon(\tilde{x}_p)F_\epsilon(\bar{x}_q)F_\epsilon(\tilde{x}_q) \tag{5.29}$$

Combining (5.11) and (5.29) leads to:

$$\begin{aligned}\mathrm{E}_{\mathcal{F}_\epsilon}\left[\mathcal{I}_1(t,l)\right] = \quad & \int\int\int\int \psi\left(\bar{x}_p(t+l)\right) \mathbb{1}_{(\tilde{x}_p(t+l))} \psi\left(\bar{x}_q(t-l)\right) \mathbb{1}_{(\tilde{x}_q(t-l))} \times \\ & \mathrm{d}F_\epsilon(\bar{x}_p)\mathrm{d}F_\epsilon(\tilde{x}_p)\mathrm{d}F_\epsilon(\bar{x}_q)\mathrm{d}F_\epsilon(\tilde{x}_q)\end{aligned} \tag{5.30}$$

Inserting (5.30) in (5.2) gives:

$$\begin{aligned}\mathrm{IF}(z_p, z_q; \mathcal{I}_1(t,l), F) \ &= \ \left.\frac{\partial \mathrm{E}_{\mathcal{F}_\epsilon}\left[\mathcal{I}_1(t,l)\right]}{\partial\epsilon}\right|_{\epsilon\downarrow 0} \\ &= \mathrm{E}\left[\left.\frac{\partial\left[\psi\left(\bar{x}_p(t+l)\right) \mathbb{1}_{(\tilde{x}_p(t+l))} \psi\left(\bar{x}_q(t-l)\right) \mathbb{1}_{(\tilde{x}_q(t-l))}\right]}{\partial\epsilon}\right|_{\epsilon\downarrow 0}\right] \\ &\quad -4\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right] \\ &\quad +\psi\left(\bar{z}_p\right)\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right] \\ &\quad +\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathbb{1}_{(\tilde{z}_p)}\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right] \\ &\quad +\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\psi\left(\bar{z}_q\right)\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right] \\ &\quad +\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathbb{1}_{(\tilde{z}_q)}\end{aligned} \tag{5.31}$$

The first term in the above Eq., when further evaluated yields the following:

$$\begin{aligned}&\frac{\partial\left[\psi\left(\bar{x}_p(t+l)\right) \mathbb{1}_{(\tilde{x}_p(t+l))} \psi\left(\bar{x}_q(t-l)\right) \mathbb{1}_{(\tilde{x}_q(t-l))}\right]}{\partial\epsilon} = \\ &\quad \frac{\partial\psi\left(\bar{x}_p(t+l)\right)}{\partial\epsilon}\mathbb{1}_{(\tilde{x}_p(t+l))}\psi\left(\bar{x}_q(t-l)\right)\mathbb{1}_{(\tilde{x}_q(t-l))} \\ &\quad +\psi\left(\bar{x}_p(t+l)\right)\frac{\partial\mathbb{1}_{(\tilde{x}_p(t+l))}}{\partial\epsilon}\psi\left(\bar{x}_q(t-l)\right)\mathbb{1}_{(\tilde{x}_q(t-l))} \\ &\quad +\psi\left(\bar{x}_p(t+l)\right)\mathbb{1}_{(\tilde{x}_p(t+l))}\frac{\partial\psi\left(\bar{x}_q(t-l)\right)}{\partial\epsilon}\mathbb{1}_{(\tilde{x}_q(t-l))} \\ &\quad +\psi\left(\bar{x}_p(t+l)\right)\mathbb{1}_{(\tilde{x}_p(t+l))}\psi\left(\bar{x}_q(t-l)\right)\frac{\partial\mathbb{1}_{(\tilde{x}_q(t-l))}}{\partial\epsilon}\end{aligned} \tag{5.32}$$

Due to the dependence of $\psi(x)$ and $\mathbb{1}_{(x)}$ on $\hat{\sigma}_x(t)$

$$\frac{\partial\psi\left(\bar{x}_p(t+l)\right)}{\partial\epsilon}=\frac{\partial\psi\left(\bar{x}_p(t+l)\right)}{\partial\hat{\sigma}_{\bar{x}_p}(t)}\frac{\partial\hat{\sigma}_{\bar{x}_p}(t)}{\partial\epsilon},\qquad \frac{\partial\mathbb{1}_{(\tilde{x}_p(t+l))}}{\partial\epsilon}=\frac{\partial\mathbb{1}_{(\tilde{x}_p(t+l))}}{\partial\hat{\sigma}_{\tilde{x}_p}(t)}\frac{\partial\hat{\sigma}_{\tilde{x}_p}(t)}{\partial\epsilon},$$
$$\frac{\partial\psi\left(\bar{x}_q(t-l)\right)}{\partial\epsilon}=\frac{\partial\psi\left(\bar{x}_q(t-l)\right)}{\partial\hat{\sigma}_{\bar{x}_q}(t)}\frac{\partial\hat{\sigma}_{\bar{x}_q}(t)}{\partial\epsilon},\qquad \frac{\partial\mathbb{1}_{(\tilde{x}_q(t-l))}}{\partial\epsilon}=\frac{\partial\mathbb{1}_{(\tilde{x}_q(t-l))}}{\partial\hat{\sigma}_{\tilde{x}_q}(t)}\frac{\partial\hat{\sigma}_{\tilde{x}_q}(t)}{\partial\epsilon}. \quad (5.33)$$

Now inserting the derivatives from Eq. (5.33) in Eq. (5.32) and applying the definition of influence function of the scale estimator $\mathrm{IF}(z;\hat{\sigma}_x,F)=\frac{\partial\hat{\sigma}_x}{\partial\epsilon}\big|_{\epsilon\downarrow 0}$ (see e.g. [122] p. 107 for the influence function of the normalized median absolute deviations scale estimator, which is used in Section 5.4), we obtain:

$$\begin{aligned}
&\mathrm{E}_F\left[\frac{\partial\left[\psi\left(\bar{x}_p(t+l)\right)\mathbb{1}_{(\tilde{x}_p(t+l))}\psi\left(\bar{x}_q(t-l)\right)\mathbb{1}_{(\tilde{x}_q(t-l))}\right]}{\partial\epsilon}\right]=\\
&\mathrm{E}_F\left[\left.\frac{\partial\psi\left(\bar{x}_p(t+l)\right)}{\partial\hat{\sigma}_{\bar{x}_p}(t)}\right|_{\epsilon\downarrow 0}\right]\mathrm{IF}(\bar{z}_p;\hat{\sigma}_{\bar{x}_p}(t),F)\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right]\\
&+\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathrm{E}_F\left[\left.\frac{\partial\mathbb{1}\left(\tilde{x}_p(t+l)\right)}{\partial\hat{\sigma}_{\tilde{x}_p}(t)}\right|_{\epsilon\downarrow 0}\right]\mathrm{IF}(\tilde{z}_p;\hat{\sigma}_{\tilde{x}_p}(t),F)\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right]\\
&+\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\mathrm{E}_F\left[\left.\frac{\partial\psi\left(\bar{x}_q(t-l)\right)}{\partial\hat{\sigma}_{\bar{x}_q}(t)}\right|_{\epsilon\downarrow 0}\right]\mathrm{IF}(\bar{z}_q;\hat{\sigma}_{\bar{x}_q}(t),F)\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_q(t-l))}\right]\\
&+\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]\mathrm{E}_F\left[\psi\left(\bar{x}_q(t-l)\right)\right]\mathrm{E}_F\left[\left.\frac{\partial\mathbb{1}(\tilde{x}_q(t-l))}{\partial\hat{\sigma}_{\tilde{x}_q}(t)}\right|_{\epsilon\downarrow 0}\right] \quad (5.34)
\end{aligned}$$

Moreover, the computation of partial derivatives in the above Eq. results in

$$\begin{aligned}
\frac{\partial\psi\left(\bar{x}_p(t+l)\right)}{\partial\hat{\sigma}_{\bar{x}_p}(t)} &= \bar{x}_p(t+l)\left[\delta\left(\bar{x}_p(t+l)+3\hat{\sigma}_{\bar{x}_p}(t)\right)-\delta\left(\bar{x}_p(t+l)-3\hat{\sigma}_{\bar{x}_p}(t)\right)\right]\\
\frac{\partial\mathbb{1}_{(\tilde{x}_p(t+l))}}{\partial\epsilon} &= \left[\delta\left(\tilde{x}_p(t+l)+3\hat{\sigma}_{\tilde{x}_p}(t)\right)-\delta\left(\tilde{x}_p(t+l)-3\hat{\sigma}_{\tilde{x}_p}(t)\right)\right]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\psi\left(\bar{x}_q(t-l)\right)}{\partial\epsilon} &= \bar{x}_q(t-l)\left[\delta\left(\bar{x}_q(t-l)+3\hat{\sigma}_{\bar{x}_q}(t)\right)-\delta\left(\bar{x}_q(t-l)-3\hat{\sigma}_{\bar{x}_q}(t)\right)\right]\\
\frac{\partial\mathbb{1}(\tilde{x}_q(t-l))}{\partial\epsilon} &= \left[\delta\left(\tilde{x}_q(t-l)+3\hat{\sigma}_{\tilde{x}_q}\right)-\delta\left(\tilde{x}_q(t-l)-3\hat{\sigma}_{\tilde{x}_p}(t)\right)\right].
\end{aligned}$$

By using (5.9), the $\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]$ in (5.31) can be written as

$$\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right]=\int_{-3\sigma}^{3\sigma}\bar{x}_p\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-\left(\bar{x}_p-\bar{\mu}_p(t+l)\right)^2}{2\sigma^2}}\mathrm{d}\bar{x}_p, \quad (5.35)$$

where $\bar{\mu}_p$ denotes the real part of the time-varying mean of $x_p(t+l)$ given by Eq. (5.28). Performing the integration in Eq. (5.35) yields

$$\begin{aligned}\mathrm{E}_F\left[\psi\left(\bar{x}_p(t+l)\right)\right] = &-\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{(\bar{\mu}_p(t+l)+3\sigma)^2}{2\sigma^2}}\left(-1+e^{\frac{6\bar{\mu}_p(t+l)}{\sigma}}\right)\\ &+\frac{1}{2}\bar{\mu}_p(t+l)\mathrm{Erf}\left(\frac{-\bar{\mu}_p(t+l)+3\sigma}{\sqrt{2}\sigma}\right)\\ &+\frac{1}{2}\bar{\mu}_p(t+l)\mathrm{Erf}\left(\frac{\bar{\mu}_p(t+l)+3\sigma}{\sqrt{2}\sigma}\right),\end{aligned} \tag{5.36}$$

where $\mathrm{Erf}(\cdot)$ denotes the error-function. Similarly, $\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right]$ in (5.31), by using (5.10) can be expressed as

$$\begin{aligned}\mathrm{E}_F\left[\mathbb{1}_{(\tilde{x}_p(t+l))}\right] &= \int_{-3\sigma}^{3\sigma}\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(\tilde{x}_p-\tilde{\mu}_p(t+l))^2}{2\sigma^2}}\mathrm{d}\tilde{x}_p\\ &= \frac{\mathrm{Erf}\left(\frac{-\tilde{\mu}_p(t+l)+3\sigma}{\sqrt{2}\sigma}\right)+\mathrm{Erf}\left(\frac{\tilde{\mu}_p(t+l)+3\sigma}{\sqrt{2}\sigma}\right)}{2},\end{aligned} \tag{5.37}$$

where $\tilde{\mu}_p(t+l)$ is the imaginary part of the time-varying mean of $x_p(t+l)$. Analogous computations for $\mathrm{IF}(\mathcal{I}_2), \mathrm{IF}(\mathcal{I}_3)$ and $\mathrm{IF}(\mathcal{I}_4)$ result in the influence function for the 3-σ rejection based STFD matrix estimator given in Eq. (5.12). In the case of an auto sensor element of the STFD matrix, i.e. $p = q$ in Eq. (5.8), for $l \neq 0$, the resulting expression for the expectations and the corresponding influence function is similar to Eq. (5.12). However, for $l = 0$, we obtain:

$$\hat{D}_{x_px_p}(t,f)\big|_{l=0} = \left(\psi\left(\bar{x}_p(t)\right)\right)^2\mathbb{1}_{(\tilde{x}_p(t))} + \left(\psi\left(\tilde{x}_p(t)\right)\right)^2\mathbb{1}_{(\bar{x}_p(t))}$$

and computing the influence function for this case requires to evaluate:

$$\begin{aligned}\mathrm{E}_{\mathcal{F}_\epsilon}\left[\left(\psi\left(\bar{x}_p(t)\right)\right)^2\mathbb{1}_{(\tilde{x}_p(t))}\right] &= \int\int\left(\psi\left(\bar{x}_p(t)\right)\right)^2\mathbb{1}_{(\tilde{x}_p(t))}\mathrm{d}F_\epsilon(\bar{x}_p)\mathrm{d}F_\epsilon(\tilde{x}_p)\\ \mathrm{E}_{\mathcal{F}_\epsilon}\left[\left(\psi\left(\tilde{x}_p(t)\right)\right)^2\mathbb{1}_{(\bar{x}_p(t))}\right] &= \int\int\left(\psi\left(\tilde{x}_p(t)\right)\right)^2\mathbb{1}_{(\bar{x}_p(t))}\mathrm{d}F_\epsilon(\bar{x}_p)\mathrm{d}F_\epsilon(\tilde{x}_p)\end{aligned}$$

where $\mathrm{E}_F\left[\left(\psi\left(\bar{x}_p(t)\right)\right)^2\right]$, by using (5.9) is given by

$$\mathrm{E}_F\left[\left(\psi\left(\bar{x}_p(t)\right)\right)^2\right] = \int_{-3\sigma}^{3\sigma}\left(\bar{x}_p(t)\right)^2\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(\bar{x}_p-\bar{\mu}_p(t))^2}{2\sigma^2}}\mathrm{d}\bar{x}_p$$

$$\begin{aligned}= &\frac{1}{2\sqrt{2\pi\sigma^2}}\left(2(\bar{\mu}_p(t)-3\sigma)\sigma^2e^{-\frac{(\bar{\mu}_p(t)+3\sigma)^2}{2\sigma^2}}-2e^{-\frac{(\bar{\mu}_p(t)-3\sigma)^2}{2\sigma^2}}\sigma^2(\bar{\mu}_p(t)+3\sigma)\right.\\ &+\sqrt{2\pi}\left(\sigma^3+(\bar{\mu}_p(t))^2\sigma\right)\mathrm{Erf}\left(\frac{-\bar{\mu}_p(t)+3\sigma}{\sqrt{2}\sigma}\right)\\ &\left.+\sqrt{2\pi}\left(\sigma^3+(\bar{\mu}_p(t))^2\sigma\right)\mathrm{Erf}\left(\frac{\bar{\mu}_p(t)+3\sigma}{\sqrt{2}\sigma}\right)\right).\end{aligned} \tag{5.38}$$

5.8 Appendix III: Derivation of the influence function for the M-estimator of the STFD matrix

In order to evaluate the influence function for the M-estimator of the STFD matrix, we need to compute the expectation of $\hat{D}_{x_p x_q}(t,f)$, as given in Eq. (5.14), w.r.t. F_ϵ. This can be written as

$$0 = \sum_{l=-L/2}^{L/2} \mathrm{E}_{\mathcal{F}_\epsilon} \left[\Psi(d(l))\right] \tag{5.39}$$

Starting from Eq. (5.39) and applying Eq. (5.2) gives

$$\begin{aligned} 0 = &\sum_{l=-L/2}^{L/2} \mathrm{E}_F \left[\left. \frac{\partial \Psi\left(d(t,l)\right)}{\partial \epsilon} \right|_{\epsilon \downarrow 0} \right] - 2 \sum_{l=-L/2}^{L/2} \mathrm{E}_F \left[\Psi\left(d(t,l)\right)\right] \\ &+ \sum_{l=-L/2}^{L/2} \mathrm{E}_{\mathrm{F}(x_p)} \Psi\left(d(t,l,z_q)\right) + \sum_{l=-L/2}^{L/2} \mathrm{E}_{\mathrm{F}(x_q)} \Psi\left(d(t,l,z_p)\right) \end{aligned} \tag{5.40}$$

Furthermore, due to the dependence of $\Psi(x)$ on $\hat{\sigma}_M(t)$ solving the first expression in Eq. (5.40) yields

$$\frac{\partial \Psi\left(d(t,l)\right)}{\partial \epsilon} = \frac{\partial \Psi\left(d(t,l)\right)}{\partial d(t,l)} \frac{\partial d(t,l)}{\partial \epsilon} \tag{5.41}$$

with

$$\frac{\partial d(t,l)}{\partial \epsilon} = \frac{\partial d(t,l)}{\partial \hat{D}_\epsilon} \frac{\mathrm{d}\hat{D}_\epsilon}{\mathrm{d}\epsilon} + \frac{\partial d(t,l)}{\partial \hat{\sigma}_M(t)} \frac{\mathrm{d}\hat{\sigma}_M(t)}{\mathrm{d}\epsilon},$$

where $\hat{D}_\epsilon$ and $\hat{\sigma}_M(t)$ denote the estimator of the TFD and the scale under the contaminated distribution F_ϵ. Let $\Psi'\left(d(t,l)\right) = \partial \Psi\left(d(t,l)\right) / (\partial d(t,l))$, then by using the results from Eq. (5.41) and the definition of the influence function for the scale estimator [122], the first term in Eq. (5.40) becomes

$$\begin{aligned} \mathrm{E}_F \left[\left. \frac{\partial \Psi\left(d(t,l)\right)}{\partial \epsilon} \right|_{\epsilon \downarrow 0} \right] = &\, \mathrm{E}_F \left[\Psi'\left(d(t,l)\right) \frac{\partial d(t,l)}{\partial D} \right] \mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), \mathrm{F}) \\ &+ \mathrm{E}_F \left[\Psi'\left(d(t,l)\right) \frac{\partial d(t,l)}{\partial \hat{\sigma}_M(t)} \right] \mathrm{IF}(z_p, z_q; \hat{\sigma}_M(t), \mathrm{F}) \end{aligned} \tag{5.42}$$

The expression for the influence function for the M-estimator of STFD is obtained by substituting Eq. (5.42) into Eq. (5.40) and is given in Eq. (5.20).

In the following, we will demonstrate the computation of statistical expectations in Eq. (5.20) for Huber's ρ function for which Ψ is defined in Eq. (5.18). For this, let us

define the following abbreviations for the expectations in (5.20):

$$\mathcal{I}_1(l) := \mathrm{E}_F\left[\Psi\left(d(t,l)\right)\right],\ \mathcal{I}_2(l,z_q) := \mathrm{E}_{\mathrm{F}(x_p)}\left[\Psi\left(d(t,l,z_q)\right)\right],\ \mathcal{I}_3(l,z_p) := \mathrm{E}_{\mathrm{F}(x_q)}\left[\Psi\left(d(t,l,z_p)\right)\right],$$
$$\mathcal{I}_4(l) := \mathrm{E}_F\left[\Psi'\left(d(t,l)\right)\frac{\partial d(t,l)}{\partial \hat{\sigma}_M(t)}\right], \qquad \mathcal{I}_5(l,z_p) := \mathrm{E}_F\left[\Psi'\left(d(t,l)\right)\frac{\partial d(t,l)}{\partial D}\right] \tag{5.43}$$

$\mathcal{I}_1(l)$, by using the signal model defined in Eq. (2.2), is given by

$$\begin{aligned}
\mathcal{I}_1(l) = &\left(\frac{\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)s_k^*(t-l)a_q^*(\theta_k)e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)}\right)\mathrm{E}_F\left[\mathbb{1}_{|d(l)|\le c_H}\right]\\
&+\left(\frac{\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl}}{\hat{\sigma}_M(t)}\right)\mathrm{E}_F\left[n_q^*(t-l)\mathbb{1}_{(|d(l)|\le c_H)}\right]\\
&+\left(\frac{\sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t-l)e^{-j4\pi fl}}{\hat{\sigma}_M}\right)\mathrm{E}_F\left[n_p(t+l)\mathbb{1}_{(|d(l)|\le c_H)}\right]\\
&+\frac{1}{\hat{\sigma}_M}\mathrm{E}_F\left[n_p(t+l)n_q^*(t-l)\mathbb{1}_{(|d(l)|\le c_H)}\right]\\
&+c_H\left(\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)s_k^*(t-l)a_q^*(\theta_k)e^{-j4\pi fl} - D\right)\mathrm{E}_F\left[\frac{\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right]\\
&+c_H\left(\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl}\right)\mathrm{E}_F\left[\frac{n_q^*(t-l)\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right]\\
&+c_H\left(\sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t-l)e^{-j4\pi fl}\right)\mathrm{E}_F\left[\frac{n_p(t+l)\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right]\\
&+c_H\mathrm{E}_F\left[\frac{n_p(t+l)n_q^*(t-l)\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right]
\end{aligned} \tag{5.44}$$

The second and third expectations in Eq. (5.44) are zero because n_p and n_q are zero-mean. The first term is non-zero while the last term is non-zero only for $p=q$ and $l=0$, which arise in the case of an auto-sensor TFD. Similarly, the remaining $\mathcal{I}_i(l),\ i\in\{2,5\}$ in Eq. (5.44) are given by:

$$\begin{aligned}
\mathcal{I}_2(l,z_q) = &\quad z_q^*\left(\frac{\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)}\right)\mathrm{E}_F\left[\mathbb{1}_{(|d(l,z_q)|\le c_H)}\right]\\
&+c_H z_q^*\left(\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl} - D\right)\mathrm{E}_F\left[\frac{\mathbb{1}_{|d(l,z_q)|>c_H)}}{|d(l,z_q)|}\right] + c_H z_q^*\mathrm{E}_F\left[\frac{n_p(t+l)\mathbb{1}_{|d(l,z_q)|>c_H)}}{|d(l,z_q)|}\right]
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_3(l,z_p) = &\quad z_p\left(\frac{\sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t-l)e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)}\right)\mathrm{E}_F\left[\mathbb{1}_{(|d(l,z_p)|\le c_H)}\right]\\
&+c_H z_p\left(\sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t+l)e^{-j4\pi fl} - D\right)\mathrm{E}_F\left[\frac{\mathbb{1}_{|d(l,z_p)|>c_H)}}{|d(l,z_p)|}\right] + c_H z_p\mathrm{E}_F\left[\frac{n_q^*(t-l)\mathbb{1}_{|d(l,z_p)|>c_H)}}{|d(l,z_p)|}\right]
\end{aligned}$$

$$\mathcal{I}_4(l) = -\left(\frac{\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)s_k^*(t-l)a_q^*(\theta_k) - D}{\hat{\sigma}_M^2(t)}\right)\mathrm{E}_F\left[\mathbb{1}_{(|d(l)|\leq c_H)}\right]$$
$$-\frac{\sum_{k=1}^{K} a_p(\theta_k)s_k(t+l)e^{-j4\pi fl}}{\hat{\sigma}_M^2(t)}\mathrm{E}_F\left[n_q^*(t-l)\mathbb{1}_{(|d(l)|\leq c_H)}\right]$$
$$-\frac{\sum_{k=1}^{K} a_q^*(\theta_k)s_k^*(t-l)e^{-j4\pi fl}}{\hat{\sigma}_M^2(t)}\mathrm{E}_F\left[n_p(t+l)\mathbb{1}_{(|d(l)|\leq c_H)}\right]$$
$$-\frac{1}{\hat{\sigma}_M^2(t)}\mathrm{E}_F\left[n_p(t+l)n_q^*(t-l)\mathbb{1}_{(|d(l)|\leq c_H)}\right]$$
$$\mathcal{I}_5(l) = -\left(\frac{1}{\hat{\sigma}_M(t)}\right)\mathrm{E}_F\left[\mathbb{1}_{(|d(l)|\leq c_H)}\right]$$

Let's now make the following definitions:

$$\mathcal{E}_{11} := \mathrm{E}_F\left[\mathbb{1}_{|d(l)|\leq c_H}\right],\ \mathcal{E}_{12} := \mathrm{E}_F\left[n_q^*(t-l)\mathbb{1}_{(|d(l)|\leq c_H)}\right],\ \mathcal{E}_{13} := \mathrm{E}_F\left[n_p(t+l)\mathbb{1}_{(|d(l)|\leq c_H)}\right]$$
$$\mathcal{E}_{14} := \mathrm{E}_F\left[n_p(t+l)n_q^*(t-l)\mathbb{1}_{(|d(l)|\leq c_H)}\right],\ \mathcal{E}_{15} := \mathrm{E}_F\left[\frac{\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right],\ \mathcal{E}_{16} := \mathrm{E}_F\left[\frac{n_q^*(t-l)\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right]$$
$$\mathcal{E}_{17} := \mathrm{E}_F\left[\frac{n_p(t+l)\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right],\ \mathcal{E}_{18} := \mathrm{E}_F\left[\frac{n_p(t+l)n_q^*(t-l)\mathbb{1}_{|d(t,l)|>c_H}}{|d(t,l)|}\right]$$
$$\mathcal{E}_{21} := \mathrm{E}_F\left[\mathbb{1}_{(|d(l,z_q)|\leq c_H)}\right]\ \ \mathcal{E}_{22} := \mathrm{E}_F\left[\frac{\mathbb{1}_{|d(l,z_q)|>c_H)}}{|d(l,z_q)|}\right]\quad \mathcal{E}_{23} := \mathrm{E}_F\left[\frac{n_p(t+l)\mathbb{1}_{|d(l,z_q)|>c_H)}}{|d(l,z_q)|}\right]$$
$$\mathcal{E}_{31} := \mathrm{E}_F\left[\mathbb{1}_{(|d(l,z_p)|\leq c_H)}\right]\quad \mathcal{E}_{32} := \mathrm{E}_F\left[\frac{\mathbb{1}_{|d(l,z_p)|>c_H)}}{|d(l,z_p)|}\right]\quad \mathcal{E}_{33} := \mathrm{E}_F\left[\frac{n_q^*(t-l)\mathbb{1}_{|d(l,z_p)|>c_H)}}{|d(l,z_p)|}\right] \tag{5.45}$$

Inserting the expressions for the statistical expectations in Eq. (5.20) yields the fol-

lowing expression for the influence function

$$
\begin{aligned}
\mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}(t,f), \mathrm{F}) = &\Bigg(2 \sum_{l=-L/2}^{L/2} \Bigg(\left(\frac{\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) s_k^*(t-l) a_q^*(\theta_k) e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)} \right) \mathcal{E}_{11} \\
&+ \left(\frac{\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl}}{\hat{\sigma}_M(t)} \right) \mathcal{E}_{12} + \left(\frac{\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl}}{\hat{\sigma}_M(t)} \right) \mathcal{E}_{13} \\
&+ \frac{1}{\hat{\sigma}_M(t)} \mathcal{E}_{14} + c_H \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) s_k^*(t-l) a_q^*(\theta_k) e^{-j4\pi fl} - D \right) \mathcal{E}_{15} \\
&+ c_H \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right) \mathcal{E}_{16} + c_H \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right) \mathcal{E}_{17} + c_H \mathcal{E}_{18} \Bigg) \\
&- \sum_{l=-L/2}^{L/2} \Bigg(z_q^* \left(\frac{\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)} \right) \mathcal{E}_{21} \\
&+ c_H z_q^* \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} - D \right) \mathcal{E}_{22} + c_H z_q^* \mathcal{E}_{23} \Bigg) \\
&- \sum_{l=-L/2}^{L/2} \Bigg(z_p \left(\frac{\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} - D}{\hat{\sigma}_M(t)} \right) \mathcal{E}_{31} \\
&+ c_H z_p \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t+l) e^{-j4\pi fl} - D \right) \mathcal{E}_{32} + c_H z_p \mathcal{E}_{33} \Bigg) \\
&- \Bigg(\left(\frac{\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) s_k^*(t-l) a_q^*(\theta_k) - D}{\hat{\sigma}_M^2(t)} \right) \mathcal{E}_{11} \\
&+ \frac{\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl}}{\hat{\sigma}_M^2(t)} \mathcal{E}_{12} \\
&+ \frac{\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl}}{\hat{\sigma}_M^2(t)} \mathcal{E}_{13} + \frac{1}{\hat{\sigma}_M^2(t)} \mathcal{E}_{14} \Bigg) \mathrm{IF}(z_p, z_q; \hat{\sigma}_M(t), \mathrm{F}) \Bigg) \\
&\times \left(\left(-\frac{1}{\hat{\sigma}_M(t)} \right) \sum_{l=-L/2}^{L/2} \mathcal{E}_{11} \right)^{-1}
\end{aligned}
\tag{5.46}
$$

Evaluating $\mathcal{E}_i$ $i \in \{11, 18\}$ requires an integration over a region in $\mathbb{R}^4$ which is either defined by $|d(t,l)| \leq c_H$ or by $|d(t,l)| > c_H$ (see Eq. (5.18)). Computing $\mathcal{E}_{21}, \mathcal{E}_{22}, \mathcal{E}_{23}, \mathcal{E}_{31}, \mathcal{E}_{32}$ and $\mathcal{E}_{33}$ require an integration over a region in $\mathbb{R}^2$ defined by $|d(t,l,z_q)| \leq c_H$ and/or $|d(t,l,z_p)| > c_H$, respectively. In the case of auto sensor STFD and $l = 0$, $\mathcal{E}_i$, $i \in \{11, 18\}$ are required to be computed over a region in $\mathbb{R}^2$. The above expectations are computed by first transforming the complex random variables to their polar representation and then solving the integrals numerically. To give an

example, the expectation $\mathcal{E}_{11}$ is computed by

$$\mathrm{E}_F\left[\mathbb{1}_{|d(l)|\leq c_H}\right] = \int_{\mathbb{R}^4} \mathbb{1}_{(|d(l)|\leq c_H)} f_{\bar{n}_p}(\bar{n}_p) f_{\tilde{n}_p}(\tilde{n}_p) f_{\bar{n}_q}(\bar{n}_q) f_{\tilde{n}_q}(\tilde{n}_q) \mathrm{d}\bar{n}_p \mathrm{d}\tilde{n}_p \mathrm{d}\bar{n}_q \mathrm{d}\tilde{n}_q,$$

where $f_{\bar{n}_p}(\bar{n}_p), f_{\tilde{n}_p}(\tilde{n}_p), f_{\bar{n}_q}(\bar{n}_q), f_{\tilde{n}_p}(\tilde{n}_q)$ are the probability density functions of $\bar{n}_p, \tilde{n}_p, \bar{n}_q, \tilde{n}_q$, respectively. $\bar{n}_p, \tilde{n}_p$ and $\bar{n}_q, \tilde{n}_q$ are the real and the imaginary part of the noise at the p^{th} and the q^{th} sensors. With Eq. (2.2), $d(t,l)$ from Eq. (5.15) becomes

$$d(t,l) = \frac{1}{\hat{\sigma}_M(t)} \times \left(\left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) + n_p(t+l) \right) \left(\sum_{k=1}^{K} a_q(\theta_k) s_k(t-l) + n_q(t-l) \right)^* e^{-j4\pi fl} - D \right).$$

Transforming into polar co-ordinates yields $n_p = r_p e^{j\phi_p}$ and $n_q = r_q e^{j\phi_q}$, where

$$r_p = \sqrt{\bar{n}_p^2 + \tilde{n}_p^2},\ \phi_p = \angle n_p,\ r_q = \sqrt{\bar{n}_q^2 + \tilde{n}_q^2},\ \phi_q = \angle n_q.$$

Since $\bar{n}_p$, $\tilde{n}_p$, $\bar{n}_q$ and $\tilde{n}_q$ are i.i.d. random variables, r_p, r_q, ϕ_p, ϕ_q are also independently distributed [125]. Moreover, ϕ_p, ϕ_q are uniformly distributed between $\{-\pi, \pi\}$ and r_p, r_q are Rayleigh distributed. $\mathbb{1}_{|d(l)|\leq c_H}$ defines the hyper-surface where the integrations are evaluated. For the computation of the resulting influence function in Eq. (5.46), simplifications are used to obtain decoupled integrals which are then evaluated numerically.

5.9 Appendix IV: Derivation of the influence function for the one-step re-wighted STFD estimator

In this Section, we derive the influence function expression for the one step re-weighted STFD estimator defined in Eq. (5.21). In order to do so, we must compute the expectation of $\hat{D}_{x_p x_q}(t,f)$ w.r.t. the contaminated distribution

$$\begin{aligned}\mathrm{E}_{\mathcal{F}_\epsilon}\left[\hat{D}_{x_p x_q}(t,f)\right] &= \int\int w(d(t,l)) x_p(t+l) x_q^*(t-l) e^{-j4\pi fl} \left((1-\epsilon)^2 \mathrm{dF}(x_p)\mathrm{dF}(x_q)\right. \\ &\quad \left. + \mathrm{dF}(x_p)\delta_{z_q} + \delta_{z_p}\mathrm{dF}(x_q)\right) \end{aligned} \tag{5.47}$$

For the signal model given in Eq. (2.2), applying the definition from Eq. (5.2), results in the following expression for the influence function

$$\begin{aligned}\mathrm{IF}\left(z_p, z_q; \hat{D}_{x_p x_q}(t,f), \mathrm{F}\right) = \sum_{l=-L/2}^{L/2} \Bigg(& \mathrm{E}_F\left[\left.\frac{\partial w(d(t,l))}{\partial \epsilon}\right|_{\epsilon\downarrow 0} x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}\right] \\ & + \mathrm{E}_F\left[w(d(t,l))x_p(t+l)x_q^*(t-l)\right]e^{-j4\pi fl} + \mathrm{E}_F\left[w(d(t,l,z_q))x_p(t+l)\right]e^{-j4\pi fl}z_q^* \\ & z_p\mathrm{E}_F\left[w(d(t,l,z_p))x_q^*(t-l)\right]e^{-j4\pi fl}\Bigg), \end{aligned} \tag{5.48}$$

where

$$\frac{\partial w(d(t,l))}{\partial \epsilon} = \frac{\partial w(d(t,l))}{\partial d(t,l)}\left(\frac{\partial d(t,l)}{\partial \hat{\sigma}_W(t)}\frac{\partial \hat{\sigma}_W(t)}{\partial \epsilon} + \frac{\partial d(t,l)}{\partial \hat{\mu}_W(t)}\frac{\partial \hat{\mu}_W(t)}{\partial \epsilon}\right).$$

By substituting the definitions of the influence function for the location and scale estimates $\hat{\mu}_W(t)$ and $\hat{\sigma}_W(t)$, the first term of Eq. (5.48) can be written as:

$$\begin{aligned} & \mathrm{E}_F\left[\left.\frac{\partial w(d(t,l))x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}}{\partial \epsilon}\right|_{\epsilon\downarrow 0}\right] = \\ & \mathrm{E}_F\left[\frac{\partial w(d(t,l))}{\partial d(t,l)}\frac{\partial d(t,l)}{\partial \hat{\sigma}_W(t)}x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}\right]\mathrm{IF}\left(z_p, z_q; \hat{\sigma}_W(t), \mathrm{F}\right) + \\ & \mathrm{E}_F\left[\frac{\partial w(d(t,l))}{\partial d(t,l)}\frac{\partial d(t,l)}{\partial \hat{\mu}_W(t)}x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}\right]\mathrm{IF}\left(z_p, z_q; \hat{\mu}_W(t), \mathrm{F}\right) \end{aligned} \tag{5.49}$$

where

$$\frac{\partial w(d(t,l))}{\partial d(t,l)} = \frac{4|d(t,l)|}{c_T^2}\left[1 - \frac{|d(t,l)|^2}{c_T^2}\right]$$

Let us define the following abbreviations for the terms in the expression for the influence function

$$\begin{aligned} \mathcal{I}_1(l) &:= \mathrm{E}_F\left[\frac{\partial w(d(t,l))}{\partial d(t,l)}\frac{\partial d(t,l)}{\partial \hat{\sigma}_W(t)}x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}\mathbb{1}_{|d(t,l)|\le c_T}\right] \\ \mathcal{I}_2(l) &:= \mathrm{E}_F\left[\frac{\partial w(d(t,l))}{\partial d(t,l)}\frac{\partial d(t,l)}{\partial \hat{\mu}_W(t)}x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}\mathbb{1}_{|d(t,l)|\le c_T}\right] \\ \mathcal{I}_3(l) &:= \mathrm{E}_F\left[w(d(t,l))x_p(t+l)x_q^*(t-l)e^{-j4\pi fl}\mathbb{1}_{|d(t,l)|\le c_T}\right] \\ \mathcal{I}_4(l) &:= \mathrm{E}_F\left[w(d(t,l,z_q))x_p(t+l)\mathbb{1}_{|d(t,l)|\le c_T}\right]e^{-j4\pi fl}z_q^* \\ \mathcal{I}_5(l) &:= z_p\mathrm{E}_F\left[w(d(t,l,z_p))x_q^*(t-l)\mathbb{1}_{|d(t,l)|\le c_T}\right]e^{-j4\pi fl} \end{aligned} \tag{5.50}$$

Each of the $\mathcal{I}$'s can be written as follows

$$\begin{aligned}\mathcal{I}_1(l) = \frac{1}{\hat{\sigma}_W^2(t)} \Bigg(& \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right)^2 \\ & \times \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right)^2 \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) (n_q^*(t-l))^2 \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l)^{-j4\pi fl} \right)^2 \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) (n_p(t+l))^2 \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) (n_p(t+l) n_q^*(t-l))^2 \mathbb{1}_{|d(t,l)| \le c_T} \right] \Bigg) \end{aligned} \tag{5.51}$$

$$\begin{aligned}\mathcal{I}_2(l) = \frac{1}{\hat{\sigma}_W^2(t)} \Bigg(& \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right) \\ & \times \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right) \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l)^{-j4\pi fl} \right) \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) n_p(t+l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) n_p(t+l) n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \Bigg) \end{aligned} \tag{5.52}$$

$$\begin{aligned}\mathcal{I}_3(l) = \Bigg(& \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right) \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right) \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l)^{-j4\pi fl} \right) \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 n_p(t+l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \\ & + \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 n_p(t+l) n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \Bigg) \end{aligned} \tag{5.53}$$

$$\mathcal{I}_4(l) = z_q^* \left(\left(\sum_{k=1}^{K} a_p(\theta_k) s_k(t+l) e^{-j4\pi fl} \right) \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 \mathbb{1}_{|d(t,l,z_q)| \le c_T} \right] + \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 n_p(t+l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \right) \tag{5.54}$$

$$\mathcal{I}_5(l) = z_p \left(\left(\sum_{k=1}^{K} a_q^*(\theta_k) s_k^*(t-l) e^{-j4\pi fl} \right) \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 \mathbb{1}_{|d(t,l,z_p)| \le c_T} \right] + \mathrm{E}_F \left[\left(1 - \frac{|d(t,l,z_p)|^2}{c_T^2} \right)^2 n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right] \right) \tag{5.55}$$

Let us abbreviate the expectations in the above Eqs. as follows

$$\mathcal{E}_{11} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{12} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) (n_q^*(t-l))^2 \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{13} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) (n_p(t+l))^2 \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{14} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) (n_p(t+l) n_q^*(t-l))^2 \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{21} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{22} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{23} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) n_p(t+l) \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{24} := \mathrm{E}_F \left[\frac{4|d(t,l)|}{c_T^2} \left(1 - \frac{|d(t,l)|^2}{c_T^2} \right) n_p(t+l) n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{31} := \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{32} := \mathrm{E}_F \left[\left(1 - \frac{|d(t,l)|^2}{c_T^2} \right)^2 n_q^*(t-l) \mathbb{1}_{|d(t,l)| \le c_T} \right]$$

$$\mathcal{E}_{33} := \mathrm{E}_F\left[\left(1-\frac{|d(t,l)|^2}{c_T^2}\right)^2 n_p(t+l)\mathbb{1}_{|d(t,l)|\leq c_T}\right]$$
$$\mathcal{E}_{32} := \mathrm{E}_F\left[\left(1-\frac{|d(t,l)|^2}{c_T^2}\right)^2 n_p(t+l)n_q^*(t-l)\mathbb{1}_{|d(t,l)|\leq c_T}\right]$$
$$\mathcal{E}_{41} := \mathrm{E}_F\left[\left(1-\frac{|d(t,l)|^2}{c_T^2}\right)^2 \mathbb{1}_{|d(t,l,z_q)|\leq c_T}\right]$$
$$\mathcal{E}_{42} := \mathrm{E}_F\left[\left(1-\frac{|d(t,l)|^2}{c_T^2}\right)^2 n_p(t+l)\mathbb{1}_{|d(t,l)|\leq c_T}\right]$$
$$\mathcal{E}_{51} := \mathrm{E}_F\left[\left(1-\frac{|d(t,l)|^2}{c_T^2}\right)^2 \mathbb{1}_{|d(t,l,z_p)|\leq c_T}\right]$$
$$\mathcal{E}_{52} := \mathrm{E}_F\left[\left(1-\frac{|d(t,l,z_p)|^2}{c_T^2}\right)^2 n_q^*(t-l)\mathbb{1}_{|d(t,l)|\leq c_T}\right] \tag{5.56}$$

The final expression for the influence function of the one-step re-wighted STFD estimator is given in Eq. (5.24). Here, $\mathcal{E}_i$'s requires integration over a region in $\mathbb{R}^4$ or in over a region in $\mathbb{R}^2$, e.g.

$$\mathcal{E}_{11} := \int_{\mathbb{R}^4} \frac{4|d(t,l)|}{c_T^2}\left(1-\frac{|d(t,l)|^2}{c_T^2}\right)\mathbb{1}_{|d(t,l)|\leq c_T} g(\bar{n}_p,\tilde{n}_p,\bar{n}_q,\tilde{n}_q)\mathrm{d}\bar{n}_p\mathrm{d}\tilde{n}_p\mathrm{d}\bar{n}_q\mathrm{d}\tilde{n}_q \tag{5.57}$$

where $g(\bar{n}_p,\tilde{n}_p,\bar{n}_q,\tilde{n}_q)$ is the joint density which by using the i.i.d. property of $\bar{n}_p,\tilde{n}_p,\bar{n}_q,\tilde{n}_q$ is given by

$$g(\bar{n}_p,\tilde{n}_p,\bar{n}_q,\tilde{n}_q) = g_{\bar{n}_p}(\bar{n}_p)g_{\tilde{n}_p}(\tilde{n}_p)g_{\bar{n}_q}(\bar{n}_q)g_{\tilde{n}_p}(\tilde{n}_q).$$

Analogous to the previous estimator the expectations are solved by first transforming the complex random variables n_p, n_q to polar co-ordinates and then solving the integrals numerically.

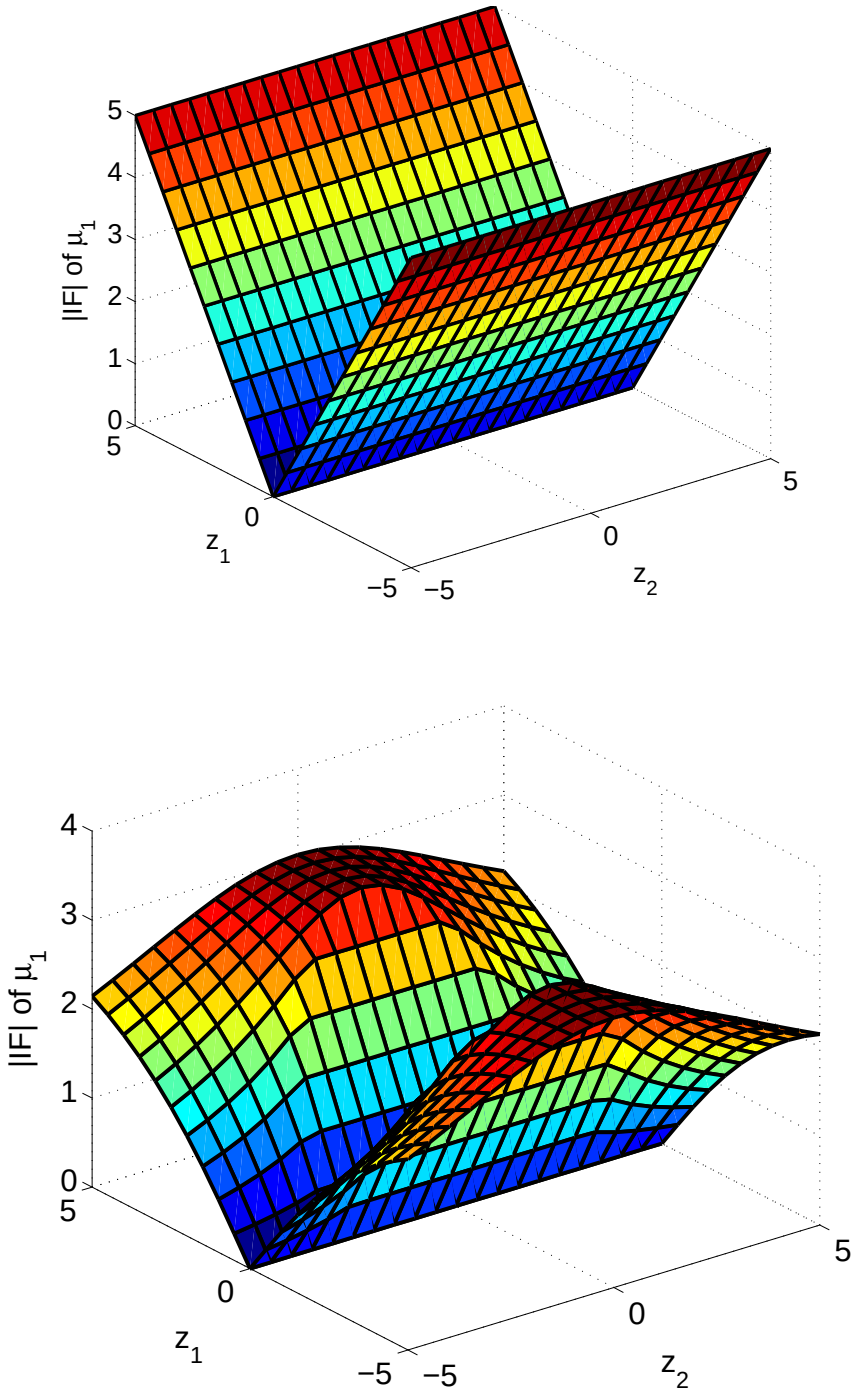

Figure 5.1. The first coordinate of the influence functions of two estimators of location for the bivariate standard normal distribution. The influence function of the sample mean is unbounded, while the influence function of the robust M-estimator is bounded and continuous, which means that the estimator is qualitatively robust.

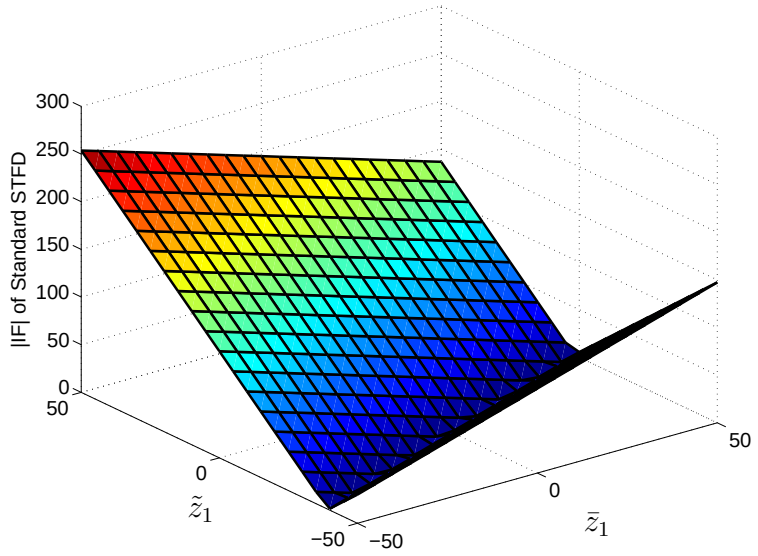

Figure 5.2. Influence function of the PWVD based STFD matrix estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

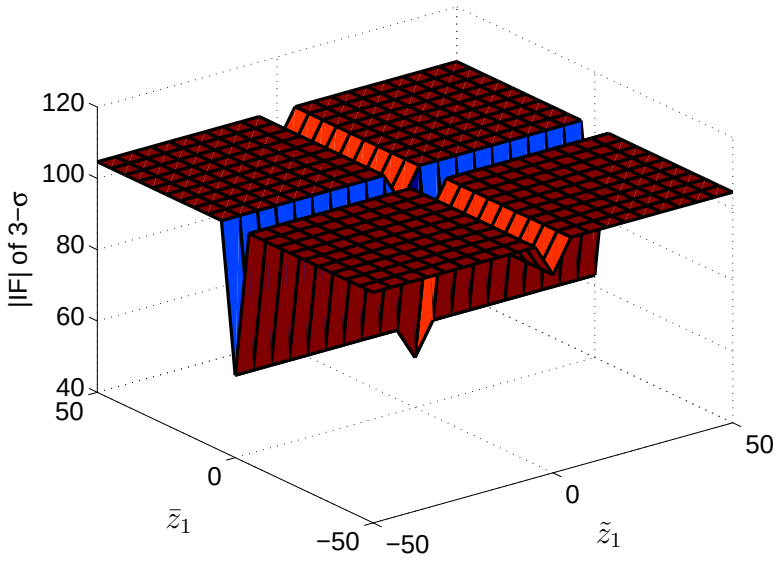

Figure 5.3. Influence function of the 3-σ rejection pre-processing based STFD matrix estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

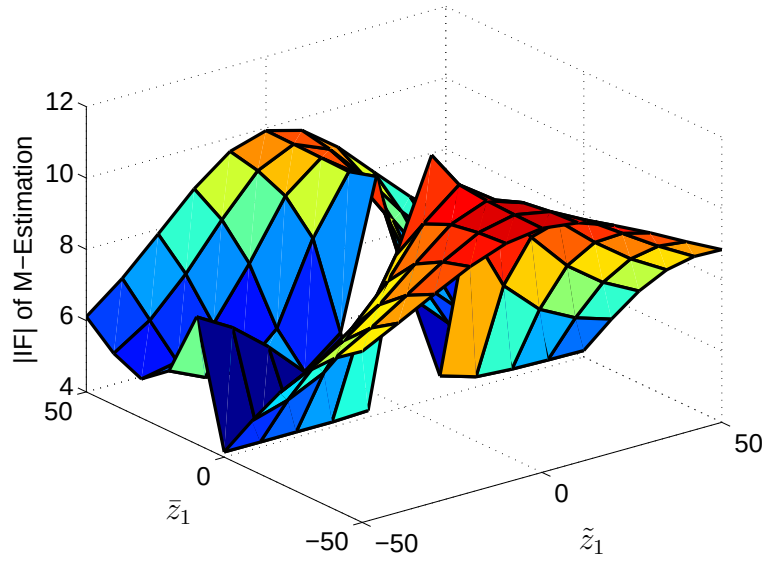

Figure 5.4. Influence function of the M-estimation based STFD matrix estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

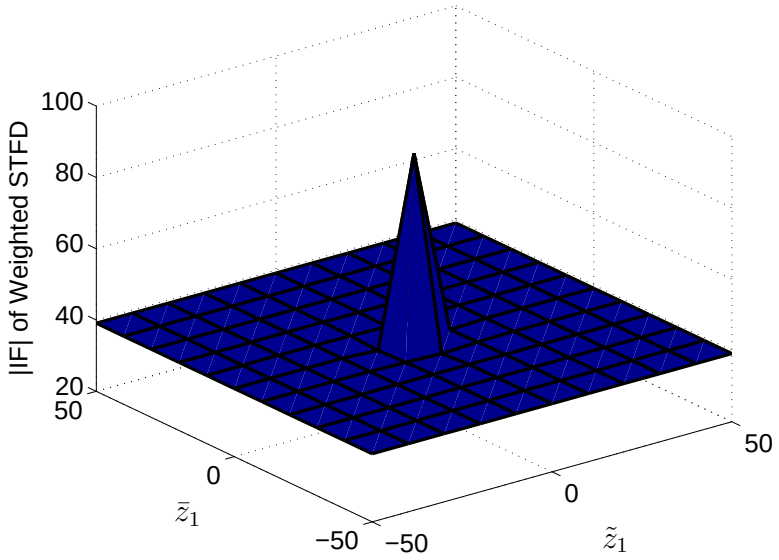

Figure 5.5. Influence function of the one-step re-weighted STFD estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

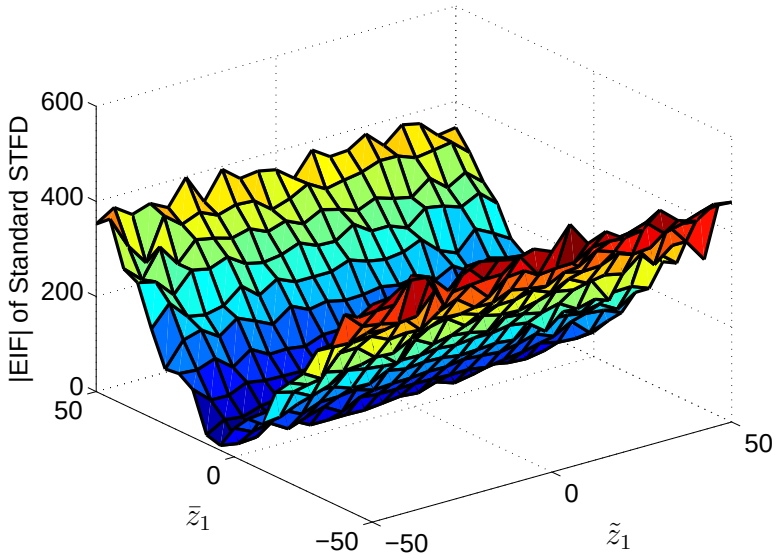

Figure 5.6. Empirical influence function of the standard STFD estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

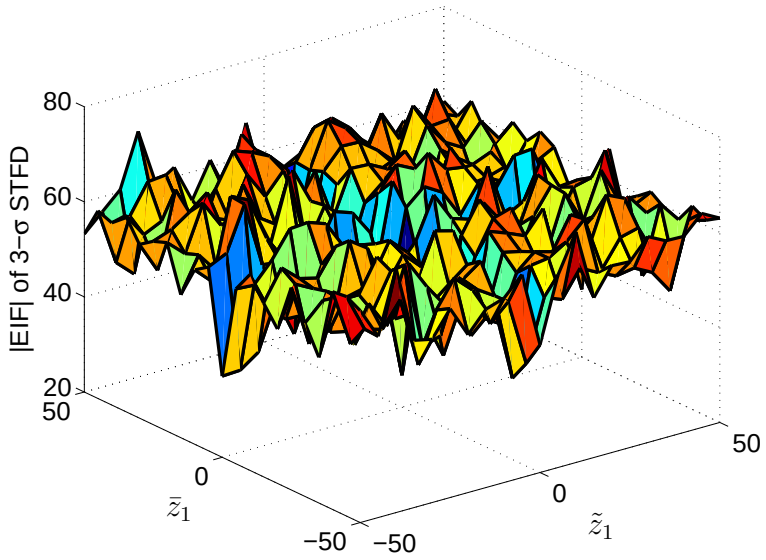

Figure 5.7. Empirical influence function of the 3-σ rejection pre-processing based STFD estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

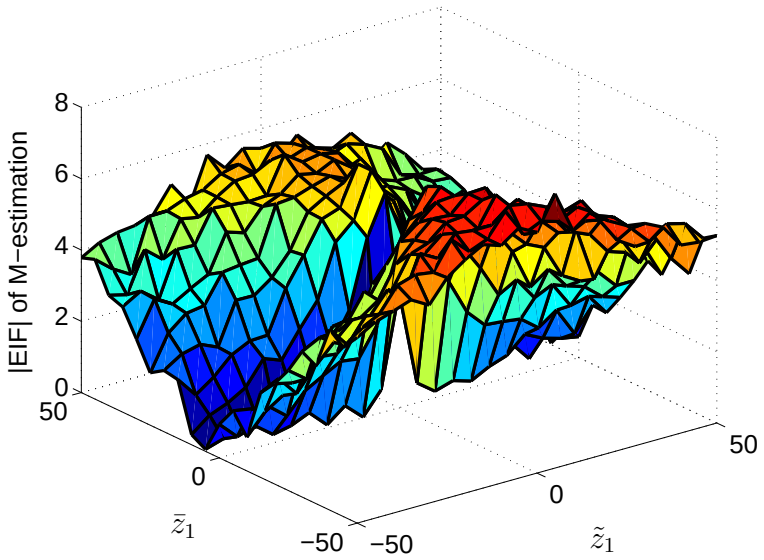

Figure 5.8. Empirical influence function of the M-estimation based STFD matrix estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

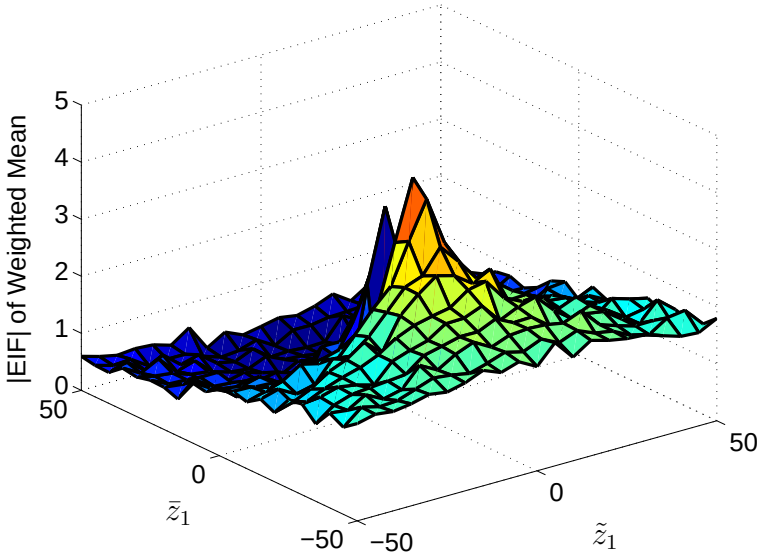

Figure 5.9. Empirical influence function of the one-step re-weighted STFD estimator for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

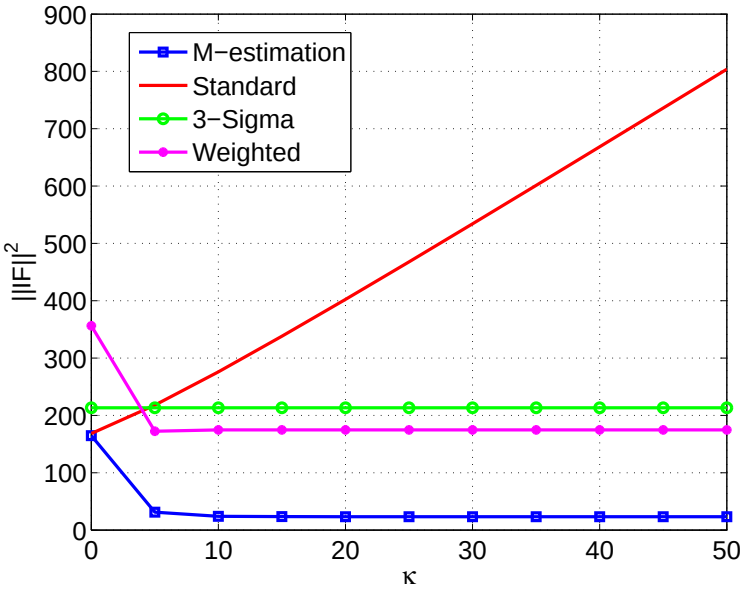

Figure 5.10. L_2 norm of influence functions for $m = 4$ sensors as a function of the outlier magnitude κ for $p = q = 1$ and with parameters as described at the beginning of Section 5.4.

Chapter 6
Conclusions and Outlook

In this thesis, DOA estimation of non-stationary signals in the presence of impulsive noise has been considered. A robust algorithm for DOA estimation which employs robust IFreq and robust STFD matrices estimation was developed. Moreover, a framework for robustness analysis of STFD matrices based on the influence function has been proposed. Conclusions are drawn in Section 6.1 and future directions of this work are provided in 6.2.

6.1 Conclusions

In this thesis, algorithm for robust DOA estimation of non-stationary signals is presented. This algorithm is based on robust IFreq and STFD matrices estimation.

For robust IFreq estimation, the proposed technique uses morphological image processing on an averaged TFD image obtained by computing robust median based MBD across sensors. Moreover, the proposed technique is non-parametric and does not require any prior knowledge about the number of sources and their IFreq signatures. Also, an automatic threshold based on iterative technique was suggested. Moreover, metrics to assess the performance of morphological image processing based technique were given.

In this dissertation, robust methods for STFD matrix estimation have been proposed. The robust methods are of three different types, i.e. the pre-processing based, the robust position based and the non-iterative robust techniques. The pre-processing based methods are computationally efficient and are a suitable candidate for robust non-stationary DOA estimation in highly impulsive noise environments. The robust position based techniques besides robustness in the case of impulsive noise also offer efficiency under nominal noise distribution. The robust position based methods are computationally more expensive than the pre-processing based approaches. Non-iterative techniques, on the other hand provide a reasonable compromise between efficiency and the computational complexity. All of the proposed robust approaches are able to mitigate the deleterious effects of impulsive noise and provide improved DOA estimation when compared to the standard methods and the robust methods from conventional array processing.

The robustness analysis of the proposed robust STFD matrix estimators on the basis of influence function was proposed. Analytical expressions for the influence functions of different kinds of STFD matrix estimators have been derived. Also, a definition for the finite sample counterpart of the influence function is provided. Both the analytical and the empirical results confirm that the robust techniques have indeed bounded influence functions as compared to the standard non-robust approaches.

6.2 Outlook

The future directions for the work are discussed in this Section.

- In this work, narrow-band signal model is assumed. However, in some applications, such as sonar or in seismology the signals are wideband [4]. Therefore, robust methods in case of wideband signals can be investigated.
- TF array processing allows to perform DOA estimation and/or BSS of non-stationary signals in under-determined setup, i.e. more sources than the number of sensors can be resolved. There, it is required to find minimum bounds and conditions on the resolvability of sources in the TF domain.
- As a further robustness analysis, influence functions for eigenvalues and eigenvectors obtained by eigenvalue decomposition of STFD matrices can also be investigated. This analysis can be then further used to predict breakdown of the TF MUSIC similar to performance bound of conventional MUSIC [126].

List of Acronyms

DOA	Direction-of-arrival
TF	Time-frequency
TFD	Time-frequency distribution
STFD	Spatial time-frequency distribution
i.i.d.	independent and identically distributed
WVD	Wigner-Ville distribution
IFreq	Instantaneous frequency
ML	Maximum-likelihood
MT	Modulus transformation
MMT	Modified Modulus Transformation
FLOM	Fractional-lower-order-moment
RMSE	Root Mean Square Error
BSS	Blind Source Separation
pdf	probability density function
PWVD	Pseudo-Wigner-Ville distribution
MBD	Modified-B-Distribution
MCD	Minimum Covariance Determinant
MUSIC	Multiple Signal Classification
WHT	Wigner Hough Transform
PWHT	Pseudo-Wigner Hough Transform
SNR	Signal-to-Noise Ratio
EIF	Empirical Influence Function
MBC	Maximum Bias Curve

List of Symbols

$\boldsymbol{x}$	The array observation vector
$\boldsymbol{a}(\theta)$	The array steering vector for direction θ
x_p	The sensor signal for the p^{th} sensor
$\boldsymbol{A}$	Array response matrix
$\boldsymbol{s}(t)$	sources' vector at time t
$f_k(t)$	Instantaneous frequency of k^{th} source
$\boldsymbol{n}(t)$	The noise vector at time t
$n_p(t)$	The noise at the p^{th} sensor at time t
$\boldsymbol{\mu}$	Vector describing the mean of a multivariate distribution
$\boldsymbol{\sigma^2}$	The variance of noise
$\mathcal{N}_{\mathbb{C}}^{m}(\boldsymbol{\mu}, \sigma^2\boldsymbol{I})$	m-dimensional circular complex Gaussian distribution
$\boldsymbol{R}$	The covariance matrix
E	structuring element
IF	influence function
E	Statistical expectation
F	The noise distribution
ϵ	The fraction of contamination
κ	The multiplying factor for the power of an outlier
F_ϵ	The contaminated distribution
f	the frequency
$\boldsymbol{D}(t, f)$	Time frequency distribution matrix at time t and frequency f
$[\cdot]$	The matrix
$\|\|\|$	Frobenious norm
k	The variable denoting the source number
K	Total number of sources
L	The window length
$D_{x_p x_p}(t, f)$	The auto-TFD
RD	The robust distances
m	Total number of sensors
N	total number of observations
SNR	signal to noise ratio
$\lvert \cdot \rvert$	Absolute value
$\mathbb{C}$	The set of complex numbers

$\mathbb{R}$	The set of real numbers
MD	Mahalanobis distance
$\text{sign}(\cdot)$	The sign function
$a_p(\theta)$	element of steering vector for p^{th} sensor
P_{MUSIC}	MUSIC spectrum
$\rho(\cdot)$	The ρ function
c_T	constant for Tukey's bi-weight function
c_H	constant for Huber's ρ-function
$\psi_H(\cdot)$	psi function for Huber's ρ-function
$\boldsymbol{\eta}$	Robust mean of observations
$\varphi(m,l)$	Kernel function for TF distributions
α_p^k	p^{th} parameters for IFreq of k^{th} signal
$\boldsymbol{I}$	Identity matrix of $m \times m$
$\mathbb{I}(\cdot)$	Indicator function
$\boldsymbol{R_{ss}}$	Sources' covariance matrix
$\boldsymbol{D_{ss}}(t,f)$	Sources' TFD matrix
$\boldsymbol{u}_i$	eigenvectors of covariance/STFD matrix
λ_i	eigenvalues of covariance/STFD matrix
$\boldsymbol{E}_s$	Signal subspace
$\boldsymbol{E}_n$	Noise subspace
$P_{BF}(\theta)$	Beamformer spectrum
σ_p	scale of noise across p^{th} sensor
$\hat{\boldsymbol{R}}_{rob}$	Robust covariance matrix estimator
$\log_{10}$	logarithm of base 10
W_1	weight function for M-estimator of mean
W_2	weight function for M-estimator of covariance
$\phi_k(t)$	Instantaneous phase of k^{th} signal
m_s	mean of TF point belonging to signals
m_n	mean of TF point belonging to noise
median	median
$MADN$	Normalized median absolute deviations
$\Re(\cdot)$	real value
$\Im(\cdot)$	Imaginary value
$\psi(x)$	3-σ operation
$\hat{D}_\infty$	Asymptotic estimate

$\boldsymbol{z}$	vector indicating an outlier
F_ϵ	ϵ-contaminated distribution
δ_{z_p}	Point mass distribution
Erf	error function
r_p	magnitude of p^{th} sensor signal
ϕ_p	Phase of p^{th} sensor signal
$f_n(n)$	probability density function of n
$\mathrm{E_F}$	Nominal distribution
$\mathrm{IF}(z_p; \hat{D}_{x_p x_p}, \mathrm{F})$	Influence function of auto-sensor TFD
$\mathrm{IF}(z_p, z_q; \hat{D}_{x_p x_q}, \mathrm{F})$	Influence function of cross-sensor TFD
γ	dispersion parameter of Cauchy distribution
$\boldsymbol{\Sigma}$	Covariance matrix
EIF	Empirical influence function
$\mathcal{E}$	Expected value
$\boldsymbol{D}_{MCD}$	MCD based STFD matrix
$\bar{x}_p$	The real value of x_p
$\tilde{x}_p$	The imaginary value of x_p
$\acute{x}$	pre-processed observation
$\mathcal{S}$	The set of TF points
$\#\mathcal{S}$	Number of points in the set
$\bigcup$	Union operation
$\odot$	element-wise multiplication
I	image
x^*	Complex conjugate
P_{WHT}	The cost function for WHT
β	kernel parameter for MBD
l	window index variable
$\mathcal{M}$	Set of STFD matrices

Bibliography

[1] B. Boashash, *Time Frequency Signal Analysis and Processing - A Comprehensive Reference*, Elsevier, 2003.

[2] A. M. Zoubir and D. R. Iskander, "Bootstrap modeling of a class of nonstationary signals," *IEEE Transactions on Signal Processing*, vol. 48, no. 2, pp. 399–408, 2000.

[3] A. W. Rihaczek, *Principles of High-Resolution Radar*, Boca Raton, FL Peninsula, 1985.

[4] R. Altes, "Sonar for generalized target description and its similarity to animal echolocation systems," *J. Acoust. Soc. Amer.*, vol. 59, pp. 97–105, 1976.

[5] H Krim and M. Viberg, "Two decades of array processing research: The parametric approach," *IEEE Signal Processing Magazine*, 1996.

[6] T. E. Tuncer and B. Friedlander, *Classical and modern direction-of-arrival estimation*, Academic Pr, 2009.

[7] L. Cohen, "Time-frequency distributions-a review," *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941–981, 1989.

[8] L. Cohen, *Time-frequency analysis*, Prentice-Hall, 1995.

[9] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representations," *IEEE Transactions on Signal Processing*, vol. 46, no. 11, pp. 2888–2897, 1998.

[10] M. G. Amin and Y. Zhang, "Direction finding based on spatial time-frequency distribution matrices," *Digital Signal Processing*, vol. 10, no. 4, pp. 325–359, Oct 2000.

[11] Y. Zhang and M. G. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2892–2902, 2000.

[12] A. M. Zoubir L. A. Cirillo and M. G. Amin, "Direction finding of nonstationary signals using a time-frequency Hough transform," in *Proc. of IEEE International Conference on Acoustic Speech and Signal Processing*, Philadelphia, USA, Mar. 2005.

[13] P. Heidenreich, L. A. Cirillo, and A. M. Zoubir, "Direction Finding of Nonstationary Signals using Spatial Time-Frequency Distributions and Morphological Image Processing," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*. Apr. 2007, pp. III–1137–III–1140, Ieee.

[14] P. Heidenreich, L. A. Cirillo, and A. M. Zoubir, "Morphological image processing for FM source detection and localization," *Signal Processing*, vol. 89, no. 6, pp. 1070–1080, June 2009.

[15] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. M. Zoubir, "Blind Separation of Nonstationary Sources," *IEEE Signal Processing Letters*, vol. 11, no. 7, pp. 605–608, July 2004.

[16] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1830–1847, 2004.

[17] K. Kim and G. Shevlyakov, "Why gaussianity?," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 102 –113, march 2008.

[18] D. M. Kriztman T. K. Blankenship and T. S. Rappaport, "Measurements and simulation of radio frequency impulsive noise in hospitals and clinic.," in *IEEE 47th Vehicular Technology Conference (VTC)*. 1997, IEEE.

[19] Y. I. Abramovich and P. Turcaj., "Impulsive noise mitigation in spatial and temporal domains for surface-wave over-the- horizon radar.," 1999.

[20] X. Guo, H. Sun, T. S. Yeo, and Y. Lu, "Lightning Interference Cancellation in High-Frequency Surface Wave Radar," *2006 CIE International Conference on Radar*, vol. 1, pp. 1–4, Oct. 2006.

[21] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods an results for class A and class B noise models," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[22] K.L. Blackard, T.S. Rappaport, and C.W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 11, no. 7, pp. 991 –1001, sep 1993.

[23] H.P. Hsu, R.M. Storwick, D.C. Schlick, and G.L. Maxam, "Measured amplitude distribution of automotive ignition noise," *Electromagnetic Compatibility, IEEE Transactions on*, vol. EMC-16, no. 2, pp. 57 –63, may 1974.

[24] M.D. Button, J.G. Gardiner, and I.A. Glover, "Measurement of the impulsive noise environment for satellite-mobile radio systems at 1.5 ghz," *Vehicular Technology, IEEE Transactions on*, vol. 51, no. 3, pp. 551 –560, may 2002.

[25] X. Dong, H. Weng, D. G. Beetner, T. H. Hubing, D. C. Wunsch, M. Noll, H. Goksu, and B. Moss, "Detection and identification of vehicles based on their unintended electromagnetic emissions," *Electromagnetic Compatibility, IEEE Transactions on*, vol. 48, no. 4, pp. 752 –759, nov. 2006.

[26] M. G. Sanchez, L. de Haro, M. C. Ramon, A. Mansilla, C. M. Ortega, and D. Oliver, "Impulsive noise measurements and characterization in a uhf digital tv channel," *Electromagnetic Compatibility, IEEE Transactions on*, vol. 41, no. 2, pp. 124 –136, may 1999.

[27] –, "Noaa space weather scale," June 2009.

[28] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurements," *Communications, IEEE Transactions on*, vol. 21, no. 11, pp. 1232 – 1241, nov 1973.

[29] R. A. Shepherd, "Measurements of amplitude probability distributions and power of automobile ignition noise at hf," *Vehicular Technology, IEEE Transactions on*, vol. 23, no. 3, pp. 72 – 83, aug 1974.

[30] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust Estimation in Signal Processing," *IEEE Signal Processing Magazine*, 2012, To appear.

[31] I Djurovic, L Stankovic, and F Boehme, "Robust L -Estimation Based Forms of Signal Transforms and Time-Frequency Representations," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1753–1761, 2003.

[32] L. Stankovic, I. Djurovic, and S. Stankovic, "The robust Wigner distribution," in *Proceedings of the Acoustics, Speech, and Signal Processing, 2000. on IEEE International Conference-Volume 01*. 2000, number 381, pp. 77–80, IEEE Computer Society.

[33] I. Djurovic and L. Stankovic, "Robust Wigner distribution with application to the instantaneous frequency estimation," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 2985–2993, 2001.

[34] M. Sahmoudi and K. Abed-Meraim, "Robust quadratic time-frequency distributions for the analysis of mulicomponent FM signals in heavy-tailed noise," *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*, vol. 2, pp. 871–876, 2005.

[35] W. Sharif, P. Heidenreich, and A.M. Zoubir, "Robust Direction-Of-Arrival Estimation For FM Sources In The Presence Of Impulsive Noise," in *In 35th IEEE Conference on Audio, Acoustics Speech and Signal Processing (ICASSP)*, Dallas, USA, 2010, vol. 1, pp. 3662–3665.

[36] W. Sharif, Y. Chakhchoukh, and A. M. Zoubir, "Direction-of-arrival estimation of FM sources based on robust spatial time-frequency distribution matrices," in *Proc. of IEEE Workshop on SSP*, Nice, France, June 2011, pp. 537–540.

[37] P Huber, *Robust Statistics*, Wiley, 2009.

[38] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust statistics: Theory and methods*, Wiley Series in Probability and Statistics, 2006.

[39] H. Krim and M. Viberg, "Two Decades Of Array Signal Processing Research: The Parametric Approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[40] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, 1990.

[41] R. L. Burden, J. D. Faires, and A. C. Reynolds, *Numerical Analysis*, Number Second Edition. Prindle, Weber, & Schmidt, Boston, 1981.

[42] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proceedings of the IEEE*, vol. 86, no. 2, pp. 421–441, 1998.

[43] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proceedings of the IEEE*, vol. 86, no. 2, pp. 421–441, 1998.

[44] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[45] P. Stoica and A Nehorai, "MUSIC, Maximum Likelihood, and Cramer-Rao Bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 5, pp. 720–740, 1989.

[46] B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2154–2163, 1995.

[47] J. Li, D. Zheng, and P. Stoica, "Angle and waveform estimation via RELAX," *IEEE Transactions on Signal Processing*, vol. 33, no. 3, pp. 1077–1087, jul 1997.

[48] B. M. Sadler and R. J. Kozick, "Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3520–3535, 2000.

[49] P. Tsakalides and C. L. Nikias, "Maximum likelihood localization of sources in noise modeled as a stable," *IEEE Transactions on Signal Processing*, 1995.

[50] S. Visuri, H. Oja, and V. Koivunen, "Subspace-Based Direction-Of-Arrival Estimation Using Nonparametric Statistics," vol. 49, no. 9, pp. 2060–2073, 2001.

[51] D. Middleton, "Non-{G}aussian Noise Models in Signal Processing for Telecommunications: New Methods and Results for Class {A} and Class {B} Noise Models," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.

[52] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *Signal Processing, IEEE Transactions on*, vol. 42, no. 10, pp. 2664 –2677, oct 1994.

[53] B. M. Sadler, G. B. Giannakis, and S. Shamsunder, "Noise subspace techniques in non-Gaussian noise using cumulants," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 3, pp. 1009–1018, 1995.

[54] D. B. Williams and D. H. Johnson, "Robust Estimation ofStructured Covariance Matrices," *IEEE Transactions on Signal Processing*, vol. 41, no. 9, pp. 2891–2906, 1993.

[55] C-H. Lim, S. C-M. See, A. M. Zoubir, and B. P. Ng, "Robust Adaptive Trimming for High-Resolution Direction Finding," *IEEE Signal Processing Letters*, vol. 16, no. 7, pp. 580–583, July 2009.

[56] C.-H. Lim, J. P. Lie, C.-M. S. See, and A. M. Zoubir, "Modified modulus transformation for high resolution direction finding," in *17th European Signal Processing Conference (EUSIPCO)*, Glasgow-Scotland, 2009, number Eusipco, pp. 1126–1130.

[57] N. R. John and J. A. Draper, "An alternative family of transformations," *Applied Statistics*, vol. 29, no. 2, pp. 190–197, 1980.

[58] D.B. Williams and D.H. Johnson, "Robust maximum-likelihood estimation of structured covariance matrices," *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, pp. 2845–2848, 1988.

[59] E. Ollila and V. Koivunen, "Robust antenna array processing using m-estimators of pseudo-covariance," in *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, sept. 2003, vol. 3, pp. 2659 – 2663 vol.3.

[60] M. Mahot, F. Pascal, P. Foster, and J. P. Ovarlez, "Robust covariance matrix estimates with attractive asymptotic properties," in *In 4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, San Juan, Puerto Rico, Dec. 2011.

[61] D. E. Tyler, "Robustness and efficiency properties of scatter matrices," *Biometrika*, vol. 70, no. 2, pp. 411–420, 1983.

[62] H. P. Lopuhaa, "On the relation between S-estimators and M-estimators of multivariate location and covariance," *The Annals of Statistics*, pp. 1662–1683, 1989.

[63] N. A. Campbell, H. P. Lopuhaä, and P. J. Rousseeuw, "On the calculation of a robust S-estimator of a covariance matrix.," *Statistics in medicine*, vol. 17, no. 23, pp. 2685–95, Dec. 1998.

[64] H. P. Lopuhaä, "Asymptotic expansion of s-estimators of location and covariance," 1995.

[65] P. L. Davies, "Asymptotic behaviour of s-estimates of multivariate location parameters and dispersion matrices," *The Annals of Statistics*, vol. 15, no. 3, pp. 1269–1292, Sept. 1987.

[66] M. Salibian-barrera and V. J. Yohai, "A fast algorithm for s-regression estimates," *J. Computat. Graphic. Statist*, vol. 15, pp. 414–427, 2006.

[67] K. S. Tatsuoka and D. E. Tyler, "The uniqueness of s and m-functionals under non-elliptical distributions," *The Annals of Statistics*, vol. 28, pp. 1219–1243, 2000.

[68] M. Salibian-Barrera, S. Van Aelst, and G. Willems, "PCA based on Multivariate MM-estimators with Fast and Robust Bootstrap," *Journal of the American Statistical Association*, , no. 101, pp. 1198 – 1211, 2005.

[69] P. J. Rousseeuw and K. Van Driessen, "A fast algorithm for the minimum covariance determinant," *Technometrics*, vol. 43, no. 3, 1999.

[70] F. Gini, M. Montanari, and L. Verrrazzani, "Estimation of Chirp Radar Signals in Approach: A Cyclostationary Approach," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1029–1039, 2000.

[71] T. Chen, W. Jin, and J. Li, "Feature extraction using surrounding-line integral bispectrum for radar emitter signal," *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, , no. 1, pp. 294–298, June 2008.

[72] L. Cohen, "Time-Frequency Distributions-A Review," *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941–981, 1989.

[73] M. G. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2892–2902, 2000.

[74] A. Belouchrani and M. G. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 6, no. 5, pp. 109–110, May 1999.

[75] L. Cirillo, A. M. Zoubir, and M. G. Amin, "Parameter Estimation for Locally Linear FM Signals Using a Time-Frequency Hough Transform," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4162–4175, Sept. 2008.

[76] Y. Zhang, W. Mu, and M. G. Amin, "Maximum likelihood methods for array processing based on time-frequency distributions," *Advanced signal processing algorithms, architectures, and implementations IX*, pp. 502–513, 1999.

[77] Y. Zhang, W. Mu, , and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 747–759, Apr 2001.

[78] A. Belouchrani, M. G. Amin, and K. Abed-Meraim, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices," *IEEE Signal Processing Letters*, vol. 4, no. 9, pp. 266–268, 1997.

[79] C. Fevotte and F.J. Theis, "Pivot selection strategies in Jacobi joint block-diagonalization," *Lecture Notes in Computer Science*, vol. 4666, pp. 177, 2007.

[80] I. Djurović, V. Katkovnik, and L. J. Stankovi, "Median filter based realizations of the robust time-frequency distributions," *Signal Processing*, vol. 81, no. 8, pp. 1771–1776, 2001.

[81] B. Barkat and L. Stanković, "Analysis of polynomial fm signals corrupted by heavy-tailed noise," *Signal Processing*, vol. 84, no. 1, pp. 69–75, Jan. 2004.

[82] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal: A tutorial review-Part 2: algorithms and applications," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 540–568, 1992.

[83] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 540–568, apr 1992.

[84] Y. Gao, Y. Guo, Y. L. Chi, and S. R. Qin, "Order Tracking Based on Robust Peak Search Instantaneous Frequency Estimation," *Journal of Physics: Conference Series*, vol. 48, pp. 479–484, 2006.

[85] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multi-component FM signals," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on.* 2000, vol. 2, pp. II657–II660, IEEE.

[86] A. Akan, M. Yalcin, and L. F. Chaparro, "An iterative method for instantaneous frequency estimation," *ICECS 2001. 8th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.01EX483)*, vol. 4, pp. 1335–1338.

[87] L Rankine, M Mesbah, and B Boashash, "IF estimation for multicomponent signals using image processing techniques in the timefrequency domain," *Signal Processing*, vol. 87, no. 6, pp. 1234–1250, 2007.

[88] B-h. Wang and J-g. Huang, "Instantaneous frequency estimation of multicomponent Chirp signals in noisy environments," *Journal of Marine Science and Application*, vol. 6, no. 4, pp. 13–17, Apr. 2008.

[89] B. Porat and B. Friedlander, "Estimation of spatial and spectral parameters of multiple sources," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 412–425, 1983.

[90] J. Angeby, "Structured autoregressive instantaneous phase and frequency estimation," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1768–1771, 1995.

[91] P. Shan and A. A. L. Beex, "High-resolution instantaneous frequency estimation based on time-varying AR modeling," *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis (Cat. No.98TH8380)*, pp. 109–112.

[92] B. Barkat and B. Barakat, "IF Estimation Of Higher-Order Polynomial FM Signals Corrupted By Multiplicative And Additive Noise," , no. 2, 2001.

[93] S. Barbarossa and A. Zanalda, "A combined Wigner-Ville and Hough transform for cross-terms suppression and optimal detection and parameter estimation," *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 173–176, 1992.

[94] S Barbarossa, "Analysis of Multicomponent LFM Signals by a Combined Wigner-Hough Transform," *IEEE Transaction on Signal Processing*, vol. 43, no. 6, pp. 1511–1515, 1995.

[95] S. Barbarossa and O. Lemoine, "Analysis of Non-linear FM Signals by Pattern Recognition of Their Time-Frequency Representation," *IEEE Signal Processing Letters*, vol. 3, no. 4, pp. 112–115, 1996.

[96] X. Li, G. Bi, S. Stankovic, and A. M. Zoubir, "Local polynomial fourier transform: A review on recent developments and applications," *Signal Process.*, vol. 91, no. 6, pp. 1370–1393, June 2011.

[97] A. K. Chan and C. K. Chui, "Linear frequency-modulated signal detection using Radon-ambiguity transform," *IEEE Transactions on Signal Processing*, vol. 46, no. 3, pp. 571–586, Mar. 1998.

[98] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[99] S. Barbarossa, A. Scaglione, and G. B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial-phase signal modeling," *IEEE Transactions on Signal Processing*, vol. 46, no. 3, pp. 691–708, Mar. 1998.

[100] D. S. Pham and A. M. Zoubir, "Analysis of multicomponent polynomial phase signals," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 56–65, Jan. 2007.

[101] S. Stankovic, N. Zaric, and C. Ioana, "General form of time-frequency distribution with complex-lag argument," *Electronics Letters*, vol. 44, no. 11, pp. 699 –701, 22 2008.

[102] "Picture Thresholding Using an Iterative Selection Method," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, no. 8, pp. 630–632, Aug. 1978.

[103] T.N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 901 –914, apr 1992.

[104] J. Brown and A. Hoger, "A morphological point thinning algorithm," *Pattern Recogn. Lett.*, vol. 17, no. 2, pp. 197–207, Feb. 1996.

[105] B. Barkat and L. Stankovic, "Analysis of polynomial fm signals corrupted by heavy-tailed noise," *Signal Processing*, vol. 84, pp. 69–75, 2004.

[106] D. S. Pham and A. M. Zoubir, "Estimation of multicomponent polynomial phase signals with missing observations," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 1710–1715, apr 2008.

[107] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2232–2239, 2001.

[108] B. Friedlander and B. Porat, "A general lower bound for parametric spectrum estimation," *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on*, vol. 32, no. 4, pp. 728–733, 1984.

[109] W. Sharif, Y. Chakhchoukh, and A. M. Zoubir, "Robust spatial time-frequency distribution matrix estimation with application to direction-of-arrival estimation," *Signal Processing*, vol. 91, no. 11, pp. 2630–2638, November 2011.

[110] A. Belouchrani, M. G. Amin, and K. Abed-Meraim, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices," *IEEE Signal Processing Letters*, vol. 4, no. 9, pp. 266–268, 1997.

[111] Y. Zhang, W. Mu, and M. G. Amin, "Time-frequency maximum likelihood methods for direction finding," *Journal of the Franklin Institute*, vol. 337, no. 4, pp. 483–497, 2000.

[112] M. G. Amin and Y. Zhang, "Direction Finding Based on Spatial Time-Frequency Distribution Matrices," *Digital Signal Processing*, vol. 10, no. 4, pp. 325–339, 2000.

[113] Y. Zhang, "Spatial polarimetric time-frequency distributions and applications to direction-of-arrival estimation," *Proceedings of SPIE*, pp. 75–85, 2003.

[114] L. A. Cirillo, A. M. Zoubir, and M. G. Amin, "Blind Source Separation in the Time-Frequency Domain Based on Multiple Hypothesis Testing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2267–2279, June 2008.

[115] M. Djeddi and M. Benidir, "Robust Polynomial Wigner-Ville Distribution For The Analysis Of Polynomial Phase Signals In α-Stable Noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2004, number 3, pp. 613–616.

[116] S. Visuri, H. Oja, and V. Koivunen, "Subspace-based direction-of-arrival estimation using nonparametric statistics," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 2060–2073, 2001.

[117] M. Sahmoudi and K. Abed-Meraim, "Robust quadratic time-frequency distributions for the analysis of mulicomponent FM signals in heavy-tailed noise," *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*, vol. 2, pp. 871–876, 2005.

[118] D. Peña and F. J. Prieto, "Multivariate Outlier Detection and Robust Covariance Matrix Estimation," *Technometrics*, vol. 43, no. 3, pp. 286–310, Aug. 2001.

[119] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley-Interscience, New York, 1987.

[120] W. Sharif, M. Muma, and A. M. Zoubir, "Robustness analysis of spatial time-frequency distributions based on the influence function," *IEEE Transactions on Signal Processing*, 2013.

[121] F. R. Hampel, "The influence curve and its role in robust estimation," *J. Am. Statist. Assoc.*, vol. 69, pp. 383393, 1974.

[122] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust statistics: the approach based on influence functions*, Wiley Series in Probability and Statistics, 1986.

[123] E. Ollila and V. Koivunen, "Influence function for array covariance matrix estimators," in *Proc. of IEEE Workshop on SSP*, Missouri, USA, Sep 2003.

[124] E. Ollila and V. Koivunen, "Influence function and asymptotic efficiency of scatter matrix based array processors: Case MVDR beamformer," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 247–259, 2009.

[125] M. D. Springer, *The algebra of random variables*, A volume in Wiley series in probability and mathematical statistics, 1979.

[126] B.A. Johnson, Y.I. Abramovich, and X. Mestre, "Music, g-music, and maximum-likelihood performance breakdown," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3944 –3958, aug. 2008.

Curriculum vitae

Name:	Waqas Sharif
Date of birth:	13.08.1981
Place of birth:	Gujrat, Pakistan
Family status:	Married, a son

Education

10/2004 - 11/2007	Technische Universität Darmstadt M.Sc. in Information & Communication Engineering
01/2000 - 02/2004	University of Engg. & Tech. Taxila, Pakistan B.Sc. in Electrical Engineering
08/1999	Govt. Zamindar College Gujrat, Pakistan Higher Secondary School Certificate

Work experience

03/2009 - 04/2012	Technische Universität Darmstadt Signal Processing Group Ph.D. Student
05/2006 - 02/2009	Nokia Siemens Networks Development Center Düsseldorf Systems Engineer
05/2007 - 11/2007	Honda Research Institute Europe Offenbach, Germany Master Thesis