MAXIMUM LIKELIHOOD PARAMETER ESTIMATION UNDER SIGNAL INTERMITTENCY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ŞAFAK BILGI AKDEMIR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

DECEMBER 2023

Approval of the thesis:

MAXIMUM LIKELIHOOD PARAMETER ESTIMATION UNDER SIGNAL INTERMITTENCY

submitted by **ŞAFAK BILGI AKDEMIR** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. İlkay Ulusoy
Head of Department, **Electrical and Electronics Engineering** _____

Prof. Dr. Çağatay Candan
Supervisor, **Electrical and Electronics Engineering** _____

Examining Committee Members:

Prof. Dr. T. Engin Tuncer
Electrical and Electronics Engineering, METU _____

Prof. Dr. Çağatay Candan
Electrical and Electronics Engineering, METU _____

Prof. Dr. Umut Orguner
Electrical and Electronics Engineering, METU _____

Prof. Dr. Sinan Gezici
Electrical and Electronics Engineering, Bilkent University _____

Prof. Dr. Orhan Arıkan
Electrical and Electronics Engineering, Bilkent University _____

Date:25.12.2023

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Şafak Bilgi Akdemir

Signature :

ABSTRACT

MAXIMUM LIKELIHOOD PARAMETER ESTIMATION UNDER SIGNAL INTERMITTENCY

Bilgi Akdemir, Şafak
Ph.D., Department of Electrical and Electronics Engineering
Supervisor: Prof. Dr. Çağatay Candan

December 2023, 115 pages

In this thesis, the parameter estimation problem under signal intermittency is explored. Firstly, the maximum-likelihood (ML) direction of arrival (DOA) estimation problem is studied for the intermittent jamming scenario. The intermittent jamming modality is based on the assumption that only a subset of the collected snapshots are contaminated by the jammer while the other snapshots are jammer-free; but the receiver does not know which is which. Solving problem analytically using maximum likelihood approach is intractable and for this reason an expectation maximization (EM) method based solution is developed. Secondly, complex exponential frequency estimation of an intermittent signal problem is addressed. As a reason for this intermittency in the signal, one may assert the operational state of frequency measuring sensor, say healthy and unhealthy state. In the healthy state, the sensor produces high quality output meaning that the signal to noise ratio is high; whereas in the faulty state the exact opposite of this situation is true. This phenomena is represented by an additive noise which is modeled as a correlated mixture of two Gaussian distributions. A computationally efficient method for the frequency estimation under this setting is proposed by utilizing again EM method. Cramer-Rao type performance

lower bounds are also derived to compare the performance of the suggested estimators. Finally, the performance improvement achieved by the proposed methods over the competing methods are shown by numerical simulations.

Keywords: Direction of arrival estimation, Jamming, Intermittent interference, Mixture of Gaussians, Frequency estimation, Maximum-likelihood estimation, Expectation-maximization algorithm, Cramer-Rao bound

ÖZ

KESİKLİ SİNYAL ALTINDA EN YÜKSEK OLABİLİRLİKLE PARAMETRE KESTİRİMİ

Bilgi Akdemir, Şafak
Doktora, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. Çağatay Candan

Aralık 2023 , 115 sayfa

Bu tezde kesikli sinyal altında parametre tahmini problemi incelenmiştir. İlk olarak, aralıklı karıştırma senaryosu için en büyük olasılıklı (ML) varış yönü (DOA) tahmin problemi çalışılmıştır. Aralıklı karıştırma yaklaşımı, toplanan örneklerin yalnızca bir alt kümesinin bozucu tarafından kirletildiği, diğer örneklerde ise bozucu sinyal olmadığı varsayımına dayanmaktadır; ancak alıcı hangisinin hangisi olduğunu bilmemektedir. Problemin en büyük olabilirlik yöntemi ile analitik olarak çözümü zordur ve bu nedenle beklenti büyütme (EM) yöntemine dayalı bir çözüm geliştirilmiştir. İkinci olarak, kesikli bir sinyalde karmaşık üstel frekans tahmini problemi ele alınmaktadır. Bu sinyal kesikliliğinin nedeni olarak frekans ölçüm sensörünün çalışma koşulları ileri sürülebilir. Sağlıklı çalışma durumunda sensör yüksek kalitede çıktı üretir, bu da sinyal gürültü oranının yüksek olduğu anlamına gelir. Arızalı durumda ise bunun tam tersi doğrudur. Bu durum, iki bileşenli Gauss dağılımına sahip ve kendi aralarında ilişkili olarak modellenen toplamsal bir gürültü ile temsil edilmektedir. Bu şartlarda, frekans tahmini için yine EM yöntemini kullanarak hesaplama açısından verimli bir yöntem önerilmektedir. Önerilen kestiricilerin performansını karşılaştırmak

için Cramer-Rao tipi performans alt sınırları da türetilmiştir. Son olarak, önerilen yöntemlerin rakip yöntemlere göre elde ettiği performans artışı sayısal simülasyonlarla gösterilmiştir.

Anahtar Kelimeler: Geliş yönü kestirimi, Karıştırma, Kesikli girişim, Gaussian karışımları, Frekans kestirimi, En yüksek olabilirlik kestirimi, Beklenti en büyütme yöntemş, Cramer Rao sınırı

To My Grandma, From Whom I Learned Resillience,

To My Mum,

To My Spouse Özgür,

To My Lovely Children Zeynep Işık and Çınar,

and also

To the Memory of My Dad...

ACKNOWLEDGMENTS

First of all, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Çağatay Candan for his valuable guidance, suggestions and innovative ideas throughout my thesis study. Especially I would like to thank by heart for his patience and constant support. It is a very big chance for me to work with him.

Besides my advisor, I would like to thank to the thesis monitoring committee members Prof. Dr. Umut Orguner and Prof. Dr. Sinan Gezici for taking time off for my thesis study and for their insightful comments in all of the thesis monitoring meetings. Also, I would like to thank other thesis jury members, Prof. Dr. T. Engin Tuncer and Prof. Dr. Orhan Arıkan, for attending my jury and for their invaluable comments.

I would like to thank to Erdal Kılıçaslan (Toprak Işık), not only for sharing his knowledge about radar systems and electronic warfare, but also for all of our conversations about life, science, philosophy, history, academic life and daily stuff. Besides these, I am grateful to him for supporting me at the hardest times of this thesis study.

I am especially indebted to my mum for raising up my two children. Her endless efforts for taking care of my son along with my bedridden dad allowed me to take time off for my studies at the beginning of my Ph.D. program. She had also looked after my daughter during my master studies making it possible for me to complete the program at that time. Unfortunately, my dad passed away five years ago and could not see the completion of my thesis studies. Afterwards, my mum couldn't get back to her old healthy days and life has also changed for me forever. I am grateful to both of my mum and dad for their love and trust on me throughout my life.

Finally, I would like to thank my spouse Özgür, my daughter Zeynep Işık and my son Çınar for their love and understanding. I spent time that I should have spent with them and for them, at my studying table. I hope they accept the completion of this thesis as an appreciation for those times.

I also own special thanks to my spouse because of his picking me up from work before every single class and driving all the way to the department during all of my M.Sc. and Ph.D. studies. Those sacrifices of him helped me a lot to complete my graduate programs.

I can't help but mention that I have to thank and congratulate myself too, for my endless efforts to push my boundaries and learn continuously. Completion of this thesis as a wife and mother with two children, an employee having a full-time job with high responsibilities, a daughter of an old mother at home seeking for care is the biggest sign of my resilliance and doggedness.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

2D	2 Dimensional
3D	3 Dimensional
ADC	Analog-to-Digital Converter
AWGN	Additive White Gaussian Noise
BW	Band Width
COSRO	Conical Scan on Receive Only
CRB	Cramer Rao Bound
CRLB	Cramer Rao Lower Bound
DFT	Discrete Fourier Transform
DIFM	Digital Instantaneous Frequency Measurement
DOA / DoA	Direction of Arrival
DTFT	Discrete Time Fourier Transform
ECM	Electronic Counter Measures
EHF	Extremely High Frequency
ELINT	Electronic Intelligence
EM	Expectation Maximization Algorithm
ESM	Electronic Support Measures
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Techniques
EW	Electronic Warfare
FHSS	Frequency Hopped Spread Spectrum
FIM	Fisher Information Matrix
FPGA	Field Programmable Gate Array
HCRB	Hybrid Cramer Rao Bound

HMM	Hidden Markov Model
IFM	Instantaneous Frequency Measurement
IoT	Internet of Things
LORO	Lobe Switching on Receive Only
LS	Least Squares
MCB	Miller Chang Bound
MCRB	Modified Cramer Rao Bound
MFIM	Modified Fisher Information Matrix
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimate/Estimator
MVDR	Minimum Variance Distortionless Response
NLS	Nonlinear Least Squares
pdf	Probability Density Function
POI	Probability of Intercept
PRF	Pulse Repetition Frequency
RF	Radio Frequency
RWR	Radar Warning Receiver
SOI	Signal Of Interest
SMI	Sample Matrix Inversion
TDOA	Time Difference of Arrival
TOA	Time of Arrival
TWS	Track While Scan

CHAPTER 1

INTRODUCTION

In this thesis, we study the parameter estimation problem under signal intermittency. Conventional parameter estimation problems assume that the signal of interest is always available but observed under noise. Causes of signal intermittency in the problems examined within the scope of this thesis are of two types: In the first type, the presence of the signal in the collected observations is not certain. The intermittency is a property of the signal of interest (SOI) itself. In the second type, SOI exists in all snapshots but the sensor operational health state is intermittent causing intermittency in the received signal.

The first examined problem naturally occurs in many applications. For example, in the direction of arrival application, the signal return from the target of interest can be corrupted by the jammer at times. More generally, the performance of the sensing device can fluctuate between nominal operation and poor operation during the data collection time. Let us say that the probability of the signal being corrupted by jammer is non-zero. The receiver knows neither this probability nor jammer power and also receiver does not know which sample is corrupted by jammer. We study the effect of jammer signal intermittency in direction of arrival estimation of a target signal.

As a second application, we study the complex exponential frequency estimation of an intermittent signal. As a reason for this intermittency, one may assert the operational health of the frequency measuring sensor. It is assumed that the sensor has two states namely healthy and faulty. Similar to the first application, health state of the sensor may evolve on its own, independent of all other effects. In the healthy state, the sensor produces high quality output meaning that, the signal to noise ratio is high;

whereas in the faulty state the exact opposite of this situation is true. Different from the first application, we allow correlation between consecutive observations which is represented by a Hidden Markov Model (HMM). In this problem, the sensor state is hidden variable not available to us and our goal is to estimate the frequency in this set-up.

1.1 Motivation and Problem Definition

Direction of arrival (DOA) and frequency estimation are two important and basic problems of statistical signal processing which have wide application areas in electronic warfare, radar signal processing, wireless communications, sonar and so on. Due to importance and heavy use of these applications, intensive research has been going on for decades making this research area a mature field. Although signal models, unknown parameters, and statistical properties of signal components of the mentioned problems have a wide range of variety based on the application area, majority of the work in the field considers the case of persistent signals. Besides this, the presence of intermittent signals, alone or aside a persistent signal is also often encountered in practice.

Intentional or unintentional interference sources that occupy the spectrum together with the desired signal can be given as an example to intermittent signal aside a persistent signal. Radar/communication jammers are examples of intentional interferers, whereas a communication system that is jammed by a friendly short packet communication system turning on and off at random time instants can be considered as an example of unintentional interference [5].

The examined intermittent interference problem is encountered in both active and passive systems. In passive sensing, there is no signal transmission from the station; but the station listens the spectrum to detect transmissions of other platforms. For instance, the spectrum sensing in cognitive radio systems is a non-military application of this kind. On the other hand, any kind of Electronic Warfare (EW) receiver can be given as an example for military application of passive sensing systems. Passive radar systems that exploit transmissions of other emitters like terrestrial television

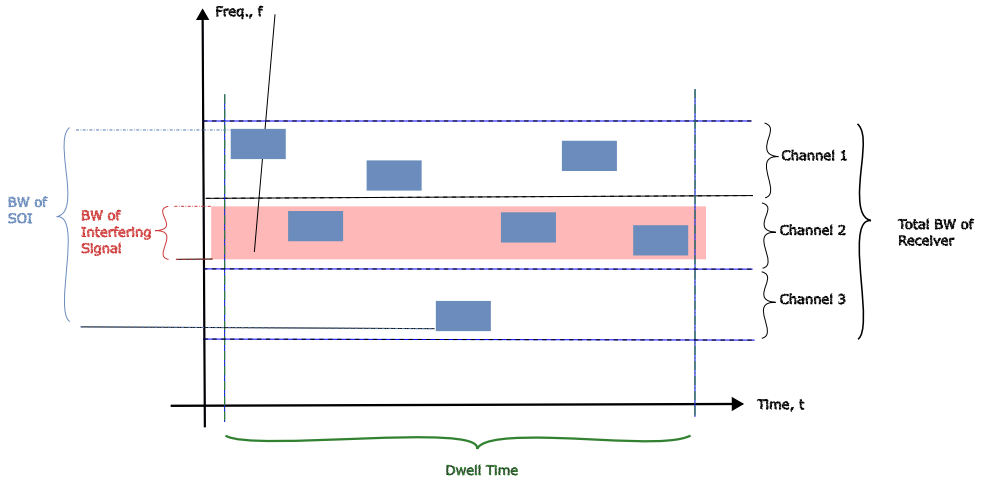
transmitters, FM radio transmitters etc. to detect targets is another application of passive sensing that is used in civil and military applications. As a specific example of a military application, one may consider a hypothetical scenario in which a channelized ESM/ELINT receiver collects signal samples from a specific frequency-agile target whose instantaneous bandwidth overlaps with an interfering signal's relatively narrow operational bandwidth as illustrated in Figure 1.1a. The output of the frequency channels may be processed jointly to improve the direction of arrival estimation performance of the system. Another example can be a relatively wideband ESM receiver that intercepts a target signal whose bandwidth is also occupied by an interfering system at random time instants from time to time as presented in Figure 1.1b. It is also known that, systems that use frequency hopped spread spectrum (FHSS) techniques like Internet of Things (IoT) and civil/military communication systems are vulnerable to intermittent interfering signals [6]. As an example to these kind of applications, a slow frequency hopping EHF (Extremely High Frequency) satellite communication system, whose operation bandwidth is occupied partly with an interferer is examined in [7] which is similar to the case represented in Figure 1.1a.

Contrary to the passive sensing systems, active sensing systems radiate energy to the environment and collect backscaterred energy from objects to extract target-related information. Radar is a typical active sensing system that is widely used in military and civil applications. After the detection process in a radar receiver, target related information like the range and velocity is extracted by comparing the received signal with the signal that was transmitted. To give an example of an intermittent interference at an active sensor, a narrow band swept noise jammer sweeping a wide range of frequencies whose bandwidth overlaps with the operational bandwidth of a radar system from time to time [8] can be considered in Figure 1.1c. This example may correspond to a hypotethical scenario in which a surveillance radar is affected by a jammer system that is not especially aimed at the specific radar under concern. Large antenna rotation periods of surveillance radars, on the order of ten seconds, may result in long coherent processing intervals making it possible for surrounding emitters interfere with the received signals by the radar system. It is known in the open literature that antenna scanning periods, beam widths and pulse repetition frequencies (PRF) of surveillance radars can be up to 15 seconds, 5 degrees and 750 Hz respectively. For

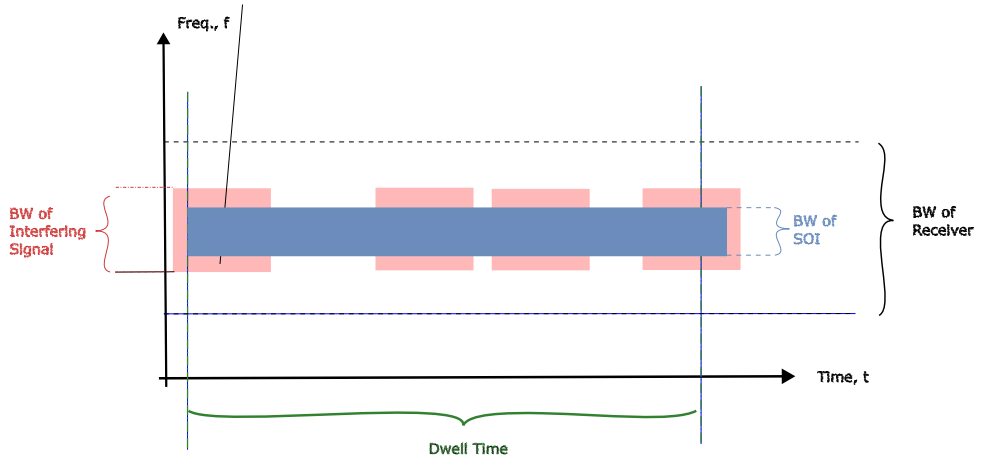
example, if a hypotethical surveillance radar with some kind of multifunction capability, which has beamwidth of 4 degrees, PRF of 500 Hz and antenna scanning period of 10 seconds, rotates with a resolution of 2 degrees, it stays at an angle for 55 ms and as a result transmits approximately 27 pulses at every angle. On the other hand, state-of-the-art Electronic Counter Measure (ECM) transmitters of today can apply ECM techniques of duration from on the order of miliseconds up to tens of seconds at a single target only or at multiple targets at once making it possible to interfere with the mentioned radar for very short durations of time.

Incoming direction of the SOI is a basic and important information that is extracted by the aforementioned receivers. It is obvious that a small improvement in DOA estimation performance may be very crucial in some applications, since it corresponds to a big distance at long ranges, say after polar coordinate to cartesian coordinate conversion. We would like to note that the running assumption in virtually all DOA estimation methods is the coexistence of multiple targets or target- jammer/interference signals in *all* snapshots. Different from these studies, we assume that jamming is intermittent; hence, not all but some snapshots are contaminated. In spite of its importance in theory and practice, as expressed by the previous examples, the topic has attracted limited attention in open literature. The main purpose of this study is to examine the amount of performance gain in DOA estimation when the intermittency is taken into account in the DOA estimation problem formulation.

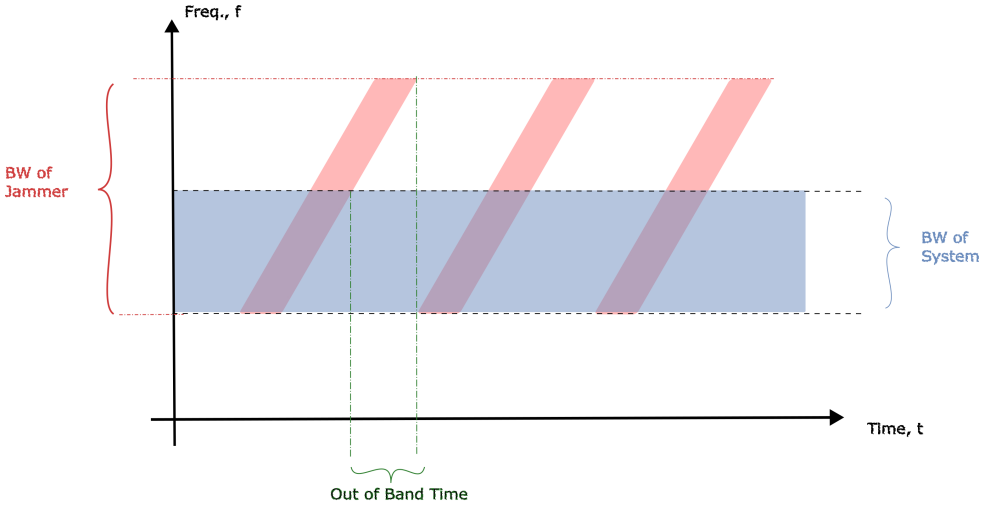
Besides DOA information, frequency is one of the most valuable informations that is exploited by military and commercial systems. As an example from military applications; frequency information is utmost important to ESM/ELINT systems since it can be the basic distinguishing property between a surveillance radar and an acquisition radar, which is a part of a fire control system. Radars also use frequency to estimate and track Doppler shift of the targets and this information is used to seperate stationary clutter from moving targets and classify targets according to their movement types. As far as we know, if not all, the majority of the existing literature concentrates on improving accuracy of estimated frequencies of single or multiple signals in white noise assuming that noise statistics do not change from sample to sample. But, in real life, due to various failures in sensors, received signal samples at the output of a sensor may not be identical in terms of assumed noise and the noise may be correlated



(a) A channelized receiver whose one of channels is occupied by interfering signal



(b) An interfering signal which is intermittent in time



(c) A receiver whose band width is partially overlapped with a swept noise jammer

Figure 1.1: Examples of a persistent signal together with an intermittent signal

from sample to sample. In other words, noise model during data collection interval may change. This change in noise model over time may represent cycles of hardware failure (signal interruption) or intentional/unintentional jammer activities on the sensor followed by recovery times. This bimodal behaviour may be considered as an intermittent sensor health that causes signal intermittency. Another purpose of this study is to examine the amount of performance gain in frequency estimation when the mentioned intermittency is taken into account.

1.2 Proposed Methods and Models

In this thesis, firstly, multiple-snapshot maximum-likelihood (ML) direction of arrival (DOA) estimation problem is studied for the intermittent jamming scenario. This type of jamming is frequently encountered in practice either inadvertently, say due to the sporadic activity of a non-hostile system; or intentionally, say due to the activity of an adversary sweeping the operational bandwidth of the receiver as mentioned in Section 1.1.

The *intermittent jamming model* assumes that each one of the collected snapshot vectors is jammer corrupted with some probability. Stated differently, a subset of collected snapshots are contaminated by jammer and the others are jammer-free; but the receiver does not know which is which. It is obvious that all the available information to the receiver should be used to improve the DOA estimation performance. So, in this study it is assumed that DOA of the jammer is known by the receiver prior to operation. We think that this is a reasonable assumption, since it is possible for the system to know the exact locations of at least friendly interferers. Also, for the case of hostile interferers, DOA of a stationary interferer may have been estimated and marked previously by the receiver.

Exact maximum likelihood solution for the problem is analytically intractable and an expectation maximization (EM) method based solution is developed for coherent and non-coherent signal models. Coherent signal model assumes that the phase difference between the coefficients of two consecutive snapshots are known a-priori which is an assumption compatible with the Swerling-1/3 target models in the radar signal

processing literature. Non-coherent signal model does not have such an assumption and it is suitable for Swerling-2/4 targets. The suggested EM based solution is shown to yield an important estimation accuracy improvement over conventional maximum-likelihood solution which ignores the intermittency of jammer in some scenarios. A related jamming scenario is studied by Besson *et al* where the performance of ESPRIT algorithm under intermittent jamming is compared with the Cramer-Rao bound [7]. Different from this study, we do not only examine the performance of well known methods under intermittent jamming; but also develop the maximum likelihood estimator for the problem.

As a complementary work to DOA estimation under interference signals, we focus on a signal received by a sensor with two operational states namely healthy and faulty. Being in a healthy state corresponds to a high SNR received signal and being in a faulty state corresponds to a low SNR received signal. A two-component Gaussian mixture noise is used to model the states. Besides this, transitions between these states are assumed to obey a hidden Markov model. This work suggests a computational efficient method for frequency estimation of this signal under the defined conditions. The formulated solution combines well-known signal processing techniques such as forward-backward recursion of the hidden Markov model structure and two-stage iterative frequency estimation techniques designed for stationary white noise.

Cramer-Rao type performance lower bounds for the aforementioned problems are also provided to illustrate the efficacy of the suggested estimators.

1.3 Contributions and Novelties

Our contributions are as follows:

- Development of EM-based maximum-likelihood DOA estimator under intermittent jamming for two different target fluctuation models (coherent and non-coherent target signal models [9]),
- Development of EM-based maximum-likelihood frequency estimation of complex exponential in an intermittent signal via correlated mixture of Gaussians

noise model

- Development of Cramer-Rao type performance bounds (MCRB) for the aforementioned signal models
- Performance comparison with alternative techniques to illustrate the value of the suggested methods

As a result of our study, we have published one journal article and one conference paper whose details are presented as follows:

- Şafak Bilgi Akdemir, Çağatay Candan, "Maximum-likelihood direction of arrival estimation under intermittent jamming", Elsevier Digital Signal Processing, Volume 113, 2021, 103028, ISSN 1051-2004, https://doi.org/10.1016/j.dsp.2021.103028. [10]
- Şafak Bilgi Akdemir, Çağatay Candan, "Study of angle of arrival estimation performance of MVDR and Capon methods under intermittent interference," 26th Signal Processing and Communications Applications Conference (SIU), 2018, pp. 1-4, doi: 10.1109/SIU.2018.8404731.[11]

In addition, an article whose information is presented below has been submitted to IEEE Signal Processing Letters.

- Ömer Çayır, Şafak Bilgi Akdemir, Çağatay Candan, "Complex Exponential Frequency Estimation of an Intermittent Signal via Correlated Mixture of Gaussians Noise Model".

1.4 The Outline of the Thesis

This thesis is comprised of 5 chapters.

In Chapter 2, direction of arrival estimation (DOA) literature is summarized in different perspectives. Behaviour of well known beamformers under signal intermittency and conventional maximum likelihood estimators are presented.

In Chapter 3, a novel algorithm that estimates the direction of a signal source under intermittent jamming is derived together with the Modified Cramer Rao Bound expressions. Besides this, simulation results that compare the performance of the proposed algorithms with bounds and conventional maximum likelihood estimator are presented.

In Chapter 4, a novel algorithm to estimate frequency of an intermittent signal and the Modified Cramer Rao bound are derived. Simulation results are presented.

Finally in Chapter 5, conclusions derived from this study are presented along with a short summary.

There is also an appendix that accompany the main material in the thesis. In Appendix A, derivations of an additional algorithm developed for DOA estimation under intermittent jamming with the knowledge of possible jamming power levels are presented.

Notation :

Lowercase, bold lowercase and bold uppercase symbols are used to represent scalars, vectors and matrices respectively. The superscripts $(\cdot)^T$ and $(\cdot)^H$ are used to respectively denote transpose and conjugate transpose of a vector and a matrix. We use $\boldsymbol{A}^{-1}$ to show inverse of a matrix and the identity matrix is represented by the symbol $\boldsymbol{I}$. $|\cdot|$ denotes either modulus of a complex number or absolute value of a real number or determinant of a matrix depending on the context. $\det(\cdot)$ is also used to show the determinant of a matrix whenever it is necessary to distinguish between modulus and determinant. $\|\cdot\|$ represents ℓ_2 norm of a vector. $\stackrel{c}{=}$ refers the equality of both sides apart from constant terms not affecting subsequent calculations. The notation $\mathcal{CN}(x;\mu,\sigma^2)$ denotes complex Gaussian density of a random variable x, with mean μ and variance σ^2. $\mathcal{CN}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes complex Gaussian density of random vector $\boldsymbol{x}$, with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

CHAPTER 2

DIRECTION OF ARRIVAL ESTIMATION UNDER SIGNAL INTERMITTENCY

2.1 Literature Overview

Direction of arrival (DOA) literature can be organized into two categories [12]. The methods in the first category are called classical methods, namely amplitude comparison [5], time-difference of arrival, interferometer [13] and mono-pulse [14] methods. These methods generate a single DOA estimate. The methods in the second category, such as MUSIC [15], ESPRIT [16] and their extensions [17–23], generate multiple DOA estimates for multiple targets. In fact, methods of both categories are reduced complexity versions of the maximum likelihood estimators under Gaussian noise/interference model for single and multiple targets [24–28]. Expectation Maximization method (EM) is another approach for the complexity reduction of multiple DOA estimation problem [29, 30]. EM introduces some latent random variables to the problem setting and aims to maximize the complete-data likelihood instead of the marginal distribution likelihood [31]. In DOA literature, EM is applied to generate DOA estimates for the multiple superposed signals by defining each signal term in superposition as a latent signal. A more recent application of EM is its utilization in the non-uniform array DOA estimation problem where a hypothetical uniform array output is chosen as the complete data [32].

In the last two decades, several high resolution techniques have been suggested for the DOA estimation in the context of sparse signal representation [33–37]. Different from the subspace-based high resolution techniques, such as MUSIC and ESPRIT; recently developed methods are based on convex optimization, such as ℓ_1-SVD [33], convex-

regularized optimization [38] and sparse Bayesian estimation [39–43]. For instance, an approach exploiting the properties of covariance matrix under the sparsity constraint in the application of sparse Bayesian learning methods is given in [44]. This approach can be interpreted as establishing a connection between the subspace methods and sparse Bayesian learning methods. Solutions exploiting sparsity in Bayesian learning have also been developed for the coherent multipath DOA estimation problem in [45, 46]. In [46], a deterministic method which uses alternating convex search for DOA estimation is formulated. The majority of these methods operate with a fixed grid for the unknown target angular location and several methods have been suggested to improve the performance due to the grid mismatch with the true target location, the off-grid target problem [39, 42, 47–50]. For example, [42] proposes two off-grid DOA estimation methods, one for narrowband signals and one for wideband signals, which are the extended versions of sparse Bayesian learning based relevance vector machine (SBLRVM) algorithm [51]. For the solution of off-grid target problem, the method of Candés and Fernandez-Granda [48] stands out with its gridless, i.e. discretization-free, framework that utilizes the total variation norm which is the continuous analog of ℓ_1 norm. This work has been extended by the atomic norm formulation induced by the complex exponentials in [52–54]. More recently, the atomic norm formulation for DOA estimation has been generalized to multiple snapshots and the problem is renamed as joint sparse frequency recovery problem [55, 56].

Atomic norm based estimaion is one of most important recent breakthroughs in frequency estimation [57]. These estimators use the classical relation between the auto-correlation sequence and power/energy spectral density in the problem definition and optimize over the auto-correlation sequence [52–54]. The change of domains from the continuum of frequency domain (energy spectrum) to the discrete-time sequences (auto-correlation sequence) allows a feasible optimization problem with finitely many unknowns. Once the optimization over the auto-correlation sequence (also called the dual polynomial) is completed, the maxima locations in the Fourier spectrum of the optimized sequence yield the spatial/temporal frequency estimates [53, 54]. The initial formulation of the atomic norm based estimators requires semi-definite programming (SDP) solvers with linear matrix inequalities (LMI) constraints which are not efficient to implement for large dimensional problems. In [54, 56], a solution based

on alternating direction method of multipliers (ADMM) method have been suggested for computational efficiency. For more information on atomic norm based estimators, readers are invited to consult [57].

In addition to the theoretical aspects of the problem, the array imperfections such as mutual coupling, gain/phase uncertainty, sensor location errors can cause performance degradations in practical systems. Due to their impact on DOA estimation accuracy, self calibration methods have also been studied [58–61]. [62] approaches this problem by sparse Bayesian perspective in a unified framework for both array calibration and DOA estimation.

Another classification can be given according to the signal of interest (SOI) bandwidth. The narrowband DOA estimation methods that have been mentioned so far can be extended to the wideband DOA estimation setting by different approaches. In the noncoherent processing approach, DOAs of each SOI frequency component is estimated by narrow-band processing methods and then the estimates are combined. The coherent processing approach uses a focusing matrix to transform the covariance matrices at different frequencies to a common frequency. Coherent Signal Subspace Method (CSSM) [63] and Weighted Average of Signal Subspaces (WAVES) [64] are well known examples for the coherent approach. Test of Orthogonality of Projected Subspaces (TOPS) [65] is another wideband DOA estimation method that is based on the orthogonality relation between the signal and the noise subspaces for multiple frequency components. Wideband Covariance Matrix Sparse Representation (W-CMSR) [66] exploits a priori information on the signal spectrum and it is shown to be capable of detecting higher number of sources than the number of sensors for some array geometries.

In many applications, intentional or unintentional interference sources occupy the spectrum together with the desired signal. Radar/communication jammers are examples of intentional interferers. Several methods have been developed to detect jamming and also to improve DOA estimation accuracy under jamming. Solutions for the detection and angle tracking under interference for monopulse radars have been studied in [67, 68]. Similarly, main lobe and side lobe jammer cancelers for monopulse DOA estimation are studied in [69]. In [70], several techniques for jammer and clut-

ter cancellation in adaptive airborne radar are proposed. More recent works related to detection and estimation under jamming are given in [71, 72]. In [71], an adaptive beamformer orthogonal rejection test (ABORT)-like detector that can blank coherent ECM signals is proposed. In [72], the invariant detectors have been developed in a multiple pulsed jamming environment. An unintentional jamming scenario, inter-radar interference between automotive FMCW radars, is examined in [73] and an adaptive interference suppression windowing width control is suggested for the detection of interference. Jamming is also of major concern for many modern civilian or military communication systems. For example, [74] deals with the estimation of fast fading and/or frequency-selective fading channels problem under interference in orthogonal frequency division multiplexing (OFDM) multiple-input multiple-output (MIMO) systems and it uses compressive sensing approach to circumvent the effects of unknown multiple asynchronous narrow-band interference (NBI). In [75], the detection of jammers in direct sequence spread spectrum (DSSS) wireless communication systems is studied and the statistics of jamming-free symbols is utilized to discriminate jammed packets. Similarly, the interference due to the Narrowband Internet-of-Things (NB-IoT) on existing broadband systems is also examined in [76].

We would like to note again that the running assumption in virtually all DOA estimation methods is the coexistence of multiple targets or target- jammer/interference signals in *all* snapshots. Different from these studies, we assume that jamming is intermittent; hence, not all but some snapshots are contaminated. The main purpose of this study is to examine the amount of performance gain when the intermittency is taken into account in the problem formulation.

2.2 Overview of Direction of Arrival Estimation Methods

In this section, widely used basic DOA estimation methods is overviewed for the completeness of this chapter and to provide a solid background for the methods used for comparison with the proposed algorithms in the upcoming chapters.

2.2.1 Amplitude Based Methods

2.2.1.1 Use of Single Spinning DF Antenna

A fundamental approach for DOA estimation is to use a narrow beam, high gain, low sidelobes rotating antenna. This method is known as *Spinning DF Antenna Method*. With the basic assumption that the signal amplitude does not change rapidly due to reasons related to the source, but an amplitude modulation occurs on the signal only due to rotation and pattern of the DOA measuring antanna, this method compares signal amplitudes or power at each position of the rotating antenna and marks the angle with the highest signal amplitude as DOA angle of the source. High gain of the antenna allows detection and DOA estimation of very weak signals as such coming from antenna sidelobes of an emitter that is far away from the DOA measuring system. The method's resolution and accuracy depend on beamwidth of the antenna as well as position and relative power of the signal sources in the medium. It can provide accurate DOA estimates whose error is less than 1° rms all over the 2-18 GHz RF band [14]. To distinguish closely spaced targets and to increase the DOA estimation accuracy either the beamwidth of the antenna should be decreased by increasing the aperture/size of the antenna or nulls of the antenna should be used [77]. Contrary to searching for the maximum amplitude/power of signals, one searches for the point where no output power is received when using nulls of the antenna. Using nulls is advantageous because low gain, small aperture antennas could be used on small-size platforms since output power changes rapidly near nulls in contrast to peak of the main beam. Because of this advantage, this method is preferred mainly by ELINT systems. Despite this advantage, it cannot find application area in ESM systems much, because its probability of intercept with the source is small and it cannot produce DOA information from a single snapshot of the signal.

2.2.1.2 Use of Multiple Stationary DF Antennas

Being aware of the surrounding environment continuously is an important feature of an EW system. Hence, a multiple-antenna-structure that always listens environment with 360° coverage is a widely preferred architecture in EW systems particularly. In

Radar Warning Receivers (RWR), which is a very basic kind of ESM systems located on fighter aircrafts, amplitude comparison method is almost always implemented by using multiple antennas located at an angle to each other [78]. In this architecture 4, 6, 8 or more antennas are placed at angles, called squint angles, to each other on the same plane. To determine the DOA of the signal source, two antennas with the strongest received signal powers are selected and the ratio of these power levels is calculated. The power ratio is mapped to a DOA estimate which is pre-defined in a look-up-table. It is known that total accuracy of about 2.95° can be achieved by using this method with 12 antennas in a RWR receiver [72]. Accuracy of this method seems worse compared to the accuracy of 1° of aforementioned method in Section 2.2.1.1. So, in practice this method is used together with other methods such as phase comparison to improve accuracy of the estimation and to avoid ambiguities in direction.

2.2.1.3 Centroiding in Scanning Radars

In active sensing systems, or in the context of this thesis say in radars, angle information is basically extracted by matching the target detections with antenna position. In practice, antenna beam is rotated to scan the predefined search volume and direction of the received echo signal is marked as if the target is at the boresight of the antenna at every detection. During this movement, amplitudes of the target echo signals are usually encountered at successive beam positions, not only because of the antenna movement resolution or antenna beamwidth, but also some of the targets having large RCS that extends over 3 dB beamwidth of the antenna pattern [79]. So, in scanning radars to increase angular estimation accuracy, centroiding is used to estimate relative angle of the target with respect to radar system. The centroid of the relative power measurements at the receiver is calculated as

$$C_p = \frac{\sum_{n=n_1}^{n_2} nP(n)}{\sum_{n=n_1}^{n_2} P(n)} \tag{2.1}$$

where $P(n)$ is the received power; n_1 and n_2, are the first beam index and the last beam index where the target is detected respectively. Then, by considering the angular orientation and spacing of the measurements, the calculated centroid is mapped to

a target angle [9]. This technique is typically used in long-range mechanical scanning surveillance radars which perform range and bearing measurements against airborne platforms. Different types of centroiding techniques and their applications in surveillance radars are summarized in [80] and [81] and the references therein. In order to give an idea about the accuracy of the method; Cramer-Rao Lower Bound (CRB) for AOA estimation error of centroiding method that uses binary integration approach is given as [82]

$$\sigma_{\hat{\theta}}^2 = \frac{\theta_3^2}{N \cdot SNR \cdot \alpha^2} \tag{2.2}$$

where θ_3, N, SNR are 3 dB beamwidth of the antenna, number of pulses transmitted within the antenna beamwidth and SNR of the received pulses respectively. α is a constant that depends on ratio of observation interval of the target and 3 dB antenna beamwidth.

2.2.1.4 Conical Scan

One of the oldest methods used in mechanical scanning tracking radars to increase accuracy of the DOA estimation in tracking is conical scan [14]. In this method, not the whole antenna but only the feed is rotated at an offset to the nominal antenna boresight, making the beam axis rotate around the target. Especially in some missile seekers, the same effect is achieved by using a tilted reflector rotated at the desired conical scan rate. This beam axis rotation causes a sinusoidal amplitude modulation in the received echo signal. The frequency of modulation is equal to antenna rotation frequency and depth of modulation is directly proportional to the target's distance from the boresight of the antenna. The angular position of the peak of this modulated signal is used together with two-way antenna pattern to estimate angle of the target [9] with respect to the antenna boresight.

2.2.1.5 Lobing or Lobe Switching

Multi purpose electronically scanning radars or other types of radars that require highly accurate angle estimates may use different lobing techniques which are based on moving antenna beam around target in discrete steps and comparing measured

return powers.

The basic type of lobing techniques especially used in older radar systems is sequential lobing. Although in this method beam positions are selected electrically, mechanical scanning is also needed to place the target in the center of the scanning area [9]. In this method, if a target is detected when the boresight of the antenna is directed at an angle θ_0, the beam is steered to each opposite side of this original boresight by the same amount, say $\Delta\theta$ degrees, sequentially and the received echo powers are measured. Then, the beam is steered to a position that is closer to the side at which higher power is measured. Usually, this procedure is repeated with shrinking $\Delta\theta$ until the desired accuracy is achieved. The beam may be steered in one or two dimensions depending on the design and mode of the radar system. At each beam position a sequence of pulses is transmitted and the received power is measured assuming that the power level remains constant from pulse to pulse. This assumption makes the systems that use this method vulnerable to pulse-to-pulse RCS fluctuations [79].

As a side information, transmissions of conical scan and lobing radars give much information about associated radar systems to ESM systems and cause ESM systems to sign them as threats and to inform onboard ECM systems, since almost always tracking radars operate in conjunction with weapon systems. To overcome this weakness and to increase resistance against jamming and deceiving, conical-scan-on-receive-only (COSRO) and lobe-switching-on-receive-only (LORO) systems have been developed which have more complexity and cost. These systems perform lobing and conical scanning like operations at their receive paths instead of transmit paths as their names suggest which makes them harder to be recognised and classifed by ESM systems.

2.2.2 Monopulse Technique

Monopulse technique is a high accuracy DOA estimation technique that is used in modern radars. Although monopulse is a simultaneous lobing technique that use two or more antenna beams at the same time, it is discussed under a separate heading from Section 2.2.1.5 since both amplitude and phase comparison monopulse methods are present in the literature. This technique is superior to sequential lobing and conical

scan techniques since it extracts the angle information more accurately by only using a single pulse and more resistant to power fluctuations in the received signal. Another advantage of monopulse systems is their resistance to deceiving and jamming [9] since no modulation exists on transmitted signal. Despite the name monopulse, these systems may use multiple pulses to improve angle estimation accureacy. The important thing here is that angle estimates are obtained for every single pulse to name the system as monopulse [79]. Monopulse beams may be either formed by mechanical construction of antenna in mechanical scanning systems or beamforming in electronical scanning sytems.

In amplitude comparison monopulse technique, two or four squinted beams, which share a common phase center, are used to extract angle information in one (azimuth) or two (azimuth and elevation) dimensions respectively. In the simplest form of this technique, sum and difference channels are formed by adding and subtracting echo signals received with each squinted beam. A monopulse ratio is calculated by dividing difference of the received signals by their sum. Then, real part of monopulse ratio is used to estimate angle of the target [79] for example by comparing the ratio with a look-up-table.

In contrary to amplitude comparison monopulse, phase centers of different beams are displaced in phase comparison monopulse systems by an amount of d . The distance between phase centers causes a delay in the received signals which corresponds to a phase difference. The measured phase difference is used to estimate target angle in a similar way to interferometry in Section 2.2.4. The phase comparison monopulse systems are also named as interferometer radar [83].

In [84], it is mentioned that to prevent ambiguities in the estimated direction, distance between beam phase centers must satisfy the following condition,

$$\frac{\lambda}{8d} > |\sin\theta| \tag{2.3}$$

in which θ and λ denote the angle of the target and wavelength of the transmitted signal respectively.

2.2.3 Time Difference of Arrival Method

Time Difference of Arrival (TDOA) method requires two antennas/sensors positioned far from each other. Because of this requirement, it is suitable to be used by ESM systems that have antennas located at the extremities of big platforms, for instance at the tip points of wings of large aircrafts. This method can be also used by systems that are located on moving platfoms that takes measurements at different positions. The time of arrival (TOA) of signals arriving at each sensor is measured and the difference between these measurements is calculated which corresponds to

$$\tau = \frac{d\sin(\theta)}{c}. \tag{2.4}$$

Using the above relation, DOA of the signal/target can be calculated as

$$\hat{\theta} = \arcsin\left(\frac{\tau c}{d}\right) \tag{2.5}$$

where τ, c and d are time difference of arrival of the signals, speed of light and distance between measurement points respectively. In systems used today, high speed analog-to-digital converters (ADC) are used to convert incoming signal to digital domain and the signal processing takes place usually in field programmable gate arrays (FPGAs). Sampling speed of ADC and the spacing of measurement points has a noticeable effect on the resolution of method. For example, the minimum delay that can be measured by an ADC which has a sampling rate of 2 GS/s (giga samples per second) is $1/(2 \times 10^9) = 0.5$ ns. If there is a spacing of 10 m between signal measurement points, resolution becomes $\Delta\theta = \arcsin\left(10^{-1} \cdot 0.5 \cdot 10^{-9} \cdot 3 \cdot 10^8\right) = 0.86°$. On the other hand, if the sampling speed of ADC were 250 MS/s (mega samples per second), the resolution would be $\Delta\theta = \arcsin\left(10^{-1} \cdot 4 \cdot 10^{-9} \cdot 3 \cdot 10^8\right) = 6.89°$. If spacing between measurement points is increased to 30 m, a resolution of $\Delta\theta = \arcsin\left((1/30) \cdot 4 \cdot 10^{-9} \cdot 3 \cdot 10^8\right) = 2.3°$ can be achieved. These examples show that resolution is directly proportional to sampling rate and separation of measurement points. We would like to also mention that this technique is one of the common techniques used by localization systems.

2.2.4 Interferometer Method

Interferometer method is a phase comparison method used in applications like ESM systems that require high degree of accuracy. To extract only one dimensional DOA information, linear arrays consisting of at least two antennas are required. On the other hand, to estimate direction in both azimuth and elevation, planar arrays that have minimum of three antennas are used. Let us consider a two-element direction finding system. The phase difference between the arriving signals corresponds to DOA of the source as described in the following equation

$$\hat{\theta} = \arcsin\left(\frac{\lambda}{2\pi d}\phi\right) \tag{2.6}$$

where λ, d and ϕ are wavelength of the received signal, distance between array elements and phase difference of the signals acquired by the sensors respectively. To avoid ambiguity in this setting, the distance between array elements should satisfy the following inequality

$$d \leq \frac{\lambda}{2}. \tag{2.7}$$

To achieve higher accuracy without causing ambiguity, more-element arrays are used. In these systems, it is sufficient to avoid uncertainty if only the distance between a pair of elements satisfies the above inequality. Other elements are placed at wavelengths of 2, 4, 8 etc. These systems are called multiple baseline interferometer in the literature.

2.2.5 MVDR and Capon Beamformers

Minimum Variance Distortionless Response (MVDR) and Capon are two basic methods which are used for DOA estimation besides calculating antenna or transmitting/receiving channel coefficients in beamforming applications. MVDR beamformer minimizes both power of noise at the receiver and power of signals coming from directions other than the intended direction of the beam. At the same time, it allows the signals from intended direction to remain undistorted.

Let us assume that we have a set of L vectors, which are obtained from output of an N element sensor array as

$$\boldsymbol{x}_i = \underbrace{\gamma_{t,i}\boldsymbol{a}(\theta_t)}_{\text{SOI part}} + \underbrace{\boldsymbol{e_i}}_{\substack{\text{interference} \\ \text{+ noise}}}, \quad i = \{1, \ldots, L\}. \tag{2.8}$$

where $\boldsymbol{a}(\theta_t)$ is the $N \times 1$ array manifold vector corresponding to SOI in the environment and $\gamma_{t,i}$ denotes the amplitude of the received signal. $\boldsymbol{e_i}$ part contains all the unwanted interfering signals and noise.

The mathematical expression of MVDR beamformer is as follows:

$$\underset{\boldsymbol{h}}{\arg\min} \quad \boldsymbol{h}^H \boldsymbol{R_e} \boldsymbol{h} \quad s.t. \quad \boldsymbol{h}^H \boldsymbol{a_t} = 1 \tag{2.9}$$

where $\boldsymbol{h}$ represents beamformer filter coefficients, $\boldsymbol{a_t}$ is the short-hand notation for $\boldsymbol{a}(\theta_t)$, $\boldsymbol{R_e} = E\{\boldsymbol{e}\boldsymbol{e}^H\}$ is the autocorrelation matrix of interference and noise components. It is well known that the optimization problem in Eq.(2.9) can be solved by method of Lagrange multipliers and MVDR filter coefficients can be found to be

$$\boldsymbol{h}_{mvdr} = \frac{\boldsymbol{R_e^{-1}} \boldsymbol{a_t}}{\boldsymbol{a_t^H} \boldsymbol{R_e^{-1}} \boldsymbol{a_t}}. \tag{2.10}$$

In Eq.(2.10) MVDR beamformer coefficients are calculated according to the direction that the beam is wanted to be directed at. When MVDR method is used for DOA estimation, filter coefficients, $\boldsymbol{h}_{mvdr}$, are calculated at different uniformly separated angle values and the angle that corresponds to the maximum expected output power, i.e.,

$$P(\theta) = \boldsymbol{E}\left\{|\boldsymbol{h_\theta^H} \boldsymbol{x}|^2\right\} = \boldsymbol{h_\theta^H} \boldsymbol{R_x} \boldsymbol{h_\theta} \tag{2.11}$$

is marked as DOA of SOI. Note that in Equation (2.11) $\boldsymbol{R_x} = E\{\boldsymbol{x}\boldsymbol{x}^H\}$ is the autocorrelation matrix of the received signal. From this point forward, we name the maximum expected output power as angular power spectrum shortly. The mathematical representation of this problem is

$$\hat{\boldsymbol{\theta}}_t = \underset{\theta}{\arg\max} \quad P(\theta) = \underset{\theta}{\arg\max} \quad \boldsymbol{h_\theta^H} \boldsymbol{R_x} \boldsymbol{h_\theta} \tag{2.12}$$

where $\boldsymbol{a_\theta} = \boldsymbol{a}(\theta)$ and

$$\boldsymbol{h_\theta} = \frac{\boldsymbol{R_e^{-1}} \boldsymbol{a_\theta}}{\boldsymbol{a_\theta^H} \boldsymbol{R_e^{-1}} \boldsymbol{a_\theta}}. \tag{2.13}$$

Although being a powerful beamformer in theory, it is not possible to implement MVDR in real world applications since it requires exact interference plus noise correlation matrix. For this reason, a powerful yet easy to implement beamformer called Capon beamformer is used in practice. Capon beamformer minimizes the total power at the array output while keeping SOI, at the angle the beam is directed at, undistorted. Mathematical expression for this approach is

$$\underset{\boldsymbol{h}}{\arg\min} \quad \boldsymbol{h}^H \boldsymbol{R_x} \boldsymbol{h} \quad s.t. \quad \boldsymbol{h}^H \boldsymbol{a_t} = 1 \tag{2.14}$$

By using Lagrange multipliers methods, Capon beamformer coefficients that solve the above minimization problem can be found as

$$\boldsymbol{h}_{capon} = \frac{\boldsymbol{R_x}^{-1} \boldsymbol{a_t}}{\boldsymbol{a_t}^H \boldsymbol{R_x}^{-1} \boldsymbol{a_t}}. \tag{2.15}$$

Since exact autocorrelation matrix of the array output is not available in real life applications, sample autocorrelation matrix which is calculated as $\hat{\boldsymbol{R}}_{\boldsymbol{x}} = 1/L \sum_{i=1}^{L} \boldsymbol{x_i} \boldsymbol{x_i^H}$, can be used instead of the exact one in Eq.(2.15). This approach is named as Sample Matrix Inversion (SMI) beamformer.

When Capon method is used for DOA estimation, filter coefficients, $\boldsymbol{h}_{capon}$, are calculated at different uniformly separated angle values and the angle that corresponds to the maximum expected output power which is given in Eq. (2.11) is marked as DOA of SOI. The mathematical expression for Capon DOA estimation algorithm can be represented as:

$$\hat{\boldsymbol{\theta}}_t = \underset{\theta}{\arg\max} \quad P(\theta) = \underset{\theta}{\arg\max} \quad \frac{1}{\boldsymbol{a_\theta}^H \boldsymbol{R_x}^{-1} \boldsymbol{a_\theta}}. \tag{2.16}$$

Similarly when SMI beamformer coefficients are used, the mathematical expression for SMI DOA estimation algorithm can be represented as:

$$\hat{\boldsymbol{\theta}}_t = \underset{\theta}{\arg\max} \quad \hat{P}(\theta) = \underset{\theta}{\arg\max} \quad \frac{1}{\boldsymbol{a_\theta}^H \hat{\boldsymbol{R}}_{\boldsymbol{x}}^{-1} \boldsymbol{a_\theta}}. \tag{2.17}$$

Note that the names Capon, MVDR and SMI beamformers are used interchangebly in the literature to mention the DOA estimation method formulated in Eq. (2.17) [85].

2.2.6 MUSIC - Multiple Signal Classification

Multiple Signal Classification Method, which is proposed in [15], is a well-known and widely used method on which many other studies are based. It is applicable to

any array geometry, and also there is no need of array elements being identical or matched in terms of directional characteristics, i.e., amplitude and phase other than polarization characteristics. It uses the properties of the autocorrelation matrix of the array output instead of the measured received signal parameters or probability density function of the received signals like the methods previously mentioned.

The signal at the output of an N element array in the form of a $N \times 1$ vector $\boldsymbol{x_i}$ can be written as

$$\boldsymbol{x}_i = \boldsymbol{A}\boldsymbol{y_i} + \boldsymbol{n_i} \qquad i = 1, \ldots, L \tag{2.18}$$

or in its open form

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}(\theta_{t,1}) & \boldsymbol{a}(\theta_{t,2}) & \ldots & \boldsymbol{a}(\theta_{t,D}) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_D \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \end{bmatrix} \tag{2.19}$$

where $\boldsymbol{A}$ is the $N \times D$ array manifold matrix whose columns are array manifold vectors corresponding to D sources existing in the environment, $\boldsymbol{y_i}$ is the $D \times 1$ vector whose entries, y_i are the amplitudes of these sources, and $\boldsymbol{n_i}$ is $N \times 1$ noise vector. It is well known that the autocorrelation matrix of the array output can be written as

$$\boldsymbol{R_x} = \boldsymbol{A}\boldsymbol{R_y}\boldsymbol{A}^H + \boldsymbol{R_n} \tag{2.20}$$

where $\boldsymbol{R_x}$, $\boldsymbol{R_y}$ and $\boldsymbol{R_n}$ are the autocorrelation matrices of array output, signal sources and noise at the receiver respectively, under the assumption that signal sources and noise are uncorrelated. The main idea of this method is the orthogonality of the subspace spanned by array manifold vectors, i.e., columns of $\boldsymbol{A}$ and eigenvectors of noise subspace. To understand the idea behind this notion, we should look at the eigenstructure of $\boldsymbol{R_x}$ more closely. First of all, it is well known that eigenvalues of $\boldsymbol{R_x}$ are real nonnegative scalars and eigenvectors associating these eigenvalues form an orthonormal set since $\boldsymbol{R_x}$ is Hermitian symmetric as being an autocorrelation matrix. Secondly, we know from linear algebra that rank of product of two matrices are equal

to or less than the minimum rank of these matrices. Using this fact, one can conclude that when the number of sources is less than the number of array elements, $D < N$ case, the rank of term $\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H$ in Eq. (2.20), is at most D and as a result $N \times N$ matrix $\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H$ is singular meaning that $\left|\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H\right| = |\boldsymbol{R}_{\boldsymbol{x}} - \boldsymbol{R}_{\boldsymbol{n}}| = 0$ where $|\cdot|$ denotes determinant of a matrix. To simplify the problem at hand, we may assume further that noise samples are spatially and temporarily white, which results in an $N \times N$ diagonal covariance matrix $\boldsymbol{R}_{\boldsymbol{n}} = \sigma^2\boldsymbol{I}$. This assumption turns Eq. (2.20) into the form $\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H + \sigma^2\boldsymbol{I}$. It is well known that eigenvalues of $\boldsymbol{R}_{\boldsymbol{x}}$ are equal to the sum of each eigenvalue of $\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H$ and σ^2. Singularity of $\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H$ also means it has only D nonzero eigenvalues. The remaining $N - D$ eigenvalues of $\boldsymbol{R}_{\boldsymbol{x}}$ are equal to σ^2. Note that $|\boldsymbol{R}_{\boldsymbol{x}} - \lambda\boldsymbol{I}| = 0$ is hold for every eigenvalue of $\boldsymbol{R}_{\boldsymbol{x}}$ and σ^2 is an eigenvalue of it, so the characteristic equation of $\boldsymbol{R}_{\boldsymbol{x}}$ can be written as

$$\left(\boldsymbol{R}_{\boldsymbol{x}} - \sigma^2\boldsymbol{I}\right)\boldsymbol{g}_{\boldsymbol{i}} = 0 \tag{2.21}$$

where $\boldsymbol{g}_{\boldsymbol{i}}$ denotes the eigenvector that corresponds to the eigenvalue σ^2. Moreover, since $\boldsymbol{R}_{\boldsymbol{x}} - \sigma^2\boldsymbol{I}$ is nothing but equal to $\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H$, the equation in (2.21) can be rewritten as $\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{A}^H\boldsymbol{g}_{\boldsymbol{i}} = 0$. This leads us to the important observation that, the eigenvectors $\boldsymbol{g}_{\boldsymbol{i}}, i = \{1, \cdots, N - D\}$ corresponding to the eigenvalues σ^2 are orthogonal to the space spanned by the columns of $\boldsymbol{A}$, or in other words, these eigenvectors belong to the null space of $\boldsymbol{A}^H$. Note also that since there are D sources, the rank of $\boldsymbol{A}$ is D, leading to the fact that dimension of $\mathcal{N}(\boldsymbol{A}^H)$ is equal to $N - D$. If we form the matrix $\boldsymbol{G}$ whose columns are the eigenvectors corresponding to the repeated eigenvalue σ^2, that is $\boldsymbol{G} = [\boldsymbol{g_1} \quad \boldsymbol{g_2} \quad \cdots \quad \boldsymbol{g_{N-D}}]$, it is obvious that dimension of range space of $\boldsymbol{G}$ is also equal to $N - D$ meaning that

$$\mathcal{R}(\boldsymbol{G}) = \mathcal{N}(\boldsymbol{A}^H) \tag{2.22}$$

where $\mathcal{R}(.)$ and $\mathcal{N}(.)$ denote the range and null spaces of a matrix respectively. This is the main idea behind MUSIC algorithm as mentioned previously. At this point, let us define the matrix $\boldsymbol{S} = [\boldsymbol{s}_1 \quad \boldsymbol{s}_2 \quad \cdots \quad \boldsymbol{s}_D]$ whose vectors are the eigenvectors of $\boldsymbol{R}_{\boldsymbol{x}}$ belonging to the eigenvalues that are greater than σ^2. Since eigenvectors of $\boldsymbol{R}_{\boldsymbol{x}}$ are orthonormal, we can also say that $\boldsymbol{S}^H\boldsymbol{G} = 0$. This leads us to the equality

$$\mathcal{R}(\boldsymbol{G}) = \mathcal{N}(\boldsymbol{S}^H). \tag{2.23}$$

If we compare Eq. (2.22) and (2.23), we can reach the following equation:

$$\mathcal{N}(\boldsymbol{A}^H) = \mathcal{N}(\boldsymbol{S}^H). \tag{2.24}$$

Note that, since range space of $\boldsymbol{S}$ and range space of $\boldsymbol{A}$ are orthogonal complements to null spaces of $\boldsymbol{S}^H$ and $\boldsymbol{A}^H$, we can conclude that

$$\mathcal{R}(\boldsymbol{S}) = \mathcal{R}(\boldsymbol{A}). \tag{2.25}$$

As a result, angles of the incoming signals correspond to the D highest peaks of the pseudospectrum which is defined as

$$\hat{\theta}_t = \arg\max_{\theta} \quad \frac{1}{\boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{G}\boldsymbol{G}^H \boldsymbol{a}_{\boldsymbol{\theta}}}. \tag{2.26}$$

In practice, to implement MUSIC algorithm, firstly sample autocorrelation matrix, then eigenvalues and related eigenvectors are calculated. The eigenvectors associated to the smallest $N-D$ eigenvalues are assumed to belong to the noise subspace. These eigenvectors are used to calculate estimate of the matrix $\boldsymbol{G}$, or in other words $\hat{\boldsymbol{G}}$. Finally, DOA estimates are obtained as the D highest peak locations of the following function:

$$\hat{\theta}_t = \arg\max_{\theta} \quad \frac{1}{\boldsymbol{a}_{\boldsymbol{\theta}}^H \hat{\boldsymbol{G}}\hat{\boldsymbol{G}}^H \boldsymbol{a}_{\boldsymbol{\theta}}}. \tag{2.27}$$

2.2.7 Maximum Likelihood Estimators

Maximum Likelihood (ML) DOA estimator, although it is generally not the first choice because of its computational load, is still used in applications where high degree of accuracy is required. In practice, an initial estimate is found by using one of other methods, then around vicinity of this estimate, necessary search, finding maxima/mimina operations of ML method are carried out depending on the form of the solution. The biggest advantage of this method is its being asymptotically efficient.

ML estimation is based on maximization of the likelihood function which is joint probability of fixed observed data viewed as a function of signal model's unknowns. In the literature there are two types of ML approaches, namely deterministic maximum likelihood estimation which depends on so-called conditional signal model and stochastic maximum likelihood estimation that depends on unconditional signal

model. In the first model, received signals are assumed to be unknown deterministic whereas in the latter one, they are assumed random. In [27], it is mentioned that if DOAs of multiple signals are being estimated, stochastic approach performs better for highly correlated signals whereas there is no significant performance difference between two methods in the presence of weakly correlated signals. The derivation for the conventional ML estimator is given in Section 2.3.1.

2.3 Behaviour of Widely Used Methods Under Signal Intermittency

Let us assume that we have a set of L vectors, which are obtained from output of an N element sensor array as

$$\boldsymbol{x}_i = \underbrace{\gamma_{t,i}\boldsymbol{a}(\theta_t)}_{\text{SOI}} + \underbrace{I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)}_{\text{jammer}} + \underbrace{\boldsymbol{n}_i}_{\text{noise}}, \quad i = \{1, \ldots, L\}. \tag{2.28}$$

The vector $\boldsymbol{x}_i$ is composed of three components which are signal of interest (target signal), jamming signal and noise. In Eq.(2.28), θ_t, $\gamma_{t,i}$ denote the DOA and complex amplitude of SOI respectively. Similar to the signal component, in the jammer component θ_j and $\gamma_{j,i}$ represent DOA and complex amplitude of interfering signal respectively. $\boldsymbol{a}(\theta_t)$ and $\boldsymbol{a}(\theta_j)$ are $N \times 1$ array steering vectors belonging to target and jammer respectively.

At this point, we would like to note that the sensor arrays used in derivations and analysis throughout this chapter and Chapter 3.1 are the uniform linear array (ULA) whose inter-element spacing is d. The structure of the array is presented in Figure 2.1. It is well known that of the array steering vector that belongs to ULA is in the following form:

$$\boldsymbol{a}(\theta) = \begin{bmatrix} 1 & e^{-j(2\pi d/\lambda)\sin(\theta)} & \ldots & e^{-j(2\pi d/\lambda)\sin(\theta)(N-1)} \end{bmatrix}^T. \tag{2.29}$$

In Eq.(2.28), I_i is the indicator variable which represents the presence or absence of the jammer component. I_i being equal to 1 means jammer component exists in the ith snapshot whereas I_i being zero indicates the opposite condition. Lastly, $N \times 1$ vector $\boldsymbol{n}_i$ in Eq.(2.28) represents the noise which is assumed to have circularly symmetric complex Gaussian distribution.

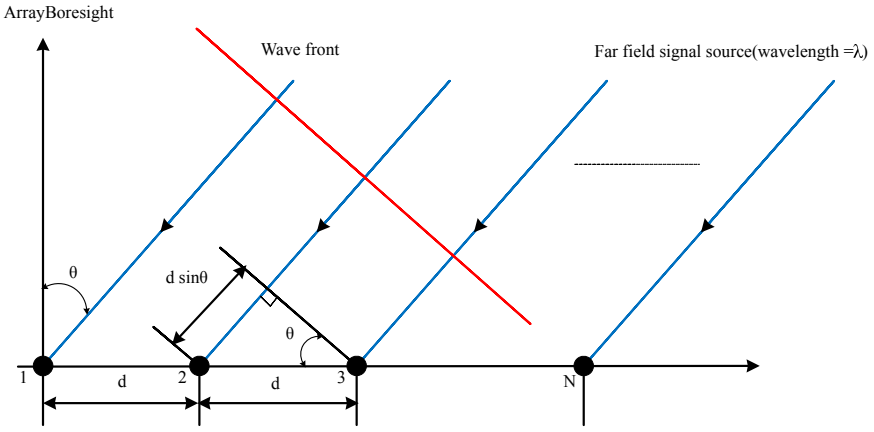


Figure 2.1: Uniform linear array

In addition, the following assumptions are made in the defined signal model:

- Complex signal amplitudes, $\gamma_{t,i},\ i = \{1, \ldots, L\}$, are non-random constants,
- Complex jammer signal amplitude, $\gamma_{j,i}$ is a zero mean complex Gaussian random variable with a variance σ_j^2, i.e., $\gamma_{j,i} \sim \mathcal{CN}(0, \sigma_j^2)$,
- Angle of the jammer, θ_j, is known,
- Circularly symmetric complex Gaussian noise, $\boldsymbol{n}_i$ is distributed with $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{I})$, meaning that the noise at different sensors are independent from sensor to sensor and from snapshot to snapshot,
- $P(I_i = 1) = \alpha_1$ and $P(I_i = 0) = \alpha_0 = 1 - \alpha_1$.

Let $\boldsymbol{R_x}$ and $\boldsymbol{R_e}$ be the autocorrelation matrices of the received signal and interference. Interference consists of jammer signal and noise components, i.e., $\boldsymbol{e}_i = I_i \gamma_{j,i} \boldsymbol{a}(\theta_j) + \boldsymbol{n}_i$ respectively. $\boldsymbol{R_e}$ can be calculated as

$$\begin{aligned}
\boldsymbol{R_e} &= \boldsymbol{E}\Big\{[I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) + \boldsymbol{n}_i][I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) + \boldsymbol{n}_i]^H\Big\} \\
&= \boldsymbol{E}\Big\{I_i^2|\gamma_{j,i}|^2\boldsymbol{a}(\theta_j)\boldsymbol{a}(\theta_j)^H\Big\} + \boldsymbol{E}\Big\{\boldsymbol{n}_i\boldsymbol{n}_i^H\Big\} \\
&= \boldsymbol{E}\Big\{I_i^2\Big\}\Big\{|\gamma_{j,i}|^2\Big\}\boldsymbol{a}(\theta_j)\boldsymbol{a}(\theta_j)^H + \boldsymbol{I} \\
&= \alpha_1\sigma_j^2\boldsymbol{a_j}\boldsymbol{a_j}^H + \boldsymbol{I}. \qquad (2.30)
\end{aligned}$$

Using this form, autocorrelation matrix of the received signal can be found to be

$$\begin{aligned}\boldsymbol{R_x} &= \sigma_t^2 \boldsymbol{a_t} \boldsymbol{a_t}^H + \alpha_1 \sigma_j^2 \boldsymbol{a_j} \boldsymbol{a_j}^H + \boldsymbol{I} \\ &= \sigma_t^2 \boldsymbol{a_t} \boldsymbol{a_t}^H + \boldsymbol{R_e}\end{aligned} \tag{2.31}$$

where $\boldsymbol{a_t}$ and $\boldsymbol{a_j}$ are the short-hand notations for $\boldsymbol{a}(\theta_t)$ and $\boldsymbol{a}(\theta_j)$ respectively.

2.3.1 Derivation of Conventional Maximum Likelihood Estimator

Conventional DOA estimators ignore the intermittency of the intercepted signals. Therefore, to develop a conventional ML estimator for the problem at hand, we may assume that $I_i = 1$ for $i = \{1, \ldots, L\}$ in Eq.(2.28), that is jammer is present in all snapshots. Hence, the signal model in Eq.(2.28) turns into the form

$$\boldsymbol{x}_i = \gamma_{t,i} \boldsymbol{a}(\theta_t) + \gamma_{j,i} \boldsymbol{a}(\theta_j) + \boldsymbol{n}_i. \tag{2.32}$$

The signal model parameters and their values, except I_i being equal to 1 for all snapshots, are as defined at the beginning of Section 2.3. Assume moreover that the variance of the jammer component amplitudes, $\gamma_{j,i}$ is known and equal to σ^2. As a result, the unknowns of the problem are $\boldsymbol{\theta} = \{\theta_t, \gamma_{t,1}, \ldots, \gamma_{t,L}\}$. Since the power of jammer and noise components in the received signal are known, we can redefine the signal in Eq.2.32 as

$$\boldsymbol{x}_i = \gamma_{t,i} \boldsymbol{a}(\theta_t) + \tilde{\boldsymbol{n}}_i \tag{2.33}$$

where $\tilde{\boldsymbol{n}}_i$ is distributed with $\mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_{\tilde{\boldsymbol{n}}})$ and $\boldsymbol{R}_{\tilde{\boldsymbol{n}}} = \boldsymbol{I} + \sigma^2 \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)$. The log-likelihood function of the received vectors can be written as

$$\begin{aligned}\mathcal{L}(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L; \boldsymbol{\theta}) &= \ln \left[\prod_{i=1}^{L} \frac{1}{\pi^N |\boldsymbol{R}_{\tilde{\boldsymbol{n}}}|} \right. \\ &\qquad \left. \exp\left(- \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t) \right)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t) \right) \right) \right] \\ &= -NL \ln \pi - L \ln |\boldsymbol{R}_{\tilde{\boldsymbol{n}}}| - \sum_{i=1}^{L} \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t) \right)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t) \right)\end{aligned} \tag{2.34}$$

If we expand the above equation and omit the constant terms and the terms not related to the unkowns of the problem, the log-likelihood function becomes [86]

$$\begin{aligned}\mathcal{L}^{'}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_L;\boldsymbol{\theta}\right) = -\sum_{i=1}^{L}\big[&-\gamma_{t,i}\boldsymbol{x}_i^H\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)\\ &-\gamma_{t,i}^*\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{x}_i+\gamma_{t,i}^*\gamma_{t,i}\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)\big].\end{aligned} \tag{2.35}$$

From this point on, derivation changes according to $\gamma_{t,i}$ changing sample to sample, or staying constant at every snapshot.

SOI Amplitudes Changing From Sample to Sample :

Taking derivative of Eq.2.35 with respect to $\gamma_{t,i}^*$ and equating the resultant expression to 0, we obtain:

$$\frac{\partial\mathcal{L}^{'}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_L;\boldsymbol{\theta}\right)}{\partial\gamma_{t,i}^*} = -\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{x}_i+\gamma_{t,i}\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)=0. \tag{2.36}$$

Then, $\hat{\gamma}_{t,i}$ can be found as

$$\hat{\gamma}_{t,i} = \frac{\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{x}_i}{\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)} \tag{2.37}$$

We put $\hat{\gamma}_{t,i}$ in Eq.(2.35) and maximize the likelihood function with respect to SOI angle. This gives the estimator for SOI angle as

$$\hat{\theta}_t = \arg\max_{\theta}\frac{\sum_{i=1}^{L}\left|\boldsymbol{a}^H(\theta)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{x}_i\right|^2}{\boldsymbol{a}^H(\theta)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta)}. \tag{2.38}$$

SOI Amplitudes Staying Constant at All Samples :

For this case, the likelihood function in Equation (2.35) can be written as

$$\begin{aligned}\mathcal{L}^{'}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_L;\boldsymbol{\theta}\right) = -\sum_{i=1}^{L}\big[&-\gamma_{t}\boldsymbol{x}_i^H\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)\\ &-\gamma_{t}^*\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{x}_i+\gamma_{t}^*\gamma_{t}\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)\big].\end{aligned} \tag{2.39}$$

Taking derivative of Eq.2.39 with respect to γ_t^* and equating the resultant expression to 0, we obtain:

$$\frac{\partial\mathcal{L}\left(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_L;\boldsymbol{\theta}\right)}{\partial\gamma_t^*} = \sum_{i=1}^{L}\left[-\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{x}_i+\gamma_t\boldsymbol{a}^H(\theta_t)\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)\right] \tag{2.40}$$

Then, $\hat{\gamma}_{t,i}$ can be found as

$$\hat{\gamma}_t = \frac{\boldsymbol{a}(\theta_t)^H\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\sum_{i=1}^{L}\boldsymbol{x}_i}{\sum_{i=1}^{L}\boldsymbol{a}(\theta_t)^H\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)} = \frac{1}{L}\frac{\boldsymbol{a}(\theta_t)^H\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\sum_{i=1}^{L}\boldsymbol{x}_i}{\boldsymbol{a}(\theta_t)^H\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}\boldsymbol{a}(\theta_t)} \tag{2.41}$$

We put $\hat{\gamma}_{t,i}$ in Eq.(2.39) and maximize the likelihood function with respect to SOI angle. This gives the estimator for SOI angle as [87].

$$\hat{\theta}_t = \arg\max_{\theta} \frac{\left| \boldsymbol{a}(\theta)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \sum_{i=1}^{L} \boldsymbol{x}_i \right|^2}{\boldsymbol{a}(\theta)^H \boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1} \boldsymbol{a}(\theta)} \tag{2.42}$$

In the above equations, $\boldsymbol{R}_{\tilde{\boldsymbol{n}}} = \boldsymbol{I} + \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$ is the jammer plus noise covariance matrix. For the intermittent jamming set-up, $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ matrix becomes $\boldsymbol{R}_{\tilde{\boldsymbol{n}}} = \boldsymbol{I} + \alpha_1 \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$; since the jammer is present with probability α_1. In practice, $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ is not available to the receiver and has to be estimated from target-free auxiliary data. Estimation of the covariance matrix is also an important and difficult problem on its own; but, we assume that $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ is exactly known by the conventional estimator in this study. Availability of exact $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}$ gives advantage to the conventional estimators, which is obviously not present in practice, in the performance comparisons.

2.3.2 Behaviour of MVDR Beamformer

First of all, we consider the beampattern obtained by an array. When the antenna pattern of each element in an array is ignored or assumed to be identical and equal to an isotropic antenna with a gain of 1, the beam pattern, B, for an N element uniform sensor array can be defined as the normalized array factor which

$$B(\theta) = \frac{\boldsymbol{h}^H \boldsymbol{a}_{\boldsymbol{\theta}}}{\max\{|\boldsymbol{h}^H \boldsymbol{a}_{\boldsymbol{\theta}}|\}}. \tag{2.43}$$

In the above equation, $\boldsymbol{h}^H \boldsymbol{a}_{\boldsymbol{\theta}}$ is the so called Array Factor (AF) and $\boldsymbol{h}$ is the weight vector used in the array.

For example, the beampattern for a uniform array consisting of 10 elements and interelement spacing of $\lambda/2$ is represented in Figure 2.2. To calculate the beampattern, it is assumed that all the elements of the weight vector of the array are equal to 1.

It is clear that for the MVDR beamformer, the beampattern is obtained by replacing $\boldsymbol{h}$ in Eq.(2.43) with the coefficients of MVDR beamformer, i.e., $\boldsymbol{h}_{mvdr}$ in Eq.(2.10). For comparison of a beampattern and angular power spectrum of MVDR beamformer, Figure 2.3 is presented. In the figure, there is a SOI at the angle of 0° with a power of

0 dB besides a jammer being at $35°$ with a power of 20 dB. The figure is obtained for a ULA of 10 elements. For the figure, a total of 20 snapshots is used with a probability of jammer signal component being equal to 0.5, i.e., $\alpha_1 = 0.5$. Also, true covariance matrices of noise and interference are used to calculate MVDR beamformer coefficients.

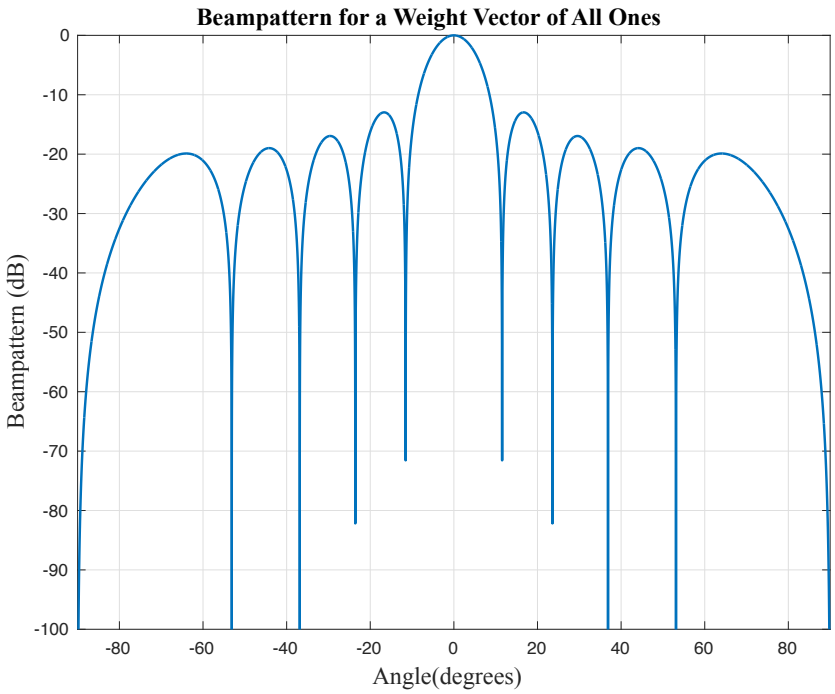


Figure 2.2: Beampattern for a weight vector of all ones

It is well known that in the beam patterns created by MVDR beamformers, there are deep nulls at the angles where the interferers are located as seen in Figure (2.3). Based on this information, one may expect the same thing in angular power spectrum of the beamformer. But, it is interesting to encounter peaks at the angles of the interferers in the power spectrum used for direction finding as illustrated in the same figure mentioned. To understand the reason behind this phenomena, let us look at the structure of the angular power spectrum in $P(\theta)$ in Eq.(2.11) more deeply.

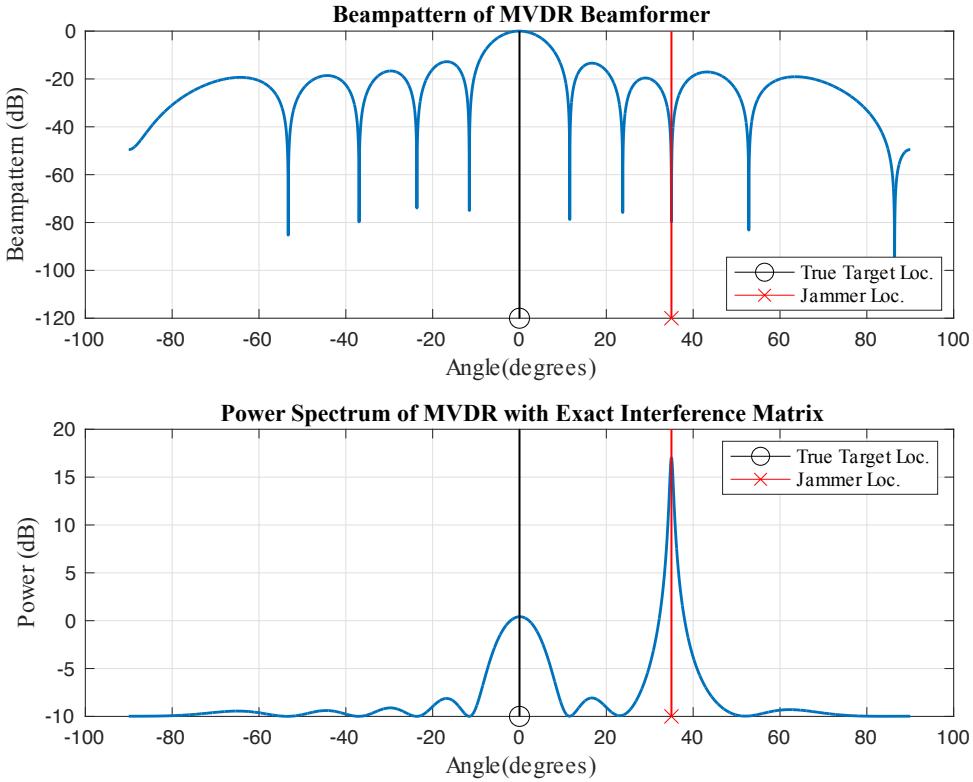


Figure 2.3: Beampattern vs power spectrum, $\theta_t = 0°$, $\theta_j = 35°$, $\alpha_1 = 0.5$

When Eq.(2.13) and Eq.(2.31) are inserted into formulation of $P(\theta)$ in Eq.(2.11), the following open form expression is obtained

$$\begin{aligned} P(\theta) &= \frac{\boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1}}{\boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{\theta}}} \left\{ \sigma_t^2 \boldsymbol{a}_{\boldsymbol{t}} \boldsymbol{a}_{\boldsymbol{t}}^H + \boldsymbol{R}_{\boldsymbol{e}} \right\} \frac{\boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{\theta}}}{\boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{\theta}}} \\ &= \frac{\sigma_t^2 \left| \boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{t}} \right|^2}{\left(\boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{\theta}} \right)^2} + \frac{1}{\boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{\theta}}}. \end{aligned} \tag{2.44}$$

Let us have a closer look to the denominator of the second term in Eq.(2.44). If Eq.(2.30) is inserted in the expression of the denominator using the matrix inversion identity $(\boldsymbol{I} + \boldsymbol{a}\boldsymbol{b}^H)^{-1} = \boldsymbol{I} - \frac{\boldsymbol{a}\boldsymbol{b}^H}{1+\boldsymbol{b}^H\boldsymbol{a}}$, the following equation is obtained:

$$\begin{aligned} \boldsymbol{a}_{\boldsymbol{\theta}}^H \boldsymbol{R}_{\boldsymbol{e}}^{-1} \boldsymbol{a}_{\boldsymbol{\theta}} &= \boldsymbol{a}_{\boldsymbol{\theta}}^H \left(\alpha_1 \sigma_j^2 \boldsymbol{a}_{\boldsymbol{j}} \boldsymbol{a}_{\boldsymbol{j}}^H + \boldsymbol{I} \right)^{-1} \boldsymbol{a}_{\boldsymbol{\theta}} \\ &= \boldsymbol{a}_{\boldsymbol{\theta}}^H \left(\boldsymbol{I} - \frac{\alpha_1 \sigma_j^2 \boldsymbol{a}_{\boldsymbol{j}} \boldsymbol{a}_{\boldsymbol{j}}^{\boldsymbol{H}}}{1 + \alpha_1 \sigma_j^2 \boldsymbol{a}_{\boldsymbol{j}}^{\boldsymbol{H}} \boldsymbol{a}_{\boldsymbol{j}}} \right) \boldsymbol{a}_{\boldsymbol{\theta}} \\ &= N - \frac{\alpha_1 \sigma_j^2}{1 + N \alpha_1 \sigma_j^2} \left| \boldsymbol{a}_{\boldsymbol{j}}^{\boldsymbol{H}} \boldsymbol{a}_{\boldsymbol{\theta}} \right|^2. \end{aligned} \tag{2.45}$$

For very big jammer power, Eq.(2.45) can be assumed to be approximately equal to $\boldsymbol{a_\theta^H R_e^{-1} a_\theta} \approx N - \left|\boldsymbol{a_j^H a_\theta}\right|^2 / N$. Additionally, the numerator of the first term in Eq.(2.44) is worthy of attention. Following the previous procedure of replacing $\boldsymbol{R_e^{-1}}$ with its open form, the numerator turns into

$$\left|\boldsymbol{a_\theta^H R_e^{-1} a_t}\right|^2 = \left|\boldsymbol{a_\theta^H a_t} - \frac{\alpha_1 \sigma_j^2}{1 + N\alpha_1\sigma_j^2} \boldsymbol{a_\theta^H a_j a_j^H a_t}\right|^2 . \tag{2.46}$$

Under the assumption of very big jammer power, the numerator of the first term in Eq.(2.44) becomes

$$\left|\boldsymbol{a_\theta^H R_e^{-1} a_t}\right|^2 = \left|\boldsymbol{a_\theta^H a_t}\right|^2 + \frac{1}{N^2} \left|K_{tj}\right|^2 \left|\boldsymbol{a_\theta^H a_j}\right|^2 - \frac{2}{N} Re\left\{K_{tj} \boldsymbol{a_\theta^H a_t a_j^H a_\theta}\right\} \tag{2.47}$$

where $K_{tj} = \boldsymbol{a_t^H a_j}$.

At this point, two different states are worthy of consideration:

State 1: If the angle of the beamformer is equal to the AOA of SOI, i.e., $\theta = \theta_t$ and $\boldsymbol{a_\theta} = \boldsymbol{a_t}$, angular power spectrum becomes

$$P(\theta)\,|_{\theta=\theta_t} = \sigma_t^2 + \frac{1}{N - \frac{1}{N}\left|K_{tj}\right|^2}. \tag{2.48}$$

Moreover, if signal and jammer array manifold vectors are nearly orthogonal to each other or in other words SOI and jammer signals are spatially well separated, angular power spectrum becomes $P(\theta_t) = \sigma_t^2 + 1/N$ since $K_{tj} \approx 0$.

State 2: If the angle of the beamformer is equal to the AOA of jammer, i.e., $\theta = \theta_j$ and $\boldsymbol{a_\theta} = \boldsymbol{a_j}$, angular power spectrum becomes

$$P(\theta)\,|_{\theta=\theta_j} = \frac{1}{\boldsymbol{a_j^H R_e^{-1} a_j}} = \frac{1 + N\alpha_1\sigma_j^2}{N} = \alpha_1\sigma_j^2 + \frac{1}{N} > 0 \tag{2.49}$$

because the term $\left|\boldsymbol{a_\theta^H R_e^{-1} a_t}\right|^2$ in Eq.(2.44) vanishes as presented in the following equation:

$$\left|\boldsymbol{a_\theta^H R_e^{-1} a_t}\right|^2 = \left|K_{tj}\right|^2 - \frac{2}{N} N \left|K_{tj}\right|^2 + \frac{1}{N^2}\left|K_{tj}\right|^2 N^2 = 0. \tag{2.50}$$

It is obvious that Eq.(2.49) explains the peak at the angle of the interferer in the power spectrum given in Figure (2.3). The difference between antenna beampattern and power spectrum is also examined in [88].

2.3.3 Performance Comparison of Capon, MUSIC and ML Estimation Methods Under Signal Intermittency

To compare the peformance of Capon, MUSIC and ML DOA estimation methods, we implemented them as described in Eqs. (2.17), (2.27) and (2.38) respectively. The angular power spectrums obtained by a 10-element ULA after collecting 20 snapshots, when there is a 0 dB SOI at an angle of $0°$ together with an interfering source of 20 dB at an angle of $10°$, is presented in Figure 2.4. Interfering source existance probability, α_1, is assumed to be equal to 0.5. It is also assumed that exact interference plus noise covariance matrix $\boldsymbol{R}_{\tilde{n}}/\boldsymbol{R}_e$ is somehow supplied to the estimators whenever necessary. Besides this sample autocorrelation matrix is used by Capon and MUSIC algorithms.

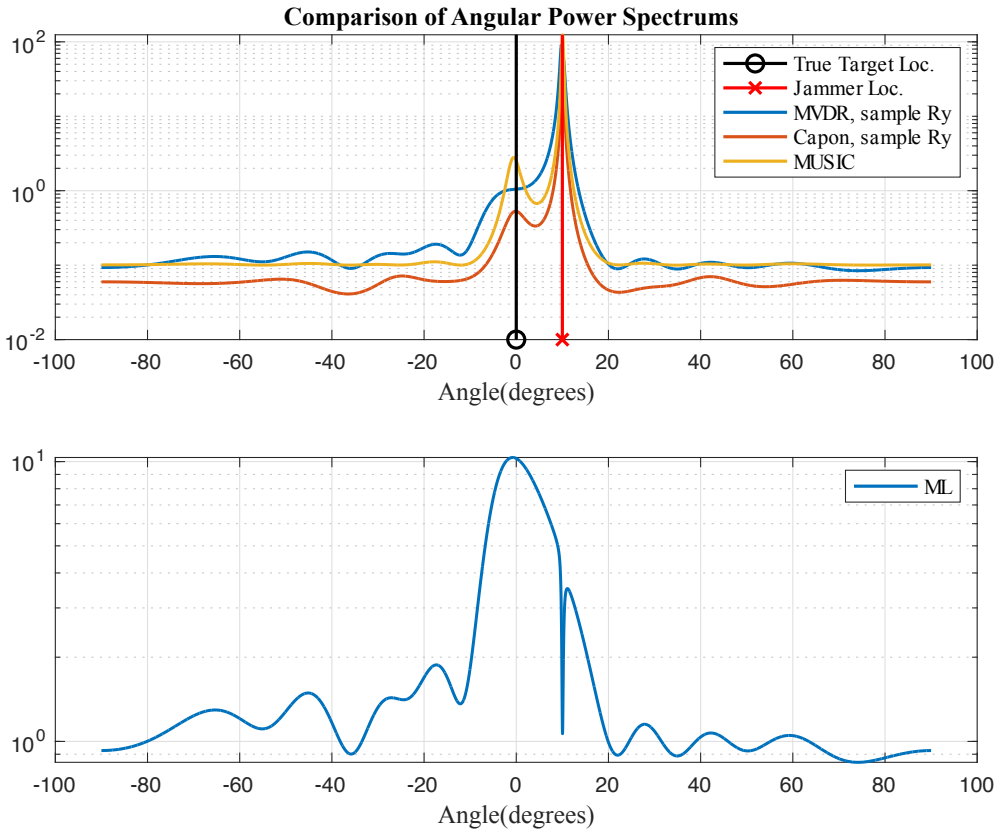


Figure 2.4: Comparison of Power Spectrums, $\theta_t = 0°$, $\theta_j = 10°$, $\alpha_1 = 0.5$

It can be seen from Figure 2.4 that Capon and MUSIC algorithms can seperate SOI and interfering source, whereas ML estimator puts a null at the location of the interferer making it easier to estimate the SOI angle at one step. The MVDR coefficients as defined in Eq. (2.10) is used together with the estimator in Eq. (2.12) directly and the obtained angular power spectrum is also presented in Figure 2.4. It can be concluded that, it is not possible to differentiate SOI and interferer by this approach and this explains why estimators that are called MVDR in literature offers the estimator in Eq. (2.17) instead of the one which uses MVDR coefficients directly.

Monte Carlo simulations are also run to compare the performance of Capon, MUSIC and Maximum Likelihood Methods under the aforementined parameter set except that the SOI power is being varied in this case. The resultant angular accuracy versus SNR graph is presented in Figure 2.5. A similar simulation is also run with the same parameters except the unwanted signal is carried to 35° and the obtained graph is presented in Figure 2.6. When the two figures are compared, it can be said that ML method outperforms both MUSIC and Capon methods as expected. Because the ML estimator has all the information needed and the SOI angle is left as the only unknown parameter. One thing to be also mentioned is that when the two sources are closer to each other, the SNR of the SOI becomes more inportant. The more seperation between the sources, easier for the method to estimate their directions at low SNR as can be seen when Figures 2.5 and 2.6 are compared. Because of the information advantage of ML estimator described in this section, it is used as a benchmark for comparison of the suggested estimator in the next section. As a final note, although MUSIC algorithm is at a disadvantage compared to ML estimator in this setting, its performance nearly achieves the performance of ML.

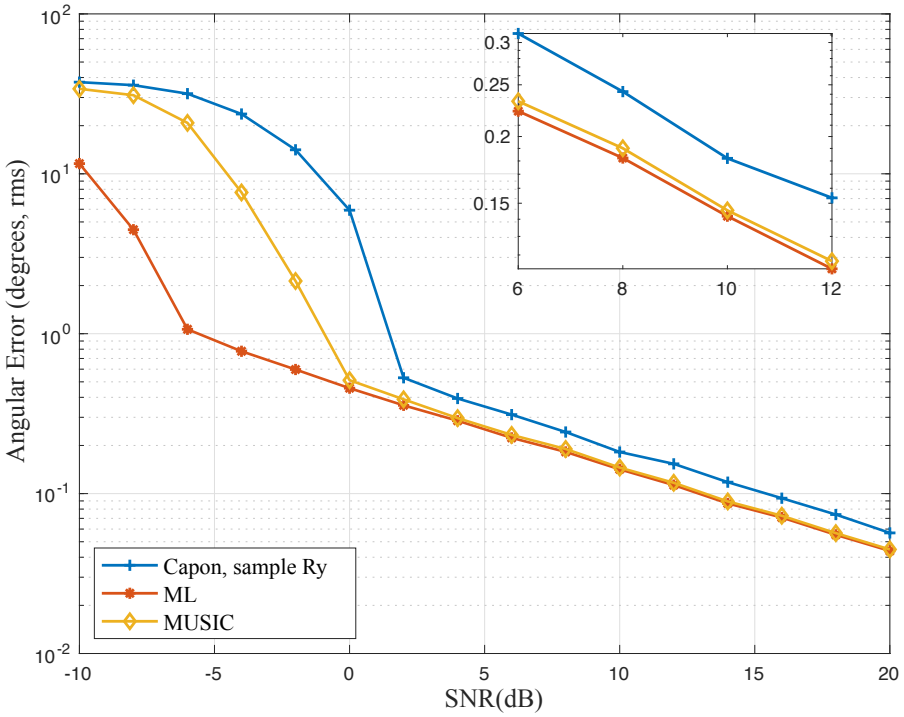


Figure 2.5: Comparison of Methods, $\theta_t = 0°$, $\theta_j = 10°$, $\alpha_1 = 0.5$, JNR = 20 dB

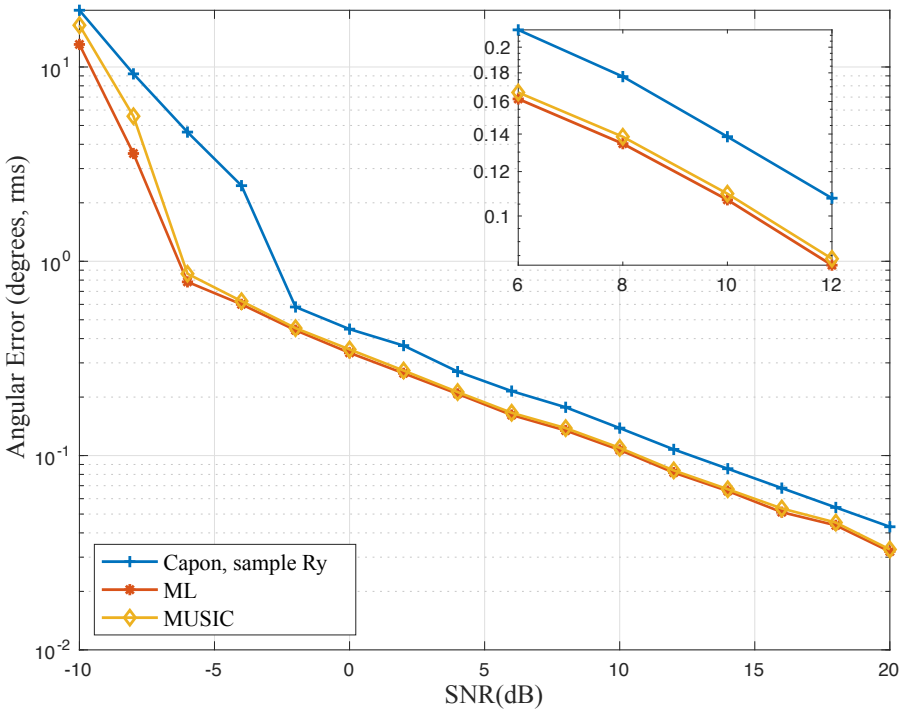


Figure 2.6: Comparison of Methods, $\theta_t = 0°$, $\theta_j = 35°$, $\alpha_1 = 0.5$, JNR = 20 dB

2.4 Theoretical Preliminaries

2.4.1 Expectation-Maximization Algorithm

For some signal models, it is not easy to obtain ML estimators due to several reasons. For instance, it may not be easy to differentiate the likelihood function with respect to unknowns of the problem because of the structure of the probability density function of the underlying signal. A very typical example to this situation is mixture noise models. Even differentiation is possible, closed form estimators cannot be reached. In those cases, EM Method is exploited.

EM Method depends on augmenting observations, $\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L\}$, by introducing latent variables , $\boldsymbol{Y} = \{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_L\}$, and maximizing the expected value of the complete data log-likelihood, $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right) = E_{\boldsymbol{Y}|\boldsymbol{X}}\{\ln p\left(\boldsymbol{X}, \boldsymbol{Y}\right) | \boldsymbol{X}; \boldsymbol{\theta}^{\text{old}}\}$, instead of log-likelihood of the snapshots. The trick of this method lies in selecting the latent variables in such a way that closed form estimators can be obtained.

The method consists of two phases called *Expectation Phase* and *Maximization Phase* that follow each other iteratively [89]. In the expectation phase, the posterior density $p\left(\boldsymbol{Y}|\boldsymbol{X}; \boldsymbol{\theta}^{\text{old}}\right)$, which is the density of latent variables $\boldsymbol{Y}$ given the observations $\boldsymbol{X}$ and $\boldsymbol{\theta}^{\text{old}}$, is computed and then the complete-data $\{\boldsymbol{X}, \boldsymbol{Y}\}$ log-likelihood function is ensemble averaged with respect to the posterior density to get $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)$. In the maximization phase, $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)$ is maximized with respect to $\boldsymbol{\theta}$ to obtain the updated non-random parameter estimates, say $\boldsymbol{\theta}^{\text{new}}$, to be utilized in the next iteration.

EM method starts with an initial estimate for the unknown variables denoted by $\boldsymbol{\theta}^{\text{old}}$. These estimates can be used to calculate initial value of likelihood. At the end of each maximization phase convergence of either the estimates or the likelihood function is checked according to a termination criterion. Expectation phase is returned unless the termination criterion is satisfied [89].

Under certain conditions, EM method is guaranteed to converge,at least to a local maximum [90, 91].

2.4.2 Cramer Rao Lower Bound [1]

To have an idea about the performance of any estimator, Cramer Rao Lower Bound (CRLB) or Cramer Rao Bound (CRB) is used as a fundamental lower limit. It is known that if the sample size and signal-to-noise ratio (SNR) are high enough, CRB is a tight bound for a wide class of estimators [1].

Let $\boldsymbol{x}(\boldsymbol{\theta}, \boldsymbol{y})$ be a noise corrupted observation where $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^T$ is a vector parameter to be estimated. Assuming an estimator $\hat{\boldsymbol{\theta}}$ is unbiased, CRLB places a bound on variance of each element of $\hat{\boldsymbol{\theta}}$.

CRLB is found as diagonal elements of the inverse of $p \times p$ Fisher Information Matrix (FIM), $\boldsymbol{I}$, or in other words,

$$\text{var}(\hat{\theta}_i) \geq \left[\boldsymbol{I}^{-1}(\boldsymbol{\theta})\right]_{ii}. \tag{2.51}$$

Note that, the $[ij]^{th}$, $i, j \in \{1, \cdots, p\}$, element of FIM is defined as

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{ij} = -E\left[\frac{\partial^2 \ln p(\boldsymbol{x};\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\right] \tag{2.52}$$

where $p(\boldsymbol{x};\boldsymbol{\theta})$ is the probability density of the observations and true values of the unknowns are used while evaluating the bound.

For a Gaussian vector, the entries of FIM can be written as [1, p. 525],

$$\begin{aligned}[\boldsymbol{I}(\boldsymbol{\theta})]_{i,j} = \text{tr}\left[\boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{C}_{\boldsymbol{x}}(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{C}_{\boldsymbol{x}}(\boldsymbol{\theta})}{\partial\theta_j}\right] \\ + 2\,\text{Re}\left[\frac{\partial \boldsymbol{\mu}^H(\boldsymbol{\theta})}{\partial\theta_i}\boldsymbol{C}_{\boldsymbol{x}}^{-1}(\boldsymbol{\theta})\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right]\end{aligned} \tag{2.53}$$

where θ_i and θ_j are i'th and j'th unknown parameters to be estimated, $C_{\boldsymbol{x}}$ is the covariance matrix, $\boldsymbol{\mu}$ is the mean vector of the observations.

As a final note, if $\boldsymbol{\theta}$ contains a mixture of real and complex unknowns, the real and imaginary part of the complex unknowns are defined separately, the real unknown vector $\boldsymbol{\xi}$ is formed by using these new real unkowns and the CRB formulation is applied.

2.4.3 Modified Cramer Rao Lower Bound [2]

In some problems, because of some random parameters in the formulation of signal model, it is very hard to derive CRB, even if possible. For these cases, different alternative bounds like Hybrid Cramer Rao Bound (HCRB), Modified Cramer Rao Bound (MCRB) and Miller-Chang Bound (MCB) are offered in literature. It is mentioned in [90] that, generally speaking, conventional CRB is tighter than HCRB and MCRB is looser than HCRB.

Throughout this thesis, MCRB is used to evaluate the performance of the proposed estimators not only for the DOA estimation problem defined in Chapter 3, but also for the frequency estimation problem solved in Chapter 4 for the following reasons:

- There are random nuisance parameters that make the distribution of observations Gaussian mixture with two components. So, it is mathematically intractable to develop CRB for the examined cases.
- Some or all random nuisance parameters of the proposed signal models are discrete not allowing the development of HCRB.

Let $\boldsymbol{x}(\boldsymbol{\theta}, \boldsymbol{y})$ be a noise corrupted observation where $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^T$ is a vector parameter to be estimated and $\boldsymbol{y} = [y_1, y_2, \cdots, y_q]^T$ is a random vector which contains nuisance parameters of the signal model. Let us assume that the following regularity conditions are hold:

1. The domain $\{\boldsymbol{x} : p_{\boldsymbol{x}}(\boldsymbol{x}; \boldsymbol{\theta}) > 0\}$ does not depend on $\boldsymbol{\theta}$.
2. The partial derivatives $\{\partial \ln p_{\boldsymbol{x}}(\boldsymbol{x}; \boldsymbol{\theta})/\partial\theta_i, 1 \leq i \leq p\}$ exist and satisfy the relations
$$\int \frac{p_{\boldsymbol{x}}(\boldsymbol{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} d\boldsymbol{x} = \frac{\partial}{\partial \boldsymbol{\theta}} \int p_{\boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{\theta}) d\boldsymbol{x} = \mathbf{0} \tag{2.54}$$
for all θ.
3. The partial derivative of $p_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ exists and has a finite second-order moment. Also, the second partial derivative exists and has finite first-order moment.

4. The joint pdf of the nuisance parameters does not depend on unknowns.

Then, Modified Cramer Rao Bound (MCRB) to the error covariance matrix $\boldsymbol{C}_{\hat{\boldsymbol{\theta}}}$ of an unbiased estimator of $\boldsymbol{\theta}$, i.e., $[\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_p]^T$, satisfies the following inequality

$$\boldsymbol{C}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{I}_M^{-1}(\theta) \geq \mathbf{0} \tag{2.55}$$

where $\geq \mathbf{0}$ means that the matrix $\boldsymbol{C}_{\hat{\boldsymbol{\theta}}} - \boldsymbol{I}_M^{-1}(\theta)$ is positive semi-definite. In Equation (2.55) , $\boldsymbol{I}_M$ is the Modified Fisher Information Matrix (MFIM) which is defined as

$$\begin{aligned} \left[\boldsymbol{I}_M\left(\boldsymbol{\theta}\right)\right]_{i,j} &= E_{\boldsymbol{x},\boldsymbol{y}}\left\{\frac{\partial \ln p_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x} \mid \boldsymbol{y};\boldsymbol{\theta})}{\partial\theta_i}.\frac{\partial \ln p_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x} \mid \boldsymbol{y};\boldsymbol{\theta})}{\partial\theta_j}\right\} \\ &= -E_{\boldsymbol{x},\boldsymbol{y}}\left\{\frac{\partial^2 \ln p_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x} \mid \boldsymbol{y};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\} \end{aligned} \tag{2.56}$$

Note that, in the above equation, expectation is taken with respect to $p_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y};\boldsymbol{\theta})$, and the derivatives are computed at the true value of the unknown vector.

CHAPTER 3

PROPOSED METHOD FOR DIRECTION OF ARRIVAL ESTIMATION UNDER INTERMITTENT JAMMING

3.1 Signal Model and Background

For the sake of completeness of this chapter, let us repeat the signal model of our problem which is defined at Section 2.3 at the beginning of the current section. Reconsider the following set of L vectors,

$$\boldsymbol{x}_i = \underbrace{\gamma_{t,i}\boldsymbol{a}(\theta_t)}_{\text{target}} + \underbrace{I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)}_{\text{jammer}} + \underbrace{\boldsymbol{n}_i}_{\text{noise}}, \quad i = \{1,\ldots,L\}. \tag{3.1}$$

where $\boldsymbol{x}_i$ is an $N \times 1$ vector denoting the i'th snapshot vector, $i = \{1,\ldots,L\}$. The vector $\boldsymbol{x}_i$ is composed of three components which are signal of interest (target signal), jamming signal and noise.

The noise component $\boldsymbol{n}_i$ in (3.1) is a complex-valued vector whose entries are independent and identically distributed (iid) with zero mean, unit variance circularly symmetric complex Gaussian distribution, $\boldsymbol{n}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I})$. The entries of the noise vector are assumed to model the electronic noise or other unmodeled phenomena affecting the output of sensors. With this definition, the noise is assumed to be independent from sensor to sensor (elements of the vector $\boldsymbol{x}_i$) and also from snapshot to snapshot.

The jamming component in (3.1) is expressed as $I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)$, which is a product of three terms: The first term, $I_i \in \{0,1\}$, is a binary valued random variable taking values 0 and 1. This variable indicates the presence or absence of jamming signal. The probability of jamming on the i'th snapshot is given as $P(I_i = 1) = \alpha_1$. The

complementary event (no jammer activity in the i'th snapshot) has the probability $P(I_i = 0) = \alpha_0 = 1 - \alpha_1$. The second term, $\gamma_{j,i}$, is a zero mean complex Gaussian random variable with variance σ^2 and represents the jammer complex amplitude for the i'th snapshot, $\gamma_{j,i} \sim \mathcal{CN}(\gamma_{j,i}; 0, \sigma^2)$. The third term is the vector $\boldsymbol{a}(\theta_j)$ which is the array manifold vector for the jammer located at the angular position of θ_j. For a uniform linear array, as in Figure (2.1), the array manifold vector becomes $\boldsymbol{a}(\theta) = [1, e^{-jkd\sin(\theta)}, \cdots, e^{-jkd\sin(\theta)(N-1)}]^T$ where d is the inter-sensor spacing and $k = 2\pi/\lambda$ is the wavenumber. The jammer-to-noise ratio is defined as the ratio of the jamming signal power to the noise power at a receive element. Since the amplitude of the jammer signal is Gaussian distributed with a variance of σ^2 and the noise power is assumed to be unity at each receive element, the jammer-to-noise ratio (JNR) coincides with σ^2, i.e. JNR $= \sigma^2$. Jammer power σ^2 and jamming probability α_1 are among the unknowns of the problem.

The target component in (3.1) is expressed as $\gamma_{t,i}\boldsymbol{a}(\theta_t)$ where $\gamma_{t,i}$ is the non-random complex-valued target signal amplitude of the i'th snapshot and θ_t is the angular position of the target. Both $\gamma_{t,i}$ and θ_t are unknowns of the problem. Target signal is assumed to be present in all snapshots. If the target amplitudes $\gamma_{t,i}$ are independent variables, that is if we have a set of L non-random unknown variables for L snapshots, we denote the target model as *non-coherent target model*. If the target amplitudes $\gamma_{t,i}$ are deterministically related, then we denote the target model as *coherent target model*. In the coherent model, one may assume that the difference between the phases of complex numbers $\gamma_{t,i}$ and $\gamma_{t,i+1}$ is deterministically known, i.e. pre-determined, which is the case for the snapshots collected with the pulse-Doppler radars for Swerling-1/3 targets [9]. Without any loss of generality, we assume that the deterministic phase progression at complex target amplitudes are undone, by multiplying both sides of equation (3.1) with the conjugate of the known phase sequence to get $\gamma_t = \gamma_{t,1} = \gamma_{t,2} = \ldots = \gamma_{t,L}$. Hence for the coherent target model, γ_t is the sole unknown that models the complex target amplitude.

In this problem, the values of the indicator variables for jammer activity I_i, $i = \{1, \ldots, L\}$ are not known to the receiver. In addition, the probability of jamming (α_1) is also an unknown of the problem. The jammer angular location θ_j is assumed to be known. This is an acceptable assumption for reconnaissance systems, since

once the jammer is detected, the jammer angular location is flagged and subsequent signal processing operations such as target detection, direction finding etc. is done with the knowledge of jammer angular location. Furthermore, in the special case of unintentional jammer, say a friendly radar system jamming DOA estimation receiver intermittently, the angular localization of jammer can be known precisely.

3.2 Proposed Method - Intermittent Jamming, Unknown Jamming Power

The unknown parameters $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$ appearing in the observation model $\boldsymbol{x}_i = \gamma_{t,i}\boldsymbol{a}(\theta_t) + I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) + \boldsymbol{n}_i, \quad i = \{1, \ldots, L\}$, are non-random variables to be determined by the likelihood maximization. The probability density function of the observation vector $\boldsymbol{x}_i$ is $p(\boldsymbol{x}_i; \boldsymbol{\theta}) = \alpha_0 \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I}) + \alpha_1 \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + \sigma^2\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j))$.

Remembering the fact that the snapshot $\boldsymbol{x}_i$ is independent from $\boldsymbol{x}_m$, $i \neq m$; the log-likelihood of L snapshots becomes

$$\begin{aligned}
&\mathcal{L}\left(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L; \boldsymbol{\theta}\right) = \\
&\sum_{i=1}^{L} \Bigg[\ln \Bigg(\alpha_0 \frac{1}{\pi^N} \exp\left(-\left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right)^H \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right) \right) + \\
&\quad \alpha_1 \frac{1}{\pi^N \left|\boldsymbol{I} + \sigma^2\boldsymbol{a}(\theta_j)\boldsymbol{a}(\theta_j)^H\right|} \\
&\qquad \exp\left(-\left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right)^H \left(\boldsymbol{I} + \sigma^2\boldsymbol{a}(\theta_j)\boldsymbol{a}(\theta_j)^H\right)^{-1} \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right) \right) \Bigg) \Bigg].
\end{aligned} \tag{3.2}$$

Since the argument of the logarithm function in (3.2) involves summation of two terms, it is not possible to have an exact analytical solution to the likelihood maximization problem. We use Expectation Maximization (EM) algorithm to iteratively maximize the likelihood via the latent variables [31].

We denote the set of observations and latent variables as $\boldsymbol{X} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L\}$ and $\boldsymbol{Y} = \{I_1, \gamma_{j,1}, I_2, \gamma_{j,2}, \cdots, I_L, \gamma_{j,L}\}$, respectively. EM methodology assumes an initial estimate for the unknown variables denoted as $\boldsymbol{\theta}^{\text{old}}$. In the expectation phase, the posterior density $p\left(\boldsymbol{Y}|\boldsymbol{X}; \boldsymbol{\theta}^{\text{old}}\right)$, which is the density of latent variables $\boldsymbol{Y}$ given the observations $\boldsymbol{X}$ and $\boldsymbol{\theta}^{\text{old}}$, is computed and then the complete-data $\{\boldsymbol{X}, \boldsymbol{Y}\}$ log-likelihood function is ensemble averaged with respect to the posterior density to get

$Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right) = E_{\boldsymbol{Y}|\boldsymbol{X}}\{\ln p\left(\boldsymbol{X}, \boldsymbol{Y}\right)|\boldsymbol{X}; \boldsymbol{\theta}^{\text{old}}\}$. In the maximization phase, $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)$ is maximized with respect to $\boldsymbol{\theta}$ to obtain the updated non-random parameter estimates, say $\boldsymbol{\theta}^{\text{new}}$, to be utilized in the next iteration. In this study, the random variables associated with the jamming activity, which are $\{I_i\}_{i=1}^{L}$ and $\{\gamma_{j,i}\}_{i=1}^{L}$, are chosen as the latent variables, that is $\boldsymbol{Y} = \{I_1, \gamma_{j,1}, I_2, \gamma_{j,2}, \cdots, I_L, \gamma_{j,L}\}$. For the sake of notational simplicity, we also denote the latent variable pair for the *i*'th snapshot as $\boldsymbol{y}_i = \{I_i, \gamma_{j,i}\}$. Hence, we have $\boldsymbol{Y} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L\}$ and $\{\boldsymbol{X}, \boldsymbol{Y}\} = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_L\}$ for the latent variables and complete-data set, respectively.

Complete-data log-likelihood :

The joint pdf for the *i*'th snapshot variables $\{\boldsymbol{x}_i, \boldsymbol{y}_i\} = \{\boldsymbol{x}_i, I_i, \gamma_{j,i}\}$ can be written as

$$\begin{aligned} p(\boldsymbol{x}_i, \boldsymbol{y}_i; \boldsymbol{\theta}) &= p(\boldsymbol{x}_i, I_i, \gamma_{j,i}; \boldsymbol{\theta}) = p(\boldsymbol{x}_i \mid I_i, \gamma_{j,i}; \boldsymbol{\theta}) p(I_i; \boldsymbol{\theta}) p(\gamma_{j,i}; \boldsymbol{\theta}) \\ &= \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t) + I_i\gamma_{j,i}\boldsymbol{a}(\theta_j), \boldsymbol{I})\alpha_1^{I_i}\alpha_0^{1-I_i}\mathcal{CN}(\gamma_{j,i}; 0, \sigma^2). \end{aligned} \tag{3.3}$$

Hence, the joint pdf of $\{\boldsymbol{X}, \boldsymbol{Y}\}$ becomes

$$\begin{aligned} p(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{\theta}) = \prod_{i=1}^{L} \Bigg[\frac{1}{\pi^N} \exp\left(-\left\|\boldsymbol{x}_i - \gamma_{t,i}a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2\right) \\ \alpha_1^{I_i}\alpha_0^{1-I_i} \frac{1}{\pi\sigma^2} \exp\left(-\frac{|\gamma_{j,i}|^2}{\sigma^2}\right) \Bigg]. \end{aligned} \tag{3.4}$$

The logarithm of (3.4), so-called complete-data log-likelihood, is then

$$\begin{aligned} \ln p(\boldsymbol{X}, \boldsymbol{Y}; \boldsymbol{\theta}) = &-(N+1)L\ln(\pi) - \sum_{i=1}^{L}\left\|\boldsymbol{x}_i - \gamma_{t,i}a(\theta_t) - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\right\|^2 \\ &- \sum_{i=1}^{L}\frac{|\gamma_{j,i}|^2}{\sigma^2} + \ln(\alpha_1)\sum_{i=1}^{L}I_i + \ln(\alpha_0)\sum_{i=1}^{L}(1-I_i) \\ &- L\ln(\sigma^2). \end{aligned} \tag{3.5}$$

Posterior Density Calculation :

For the derivation of the $Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right)$, we need to calculate the posterior density of the latent variables given observations and the current values of unknowns, or in other words $p(\boldsymbol{Y} \mid \boldsymbol{X}; \boldsymbol{\theta}^{\text{old}})$. With the application of the Bayes' theorem, the posterior

density $p(\boldsymbol{y}_i \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}})$ can be written as

$$p(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}}) = \frac{p(I_i, \gamma_{j,i}, \boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}})}{p(\boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}})} = \frac{p(\boldsymbol{x}_i | I_i, \gamma_{j,i}; \boldsymbol{\theta}^{\text{old}}) p(I_i, \gamma_{j,i}; \boldsymbol{\theta}^{\text{old}})}{p(\boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}})}$$
$$= \frac{p(\boldsymbol{x}_i | I_i, \gamma_{j,i}; \boldsymbol{\theta}^{\text{old}}) p(I_i; \boldsymbol{\theta}^{\text{old}}) p(\gamma_{j,i}; \boldsymbol{\theta}^{\text{old}})}{p(\boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}})}$$
$$= \frac{\left(\alpha_1^{\text{old}}\right)^{I_i} \left(\alpha_0^{\text{old}}\right)^{1-I_i} \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}}) + I_i \gamma_{j,i} \boldsymbol{a}(\theta_j), \boldsymbol{I}) \mathcal{CN}(\gamma_{j,i}; 0, \sigma^{2,\text{old}})}{\alpha_1^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,old} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right) + \alpha_0^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I}\right)}. \tag{3.6}$$

Setting $I_i = 0$ in Eq. (3.6), yields the following simplified expression for the probability $p(I_i = 0, \gamma_{j,i} \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}})$

$$p(I_i = 0, \gamma_{j,i} \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}}) = M_{0,i} \mathcal{CN}(\gamma_{j,i}; 0, \sigma^{2,\text{old}}) \tag{3.7}$$

where $M_{0,i}$ is given as

$$M_{0,i} = \frac{\alpha_0^{\text{old}} \mathcal{CN}(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}}), \boldsymbol{I})}{\alpha_1^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right) + \alpha_0^{\text{old}} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I}\right)}. \tag{3.8}$$

When $I_i = 1$ is inserted in Eq.(3.6), the numerator of the ratio in Eq.(3.6) becomes $p(I_i = 1, \boldsymbol{x}_i, \gamma_{j,i}; \boldsymbol{\theta}^{\text{old}}) = p(I_i = 1; \boldsymbol{\theta}^{\text{old}}) p(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1; \boldsymbol{\theta}^{\text{old}})$. Note that the density $p(I_i = 1; \boldsymbol{\theta}^{\text{old}})$ is simply α_1^{old}. The other term $p(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1; \boldsymbol{\theta}^{\text{old}})$ can be written, by considering the marginal distributions and the correlation of the random variables $\boldsymbol{x}_i$ and $\gamma_{j,i}$ under the assumption of active jammer, as follows:

$$p\left(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}\right) = \mathcal{CN}\left(\begin{bmatrix} \boldsymbol{x}_i \\ \gamma_{j,i} \end{bmatrix}; \begin{bmatrix} \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}}) \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j) & \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \\ \sigma^{2,\text{old}} \boldsymbol{a}^H(\theta_j) & \sigma^{2,\text{old}} \end{bmatrix}\right). \tag{3.9}$$

To further simplify the probability density function $p\left(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}\right)$ in Eq. (3.9), we express $p\left(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}\right)$ as

$$p\left(\boldsymbol{x}_i, \gamma_{j,i} \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}\right) = p\left(\boldsymbol{x}_i \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}\right) p\left(\gamma_{j,i}, \mid \boldsymbol{x}_i, I_i = 1; \boldsymbol{\theta}^{\text{old}}\right). \tag{3.10}$$

The term $p\left(\boldsymbol{x}_i, \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}\right)$ can be explicitly written as

$$p(\boldsymbol{x}_i \mid I_i = 1; \boldsymbol{\theta}^{\text{old}}) = \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{\text{old}} \boldsymbol{a}\left(\theta_t^{\text{old}}\right), \boldsymbol{I} + \sigma^{2,\text{old}} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right). \tag{3.11}$$

The term $p\left(\gamma_{j,i},\mid \boldsymbol{x}_i, I_i=1;\boldsymbol{\theta}^{\text{old}}\right)$ can be written as

$$p(\gamma_{j,i},\mid \boldsymbol{x}_i, I_i=1;\boldsymbol{\theta}^{\text{old}}) = \mathcal{CN}\left(\gamma_{j,i};\frac{\sigma^{2,\text{old}}}{1+N\sigma^{2,\text{old}}}\boldsymbol{a}^H(\theta_j)\left(\boldsymbol{x}_i-\gamma_{t,i}^{\text{old}}\boldsymbol{a}\left(\theta_t^{\text{old}}\right)\right),\frac{\sigma^{2,\text{old}}}{1+N\sigma^{2,\text{old}}}\right) \tag{3.12}$$

considering the relation for the posterior density calculation for jointly Gaussian distributed random variables [1].

Finally, the density $p\left(I_i=1,\gamma_{j,i}\mid \boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}\right)$, which is nothing but the product of $p(I_i=1;\boldsymbol{\theta}^{\text{old}})=\alpha_1^{\text{old}}$, $p\left(\boldsymbol{x}_i,\mid I_i=1;\boldsymbol{\theta}^{\text{old}}\right)$ and $p(\gamma_{j,i},\mid \boldsymbol{x}_i, I_i=1;\boldsymbol{\theta}^{\text{old}})$ divided by $p(\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}})$, can be calculated from Eqs. (3.11) and (3.12) as

$$p(I_i=1,\gamma_{j,i}\mid \boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}) = M_{1,i}\mathcal{CN}\left(\gamma_{j,i};\frac{\sigma^{2,\text{old}}}{1+N\sigma^{2,\text{old}}}\boldsymbol{a}^H(\theta_j)\left(\boldsymbol{x}_i-\gamma_{t,i}^{\text{old}}\boldsymbol{a}\left(\theta_t^{\text{old}}\right)\right),\frac{\sigma^{2,\text{old}}}{1+N\sigma^{2,\text{old}}}\right) \tag{3.13}$$

where $M_{1,i}$ is defined as

$$M_{1,i} = \frac{\alpha_1^{\text{old}}\mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}^{\text{old}}\boldsymbol{a}\left(\theta_t^{\text{old}}\right),\boldsymbol{I}+\sigma^{2,\text{old}}\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}{\alpha_1^{\text{old}}\mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}^{\text{old}}\boldsymbol{a}\left(\theta_t^{\text{old}}\right),\boldsymbol{I}+\sigma^{2,\text{old}}\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)+\alpha_0^{\text{old}}\mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}^{\text{old}}\boldsymbol{a}\left(\theta_t^{\text{old}}\right),\boldsymbol{I}\right)}. \tag{3.14}$$

As a final note, we can easily verify from Eqs. (3.7) and (3.13), by marginalization operation, that $p(I_i=0\mid \boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}})=M_{0,i}$ and $p(I_i=1\mid \boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}})=M_{1,i}$. As expected, we have $M_{0,i}+M_{1,i}=1$.

Expectation Phase :

Using the fact that random variables $\{\boldsymbol{x}_i,\boldsymbol{y}_i\}$ are independent and identically distributed, the expected value of complete-data log-likelihood function can be written as

$$\begin{aligned} Q(\boldsymbol{\theta},\boldsymbol{\theta}^{\text{old}}) &= \boldsymbol{E}_{\boldsymbol{Y}\mid\boldsymbol{X};\boldsymbol{\theta}^{\text{old}}}\{\ln p(\boldsymbol{X},\boldsymbol{Y};\boldsymbol{\theta})\mid \boldsymbol{X};\boldsymbol{\theta}^{\text{old}}\} \\ &= \sum_{i=1}^{L} E_{\boldsymbol{y}_i\mid\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\ln p(\boldsymbol{x}_i,\boldsymbol{y}_i;\boldsymbol{\theta})\mid \boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}\}. \end{aligned} \tag{3.15}$$

The expectation operation is over the posterior density of the latent variables $\boldsymbol{y}_i=\{I_i,\gamma_{j,i}\}$, given in Eqs. (3.7) and (3.13):

$$E_{\boldsymbol{y}_i\mid\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\ln p(\boldsymbol{x}_i,\boldsymbol{y}_i;\boldsymbol{\theta})\} = \sum_{I_i=0}^{1}\int_{\gamma_{j,i}} \ln p(\boldsymbol{x}_i,\boldsymbol{y}_i;\boldsymbol{\theta})p\left(I_i,\gamma_{j,i}\mid \boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}\right)d\gamma_{j,i}. \tag{3.16}$$

The expression for the argument of the expectation operation, $\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i; \boldsymbol{\theta})$, that is the complete log-likelihood function for the *i*'th snapshot can be written from (3.5) as

$$\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i; \boldsymbol{\theta}) \stackrel{c}{=} -\left\|\boldsymbol{x}_i - \gamma_{t,i} a(\theta_t) - I_i \gamma_{j,i} \boldsymbol{a}(\theta_j)\right\|^2 + \ln(\alpha_1) I_i + \ln(\alpha_0)(1 - I_i) - \ln(\sigma^2) - \frac{|\gamma_{j,i}|^2}{\sigma^2}$$

where $\stackrel{c}{=}$ refers the equality of both sides apart from constant terms not affecting subsequent calculations.

The evaluation of $E_{\boldsymbol{y}_i|\boldsymbol{x}_i,\boldsymbol{\theta}^{\text{old}}}\{\ln p(\boldsymbol{x}_i, \boldsymbol{y}_i|\boldsymbol{\theta})\}$ requires the calculation of three expectations operations:

$$E_1 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{I_i\}, \tag{3.17}$$

$$E_2 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{|\gamma_{j,i}|^2\}, \tag{3.18}$$

$$E_3 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{x}_i - \gamma_{t,i} a(\theta_t) - I_i \gamma_{j,i} \boldsymbol{a}(\theta_j)\right\|^2\}. \tag{3.19}$$

The first expectation is straightforward to calculate

$$E_1 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{I_i\} = p(I_i = 1 \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{\text{old}}) = M_{1,i} \tag{3.20}$$

where $M_{1,i}$ is given in Eq. (3.14).

The second expectation E_2 can be evaluated from Eqs.(3.7) and eq.(3.13) as

$$\begin{aligned} E_2 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{|\gamma_{j,i}|^2\} &= \sum_{I_i=0}^{1} \int_{\gamma_{j,i}} |\gamma_{j,i}|^2 p\left(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}\right) d\gamma_{j,i} \\ &= M_{0,i}\sigma^{2,\text{old}} + M_{1,i}(P^{\text{old}} + |\bar{\gamma}_i^{\text{old}}|^2) \end{aligned} \tag{3.21}$$

where $\bar{\gamma}_i^{\text{old}}$ and P^{old} are the mean and variance of the Gaussian density given in Eq. (3.13), whose explicit expressions are

$$\bar{\gamma}_i^{\text{old}} = \left(\frac{\sigma^{2,\text{old}}}{1 + N\sigma^{2,\text{old}}}\right) \boldsymbol{a}^H(\theta_j) \left(\boldsymbol{x}_i - \gamma_{t,i}^{\text{old}} \boldsymbol{a}(\theta_t^{\text{old}})\right), \quad P^{\text{old}} = \frac{\sigma^{2,\text{old}}}{1 + N\sigma^{2,\text{old}}}. \tag{3.22}$$

The third expectation, $E_3 = E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{x}_i - \gamma_{t,i} a(\theta_t) - I_i \gamma_{j,i} \boldsymbol{a}(\theta_j)\right\|^2\}$, can be compactly written as $E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - I_i \gamma_{j,i} \boldsymbol{a}(\theta_j)\right\|^2\}$ by introducing $\boldsymbol{c}_i = \boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t)$.

The vector $\boldsymbol{c}_i$ can be treated as a non-random vector for the sake of this calculation and E_3 can be written as:

$$
\begin{aligned}
E_3 &= E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\|^2\} && (3.23)\\
&= \sum_{I_i=0}^{1}\int_{\gamma_{j,i}} \|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\|^2\, p\left(I_i, \gamma_{j,i} \mid \boldsymbol{x}_i, \boldsymbol{\theta}^{\text{old}}\right) d\gamma_{j,i}\\
&= M_{0,i}\,\|\boldsymbol{c}_i\|^2 + M_{1,i}E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1,\boldsymbol{\theta}^{\text{old}}}\{\|\boldsymbol{c}_i - \gamma_{j,i}\boldsymbol{a}(\theta_j)\|^2\}. && (3.24)
\end{aligned}
$$

Hence, the evaluation of E_3 in Eq.(3.24) reduces to the calculation of the following term

$$
E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1;\boldsymbol{\theta}^{\text{old}}}\{\|\boldsymbol{c}_i - \gamma_{j,i}\boldsymbol{a}(\theta_j)\|^2\}, \tag{3.25}
$$

which is an expectation over the complex Gaussian random variable $\gamma_{j,i}$ given the observation vector $\mathbf{x}$ under the assumption that jamming is active, i.e. $I_i = 1$. Using expression (3.13), we can express this density as $\mathcal{CN}(\gamma_{j,i}; \bar{\gamma}_i^{\text{old}}, P^{\text{old}})$ where $\bar{\gamma}_i^{\text{old}}$ and P^{old} are as stated in (3.22) and evaluate the expectation result as follows:

$$
\begin{aligned}
&E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1;\boldsymbol{\theta}^{\text{old}}}\{\|\boldsymbol{c}_i - \gamma_{j,i}\boldsymbol{a}(\theta_j)\|^2\}\\
&\quad= E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1;\boldsymbol{\theta}^{\text{old}}}\{\left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j) - (\gamma_{j,i} - \bar{\gamma}_i^{\text{old}})\boldsymbol{a}(\theta_j)\right\|^2\}\\
&\quad= \left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1;\boldsymbol{\theta}^{\text{old}}}\{|\gamma_{j,i} - \bar{\gamma}_i^{\text{old}}|^2\}\,\|\boldsymbol{a}(\theta_j)\|^2 && (3.26)\\
&\qquad\quad - 2\text{real}\left[(\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j))^H E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1;\boldsymbol{\theta}^{\text{old}}}\{\gamma_{j,i} - \bar{\gamma}_i^{\text{old}}\}\boldsymbol{a}(\theta_j)\right]\\
&\quad= \left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + P^{\text{old}}N.
\end{aligned}
$$

Note that, we use $E_{\gamma_{j,i}|\boldsymbol{x}_i,I_i=1;\boldsymbol{\theta}^{\text{old}}}\{\gamma_{j,i} - \bar{\gamma}_i^{\text{old}}\} = 0$ in the second line of (3.26). By inserting (3.26) into (3.24), we finalize the calculation of the third expectation:

$$
\begin{aligned}
E_3 &= E_{\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}^{\text{old}}}\{\|\boldsymbol{c}_i - I_i\gamma_{j,i}\boldsymbol{a}(\theta_j)\|^2\}\\
&= M_{0,i}\,\|\boldsymbol{c}_i\|^2 + M_{1,i}(\left\|\boldsymbol{c}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right\|^2 + P^{\text{old}}N).
\end{aligned} \tag{3.27}
$$

Once the expressions for three expectations, given by equations (3.20), (3.21) and (3.27), are inserted in Eq.(3.15), the E-phase of EM algorithm is completed:

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \stackrel{c}{=} & \ln(\alpha_0) \sum_{i=1}^{L} M_{0,i}^{\text{old}} + \ln(\alpha_1) \sum_{i=1}^{L} M_{1,i}^{\text{old}} - L \ln\left(\sigma^2\right) \\ & - \frac{1}{\sigma^2} \sum_{i=1}^{L} \left[M_{0,i}^{\text{old}} \sigma^{2,\text{old}} + M_{1,i}^{\text{old}} \left(P^{\text{old}} + \left| \bar{\gamma}_i^{\text{old}} \right|^2 \right) \right] \\ & - \sum_{i=1}^{L} M_{0,i}^{\text{old}} \left\| \boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t) \right\|^2 \\ & - \sum_{i=1}^{L} M_{1,i}^{\text{old}} \left(\left\| \boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t) - \bar{\gamma}_i^{\text{old}} \boldsymbol{a}(\theta_j) \right\|^2 + P^{\text{old}} N \right). \end{aligned} \tag{3.28}$$

Maximization Phase:

In the maximization phase, we optimize the expected complete log-likelihood with respect to the unknown parameters $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$. We start with the optimization of σ^2 (jamming power). By taking the derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$ with respect to σ^2 and equating the result to 0, the updated $\hat{\sigma}^2$ estimate becomes:

$$\hat{\sigma}^2 = \frac{1}{L} \sum_{i=1}^{L} \left[M_{0,i}^{\text{old}} \sigma^{2,\text{old}} + M_{1,i}^{\text{old}} \left(P^{\text{old}} + \left| \bar{\gamma}_i^{\text{old}} \right|^2 \right) \right]. \tag{3.29}$$

Next, we optimize over the jamming probability α_1. The optimization with respect to $\alpha_m, m = \{0, 1\}$ is a constrained optimization problem, since $\alpha_0 + \alpha_1 = 1$. Forming the Lagrangian $\Lambda = Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}\right) + \lambda\left(\alpha_0 + \alpha_1 - 1\right)$ and taking its derivative with respect to $\alpha_m, m = \{0, 1\}$ leads to

$$\hat{\alpha}_0 = \frac{1}{L} \sum_{i=1}^{L} M_{0,i}^{\text{old}} \quad \text{and} \quad \hat{\alpha}_1 = \frac{1}{L} \sum_{i=1}^{L} M_{0,i}^{\text{old}}. \tag{3.30}$$

Next, we optimize over the target amplitude variables $\gamma_{t,i}$ $i = \{1, 2, \ldots, L\}$. We present two different target signal models, namely non-coherent and coherent target models.

Target Amplitude and Angle Estimation :

Noncoherent Target Signal Model :

This model assumes that $\gamma_{t,i}$ for $i = \{1, 2, \ldots, L\}$ are L independent non-random variables. This model is valid for Swerling-2/Swerling-4 target fluctuation models in radar signal processing literature [9].

By taking derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$ in (3.28) with respect to $\gamma^*_{t,i}$,

$$\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})}{\partial \gamma^*_{t,i}} = -N\gamma_{t,i} + \boldsymbol{a}^H(\theta_t)\boldsymbol{x}_i - M^{\text{old}}_{1,i}\boldsymbol{a}^H(\theta_t)\bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j) \tag{3.31}$$

and equating the result to zero, the updated estimate $\hat{\gamma}_{t,i}$ can be given as

$$\hat{\gamma}_{t,i} = \frac{1}{N}\boldsymbol{a}^H(\theta_t)\left(\boldsymbol{x}_i - M^{\text{old}}_{1,i}\bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j)\right), \quad i = \{1, 2, \ldots, L\}. \tag{3.32}$$

The only remaining unknown to determine is the target angular position θ_t. When the updated target amplitudes are inserted in (3.28), we get the compressed objective function $Q_c(\cdot, \cdot)$ as

$$Q_c(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \overset{c}{=} \frac{1}{N}\sum_{i=1}^{L}\boldsymbol{a}^H(\theta_t)\left(\boldsymbol{x}_i - M^{\text{old}}_{1,i}\bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j)\right)\left(\boldsymbol{x}_i - M^{\text{old}}_{1,i}\bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j)\right)^H\boldsymbol{a}(\theta_t). \tag{3.33}$$

The angular position θ_t maximizing the compressed objective function is the updated DOA estimate for the target signal, which is

$$\hat{\theta}_t = \arg\max_{\theta}\sum_{i=1}^{L}\left|\boldsymbol{a}^H(\theta)\left(\boldsymbol{x}_i - M^{\text{old}}_{1,i}\bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j)\right)\right|^2. \quad \text{(Non-coherent Target Model)} \tag{3.34}$$

Coherent Target Signal Model :

This model assumes that the target amplitudes $\gamma_{t,i}$ for $i = \{1, 2, \ldots, L\}$ are identically the same, after a deterministic phase correction, as mentioned before. This model is valid for Swerling-1/Swerling-3 target fluctuation models in radar signal processing literature [9].

To simplify the notation, we denote the common target amplitude with γ_t, and substitute $\gamma_t = \gamma_{t,1} = \gamma_{t,2} = \ldots = \gamma_{t,L}$ in (3.28). After the substitution, we get,

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \overset{c}{=} -\sum_{i=1}^{L}M^{\text{old}}_{0,i}\left\|\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t)\right\|^2 - \sum_{i=1}^{L}M^{\text{old}}_{1,i}\left\|\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t) - \bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j)\right\|^2. \tag{3.35}$$

By taking the derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$ with respect to γ^*_t

$$\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})}{\partial \gamma^*_t} = -NL\gamma_t + \boldsymbol{a}^H(\theta_t)\sum_{i=1}^{L}\left(\boldsymbol{x}_i - M^{\text{old}}_{1,i}\bar{\gamma}^{\text{old}}_i\boldsymbol{a}(\theta_j)\right) \tag{3.36}$$

and equating the result to zero, the updated target amplitude $\hat{\gamma}_t$ can be given as

$$\hat{\gamma}_t = \frac{1}{NL}\boldsymbol{a}^H(\theta_t)\sum_{i=1}^{L}\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right). \tag{3.37}$$

Inserting the optimized $\hat{\gamma}_t$ value in (3.35), we get the compressed objective function in the coherent case as

$$Q_c(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) \overset{c}{=} \hat{\gamma}_t \sum_{i=1}^{L} M_{0,i}^{\text{old}}\boldsymbol{x}_i^H\boldsymbol{a}(\theta_t) + \hat{\gamma}_t \sum_{i=1}^{L} M_{1,i}^{\text{old}}\left(\boldsymbol{x}_i - \bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)^H\boldsymbol{a}(\theta_t). \tag{3.38}$$

Maximizing the compressed objective function over the unknown target angle yields the updated DOA estimate for the coherent target signal model as:

$$\hat{\theta}_t = \arg\max_{\theta}\left|\boldsymbol{a}^H(\theta)\sum_{i=1}^{L}\left(\boldsymbol{x}_i - M_{1,i}^{\text{old}}\bar{\gamma}_i^{\text{old}}\boldsymbol{a}(\theta_j)\right)\right|^2. \quad \text{(Coherent Target Model)} \tag{3.39}$$

This completes the derivation of M-step of EM algorithm. The steps of the proposed estimator for both target models are given in Table 3.1 - Algorithm 1 listing.

3.3 Performance Bounds For Proposed Method

This section is devoted to CRB derivations of the estimators that are used to estimate unknown paremeters of the signal model in Eq.(3.1). The derivation of Modified CRB (MCRB) is also presented.

3.3.1 Cramer-Rao Lower Bound

Reconsider the following signal model which is originally defined in Eq. (3.1).

$$\boldsymbol{x}_i = \gamma_{t,i}\boldsymbol{a}(\theta_t) + I_i\gamma_{j,i}\boldsymbol{a}(\theta_j) + \boldsymbol{n}_i \quad i = \{1, \ldots, L\}.$$

It was mentioned in Section 3.2 that since the argument of the logarithm function in (3.2) involves summation of two terms, it is not possible to take derivative of it directly to find the CRB. To overcome this challenge and produce a meaningful performance comparison, it may be useful to assume that the jammer component is either present in all snapshots or not at all. Let us name the first assumption as jammer present assumption and the latter one as the jammer absent assumption. Under these assumptions, calculation of CRB becomes straightforward.

Table 3.1: Proposed Method for DOA Estimation Under Signal Intermittency

Algorithm 1 Intermittent Jamming with Unkown Jammer Power

Initialization: Initialize $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$

Expectation Phase:

compute $M_{0,i}$ from (3.8)

compute $M_{1,i}$ from (3.14)

compute P^{old} and $\bar{\gamma}_i^{\text{old}}$ from (3.22)

Maximization Phase:

compute $\hat{\sigma}^2$ from (3.29)

compute $\hat{\alpha}_0$ and $\hat{\alpha}_1$ from (3.30)

if Signal model = non-coherent **then**

compute $\hat{\gamma}_{t,i}$, $i = \{1, 2, \ldots, L\}$ from (3.32)

compute $\hat{\theta}_t$ from (3.34)

end if

if Signal model = coherent **then**

compute $\hat{\gamma}_t$ from (3.37)

compute $\hat{\theta}_t$ from (3.39)

end if

Termination Check:

if the convergence criteria is satisfied **then**

return $\hat{\sigma}^2$, $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\gamma}_{t,i}(\hat{\gamma}_t)$ and $\hat{\theta}_t$

else

Go to Expectation Phase

end if

It is clear that the probability density function of the obervations is Gaussian under both of these assumptions, i.e., $\boldsymbol{x}_i \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{C}_{I_i})$. where $\boldsymbol{\mu}$ and $\boldsymbol{C}_{I_i}$ denote the mean vector and covariance matrix of the distribution respectively. $\boldsymbol{C}_{I_i=0}$ is the covariance matrix of the jammer absent case and $\boldsymbol{C}_{I_i=1}$ is the covariance matrix of the jammer present case. The set of unknown parameters to be estimated of the signal model in Eq. (3.1) includes target angle, target signal amplitudes, probabil-

ity of jammer signal component being present or being absent and jammer power, i.e, $\boldsymbol{\theta} = \{\theta_t, \gamma_{t,1}, \cdots, \gamma_{t,L}, \alpha_0, \alpha_1, \sigma^2\}$ for noncoherent target signal model and $\boldsymbol{\theta} = \{\gamma_t, \theta_t, \sigma^2, \alpha_0, \alpha_1\}$ for coherent target signal model. The mean vector of the density of the observation $\boldsymbol{x_i}$ does not depend on indicator variable, I_i. So, it is $\boldsymbol{\mu} = \gamma_{t,i}\boldsymbol{a}(\theta_t)$ for both of jammer present and jammer absent cases. The structure of the covariance matrix depends on the indicator variable and it is $\boldsymbol{C}_{I_i=1} = \boldsymbol{I} + \sigma^2\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$ when $I_i = 1$ and it is equal to the identity matrix in the other case, i.e., $I_i = 0$. Since the mean vector and the covariance matrix do not depend explicitly on indicator variable or its probability values under jammer present and jammer absent assumptions, it is meaningless to calculate the CRB for α_0 or α_1, although they are members of the set θ. Similarly, it is reasonable to assume that jammer power is known for jammer present case, since there is no such unknown in the jammer absent case. Therefore, let us denote CRB parameters of noncoherent target signal model as $\boldsymbol{\theta}_{Noncoh} = \begin{bmatrix} \theta_t & \mathrm{Re}\{\gamma_{t,1}\} & \cdots & \mathrm{Re}\{\gamma_{t,L}\} & \mathrm{Im}\{\gamma_{t,1}\} & \cdots & \mathrm{Im}\{\gamma_{t,L}\} \end{bmatrix}$ and CRB parameters of coherent target signal model as $\boldsymbol{\theta}_{Coh} = \begin{bmatrix} \theta_t & \mathrm{Re}\{\gamma_t\} & \mathrm{Im}\{\gamma_t\} \end{bmatrix}$. There is also one thing to be mentioned that the first term in eq. (2.53) can be ignored. Because partial derivative of $\boldsymbol{C}_{I_i}$ with respect to any of these parameters is zero, i.e.,

$$\frac{\partial \boldsymbol{C}_{\boldsymbol{I_i=0}}}{\partial \theta_k} = \frac{\partial \boldsymbol{I}}{\partial \theta_k} = \boldsymbol{0}, \quad \forall \theta_k \tag{3.40}$$

$$\frac{\partial \boldsymbol{C}_{\boldsymbol{I_i=1}}}{\partial \theta_k} = \frac{\partial \left(\boldsymbol{I} + \sigma^2\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}{\partial \theta_k} = \boldsymbol{0}, \quad \forall \theta_k \tag{3.41}$$

where $\boldsymbol{0}$ is the $(2L+1)\times(2L+1)$ matrix whose elements are all zero and $\boldsymbol{\theta}$ is either $\boldsymbol{\theta}_{Noncoh}$ or $\boldsymbol{\theta}_{Coh}$ depending on the problem. To find the second term in Eq. (2.53), we calculate the partial derivatives of $\boldsymbol{\mu} = \gamma_{t,i}\boldsymbol{a}(\theta_t)$ as

$$\frac{\partial \boldsymbol{\mu}}{\partial \,\mathrm{Re}\{\gamma_{t,i}\}} = \boldsymbol{a}(\theta_t) \qquad \frac{\partial \boldsymbol{\mu}^H}{\partial \,\mathrm{Re}\{\gamma_{t,i}\}} = \boldsymbol{a}^H(\theta_t)$$

$$\frac{\partial \boldsymbol{\mu}}{\partial \,\mathrm{Im}\{\gamma_{t,i}\}} = j\boldsymbol{a}(\theta_t) \qquad \frac{\partial \boldsymbol{\mu}^H}{\partial \,\mathrm{Im}\{\gamma_{t,i}\}} = -j\boldsymbol{a}^H(\theta_t)$$

$$\frac{\partial \boldsymbol{\mu}}{\partial \theta_t} = -j\gamma_{t,i}\frac{2\pi}{\lambda}\cos(\theta)\boldsymbol{D}\boldsymbol{a}(\theta_t) \qquad \frac{\partial \boldsymbol{\mu}^H}{\partial \theta_t} = j\gamma_{t,i}^*\frac{2\pi}{\lambda}\cos(\theta)\boldsymbol{a}^H(\theta_t)\boldsymbol{D}$$

for $i = \{1, \ldots, L\}$. In the last equality of the above equation set $\boldsymbol{D}$ is the diagonal matrix with the diagonal entries $[0 \;\; d \;\cdots\; (N-1)d]$.

Lastly, we remind that the Fisher information matrix (FIM) for L independent snap-

shot is the sum of individual FIM for each snaphot.

Non-Coherent Target Signal Model :

This model includes $2L+1$ parameters to be estimated:

$$\boldsymbol{\theta}_{Noncoh} = \begin{bmatrix} \theta_t & \mathrm{Re}\{\gamma_{t,1}\} & \cdots & \mathrm{Re}\{\gamma_{t,L}\} & \mathrm{Im}\{\gamma_{t,1}\} & \cdots & \mathrm{Im}\{\gamma_{t,L}\} \end{bmatrix}^T.$$

For this model, the non-zero entries of FIM matrix can be given as:

The first element of FIM

$$[\boldsymbol{I}_{JP}(\boldsymbol{\theta})]_{1,1} = \sum_{i=1}^{L} 2\,\mathrm{Re}\left[|\gamma_{t,i}|^2 (2\pi/\lambda)^2 \cos^2(\theta)\, \boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],$$

$$[\boldsymbol{I}_{JA}(\boldsymbol{\theta})]_{1,1} = \sum_{i=1}^{L} 2\,\mathrm{Re}\left[|\gamma_{t,i}|^2 (2\pi/\lambda)^2 \cos^2(\theta)\, \boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],$$

For $\quad l = \{2,\ldots,L+1\}, \quad i = \{1\ldots,L\}$

$$[\boldsymbol{I}_{JP}(\boldsymbol{\theta})]_{1,l} = \left[\boldsymbol{I}_{I_i}^*(\boldsymbol{\theta})\right]_{l,1} = 2\,\mathrm{Re}\left[-j\gamma_{t,i}^* (2\pi/\lambda)\cos(\theta)\, \boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{a}(\theta_t)\right],$$

$$[\boldsymbol{I}_{JA}(\boldsymbol{\theta})]_{1,l} = \left[\boldsymbol{I}_{I_i}^*(\boldsymbol{\theta})\right]_{l,1} = 2\,\mathrm{Re}\left[-j\gamma_{t,i}^* (2\pi/\lambda)\cos(\theta)\, \boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],$$

For $\quad l = \{L+2,\ldots,2L+1\}, \quad i = \{1\ldots,L\}$

$$[\boldsymbol{I}_{JP}(\boldsymbol{\theta})]_{1,l} = \left[\boldsymbol{I}_{I_i}^*(\boldsymbol{\theta})\right]_{l,1} = 2\,\mathrm{Re}\left[\gamma_{t,i}^* (2\pi/\lambda)\cos(\theta)\, \boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{a}(\theta_t)\right],$$

$$[\boldsymbol{I}_{JA}(\boldsymbol{\theta})]_{1,l} = \left[\boldsymbol{I}_{I_i}^*(\boldsymbol{\theta})\right]_{l,1} = 2\,\mathrm{Re}\left[\gamma_{t,i}^* (2\pi/\lambda)\cos(\theta)\, \boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],$$

For $\quad l = \{2,\ldots,2L+1\}, \quad i = \{1\ldots,L\}$

$$[\boldsymbol{I}_{JP}(\boldsymbol{\theta})]_{l,l} = 2\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i}^{-1}\boldsymbol{a}(\theta_t)\right],$$

$$[\boldsymbol{I}_{JA}(\boldsymbol{\theta})]_{l,l} = 2\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{a}(\theta_t)\right] = 2N.$$

In the above equations, $\boldsymbol{I}_{JP}(\boldsymbol{\theta})$ represents the FIM when jammer component exists in all of the received snapshots, whereas $\boldsymbol{I}_{JA}(\boldsymbol{\theta})$ denotes the FIM when none of the received snapshots contain jammer component. Note also that $\boldsymbol{C}_{I_i=1}^{-1} = \boldsymbol{I} - \frac{\sigma^2}{1+N\sigma^2}\boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$.

Coherent Target Signal Model :

This model includes 3 parameters to be estimated:

$$\boldsymbol{\theta}_{Coh} = \begin{bmatrix} \theta_t & \mathrm{Re}\{\gamma_t\} & \mathrm{Im}\{\gamma_t\} \end{bmatrix}^T.$$

The elements of FIM can be written as:

$$
\begin{aligned}
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{1,1} &= 2L\,\mathrm{Re}\left[\left|\gamma_t\right|^2\left(2\pi/\lambda\right)^2\cos^2\left(\theta\right)\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{1,2} &= 2L\,\mathrm{Re}\left[-j\gamma_t^*\left(2\pi/\lambda\right)\cos\left(\theta\right)\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{1,3} &= 2L\,\mathrm{Re}\left[\gamma_t^*\left(2\pi/\lambda\right)\cos\left(\theta\right)\boldsymbol{a}^H(\theta_t)\boldsymbol{D}\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{2,1} &= 2L\,\mathrm{Re}\left[j\gamma_t\left(2\pi/\lambda\right)\cos\left(\theta\right)\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{2,2} &= 2L\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{2,3} &= 0,\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{3,1} &= 2L\,\mathrm{Re}\left[\gamma_t\left(2\pi/\lambda\right)\cos\left(\theta\right)\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right],\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{3,2} &= 0,\\
\left[\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)\right]_{3,3} &= 2L\,\mathrm{Re}\left[\boldsymbol{a}^H(\theta_t)\boldsymbol{C}_{I_i=1}^{-1}\boldsymbol{D}\boldsymbol{a}(\theta_t)\right].
\end{aligned}
$$

Similar to the noncoherent target model case, $\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)$ in the above equations represents the FIM when jammer component exists in all of the received snapshots. The FIM when none of the received snapshots contain jammer component, i.e., $\boldsymbol{I}_{JA}\left(\boldsymbol{\theta}\right)$ can be obtained by replacing $\boldsymbol{C}_{I_i=1}^{-1}$ with the identity matrix.

For both target models, CRLB on the DOA estimation error is $(1,1)$th element of the inverse of $\boldsymbol{I}_{JP}\left(\boldsymbol{\theta}\right)$ or $\boldsymbol{I}_{JA}\left(\boldsymbol{\theta}\right)$. Using the above results, FIM and its inverse can be calculated numerically at the true values of parameters and signal model unkowns. It is clear that variance of frequency estimator is greater than $\left[\boldsymbol{I}_{JP}^{-1}\left(\boldsymbol{\theta}\right)\right]_{1,1}$ or $\left[\boldsymbol{I}_{JA}^{-1}\left(\boldsymbol{\theta}\right)\right]_{1,1}$ depending on the case.

3.3.2 Modified Cramer Rao Bound

Since $\{I_i\}_{i=1}^{L}$ is a random sequence which is not known a priori, it can be considered as the nuisance random parameter set in the definition of the Modified Cramer Rao Bound (MCRB). We denote the set of unkowns to be estimated as $\boldsymbol{\theta}$ and $\boldsymbol{\beta} = \{I_1, I_2, \cdots, I_L\}$ respectively. As defined in [2] and [91], the (i,j)th entry of the modified FIM $\boldsymbol{I}_M\left(\boldsymbol{\theta}\right)$ is given by

$$
\begin{aligned}
\left[\boldsymbol{I}_M\left(\boldsymbol{\theta}\right)\right]_{i,j} &= E_{\boldsymbol{X},\boldsymbol{\beta}}\left\{\frac{\partial \ln p(\boldsymbol{X}\mid\boldsymbol{\beta};\boldsymbol{\theta})}{\partial\theta_i}\cdot\frac{\partial \ln p(\boldsymbol{X}\mid\boldsymbol{\beta};\boldsymbol{\theta})}{\partial\theta_j}\right\}\\
&= -E_{\boldsymbol{X},\boldsymbol{\beta}}\left\{\frac{\partial^2 \ln p(\boldsymbol{X}\mid\boldsymbol{\beta};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right\}
\end{aligned}
\tag{3.42}
$$

where the expectation is taken with respect to joint probability density function of $\boldsymbol{X}$ and the nuisance parameters $\boldsymbol{Y}$. Since $(\boldsymbol{x}_i, I_i)$ pair is independent of $(\boldsymbol{x}_j, I_j)$, modified FIM can be written as the sum of individual FIM of the observations, i.e.,

$$\begin{aligned}\left[\boldsymbol{I}_M(\boldsymbol{\theta})\right]_{i,j} &= \sum_{i=1}^{L} E_{\boldsymbol{x}_i, I_i}\left\{-\frac{\partial^2 \ln p(\boldsymbol{x}_i \mid I_i; \boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right\} \\ &= \sum_{i=1}^{L} E_{I_i}\left\{E_{\boldsymbol{x}_i|I_i}\left\{-\frac{\partial^2 \ln p(\boldsymbol{x}_i \mid I_i; \boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right\}\right\}.\end{aligned} \tag{3.43}$$

If the above expectation operations are calculated, the following modified FIM is reached

$$\boldsymbol{I}_{MCRB}(\boldsymbol{\theta}) = \alpha_0 \boldsymbol{I}_{JA}(\boldsymbol{\theta}) + \alpha_1 \boldsymbol{I}_{JP}(\boldsymbol{\theta}) \tag{3.44}$$

where $\boldsymbol{I}_{JA}(\boldsymbol{\theta})$ and $\boldsymbol{I}_{JP}(\boldsymbol{\theta})$ are previously given FIM for the cases of jammer absent and present cases, respectively.

3.4 Simulation Results

We present three sets of Monte Carlo simulation results for performance comparisons. In all comparisons, a uniform linear array of $N = 10$ elements with an interelement spacing of $\lambda/2$ is utilized. The ULA is illustrated in Figure 2.1. Noise variance is taken as unity. The target signal complex amplitude is denoted with $\gamma_{t,i}$ where $\gamma_{t,i} = |\gamma_{t,i}|\, \mathrm{e}^{\phi_{t,i}}$. For coherent target model, $|\gamma_{t,i}|$ and $\phi_{t,i}$ are identically the same for $i = \{1, \ldots, L\}$. For noncoherent target model, $|\gamma_{t,i}|$ is held constant for all snapshots whereas phase is independently sampled from uniform distribution at each snapshot. For both target models, SNR is defined as SNR $= |\gamma_{t,i}|^2$. The jammer existence probability (α_1) is fixed to 0.5 in the first and the third numerical experiments and it is varied in the second experiment. The jammer-to-noise-ratio (JNR) is identical to the jammer power σ^2, since the noise variance is taken as unity.

Under the intermittent interference model, the interference plus noise covariance matrix becomes $\boldsymbol{R_e} = \boldsymbol{I} + \alpha_1 \sigma^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)$ as defined in Eq.(2.30). It is assumed that the conventional method utilizes the exact $\boldsymbol{R_e}$ matrix. Hence, the conventional method has the exact knowledge of jammer power and its angular position, but is not aware that jamming is intermittent. The suggested method assumes that jamming is intermittent, but is not aware of jamming probability and jamming power. The target

and jammer angular locations are taken as at 0° and 10°, respectively. It should be noted that the 3 dB beamwidth of the uniform linear array with 10 elements is approximately 11 degrees. Hence, the jammer lies at the edge of the beamwidth which makes it difficult to suppress without a significant SNR loss via conventional methods.

In all figures, we present the Cramer-Rao bound (CRB) for the cases of jammer always absent/present and Modified Cramer-Rao Bound (MCRB). The performance metric used for angle estimation is root mean square error,

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} \left(\hat{\theta}_i - \theta_t\right)^2} \tag{3.45}$$

where N_{runs} is the number of Monte Carlo runs, which is 100,000 for the first two experiments and 10,000 for the third one.

3.4.1 Experiment 1

Figure 3.1 shows the RMS error on the target angular position estimate as JNR increases. This experiment is conducted with $L = 20$ snapshots, jamming probability of $\alpha_1 = 0.5$ and SNR = 0 dB.

The sub-figures of Figure 3.1 present the results for coherent and non-coherent signal models. The red and green colored curves indicate the CRB when the jammer is always present ($\alpha_1 = 1$) and absent ($\alpha_1 = 0$), respectively. The MCRB for $\alpha_1 = 0.5$ is shown with the purple line.

It can be noted from Figure 3.1 that the proposed and conventional method perform almost identically for sufficiently small JNR values (weak jamming signal). As expected, this operational regime is dominated by the noise, not by the jammer activity; hence the performance of the proposed method, which models the intermittent nature of the jammer, is identical to the conventional one. As JNR increases, the performance of the conventional estimator, i.e. the maximum likelihood estimator with the assumption of $\alpha_1 = 1$, degrades rapidly after a threshold value. The threshold is around JNR $= -5$ dB in Figure 3.1. The proposed method yields a better performance over the conventional estimator above this JNR threshold.

Atomic norm based technique shown in Figure 3.1 estimates both jammer and target locations via convex optimization by processing L snapshots [56]. Atomic norm method, like other high resolution methods such as MUSIC or ESPRIT, does not utilize the statistical information of the observations and have no dependence on number of targets in this formulation. We observe from Figure 3.1 that the performance of atomic norm based technique is far more worse than conventional and suggested estimator which utilize the specifics of the problem. The lack of sensitivity to the problem specifics of atomic norm estimator is also reflected by almost identical performance for both target models, while other estimators utilize the target signal model to further improve the performance.

Figure 3.2 shows the accuracy of other parameters estimated in Experiment 1. Sub-figures (a)-(b) of Figure 3.2 show the mean and standard deviation of the jammer existence probability estimate for the noncoherent case and coherent case respectively. As expected, as JNR increases, the jammer existence probability converges to the true value of 0.5. The mean and standard deviation for α_1 estimates are almost identical for both target models.

Sub-figures (a)-(b) of Figure 3.3 show the estimate for the jammer power for the non-coherent case and coherent case respectively. The true jammer power is indicated by the dashed red-line in Figures 3.3(a) and 3.3(b). The standard deviation of the jammer power estimate is also indicated by the error bars. Error bar graphs are obtained by plotting jammer power in linear scale and adding bars represening standart deviation in linear scale too. Then, y-axes of the graph is converted to logarithmic scale. [1]

3.4.2 Experiment 2

JNR value is fixed in this experiment, but the jammer existence probability α_1 is varied. The other experiment parameters are identical to the ones in Experiment 1. JNR is fixed to two values, 10 and 20 dB, to illustrate the performance under medium and heavy jamming. It can be observed from Figure 3.4 that conventional and atomic norm estimator suffer from a significant performance loss even at $\alpha_1 = 0.1$. This is

[1] There was an unintentional mistake in the usage of error bar command in Figure 3 (c) and (d) in [11]. Figure 3.3 in this thesis should replace those figures.

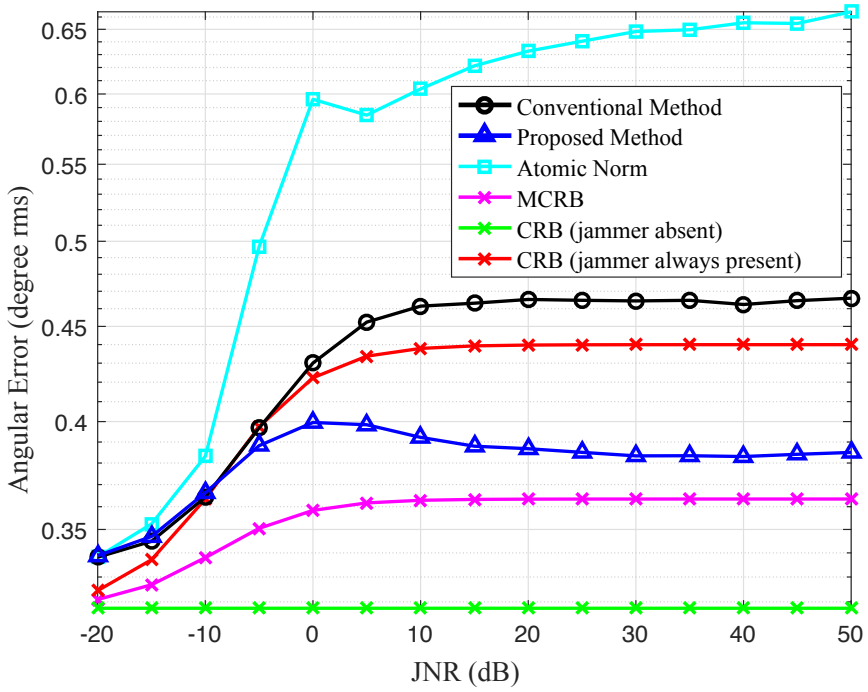
essentially due to the ignorance of signal intermittency in the problem formulation. The performance of the suggested method is identical to the conventional method for the extreme cases of jammer always absent ($\alpha_1 = 0$) and always present ($\alpha_1 = 1$) and tracks the MCRB lower bound for all α_1 values. It is interesting to note that the case of coherent signal model tracks the MCRB very closely in the heavy jamming case. We would like to underline that the suggested method is unaware of true α_1 value and estimates this parameter in the processing chain. In spite of this, the performance of the suggested method is in close vicinity of the performance bounds for all α_1 values. The estimation accuracy of other parameter estimates ($\hat{\sigma}^2$ and $\hat{\alpha}_1$) is similar to the results in Figure 3.2 and Figure 3.3, so they are not provided.

The success of the suggested method can be attributed to the successful "soft classification" of snapshots in terms of jammer contamination level. Another observed feature of the suggested method is the graceful degradation of estimation accuracy with increasing jamming probability. As can be most easily seen from Figure 3.5(b), the suggested method works as if jammer is absent for small jammer existence probability values, while the conventional estimator suffers from significant performance losses and works as if jammer is always present.

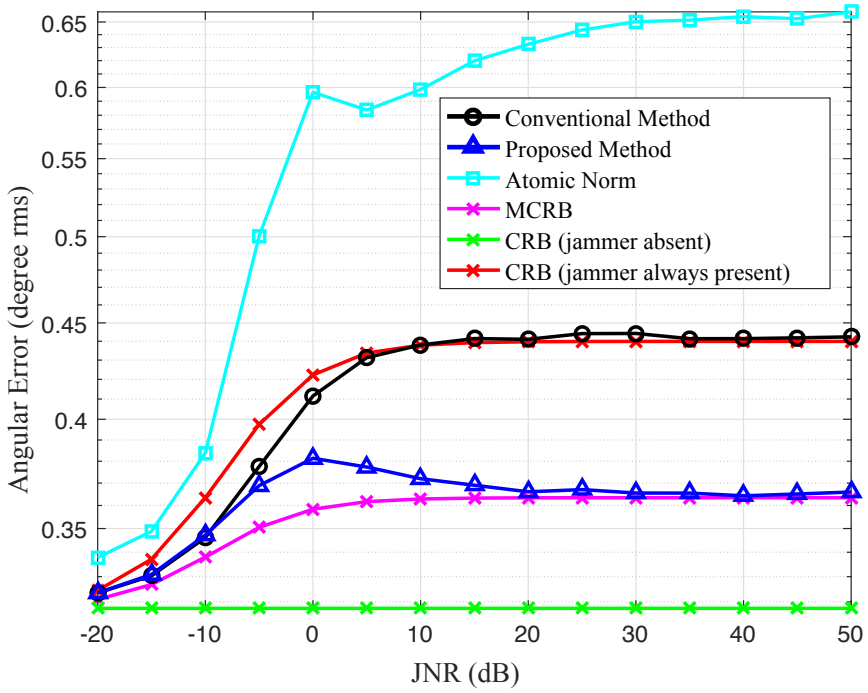
3.4.3 Experiment 3

This experiment shows the accuracy of the angle estimate for a fixed JNR as SNR changes. The experiment conditions are identical to the earlier experiments except JNR is fixed to 10 dB and jammer existence probability is set to $\alpha_1 = 0.5$. Results in Figure 3.6 are consistent with earlier experiments. In this experiment, atomic norm method cannot estimate the target angular location at low SNR and presents a single estimate (instead of two), which is the estimate for the jammer location whose true value is 10 degrees. Suggested method performs close to MCRB in the asymptotic region (high SNR region of Figure 3.6) and outperforms other alternatives.

As a final note, angular accuracy graphs presented in this section are drawn using *semilogy(.)* command of MATLAB. To zoom more in high SNR region, say for SNR $\geq$ 5dB, *loglog(.)* command is used and the resulting graphs are presented in Figure 3.7.

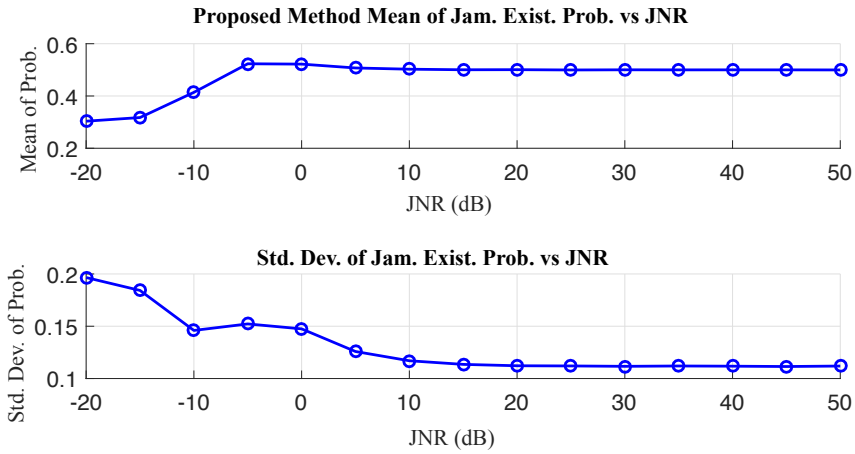


(a) Non-coherent case

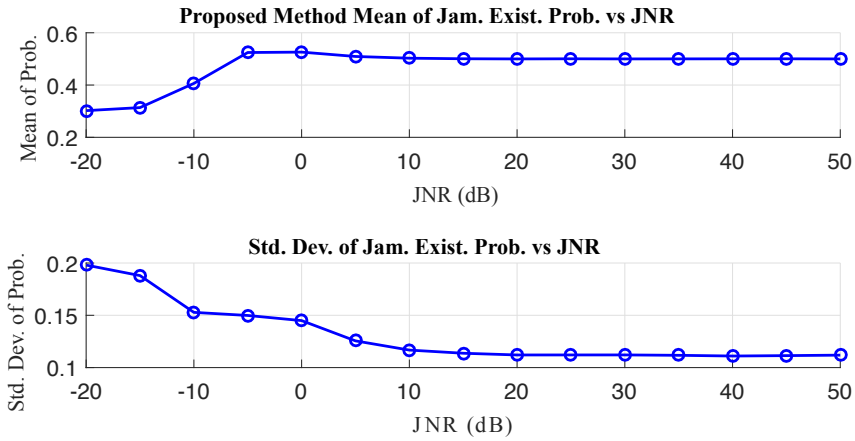


(b) Coherent case

Figure 3.1: Experiment 1 - Angle estimation accuracy vs JNR, Experiment 1,

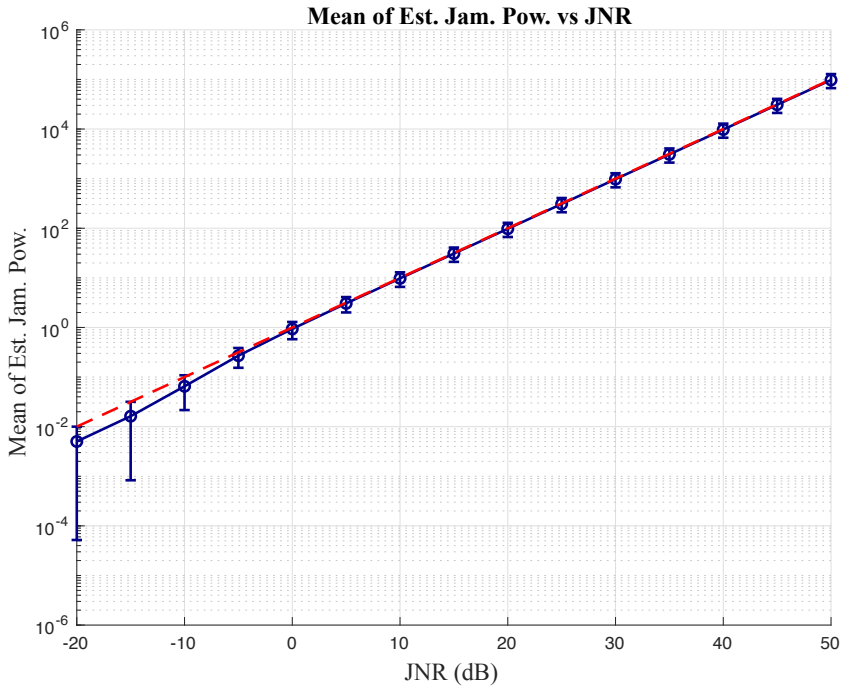


(a) Non-coherent case

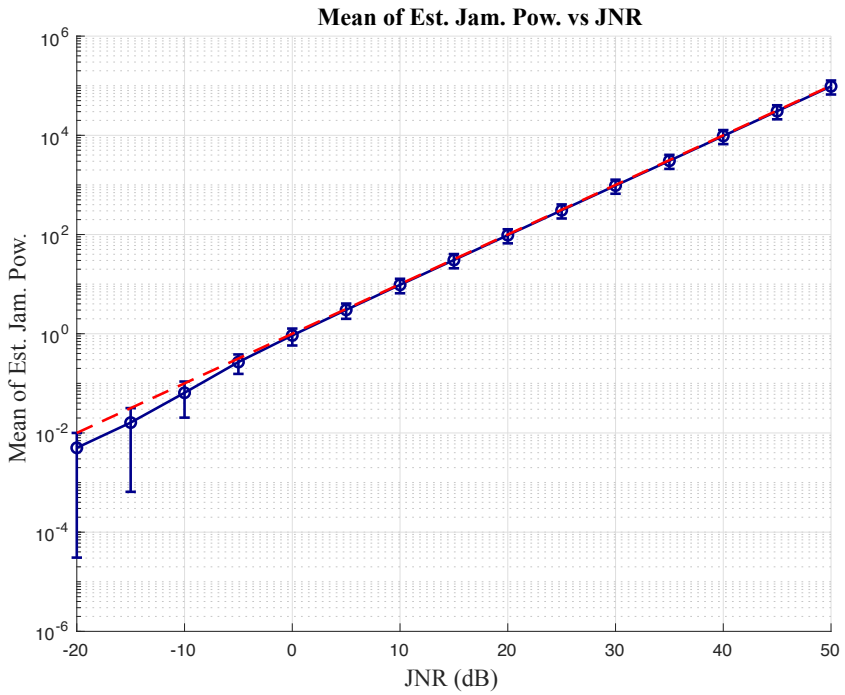


(b) Coherent case

Figure 3.2: Experiment 1 - Mean and standard deviation of jammer existence probability estimate

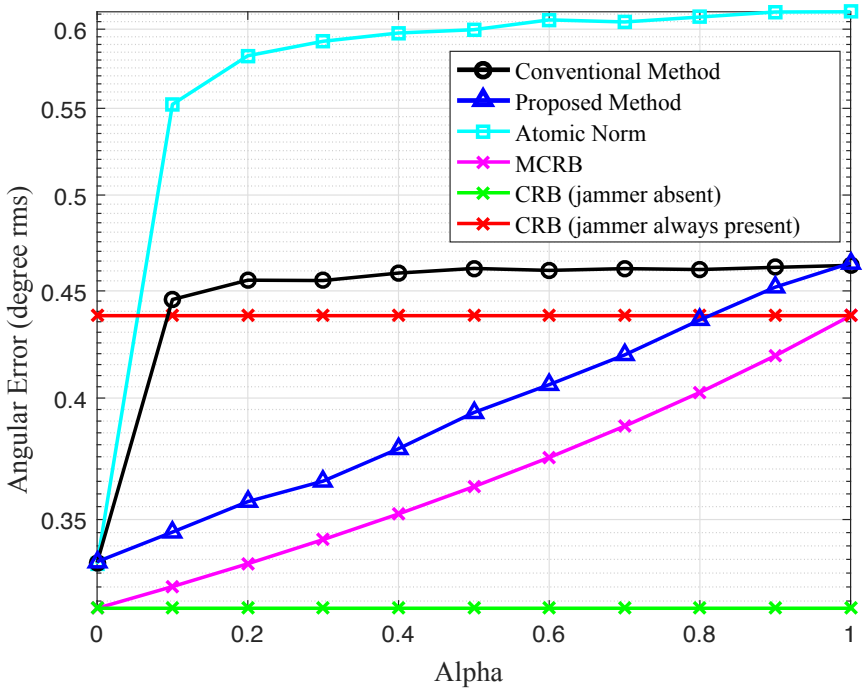


(a) Non-coherent case

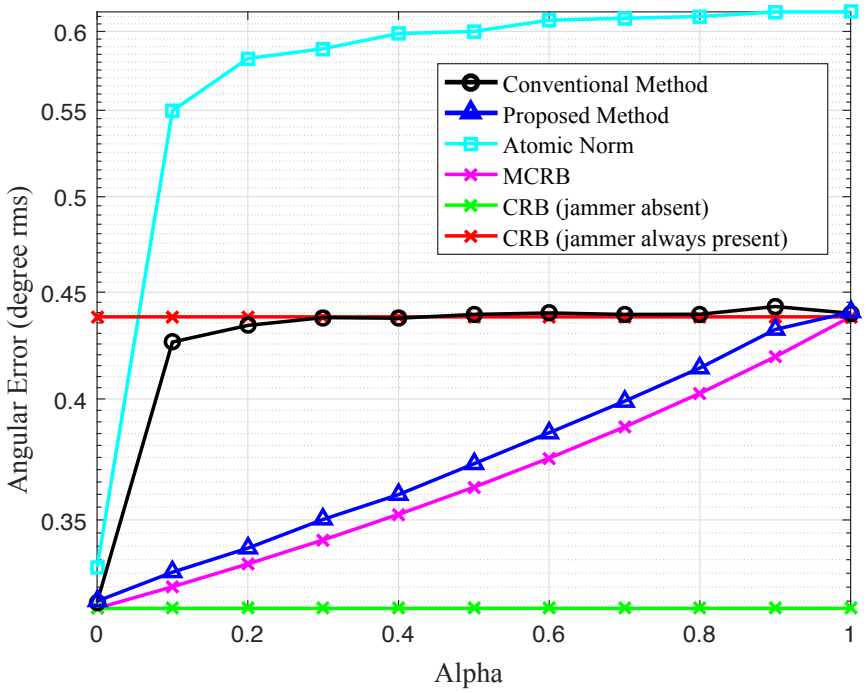


(b) Coherent case

Figure 3.3: Experiment 1 - Mean value of jammer power estimates vs JNR, Correct Representation

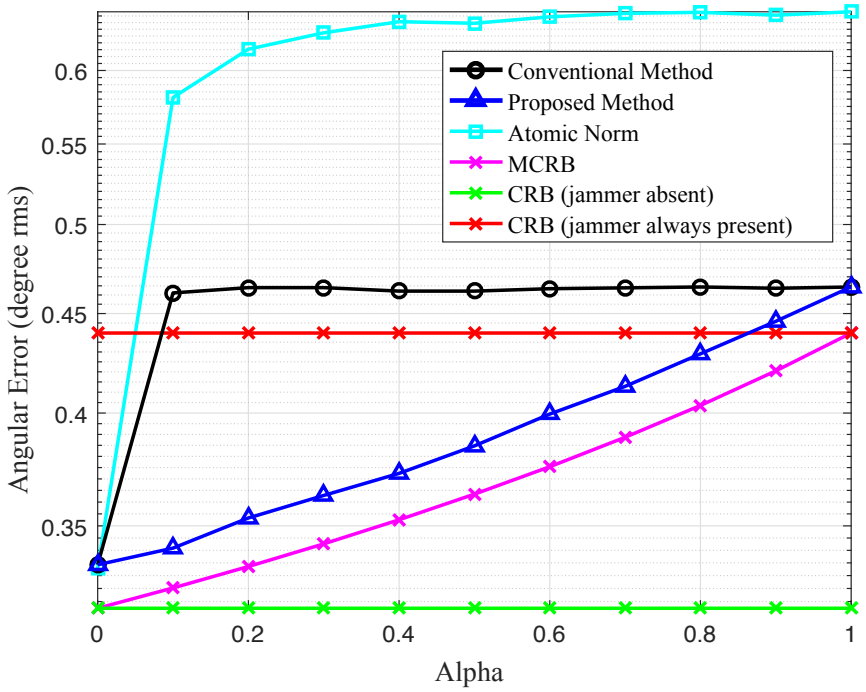


(a) Non-coherent case

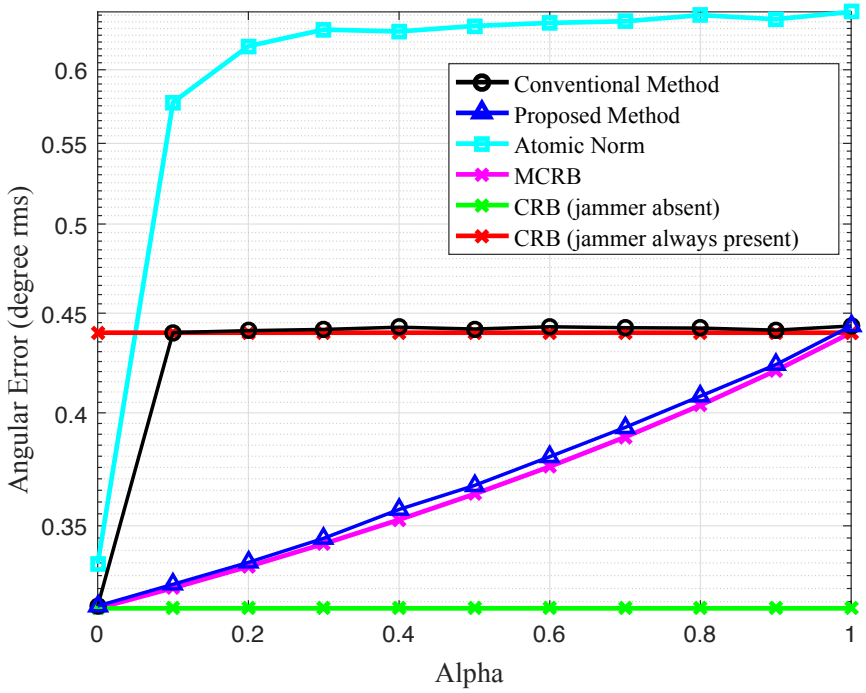


(b) Coherent case

Figure 3.4: Experiment 2 - Angle estimation accuracy vs probability, JNR = 10 dB

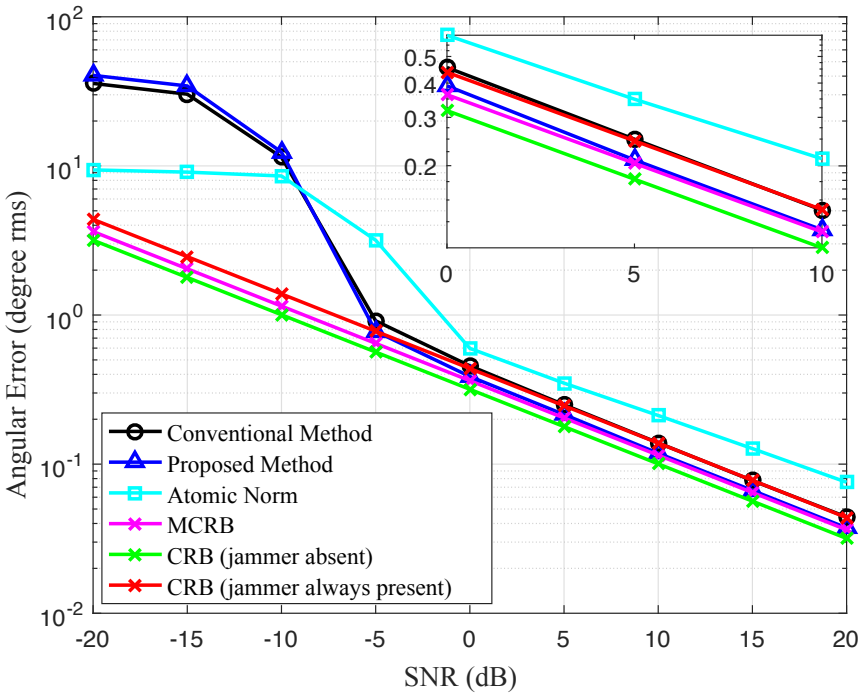


(a) Non-coherent case

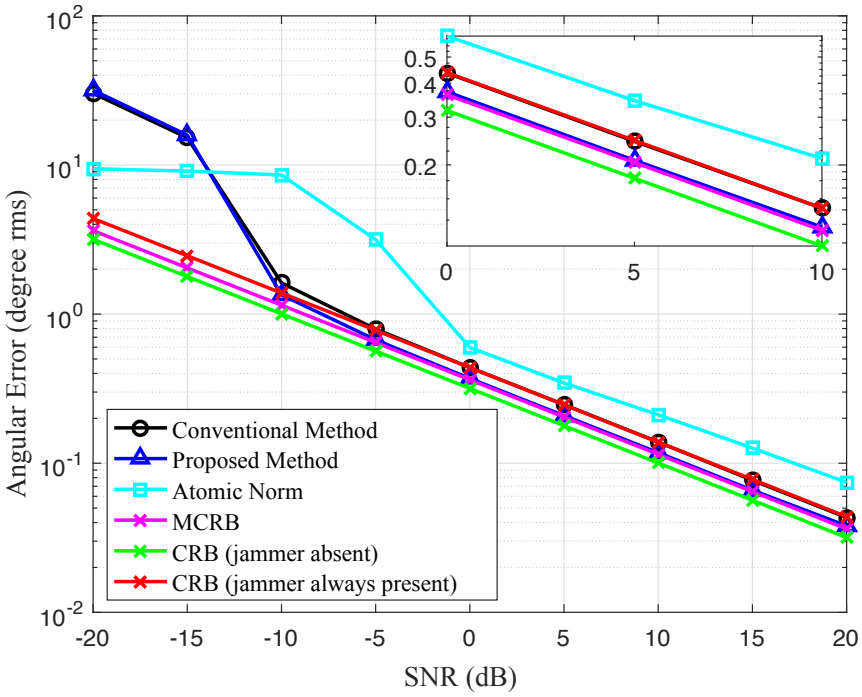


(b) Coherent case

Figure 3.5: Experiment 2 - Angle estimation accuracy vs probability, JNR = 20 dB

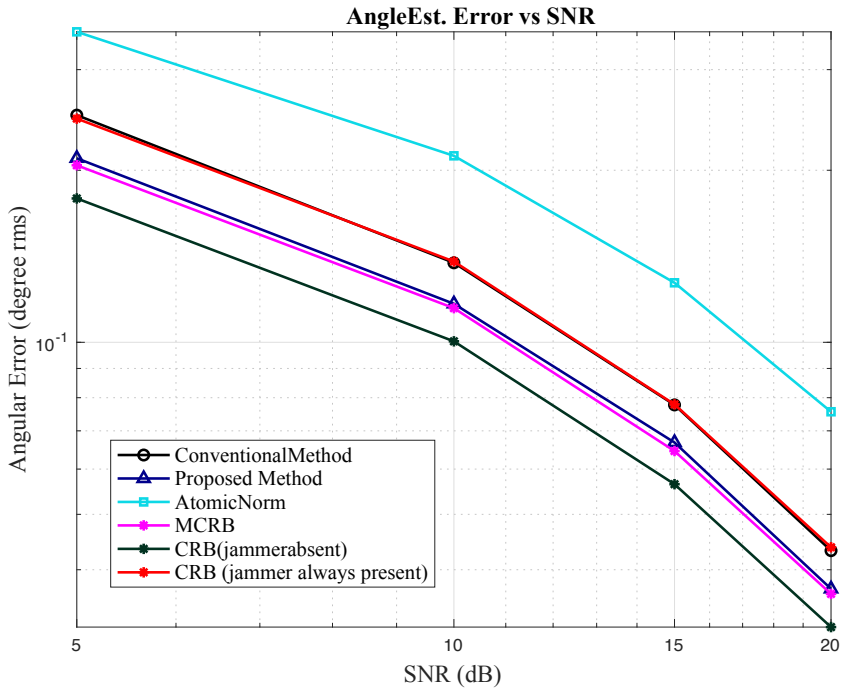


(a) Non-coherent case

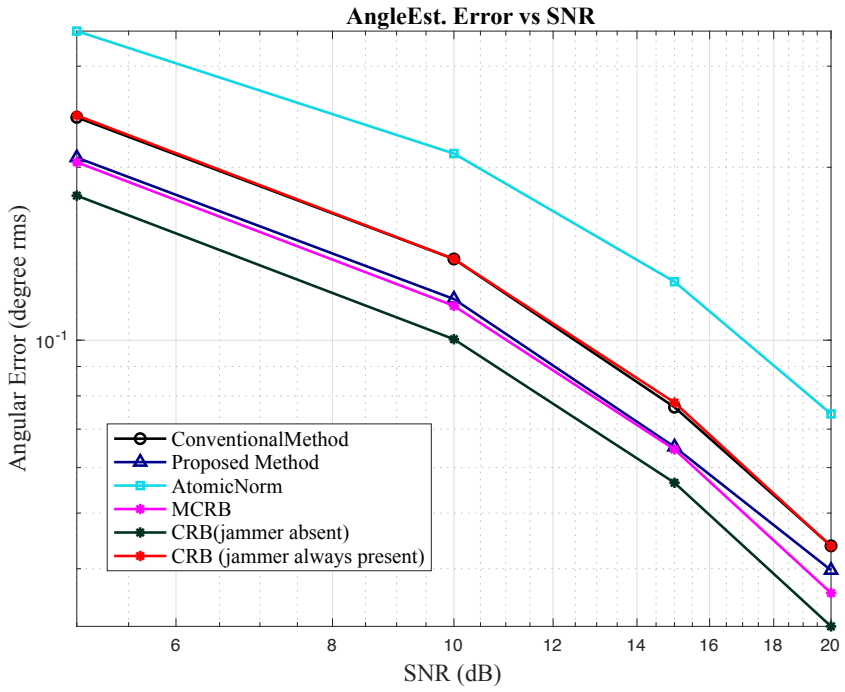


(b) Coherent case

Figure 3.6: Experiment 3 - Angle estimation accuracy vs SNR, JNR = 20 dB



(a) Non-coherent case



(b) Coherent case

Figure 3.7: Experiment 3 - Angle estimation accuracy vs SNR, Zoomed View, JNR = 20 dB

CHAPTER 4

FREQUENCY ESTIMATION OF A COMPLEX EXPONENTIAL IN CORRELATED MIXTURE OF GAUSSIANS NOISE

The frequency estimation of a complex exponential is a well-studied problem naturally emerging in many applications such as spectrum estimation [24], radar signal processing [92], and communications [93]. It is well known that acquisition and tracking of Doppler frequency is an important functionality in tracking radars which is at the heart of military applications [67]. As a civil application area, velocity estimation of cars in highways using Doppler frequency measured by radars has been a conventional approach for years. Besides this, parallel to the development of high-end smart cars, usage of automotive radars which are a part of adaptive cruise control and collision avoidance systems or which are utilized for lane change assistance and rear traffic crossing alert has been increased leading to increased research for the Doppler frequency estimation problem in recent years [94].

The ML estimator of the frequency for a complex exponential waveform under AWGN is the peak detector in the Fourier domain [1, p. 543]. When the unknown frequency is expressed in terms of N-point DFT bins, that is $f = (k_p + \delta)/N$, where $k_p \in \{0, 1, \ldots, N-1\}$ and $\delta \in (-0.5, 0.5)$, the computational load to track CRB can be significantly reduced by the sequential estimations of k_p and δ via two-stage estimation approach [3, 92, 95–99].

Approaches that provide high resolution like spectral-MUSIC, root-MUSIC and ESPRIT are also widely used in literature [24, Ch. 4]. These techniques have much higher computational load due to the need of eigen decomposition.

As far as we know, if not all, the majority of the existing literature concentrates on

improving accuracy of estimated frequencies of single or multiple signals in white noise assuming that noise statistics do not change from sample to sample. But, in real life, due to various failures in sensors, received signal samples at the output of a sensor may not be identical in terms of assumed noise and the noise may be correlated from sample to sample. There are practical scenarios which may involve a sudden change in the noise model during the data collection interval [100–103].

Here we focus on such piecewise stationary noise models representing, say, the cycles of hardware failure (signal interruption) followed by recovery or intentional/unintentional jamming activities on the sensor [10]. We name these operational states as healthy and faulty. Being in a healthy state corresponds to a high SNR received signal and being in a faulty state corresponds to a low SNR received signal. A two-component Gaussian mixture noise is used to model the states. Besides this, transitions between these states are assumed to obey a hidden Markov model.

We assume that for both healthy and faulty states, the noise is Gaussian distributed with unknown mean and variance values. In this problem setting, the state sequence, the parameters of the noise distribution, and the frequency of complex exponential signal are the unknowns of the problem. Our aim is to extend the computational advantages and high performance of two-stage estimators to the case of piecewise stationary noise. For this purpose, a method for frequency estimation is proposed under the defined conditions. The formulated solution combines well-known signal processing techniques such as forward-backward recursion of the hidden Markov model structure and two-stage iterative frequency estimation techniques designed for stationary white noise.

4.1 Maximum Likelihood Frequency Estimation

Let us consider the observations which is of the form

$$x_n = A\exp(j2\pi fn) + v_n, \quad n = \{0, \ldots, N-1\}, \tag{4.1}$$

where A is the complex amplitude which is defined as $A = A'\exp(\phi)$. The noise component, v_n, in Eq. (4.1) is assumed to be additive white Gaussian noise (AWGN) with density, $v_n \sim \mathcal{CN}(0, \sigma_v^2)$. SNR of the signal is defined as $|A|^2/\sigma_v^2$.

It is assumed that $A^{'}$, ϕ and f are unknowns.

It is well known that MLE of the frequency for a complex exponential waveform under AWGN is the peak detector of the periodogram of observations [1, p. 543]. In practice, N-point discrete Fourier transform (DFT) is calculated and the frequency at which the peak magnitude occurs is marked as the frequency of the signal. This provides resolution of $1/N$ in terms of the normalized frequency $f \in [0, 1)$. The resolution can be improved with the zero-padded DFT operation of length N_D at the expense of computational load. The CRLB for the estimation of parameters of the signal in Eq. (4.1) is given as [9]:

$$\sigma^2_{\hat{A}'} \geq \frac{\sigma_v^2}{2\,N}, \tag{4.2}$$

$$\sigma^2_{\hat{f}} \geq \frac{6}{(2\pi)^2\,N\,(N^2-1)\,\text{SNR}}, \tag{4.3}$$

$$\sigma^2_{\hat{\phi}} \geq \frac{2N-1}{N(N+1)\text{SNR}}. \tag{4.4}$$

4.2 Fine Resolution Frequency Estimation

It is clear that, if true frequency of the signal is not at the multiples of $1/N_D$ meaning that $\delta \neq 0$, true periodogram maximum and DFT maximum do not coincide resulting in unavoidable frequency estimation errors. To enhance estimation accuracy, the use of two-stage frequency estimation procedures has been adopted. The first stage, namely coarse frequency estimation, is common to all methods and is based on calculating DFT of the signal and choosing the peak location. The methods proposed in the literature differ with their approaches in the second stage, namely fine frequency estimation stage. According to the second stage, the methods can be broadly classsified as non-iterative and iterative method. Non-iterative methods use a simple function of DFT outputs whereas iterative methods employ an iterative cycle for fine frequency estimation. Some interpolation-based techniques that use two or more DFT outputs around peak position to interpolate the true frequency iteratively can be counted among iterative methods. Algorithm in [3] is a this kind of technique, which produces close to the state-of-art estimation performance, is summerized in Table 4.1.

Recently, two-stage estimators have been extended to the cases in which arbitrary

windowing function and zero-padding are applied to the received signal. For example, in [4] a new method, whose performance closely tracks the CRB without any bias at high SNR values, is proposed. This algorithm is summarized in Table 4.2. The algorithm in Table 4.2 together with the algorithm in Table 4.1 are used to compare the performance of the suggested estimator in Section 4.5.

Table 4.1: Iterative Frequency Estimation by Interpolation on Fourier Coefficients Algorithm [3, Table 1]

Algorithm 2 Aboutanios and Mulgrew (A&M) Method

Input : $\mathbf{x} = [x_0, x_1, \ldots, x_{N-1}]^T$

Initialization :

Calculate $X_k = \sum_{n=0}^{N-1} x_n \, e^{-j2\pi nk/N_{\mathrm{D}}}, \quad n \in \{0, \cdots, N_D - 1\}$

Find $\hat{k}_p = \arg\max_k \{X_k\}$

Set $\hat{\delta}_0 = 0$

Iterations :

for $i = 1$ **to** Q **do**

$$X_m = \sum_{n=0}^{N-1} x_n \, e^{-j2\pi n \frac{\hat{k}_p + \hat{\delta}_{i-1} + m}{N_{\mathrm{D}}}}, \quad m = \pm 0.5$$

$$\hat{\delta}_i = \hat{\delta}_{i-1} + h(\hat{\delta}_{i-1})$$

where

$$h(\hat{\delta}_{i-1}) = \frac{1}{2} \mathrm{Re}\left\{ \frac{X_{0.5} + X_{-0.5}}{X_{0.5} - X_{-0.5}} \right\} \quad , \text{ for A\&M 1}$$

or

$$h(\hat{\delta}_{i-1}) = \frac{1}{2} \frac{|X_{0.5}| + |X_{-0.5}|}{|X_{0.5}| - |X_{-0.5}|} \quad , \text{ for A\&M 2}$$

end for

Output :

$$\hat{f} = \frac{\hat{k}_p + \hat{\delta}_Q}{N_D}$$

Table 4.2: Interpolation on Fourier Coefficients for Windowed Data [4, Table 1]

Algorithm 3 Windowed Signal Frequency Estimation Method

Steps :

1. Define the function $g_{\mathbf{w}}(x)$ and the first derivative of it $g'_{\mathbf{w}}(x)$ as

$$g_{\mathbf{w}}(x) = \sum_{n=0}^{N-1} w_n \, e^{j2\pi nx/N_{\mathrm{D}}}, \qquad g'_{\mathbf{w}}(x) = \frac{j2\pi}{N_{\mathrm{D}}} \sum_{n=0}^{N-1} n \, w_n \, e^{j2\pi nx/N_{\mathrm{D}}}$$

2. Calculate the bias correction factor $c_{\mathbf{w}} = \dfrac{B_0^2}{A_1 B_0 + A_0 B_1}$, where

$$A_0 = g_{\mathbf{w}}(1) - g_{\mathbf{w}}(-1) \qquad A_1 = g'_{\mathbf{w}}(1) - g'_{\mathbf{w}}(-1)$$
$$B_0 = 2g_{\mathbf{w}}(0) - g_{\mathbf{w}}(1) - g_{\mathbf{w}}(-1) \quad B_1 = 2g'_{\mathbf{w}}(0) - g'_{\mathbf{w}}(1) - g'_{\mathbf{w}}(-1)$$

3. Calculate $N_D \geq N$ point DFT of the windowed input data

$$R_k = \sum_{n=0}^{N-1} w_n x_n \, e^{-j2\pi nk/N_{\mathrm{D}}}, \quad n \in \{0, \cdots, N_D - 1\}$$

4. Find index of the peak point of the spectrum

$$\hat{k}_p = \arg \max_{k \in \{0,1,\ldots,N_{\mathrm{D}}-1\}} |R_k|^2$$

5. Estimate $\hat{\delta}_1$ using $R_{\hat{k}_{p-1}}$, $R_{\hat{k}_p}$, $R_{\hat{k}_{p+1}}$ and $c_{\mathbf{w}}$ which is given in step-3

$$\hat{\delta}_1 = c_{\mathbf{w}} \, \mathrm{Re} \left\{ \frac{R_{\hat{k}_{p-1}} - R_{\hat{k}_{p+1}}}{2R_{\hat{k}_p} - R_{\hat{k}_{p-1}} - R_{\hat{k}_{p+1}}} \right\}$$

6. Form $\breve{\mathbf{r}} = [\breve{r}_0, \breve{r}_1, \ldots, \breve{r}_{N-1}]^T$, where

$$\breve{r}_n = r_n \, e^{-j2\pi \widehat{\delta}_1 n/N_{\mathrm{D}}}, \quad n = 0, 1, \ldots, N-1$$

7. Calculate the windowed N_{D}-point DFT of $\breve{\mathbf{r}}$

$$\breve{R}_k = \sum_{n=0}^{N-1} w_n \, \breve{r}_n \, e^{-j2\pi kn/N_{\mathrm{D}}}, k = 0, 1, \ldots, N_{\mathrm{D}} - 1$$

8. Estimate $\widehat{\delta}_2$ using $\breve{R}_{\hat{k}_p-1}$, $\breve{R}_{\hat{k}_p}$, $\breve{R}_{\hat{k}_p+1}$, where $c_{\mathbf{w}}$ is given in step-3

$$\widehat{\delta}_2 = c_{\mathbf{w}} \, \frac{\breve{R}_{\hat{k}_p-1} - \breve{R}_{\hat{k}_p+1}}{2\breve{R}_{\hat{k}_p} - \breve{R}_{\hat{k}_p-1} - \breve{R}_{\hat{k}_p+1}}$$

9. $\widehat{f} = (\hat{k}_p + \widehat{\delta}_1 + \widehat{\delta}_2)/N_{\mathrm{D}}$

4.3 Proposed Method For Frequency Estimation of a Complex Exponential in an Intermittent Signal via Correlated Mixture of Gaussian Noise

Consider the following signal model that represents N samples at the output of a sensor taken at regular time intervals

$$x_n = \underbrace{A\exp(j2\pi fn)}_{\text{signal of interest}} + \underbrace{v_n}_{\text{noise}}, \quad n = \{0, \ldots, N-1\}. \tag{4.5}$$

The sample x_n in eq. (4.5) has two main constituents. The signal component which is expressed as $A\exp(j2\pi fn)$ includes two parameters: The parameter A is the unknown complex amplitude of the signal of interest which is expressed as $A = A^{'}\exp(j\phi)$ where $A^{'}$ and $\phi \in [0, 2\pi)$ are unknown real scalars that represent the signal amplitude and phase respectively. $f \in [0, 1)$ is the normalized frequency of the signal and it is the main point of interest. The second component is the receiver noise v_n whose distribution is modeled by a two-component Gaussian mixture, that is

$$p(v_n) = \alpha_0 \mathcal{CN}(\mu_f; \sigma_f^2) + \alpha_1 \mathcal{CN}(\mu_h; \sigma_h^2). \tag{4.6}$$

At this point, we introduce a latent variable s_n, which is a binary valued random variable taking values 0 and 1, $s_n \in \{0, 1\}$ with probabilities $P(s_n = 0) = \alpha_0$ and $P(s_n = 1) = \alpha_1 = 1 - \alpha_0$. It can be thought as the indicator that represents the health state of a sensor, that is $s_n = 1$ means that the sensor works properly and on the other hand $s_n = 0$ indicates a failure in the sensor. When the sensor is healthy, the distribution of the noise becomes $p(v_n \mid s_n = 1) = \mathcal{CN}(\mu_h, \sigma_h^2)$ and it is equal to $p(v_n \mid s_n = 0) = \mathcal{CN}(\mu_f, \sigma_f^2)$ when there is a failure. In short hand notation this conditional density can be written as

$$p(v_n \mid s_n) = \mathcal{CN}\left(v_n; \mu_f^{(1-s_n)}\mu_h^{s_n}, (\sigma_f^2)^{(1-s_n)}(\sigma_h^2)^{s_n})\right). \tag{4.7}$$

It is assumed that s_n is generated by a Hidden Markov Model (HMM), which is presented in Figure 4.1,with known transition probabilities of $P(s_n = k|s_{n-1} = \ell) = \pi_{\ell k}$ for $k, \ell \in \{0, 1\}$. Without any loss of generality, the initial state probabilities $P(s_0 = 0) = \pi_0$ and $P(s_0 = 1) = \pi_1 = 1 - \pi_0$ are assumed to be known. Note that, this general HMM also covers the simpler case where $s_0, s_1, \cdots, s_{N-1}$ are independent and identically distributed.

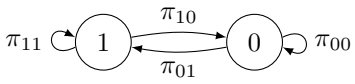


Figure 4.1: Markov model for sensor health states

Note that SNR changes according to the sensor health state. If the signal component is represented by y_n, that is $y_n = A\exp(j2\pi fn)$, then SNR $= E\left\{|y_n|^2\right\}/E\left\{|v_n|^2\right\}$ can be defined for healty and faulty states as

$$\text{SNR}_f = \frac{|A|^2}{|\mu_f|^2 + \sigma_f^2}, \qquad \text{SNR}_h = \frac{|A|^2}{|\mu_h|^2 + \sigma_h^2}. \tag{4.8}$$

We would like to also mention that, sensor health state can be well approximated by choosing $\mu_h = 0$ and a small value for variance σ_h^2. On the other hand, sensor failure can be modeled by a DC level represented by μ_f and a medium to large value for σ_f^2. Representing sensor anomalies by DC levels is one of the sensor fault models offered in [104].

The unkown parameters of the problem to be estimated are $\boldsymbol{\theta} = \{A, f, \mu_f, \sigma_f^2, \mu_h, \sigma_h^2\}$. Our aim is to find closed form estimators of the unknowns specific to this signal model by exploiting maximum likelihood approach and reveal any advantages of this effort, if any. Since the probability density function of the observations is a mixture of Gaussians, a close formed solution for the ML frequency estimator is not available. Therefore, we employ Expectation Maximization (EM) algorithm to the parameter estimation problem in a Gaussian mixture noise environment to iteratively maximize the expected value of likelihood function of the so-called complete data. Complete data in our model is chosen as the set that contains the observations $\boldsymbol{x} = \{x_0, \cdots, x_{N-1}\}$ and the latent variables $\boldsymbol{s} = \{s_0, \cdots, s_{N-1}\}$ which can be inferred from these observations through our probability density model in Eq. (4.7).

Expectation Phase :

In the Expectation Phase, the expected log-likelihood of the complete data is determined by taking expected value of complete data log-likelihood over posterior density of latent variables given the observations and current values of parameter estimates. In other words, $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) = E_{\boldsymbol{s}|\boldsymbol{x},\boldsymbol{\theta}^{old}}\{\ln p(\boldsymbol{s}, \boldsymbol{x})|\boldsymbol{x}; \boldsymbol{\theta}^{old}\}$ is calculated. By using definiton of conditional probability, joint pdf of complete data can be written as

$p(\boldsymbol{s},\boldsymbol{x}) = p(\boldsymbol{x}|\boldsymbol{s})p(\boldsymbol{s})$. Since the marginal distribution of latent sequence $p(\boldsymbol{s})$ does not depend on unknowns of the problem, the log-likelihood function can be written as the logarithm of conditional density that is $\mathcal{L}(\boldsymbol{x};\boldsymbol{\theta}) = \ln p(\boldsymbol{s},\boldsymbol{x}) \stackrel{c}{=} \ln p(\boldsymbol{x}|\boldsymbol{s})$ by omitting $p(\boldsymbol{s})$ term. It is straightforward to write the conditional pdf as $p(\boldsymbol{x}|\boldsymbol{s}) = \prod_{n=0}^{N-1} \mathcal{CN}\left(x_n; A\exp(j2\pi fn) + \mu_f^{(1-s_n)}\mu_h^{s_n}, (\sigma_f^2)^{(1-s_n)}(\sigma_h^2)^{s_n})\right)$ since given the sequence $\boldsymbol{s}$, the elements of the observation sequence, say x_n and x_m, become independent when $n \neq m$. After eliminating the constant terms that do not affect subsequent calculations, $Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old})$ turns into the form

$$\begin{aligned} Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old}) = &-\ln\sigma_f^2 \sum_{n=0}^{N-1} E_{\boldsymbol{s}|\boldsymbol{x},\boldsymbol{\theta}^{old}}\{1-s_n\} - \ln\sigma_h^2 \sum_{n=0}^{N-1} E_{\boldsymbol{s}|\boldsymbol{x},\boldsymbol{\theta}^{old}}\{s_n\} \\ &- \sum_{n=0}^{N-1} E_{\boldsymbol{s}|\boldsymbol{x},\boldsymbol{\theta}^{old}}\left\{\frac{1}{(\sigma_f^2)^{1-s_n}(\sigma_h^2)^{s_n}}\left|x_n - A\exp(j2\pi fn) - \mu_f^{(1-s_n)}\mu_h^{s_n}\right|^2\right\}. \end{aligned} \tag{4.9}$$

To calculate the conditional expected values in the above equation, we need the posterior density $p(\boldsymbol{s}|\boldsymbol{x};\boldsymbol{\theta}^{old})$. It can be seen from Eq.(4.9) that calculating $p\left(s_n \mid \boldsymbol{x};\boldsymbol{\theta}^{\text{old}}\right)$ for $n = \{0,\cdots,N-1\}$ is enough instead of trying to calculate $p\left(\boldsymbol{s}|\boldsymbol{x};\boldsymbol{\theta}^{old}\right)$ directly. This posterior density, $p\left(s_n \mid \boldsymbol{x};\boldsymbol{\theta}^{\text{old}}\right)$, is well known in the literature and can be calculated by $\alpha\beta$ recursion method [105, 106]. In this method, $\alpha(s_n)$ denotes joint probability density of n^{th} sensor state s_n and the observations from x_0 up to x_n. Whereas, $\beta(s_n)$ denotes conditional probability density of the observations from x_{n+1} up to x_{N-1} given s_n. In other words,

$$\alpha(s_n) = p(s_n, x_n, \cdots, x_0), \tag{4.10}$$

$$\beta(s_n) = p(x_{n+1}, \cdots, x_{N-1}|s_n). \tag{4.11}$$

The recursion equations of $\alpha(s_n)$ and $\beta(s_n)$ are defined as follows:

$$\alpha(s_n) = p\left(x_n|s_n;\boldsymbol{\theta}^{old}\right)\sum_{k=0}^{1} P(s_n|s_{n-1}=k)\alpha(s_{n-1}) \tag{4.12}$$

$$\beta(s_{n-1}) = \sum_{k=0}^{1} p\left(x_n|s_n=k;\boldsymbol{\theta}^{old}\right)P(s_n=k|s_{n-1})\beta(s_n=k). \tag{4.13}$$

$\alpha(s_n)$ is calculated forward for $n \geq 1$, and $\beta(s_n)$ is calculated backward for $2 \leq n \leq N-1$ after initializing α and β as

$$\alpha(s_0) = P(s_0)p\left(x_0|s_0;\boldsymbol{\theta}^{old}\right) = \pi_0 p\left(x_0|s_0;\boldsymbol{\theta}^{old}\right), \tag{4.14}$$

$$\beta(s_{N-1}) = 1. \tag{4.15}$$

As a result, the posterior density can be calculated by the following equation:

$$p\left(s_n|\boldsymbol{x};\boldsymbol{\theta}^{old}\right) = \frac{\alpha(s_n)\beta(s_n)}{\alpha(s_n=0)\beta(s_n=0)+\alpha(s_n=1)\beta(s_n=1)}. \tag{4.16}$$

In the above equations, $\boldsymbol{\theta}^{old}$ is shown explicitly to emphasize that the posterior densities in EM algorithm are evaluated at the old values of the unknown estimates. $\alpha\beta$ recursion as described here is summarized in Table 4.3 - Algorithm 4.

Finally, expectation step is completed by obtaining the $Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old})$ as

$$\begin{aligned} Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old}) = &- \ln\sigma_f^2 \sum_{n=0}^{N-1}(1-\hat{q}_n) \\ &- \frac{1}{\sigma_f^2}\sum_{n=0}^{N-1}(1-\hat{q}_n)\left|x_n - A\exp(j2\pi fn) - \mu_f\right|^2 \\ &- \ln\sigma_h^2 \sum_{n=0}^{N-1}\hat{q}_n \quad - \quad \frac{1}{\sigma_h^2}\sum_{n=0}^{N-1}\hat{q}_n\left|x_n - A\exp(j2\pi fn) - \mu_h\right|^2 \end{aligned} \tag{4.17}$$

where $\hat{q}_n$ in Eq.(4.17) denotes the posterior probability $\hat{q}_n = P(s_n=1|\boldsymbol{x};\boldsymbol{\theta}^{old})$.

Maximization Phase :

In the maximization phase, the expected complete log-likelihood is maximized with respect to the unknown parameters $\boldsymbol{\theta} = \{A, f, \mu_f, \mu_h, \sigma_f^2, \sigma_h^2\}$:

$$\hat{\boldsymbol{\theta}} = \{\hat{A}, \hat{f}, \hat{\mu}_f, \hat{\mu}_h, \hat{\sigma}_f^2, \hat{\sigma}_h^2\} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old}). \tag{4.18}$$

We would like to mention that, joint optimization of $Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old})$ with respect to all unkowns of the problem is intractable. But, we can make the observation that once the parameters $\mu_f, \mu_h, \sigma_f^2, \sigma_h^2$ are known, problem of optimizing $Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old})$ in Eq. (4.17) turns into the well known nonlinear least squares (NLS) problem for frequency estimation of a signal with unknown amplitude. Moreover, if we assume that frequency is known, ignoring the constant terms, the function $Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old})$ can be considered as the negative of the NLS criterion function.

Table 4.3: Calculation of the Posterior Probabilities

Algorithm 4 The Forward-Backward Procedure, [106]

Input : $\mathbf{x} = [x_0, x_1, \ldots, x_{N-1}]^T$, $\boldsymbol{\theta}^{old}$, $\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}, \pi_0, \pi_1$

Forward Recursion :

$\alpha(s_0) = \pi_0 p\left(x_0|s_0, \boldsymbol{\theta}^{old}\right)$

for $n = 1$ **to** $n = N-1$ **do**

$\alpha(s_n) = p\left(x_n|s_n, \boldsymbol{\theta}^{old}\right) \sum_{k=0}^{1} P(s_n|s_{n-1} = k)\alpha(s_{n-1})$

end for

Backward Recursion :

$\beta(s_{N-1}) = 1$

for $n = N-2$ **to** $n = 0$ **do**

$\beta(s_{n-1}) = \sum_{k=0}^{1} p\left(x_n|s_n = k, \boldsymbol{\theta}^{old}\right) P(s_n = k|s_{n-1})\beta(s_n = k)$

end for

Probability Calculation:

for $n = 0$ **to** $N-1$ **do**

$$\hat{q}_n = \frac{\alpha(s_{n-1} = 1)\,\beta(s_{n-1} = 1)}{\sum_{k=0}^{1} \alpha(s_{n-1} = k)\,\beta(s_{n-1} = k)}$$

end for

Output : $\boldsymbol{q} = [\hat{q}_0, \hat{q}_1, \ldots, \hat{q}_{N-1}]^T$

The NLS criterion function of the linear model $\boldsymbol{x}' = \boldsymbol{M}A + \boldsymbol{v}'$ can be written as

$$\begin{aligned} J(\boldsymbol{\theta}) = -\frac{1}{\sigma_f^2} \sum_{n=0}^{N-1} (1 - \hat{q}_n) \left|x_n - A\exp(j2\pi fn) - \mu_f\right|^2 \\ - \frac{1}{\sigma_h^2} \sum_{n=0}^{N-1} \hat{q}_n \left|x_n - A\exp(j2\pi fn) - \mu_h\right|^2. \end{aligned} \tag{4.19}$$

In vector matrix form, the error criterion is $J(\boldsymbol{\theta}) = (\boldsymbol{x}' - \boldsymbol{M}A)^H(\boldsymbol{x}' - \boldsymbol{M}A)$.

In the model, $\boldsymbol{x}^{'}$, $\boldsymbol{M}$ and $\boldsymbol{v}^{'}$ are defined as

$$\underbrace{\begin{bmatrix} \frac{\sqrt{1-\hat{q}_0}}{\sigma_f}(x_0-\mu_f) \\ \frac{\sqrt{1-\hat{q}_1}}{\sigma_f}(x_1-\mu_f) \\ \vdots \\ \frac{\sqrt{1-\hat{q}_{N-1}}}{\sigma_f}(x_{N-1}-\mu_f) \\ \frac{\sqrt{\hat{q}_0}}{\sigma_h}(x_0-\mu_h) \\ \frac{\sqrt{\hat{q}_1}}{\sigma_h}(x_1-\mu_h) \\ \vdots \\ \frac{\sqrt{\hat{q}_{N-1}}}{\sigma_h}(x_{N-1}-\mu_h) \end{bmatrix}}_{\boldsymbol{x}^{'}} = \underbrace{\begin{bmatrix} \frac{\sqrt{1-\hat{q}_0}}{\sigma_f}1 \\ \frac{\sqrt{1-\hat{q}_1}}{\sigma_f}\exp(j2\pi f) \\ \vdots \\ \frac{\sqrt{1-\hat{q}_{N-1}}}{\sigma_f}\exp(j2\pi f(N-1)) \\ \frac{\sqrt{\hat{q}_0}}{\sigma_h}1 \\ \frac{\sqrt{\hat{q}_1}}{\sigma_h}\exp(j2\pi f) \\ \vdots \\ \frac{\sqrt{\hat{q}_{N-1}}}{\sigma_h}\exp(j2\pi f(N-1)) \end{bmatrix}}_{\boldsymbol{M}} A + \boldsymbol{v}^{'}. \tag{4.20}$$

Maximizing $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ in Eq.(4.17) is equivalent to minimizing $J(\boldsymbol{\theta})$ and the estimator of A that minimizes $J(\boldsymbol{\theta})$ is known to be $\hat{A} = \left(\boldsymbol{M}^H\boldsymbol{M}\right)^{-1}\boldsymbol{M}^H\boldsymbol{x}^{'}$ which is equal to

$$\begin{aligned} \hat{A} = &\frac{\left(\sigma_f^2\right)^{-1}\sum_{n=0}^{N-1}(1-\hat{q}_n)(x_n-\mu_f)\exp(-j2\pi fn)}{\left(\sigma_f^2\right)^{-1}\sum_{n=0}^{N-1}(1-\hat{q}_n)+\left(\sigma_h^2\right)^{-1}\sum_{n=0}^{N-1}\hat{q}_n} \\ &+\frac{\left(\sigma_h^2\right)^{-1}\sum_{n=0}^{N-1}\hat{q}_n(x_n-\mu_h)\exp(-j2\pi fn)}{\left(\sigma_f^2\right)^{-1}\sum_{n=0}^{N-1}(1-\hat{q}_n)+\left(\sigma_h^2\right)^{-1}\sum_{n=0}^{N-1}\hat{q}_n}. \end{aligned} \tag{4.21}$$

When $\hat{A}$ in Eq. (4.21) is inserted in $J(\boldsymbol{\theta})$, frequency estimate that maximizes expected value of the complete data log-likelihood function, $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$, is found to be the value that maximizes $J^{'}(\boldsymbol{\theta}) = (\boldsymbol{x}^{'})^H\boldsymbol{M}\left(\boldsymbol{M}^H\boldsymbol{M}\right)^{-1}\boldsymbol{M}^H\boldsymbol{x}^{'}$. Since $\boldsymbol{M}^H\boldsymbol{M}$ is a scalar, it can be written as $J^{'}(\boldsymbol{\theta}) = \left(\boldsymbol{M}^H\boldsymbol{M}\right)^{-1}\left|\boldsymbol{M}^H\boldsymbol{x}^{'}\right|^2$ which leads us to the following updated estimate for frequency:

$$\begin{aligned} \hat{f} = \arg\max_f \Bigg| &\left(\sigma_f^2\right)^{-1}\sum_{n=0}^{N-1}(1-\hat{q}_n)(x_n-\mu_f)\exp(-j2\pi fn) \\ &+\left(\sigma_h^2\right)^{-1}\sum_{n=0}^{N-1}\hat{q}_n(x_n-\mu_h)\exp(-j2\pi fn)\Bigg|^2. \end{aligned} \tag{4.22}$$

We would like to note that, updated values of frequency, f^{new}, can be found by replacing mean and variance values of noise in Eq. (4.22) by μ_f^{old}, σ_f^{old}, μ_h^{old} and σ_h^{old}.

After obtaining updated frequency, the new value of signal amplitude, A^{new}, can be calculated by replacing f in Eq. (4.21) with f^{new}, and the other parameters with their old values. After A^{new} is obtained, we continue with the optimization of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ over the mean values of noise. If partial derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ is taken with respect to μ_f^*, μ_h^* and the resulting expression is equated to zero, we obtain the updated estimates of noise mean values as

$$\hat{\mu}_f = \left(\sum_{n=0}^{N-1} (1 - \hat{q}_n) \right)^{-1} \left(\sum_{n=0}^{N-1} (1 - \hat{q}_n) \left(x_n - A \exp(j2\pi f n) \right) \right), \tag{4.23}$$

$$\hat{\mu}_h = \left(\sum_{n=0}^{N-1} \hat{q}_n \right)^{-1} \left(\sum_{n=0}^{N-1} \hat{q}_n \left(x_n - A \exp(j2\pi f n) \right) \right). \tag{4.24}$$

We can use the updated values of A and f that are estimated at the current step of the iterative maximization procedure of EM Method, i.e., A^{new} and f^{new} to evaluate the estimators in Eq. (4.23) and Eq. (4.24). Next, we optimize over noise variance values by taking derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ with respect to σ_f^2, σ_h^2 and equating the result to zero. The updated $\hat{\sigma_f}^2$ and $\hat{\sigma_h}^2$ estimators become

$$\hat{\sigma}_f^2 = \left(\sum_{n=0}^{N-1} (1 - \hat{q}_n) \right)^{-1} \left(\sum_{n=0}^{N-1} (1 - \hat{q}_n) \left| x_n - A \exp(j2\pi f n) - \mu_f \right|^2 \right) \tag{4.25}$$

$$\hat{\sigma}_h^2 = \left(\sum_{n=0}^{N-1} \hat{q}_n \right)^{-1} \left(\sum_{n=0}^{N-1} \hat{q}_n \left| x_n - A \exp(j2\pi f n) - \mu_h \right|^2 \right) \tag{4.26}$$

Note that, the new estimates of mean noise values μ_f^{new}, μ_h^{new} together with A^{new} and f^{new} that are updated in the current step can be used to obtain estimates of noise variances $\hat{\sigma}_f$, $\hat{\sigma}_h$.

As a final note, the maximization procedure used in the above derivations can be considered as a form of Coordinate Descent Algorithm with round-robin coordinate selection approach. The proposed algorithm is summarized in Table 5. The reason why this Group Coordinate Algorithm is preferred is that joint maximization of the unknowns of the problem is not analytically possible. This problem is due to the coupling of the unknowns in the score functions that are presented in equations from (4.21) to (4.24).

Table 4.4: Proposed Method for Frequency Estimation Under Sensor Failure

Algorithm 5 Frequency Estimation Under Sensor Failure

Initialization: Initialize $\boldsymbol{\theta}$ as $\boldsymbol{\theta}^{new} = \left\{\mu_f^{old}, \mu_h^{old}, \left(\sigma_f^2\right)^{old}, \left(\sigma_h^2\right)^{old}, f^{old}, A^{old}\right\}$

Expectation Phase:

Compute the posterior probabilities $\hat{q}_n = P(s_n = 1|\boldsymbol{x}, \boldsymbol{\theta}^{old})$,

$n \in \{0, 1, \cdots, N-1\}$ by the forward backward procedure given in Table 4.3

Maximization Phase:

Compute $\hat{A}$ from (4.21)

Compute $\hat{f}$ from (4.22)

Compute $\hat{\mu}_f$ from (4.23)

Compute $\hat{\mu}_h$ from (4.24)

Compute $\hat{\sigma}_f$ from (4.25)

Compute $\hat{\sigma}_h$ from (4.26)

Termination Check:

if the convergence criteria is satisfied **then**

return $\hat{A}, \hat{f}, \hat{\mu}_f, \hat{\mu}_h, \hat{\sigma}_f, \hat{\sigma}_h$

else

Go to Expectation Phase

end if

4.4 Performance Bound for the Proposed Method

The frequency estimation problem includes 9 deterministic parameters to be estimated as:

$$\boldsymbol{\theta} = \left[\begin{array}{ccccccccc} \mathrm{Re}\{A\} & \mathrm{Im}\{A\} & f & \mathrm{Re}\{\mu_f\} & \mathrm{Im}\{\mu_f\} & \sigma_f^2 & \mathrm{Re}\{\mu_h\} & \mathrm{Im}\{\mu_h\} & \sigma_h^2 \end{array} \right]^T.$$

If partial derivatives of variance and mean of the observartions are taken with respect to the unkowns assuming that the sensor is in faulty state, the following equalities are obtained:

$$\begin{array}{lll} \dfrac{\partial C_{s_n=0}}{\partial \mathrm{Re}\{A\}} = 0 & \dfrac{\partial C_{s_n=0}}{\partial \mathrm{Re}\{\mu_f\}} = 0 & \dfrac{\partial C_{s_n=0}}{\partial \mathrm{Re}\{\mu_h\}} = 0 \\ \dfrac{\partial C_{s_n=0}}{\partial \mathrm{Im}\{A\}} = 0 & \dfrac{\partial C_{s_n=0}}{\partial \mathrm{Im}\{\mu_f\}} = 0 & \dfrac{\partial C_{s_n=0}}{\partial \mathrm{Im}\{\mu_h\}} = 0 \\ \dfrac{\partial C_{s_n=0}}{\partial f} = 0 & \dfrac{\partial C_{s_n=0}}{\partial \sigma_f^2} = 1 & \dfrac{\partial C_{s_n=0}}{\partial \sigma_h^2} = 0 \end{array}$$

$$\begin{array}{lll} \dfrac{\partial \mu_{s_n=0}}{\partial \mathrm{Re}\{A\}} = \exp(j2\pi fn) & \dfrac{\partial \mu_{s_n=0}}{\partial \mathrm{Re}\{\mu_f\}} = 1 & \dfrac{\partial \mu_{s_n=0}}{\partial \mathrm{Re}\{\mu_h\}} = 0 \\ \dfrac{\partial \mu_{s_n=0}}{\partial \mathrm{Im}\{A\}} = j\exp(j2\pi fn) & \dfrac{\partial \mu_{s_n=0}}{\partial \mathrm{Im}\{\mu_f\}} = j & \dfrac{\partial \mu_{s_n=0}}{\partial \mathrm{Im}\{\mu_h\}} = 0 \\ \dfrac{\partial \mu_{s_n=0}}{\partial f} = Aj2\pi n\exp(j2\pi fn) & \dfrac{\partial \mu_{s_n=0}}{\partial \sigma_f^2} = 0 & \dfrac{\partial \mu_{s_n=0}}{\partial \sigma_h^2} = 0 \end{array}$$

If partial derivatives of variance and mean of the observartions are taken with respect to the unkowns assuming that the sensor is in healthy state, the following equalities are obtained:

$$\begin{array}{lll} \dfrac{\partial C_{s_n=1}}{\partial \mathrm{Re}\{A\}} = 0 & \dfrac{\partial C_{s_n=1}}{\partial \mathrm{Re}\{\mu_f\}} = 0 & \dfrac{\partial C_{s_n=1}}{\partial \mathrm{Re}\{\mu_h\}} = 0 \\ \dfrac{\partial C_{s_n=1}}{\partial \mathrm{Im}\{A\}} = 0 & \dfrac{\partial C_{s_n=1}}{\partial \mathrm{Im}\{\mu_f\}} = 0 & \dfrac{\partial C_{s_n=1}}{\partial \mathrm{Im}\{\mu_h\}} = 0 \\ \dfrac{\partial C_{s_n=1}}{\partial f} = 0 & \dfrac{\partial C_{s_n=1}}{\partial \sigma_f^2} = 0 & \dfrac{\partial C_{s_n=1}}{\partial \sigma_h^2} = 1 \end{array}$$

$$\begin{array}{lll} \dfrac{\partial \mu_{s_n=1}}{\partial \mathrm{Re}\{A\}} = \exp(j2\pi fn) & \dfrac{\partial \mu_{s_n=1}}{\partial \mathrm{Re}\{\mu_f\}} = 0 & \dfrac{\partial \mu_{s_n=1}}{\partial \mathrm{Re}\{\mu_h\}} = 1 \\ \dfrac{\partial \mu_{s_n=1}}{\partial \mathrm{Im}\{A\}} = j\exp(j2\pi fn) & \dfrac{\partial \mu_{s_n=1}}{\partial \mathrm{Im}\{\mu_f\}} = 0 & \dfrac{\partial \mu_{s_n=1}}{\partial \mathrm{Im}\{\mu_h\}} = j \\ \dfrac{\partial \mu_{s_n=1}}{\partial f} = Aj2\pi n\exp(j2\pi fn) & \dfrac{\partial \mu_{s_n=1}}{\partial \sigma_f^2} = 0 & \dfrac{\partial \mu_{s_n=1}}{\partial \sigma_h^2} = 1 \end{array}$$

Using the identity in Eq. (2.53), the non-zero elements of FIM when sensor is in faulty state can be written as:

$$
\begin{aligned}
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{1,1} &= 2/\sigma_f^2, \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{1,3} &= -4\pi n\,\mathrm{Im}\{A\}(1/\sigma_f^2), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{1,4} &= (2/\sigma_f^2)\cos(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{1,5} &= (2/\sigma_f^2)\sin(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{2,2} &= 2/\sigma_f^2, \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{2,3} &= 4\pi n\,\mathrm{Re}\{A\}(1/\sigma_f^2), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{2,4} &= -(2/\sigma_f^2)\sin(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{2,5} &= (2/\sigma_f^2)\cos(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{3,1} &= -4\pi n\,\mathrm{Im}\{A\}(1/\sigma_f^2), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{3,2} &= 4\pi n\,\mathrm{Re}\{A\}(1/\sigma_f^2), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{3,3} &= 8\pi^2 n^2\,|A|^2\,(1/\sigma_f^2), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{3,4} &= -(4\pi n/\sigma_f^2)\left[\,\mathrm{Re}\{A\}\sin(2\pi fn) + \mathrm{Im}\{A\}\cos(2\pi fn)\,\right], \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{3,5} &= (4\pi n/\sigma_f^2)\left[\,\mathrm{Re}\{A\}\cos(2\pi fn) - \mathrm{Im}\{A\}\sin(2\pi fn)\,\right], \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{4,1} &= 2/\sigma_f^2\cos(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{4,2} &= -2/\sigma_f^2\sin(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{4,3} &= -(4\pi n/\sigma_f^2)\left[\,\mathrm{Re}\{A\}\sin(2\pi fn) + \mathrm{Im}\{A\}\cos(2\pi fn)\,\right], \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{4,4} &= 2/\sigma_f^2, \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{5,1} &= 2/\sigma_f^2\sin(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{5,2} &= 2/\sigma_f^2\cos(2\pi fn), \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{5,3} &= (4\pi n/\sigma_f^2)\left[\,\mathrm{Re}\{A\}\cos(2\pi fn) - \mathrm{Im}\{A\}\sin(2\pi fn)\,\right], \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{5,5} &= 2/\sigma_f^2, \\
[\boldsymbol{I}_{s_n=0}(\boldsymbol{\theta})]_{6,6} &= 1/\sigma_f^4,
\end{aligned}
$$

Similarly, the non-zero elements of FIM when sensor is in healthy state can be written as:

$$
\begin{aligned}
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{1,1} = 2/\sigma_h^2,\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{1,3} = -4\pi n\,\mathrm{Im}\{A\}(1/\sigma_h^2),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{1,7} = 2/\sigma_h^2\cos(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{1,8} = 2/\sigma_h^2\sin(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{2,2} = 2/\sigma_h^2,\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{2,3} = 4\pi n\,\mathrm{Re}\{A\}(1/\sigma_h^2),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{2,7} = -(2/\sigma_h^2)\sin(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{2,8} = (2/\sigma_h^2)\cos(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{3,1} = -4\pi n\,\mathrm{Im}\{A\}(1/\sigma_h^2),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{3,2} = 4\pi n\,\mathrm{Re}\{A\}(1/\sigma_h^2),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{3,3} = 8\pi^2 n^2\,|A|^2\,(1/\sigma_h^2),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{3,7} = -(4\pi n/\sigma_h^2)\left[\,\mathrm{Re}\{A\}\sin(2\pi fn) + \mathrm{Im}\{A\}\cos(2\pi fn)\,\right],\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{3,8} = (4\pi n/\sigma_h^2)\left[\,\mathrm{Re}\{A\}\cos(2\pi fn) - \mathrm{Im}\{A\}\sin(2\pi fn)\,\right],\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{7,1} = (2/\sigma_h^2)\cos(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{7,2} = -(2/\sigma_h^2)\sin(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{7,3} = -(4\pi n/\sigma_h^2)\left[\,\mathrm{Re}\{A\}\sin(2\pi fn) + \mathrm{Im}\{A\}\cos(2\pi fn)\,\right],\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{7,7} = (2/\sigma_h^2),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{8,1} = (2/\sigma_h^2)\sin(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{8,2} = (2/\sigma_h^2)\cos(2\pi fn),\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{8,3} = (4\pi n/\sigma_h^2)\left[\,\mathrm{Re}\{A\}\cos(2\pi fn) - \mathrm{Im}\{A\}\sin(2\pi fn)\,\right]\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{8,8} = 2/\sigma_h^2,\\
&[\boldsymbol{I}_{s_n=1}(\boldsymbol{\theta})]_{9,9} = 1/\sigma_h^4,
\end{aligned}
$$

We would like to note that, representing complex unknowns in real and imaginary parts and calculating FIM accordingly is preferred in this section and Section 3.3.1. But, it is also possible to adopt a direct approach as in [107] which allows forming a hybrid unkown vector with real and complex parameters and calculating FIM

accordingly.

Finally, MCRB can be calculated by using the following equation:

$$\begin{aligned}\left[\boldsymbol{I}_M\left(\boldsymbol{\theta}\right)\right]_{i,j} &= E_{\boldsymbol{x},\boldsymbol{s}}\left\{-\frac{\partial^2 \ln p(\boldsymbol{x} \mid \boldsymbol{s};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right\} \\ &= E_{\boldsymbol{s}}\left\{E_{\boldsymbol{x}|\boldsymbol{s}}\left\{-\frac{\partial^2 \ln p(\boldsymbol{x} \mid \boldsymbol{s};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right\}\right\}.\end{aligned} \tag{4.27}$$

4.5 Simulation Results For Proposed Method

In all experiments, the sensor noise is zero mean ($\mu_{\text{h}} = 0$), unit variance ($\sigma_{\text{h}}^2 = 1$) in the healthy state and has the bias of $\mu_{\text{f}} = 2\,e^{j\pi/4}$ and the variance of $\sigma_{\text{f}}^2 = 6$ in the faulty state. Hence, the sensor operates at 10 dB SNR loss in the faulty state, $\text{SNR}_{\text{f}} = \text{SNR}_{\text{h}}/10$. The phase of the complex amplitude A is randomly chosen from the uniform distribution over $[0, 2\pi)$. The magnitude is set as $|A| = \sqrt{(|\mu_{\text{h}}|^2 + \sigma_{\text{h}}^2)\,\text{SNR}_{\text{h}}}$ depending on SNR_{h}. The data length of $N = 128$ and the zero-padded DFT of length $N_{\text{D}} = 2N$ is used. The unknown normalized frequency $f = (k_p + \delta)/N$ is set as $k_p = 4$ and $\delta = 0.42$. The Markov chain model parameters are $\pi_{00} = 0.6$ and $\pi_{11} = 0.9$. The initial states probabilities of the chain are taken as the steady-state probabilities of $\pi_0 = 0.2$ and $\pi_1 = 0.8$. Hence, the chain output can be considered at steady-state. (The proposed method is additionally run with the transition probabilities $\pi_{\ell k} = 0.5$ for $k, \ell \in \{0, 1\}$ and the equally-likely initial states, $\pi_0 = \pi_1 = 0.5$, to study the mismatched model.) ESPRIT with the 4×4 sample covariance matrix and the A&M (Alg1 and Alg2) method [3, Table 1] with 20 iterations are used for performance comparisons.

Figure 4.2 shows a typical realization under this set-up with $\text{SNR}_{\text{h}} = 0$ dB. The top panel shows the real parts of the signal and sensor observations. The bottom panel shows the true sensor state and the state prediction (the probability that the sensor is in the healthy state).

The CRB bounds and the clairvoyant estimator (the estimator which assumes the knowledge of A, μ_{f}, μ_{h}, σ_{f}^2, σ_{h}^2 and true-states) results are given as performance bounds. Each experiment is conducted with 5000 Monte Carlo runs. The root mean square error (RMSE) values are given in terms of the DFT bins $(k_p + \delta)$ where the

normalized frequency is $f = (k_p + \delta)/N$. Therefore, an RMSE value of, say, 10^{-2} has a natural interpretation of 100 times accuracy improvement over N-point DFT estimate (periodogram).

At this point, we would like to look at the structure of the frequency estimator more closely.

Let us define η_n, ζ_n and $\tilde{x}_n$ as

$$\eta_n = (\sigma_f^2)^{old}\hat{q}_n + (\sigma_h^2)^{old}(1-\hat{q}_n), \tag{4.28}$$

$$\zeta_n = \left((\sigma_f^2)^{old}\mu_h^{old}\hat{q}_n + (\sigma_h^2)^{old}\mu_f^{old}(1-\hat{q}_n)\right)/\eta_n, \tag{4.29}$$

$$\tilde{x}_n = x_n - \zeta_n. \tag{4.30}$$

Then, it is clear that the frequency estimator in Eq. (4.22) can be rewritten as

$$\hat{f} = \arg\max_f \left| \sum_{n=0}^{N-1} \eta_n \tilde{x}_n \exp(-j2\pi f n) \right|^2. \tag{4.31}$$

If Eq. (4.31) is examined, it can be seen that the sequence η_n, $n \in \{0, \cdots, N-1\}$ can be considered as a window function making it possible for us to use a windowed fine frequency estimator [4] for a computationally efficient solution. Therefore, in the simulations updating frequencies using the function in (4.31) with $\hat{q}_n$ as a window sequence is preferred.

Before the first expectation step, we initialize the unkowns as:

$$f^{old} = \frac{1}{N_D} \arg \max_{k \in \{0,1,\cdots,N_D-1\}} \left| \sum_{n=0}^{N-1} x_n \exp(-j2\pi k n/N_D) \right|^2, \tag{4.32}$$

$$A^{old} = \frac{1}{N} \sum_{n=0}^{N-1} x_n \exp(-j2\pi f^{old} n), \tag{4.33}$$

$$\mu_f^{old} = \frac{1}{N} \sum_{n=0}^{N-1} \left(x_n - A^{old}\exp(j2\pi f^{old} n)\right), \tag{4.34}$$

$$\mu_h^{old} = 0.1\mu_0^{old}, \tag{4.35}$$

$$(\sigma_f^2)^{old} = \frac{1}{N} \sum_{n=0}^{N-1} \left|x_n - A^{old}\exp(j2\pi f^{old} n)\right|^2, \tag{4.36}$$

$$(\sigma_h^2)^{old} = 0.1(\sigma_0^2)^{old}. \tag{4.37}$$

We control convergence of marginal posterior probability sequence and terminate the iterations when $\|\boldsymbol{q}^{\text{new}} - \boldsymbol{q}^{\text{old}}\|_\infty < \epsilon$, where $\|\cdot\|_\infty$ is the ℓ_∞ norm and the vector $\boldsymbol{q}$ is defined as $\boldsymbol{q} = \{\hat{q}_0, \cdots, \hat{q}_{N-1}\}$.

4.5.1 Experiment 1

In this experiment, the state sequence s_n is generated once and the same state sequence is used in all Monte Carlo runs. The results are given in Figure 4.3. Here CRB depends on the initially determined state-sequence. We observe that the proposed method with either true or mismatched transition probabilities follows CRB at high SNR more closely than the other methods and both cases present a better performance at all SNR values.

4.5.2 Experiment 2

Different from the earlier experiment, the state sequence is generated independently at each Monte Carlo run. The results are shown in Figure 4.4. (Here the modified CRB (MCRB) [2] is used for comparisons.) As in the earlier experiment, the proposed method tracks the bounds at high SNR and performs better than the alternative methods.

4.5.3 Experiment 3

In this experiment, the true frequency is varied from 0.55 and 5.84 with the grid spacing of 0.23 units (in DFT bins) and $(\text{SNR}_\text{f}, \text{SNR}_\text{h})$ pair is fixed to $(-10, 0)$ dB. The state sequence is generated independently for each realization as in the 2nd experiment. It can be observed from Figure 4.5 that the proposed method has better accuracy than other methods at all frequencies.

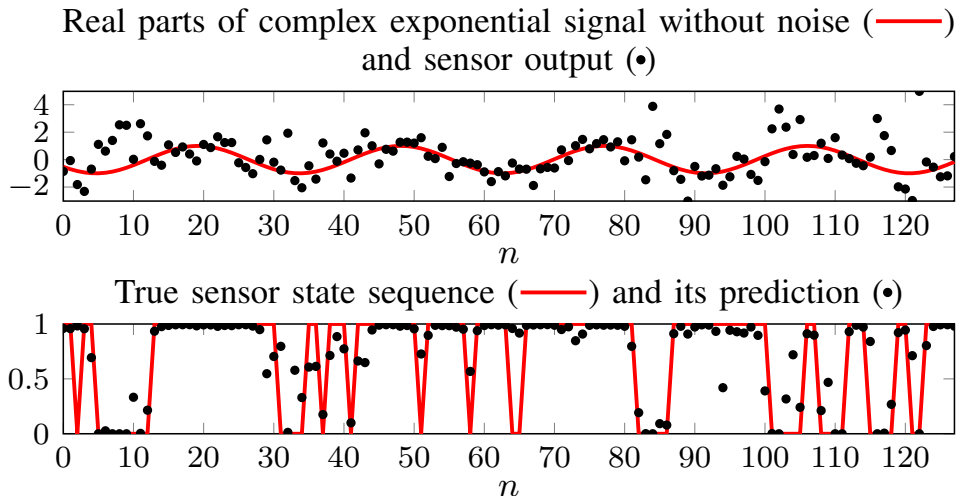


Figure 4.2: A realization for $\mathrm{SNR_h} = 0$ dB. The state prediction refers to the probability of the sensor being in the healthy state, which is given by Eq. (4.16)

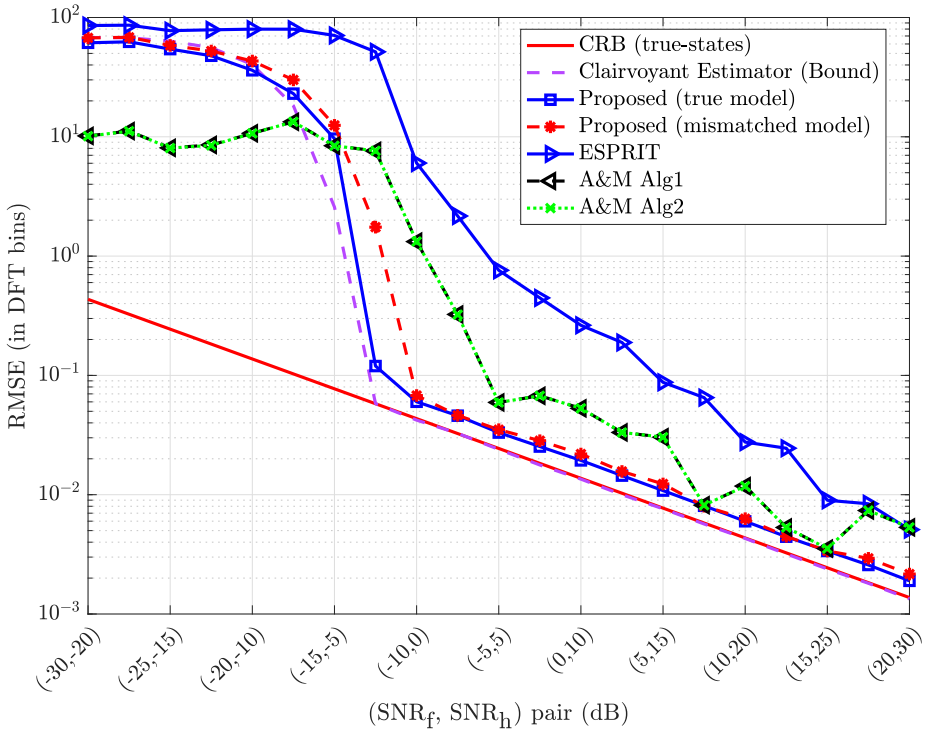


Figure 4.3: Experiment 1 – The comparison of RMSE curves.

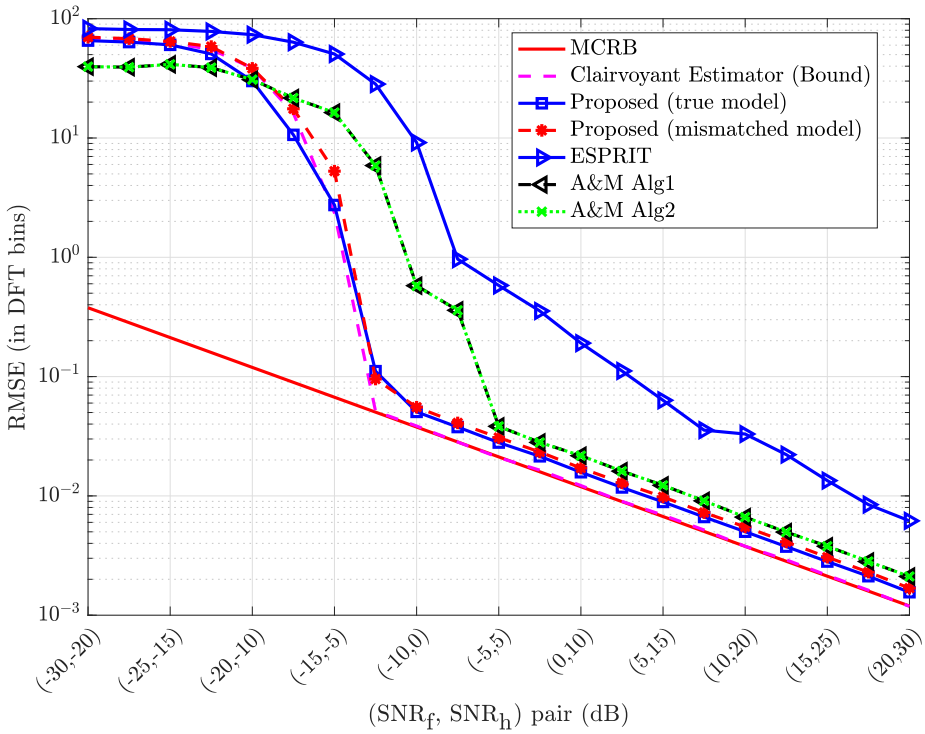


Figure 4.4: Experiment 2 – The comparison of RMSE curves.

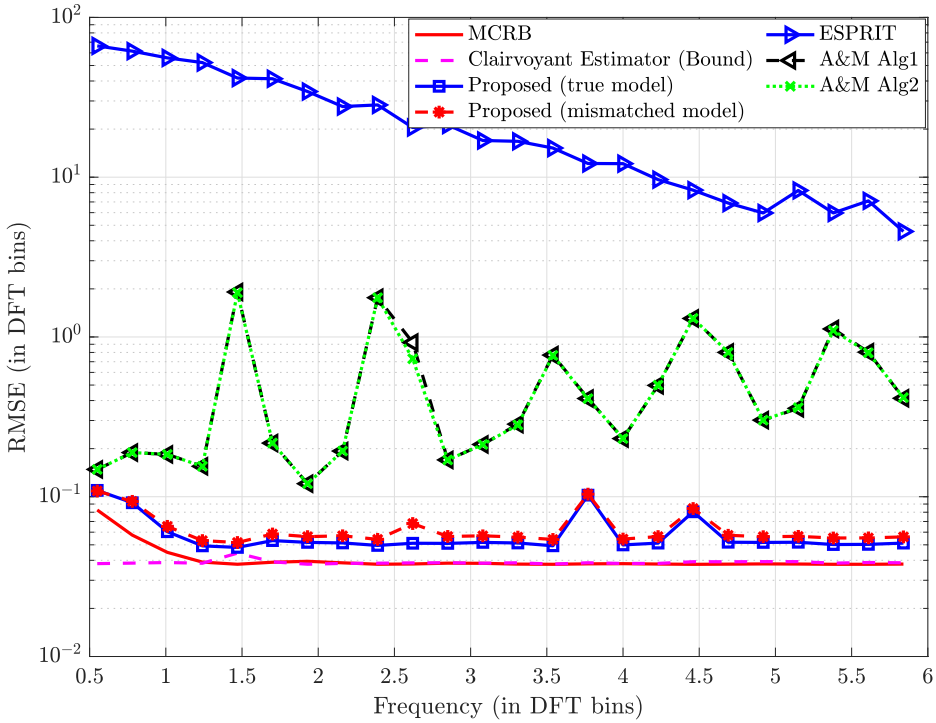


Figure 4.5: Experiment 3 – The comparison of RMSE curves.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

In this thesis, we studied parameter estimation problem under signal intermittency.

Firstly, we examined direction of arrival estimation problem under intermittent jamming. Intermittent jamming is a frequently encountered problem especially in electronic support systems [67]. In spite of its importance and practicality, DOA estimation under this jamming modality has received little attention in open literature. In this study, we presented the maximum likelihood angle estimator specific to the intermittent jamming modality using an EM based approach. In the performance comparisons, the suggested scheme is compared with the conventional estimator, atomic norm based estimator and performance bounds. The conventional estimator is assumed to have perfect knowledge of jammer parameters (jammer angular position and power), while the suggested scheme is unaware of these parameters and estimates them through the maximum likelihood framework. In spite of the additional information provided to the conventional estimator, the suggested scheme is shown to perform better in all scenarios. We note that the performance of the conventional estimator suddenly degrades with increasing jammer existence probability for strong jammers. This is not the case for the suggested estimator. (This is most clearly observed from Figure 3.5b where the conventional estimator performance rapidly degrades towards the jammer-always-present CRB with increasing jammer existing probability; while the suggested estimator presents a graceful degradation closely tracking the MCRB performance bound. At the operating point of jammer existence probability being equal to 0.1 in Figure 3.5b , the conventional estimator operates as if jammer is always present, while the suggested estimator operates as if jammer is always absent or with a minor degradation.) The atomic norm estimator operates independent of the

number of target signals and does not explicitly take into account the statistical description of the observation model such as signal intermittency. This leads to poorer performance for these generic estimator in comparison to the maximum likelihood estimators developed specifically for the problem. As a final note on the first problem, we can state that the suggested DOA estimation method for the intermittent jamming scenario significantly benefits from the inclusion of problem specific information into estimator design via the EM-formulation and this leads to significant estimation accuracy improvements over conventional or off-the-shelf estimators.

Secondly, we examined the problem of frequency estimation under correlated mixture of Gaussians noise model. This problem formulation can be considered as the extension of the first one on DOA estimation. Here, we assume signal intermittency is due to Markov process and intermittency persists according to the model. In the first problem, intermittency is modelled as a faulty snapshot which are independent and identically distributed. The noise model is utilized to represent the bimodal noise density of the sensor, which is assumed to have two states representing its nominal operation and intermittent failures. The proposed method is based on the expectation-maximization approach and utilizes efficient signal processing techniques such as forward-backward recursion and two-stage frequency estimators to develop a solution to the problem. Numerical results show that the proposed method outperforms the competing methods.

Future works can be stated as follows: As the first future work, two-dimensional DOA estimation problem can be studied in the presence of a single or multiple interferers. Besides this, the same problem can be restated as the indicator variable being correlated from sample to sample and a DOA estimation performance under this setting can be investigated. In our second problem, we focused on frequency estimation of a single tone. As an extention to this problem estimating frequencies of multiple tones in correlated Gaussian mixture noise can be considered as a future work. The suggested EM based solutions can be used for other intermittent signal/noise models as well.

APPENDIX A

DOA ESTIMATION UNDER INTERMITTENT JAMMING WITH THE KNOWLEDGE OF POSSIBLE JAMMING POWER LEVELS

In this section, an alternative solution for a related DOA estimation problem under signal intermittency is presented. In this case, contrary to the problem in Section 3.1, it is assumed that the possible jammer power levels is known, but it is not known to the receiver the jammer power sample by sample.

Reconsider the signal model in Eq. (2.32), that is

$$\boldsymbol{x}_i = \gamma_{t,i}\boldsymbol{a}(\theta_t) + \gamma_{j,i}\boldsymbol{a}(\theta_j) + \boldsymbol{n}_i, \quad i = \{1, ..., L\} \tag{A.1}$$

where $\boldsymbol{a}(\theta_t)$, $\boldsymbol{a}(\theta_j)$, $\gamma_{t,i}$, θ_t, θ_j and $\boldsymbol{n}_i$ are as defined in Section 3.1. Unlike Eq.(3.1), in eq. (A.1) there is no indicator variable that indicates the presence of the interfering signal. Yet, similar to the eq. (3.1), $\gamma_{j,i}$, the random variable that represents jammer signal amplitudes, is a zero mean complex Gaussian random variable with a variance σ_i^2, i.e., $\gamma_{j,i} \sim \mathcal{CN}(0, \sigma_i^2)$. Also, $\gamma_{j,i}$ and $\gamma_{j,k}$ are independent when $i \neq k$. Assume moreover that σ_i^2 can have one of M distinct values, i.e., $\sigma_i^2 \in \{s_1, s_2, \cdots, s_M\}$. Let us denote the probability of the variance of the interfering signal amplitude being equal to s_m with α_m, that is $P(\sigma_i^2 = s_m) = \alpha_m$. It is also clear that $\alpha_1 + \alpha_2 + \cdots + \alpha_M = 1$. One thing to be emphasized again here is that the distinct values of variances of the jammer signal amplitudes are known to the receiver. However, neither which interference signal component has which variance nor their probabilities, $\alpha_i, i = \{1, \cdots, L\}$ are known to the receiver.

Note that with these assumptions we can still model the jammer absent case by considering the limit point where $\sigma_i^2 = 0$.

The unknown non-random parameters $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \alpha_1, \cdots, \alpha_M\}$ are the

variables to be estimated by maximum likelihood approach.

It can be seen from the eq. (A.1) that, the distribution of a single observation is

$$p(\boldsymbol{x}_i;\boldsymbol{\theta}) = \sum_{m=1}^{M} \alpha_m \mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + \sigma_m^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right) \tag{A.2}$$

which is in the form of Gaussian mixture model [89]. The log-likelihood of L snapshots is

$$\mathcal{L} = \sum_{i=1}^{L} \ln \left\{ \sum_{m=1}^{M} \alpha_m \mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + \sigma_m^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right) \right\}. \tag{A.3}$$

Similar to the case in Eq. (3.2) of Section 3.2, the natural logarithm cannot be directly applied to the distribution because of the summation terms in the distribution. As a result, EM Algorithm can be employed instead of the maximum likelihood approach to obtain closed form solutions of the unknowns.

To solve the above estimation problem with EM Algorithm, we choose $\{\sigma_i^2\}_{i=1}^{L}$ as latent variables. Then, denoting observed and latent variables by $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively, our complete data set becomes $\{\boldsymbol{X}, \boldsymbol{Y}\} = \{x_1, \cdots, x_L, \sigma_1^2, \cdots, \sigma_L^2\}$. To calculate the expected value of log-likelihood of the complete data, i.e., $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ we need to write down the joint and conditional probability density functions. The marginal densities of observed and latent variables, together with their conditional probability density function can be written as

$$p(\boldsymbol{x}_i \mid \sigma_i^2;\boldsymbol{\theta}) = \mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + \sigma_i^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right), \tag{A.4}$$

$$p(\sigma_i^2;\boldsymbol{\theta}) = \sum_{k=1}^{M} \alpha_k \delta(\sigma_i^2 - s_k), \tag{A.5}$$

$$p(\boldsymbol{x}_i;\boldsymbol{\theta}) = \sum_{k=1}^{M} \alpha_k \mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + s_k \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right). \tag{A.6}$$

By using Eq. (A.4) and Eq. (A.5), the joint density of observed and latent variables can be written as

$$p(\boldsymbol{x}_i, \sigma_i^2;\boldsymbol{\theta}) = \sum_{k=1}^{M} \alpha_k \mathcal{CN}\left(\boldsymbol{x}_i;\gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + \sigma_i^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right) \delta(\sigma_i^2 - s_k). \tag{A.7}$$

Then, by using Eq. (A.6), Eq. (A.7) and applying Bayes' theorem, the density of

latent variables given observations is obtained as the following:

$$
\begin{aligned}
p(\sigma_i^2 \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{old}) &= \frac{p(\boldsymbol{x}_i, \sigma_i^2; \boldsymbol{\theta}^{old})}{p(\boldsymbol{x}_i; \boldsymbol{\theta}^{old})} \\
&= \frac{\sum_{k=1}^{M} \alpha_k^{old} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{old} \boldsymbol{a}(\theta_t^{old}), \boldsymbol{I} + \sigma_i^2 \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right) \delta(\sigma_i^2 - s_k)}{\sum_{k=1}^{M} \alpha_k^{old} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{old} \boldsymbol{a}(\theta_t^{old}), \boldsymbol{I} + s_k \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}.
\end{aligned} \tag{A.8}
$$

Let us define the responsibility $M_{m,i}^{old}$ as

$$
\begin{aligned}
M_{m,i}^{old} &= p(\sigma_i^2 = s_m \mid \boldsymbol{x}_i; \boldsymbol{\theta}^{old}) \\
&= \frac{\alpha_m^{old} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{old} \boldsymbol{a}(\theta_t^{old}), \boldsymbol{I} + s_m \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}{\sum_{k=1}^{M} \alpha_k^{old} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{old} \boldsymbol{a}(\theta_t^{old}), \boldsymbol{I} + s_k \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}.
\end{aligned} \tag{A.9}
$$

Expectation Phase:

The expected value of complete data log-likelihood function can be written as

$$
Q(\boldsymbol{\theta},\boldsymbol{\theta}^{old}) = \boldsymbol{E}_{\boldsymbol{Y}}\{\ln(p(\boldsymbol{X}, \boldsymbol{Y} \mid \theta)) \mid \boldsymbol{X} = \boldsymbol{X}_{data}; \boldsymbol{\theta} = \boldsymbol{\theta}^{old}\} \tag{A.10}
$$

$$
\begin{aligned}
&= \sum_{\sigma_L^2 = s_1}^{s_M} \cdots \sum_{\sigma_1^2 = s_1}^{s_M} \ln\left(\prod_{i=1}^{L} p\left(\boldsymbol{x}_i, \sigma_i^2; \boldsymbol{\theta}\right)\right)\left(\prod_{l=1}^{L} p\left(\sigma_l^2 \mid \boldsymbol{x}_l; \boldsymbol{\theta}^{old}\right)\right) \\
&= \sum_{i=1}^{L}\left(\sum_{\sigma_L^2 = s_1}^{s_M} \cdots \sum_{\sigma_1^2 = s_1}^{s_M} \ln p\left(\boldsymbol{x}_i, \sigma_i^2; \boldsymbol{\theta}\right)\left(\prod_{l=1}^{L} p\left(\sigma_l^2 \mid \boldsymbol{x}_l; \boldsymbol{\theta}^{old}\right)\right)\right) \\
&= \sum_{i=1}^{L}\left(\sum_{m=1}^{M} M_{m,i}^{old} \ln\left[\alpha_m \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}\boldsymbol{a}(\theta_t), \boldsymbol{I} + s_m \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)\right]\right).
\end{aligned} \tag{A.11}
$$

If we expand the probability density term in Eq. (A.11), we obtain

$$
\begin{aligned}
Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) &= -\sum_{i=1}^{L}\sum_{m=1}^{M} N M_{m,i}^{old} \ln \pi \\
&+ \sum_{i=1}^{L}\sum_{m=1}^{M} M_{m,i}^{old} \ln\left(\alpha_m\right) - \sum_{i=1}^{L}\sum_{m=1}^{M} M_{m,i}^{old} \ln\left|\boldsymbol{I} + s_m \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right| \\
&- \sum_{i=1}^{L}\sum_{m=1}^{M} M_{m,i}^{old} \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right)^H \left(\boldsymbol{I} + s_m \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)^{-1} \left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right).
\end{aligned} \tag{A.12}
$$

Maximization Phase:

To find $\hat{\alpha}_m$, $m = 1, \cdots, M$, form the Lagrangian Λ as the following:

$$
\Lambda = Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) + \lambda\left(\sum_{m=1}^{M} \alpha_m - 1\right). \tag{A.13}
$$

If we take derivative of Λ with respect to $\alpha_m, m = 1, \cdots, Ms$ we obtain

$$\frac{\partial \Lambda}{\partial \alpha_m} = \frac{1}{\alpha_m} \sum_{i=1}^{L} M_{m,i}^{old} + \lambda = 0 \quad \rightarrow \quad \alpha_m = -\frac{1}{\lambda} \sum_{i=1}^{L} M_{m,i}^{old}. \tag{A.14}$$

Since we know that

$$\sum_{m=1}^{M} \alpha_m = -\frac{1}{\lambda} \sum_{i=1}^{L} \sum_{m=1}^{M} M_{m,i}^{old} = 1, \tag{A.15}$$

we obtain $\lambda = -L$. As a result, $\hat{\alpha}_m$ is found to be

$$\hat{\alpha}_m = \frac{1}{L} \sum_{i=1}^{L} M_{m,i}^{old}. \tag{A.16}$$

If we put $\hat{\alpha}_m$ in Eq. (A.12), we obtain

$$\begin{aligned} &Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) = constants \\ &- \sum_{i=1}^{L} \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t)\right)^H \left[\sum_{m=1}^{M} M_{m,i}^{old} \left(\boldsymbol{I} + s_m \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right)^{-1}\right] \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t)\right). \end{aligned} \tag{A.17}$$

In the above equation, we can define $\boldsymbol{W}_i$ as

$$\begin{aligned} \boldsymbol{W}_i &= \sum_{m=1}^{M} M_{m,i}^{old} \left(\boldsymbol{I} + s_m \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right)^{-1} \\ &= \sum_{m=1}^{M} M_{m,i}^{old} \left(\boldsymbol{I} - \frac{s_m}{1 + N s_m} \boldsymbol{a}(\theta_j) \boldsymbol{a}^H(\theta_j)\right). \end{aligned} \tag{A.18}$$

If we put $\boldsymbol{W}_i$ in Eq. (A.17), the expected value of complete data log-likelihood can be written as

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) = constants - \sum_{i=1}^{L} \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t)\right)^H \boldsymbol{W}_i \left(\boldsymbol{x}_i - \gamma_{t,i} \boldsymbol{a}(\theta_t)\right). \tag{A.19}$$

Note that the form of Eq. (A.19) is similar to the form of Eq.(2.34) except that $\boldsymbol{R}_{\tilde{\boldsymbol{n}}}^{-1}$ in Eq.(2.34) is replaced by $\boldsymbol{W}_i$ in Eq.(A.19). From this point on, derivation changes according to $\gamma_{t,i}$ changing from snapshot to snapshot, and $\gamma_{t,i}$ being the same at all snapshots.

SOI Amplitude and Angle Estimation :

SOI Amplitudes Changing From Sample to Sample :

If we take derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ in Eq. (A.19) with respect to $\gamma_{t,i}^*$, we obtain

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})}{\partial \gamma_{t,i}^*} &= \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\left(\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}(\theta_t)\right) \\ &= \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{x}_i - \gamma_{t,i}\boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{a}(\theta_t) = 0. \end{aligned} \tag{A.20}$$

Then, $\hat{\gamma}_{t,i}$ can be found as

$$\hat{\gamma}_{t,i} = \frac{\boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{x}_i}{\boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{a}(\theta_t)}. \tag{A.21}$$

If $\hat{\gamma}_{t,i}$ is put in Eq. (A.19), $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ turns into the form:

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) &= constants - \sum_{i=1}^{L} \boldsymbol{x}_i^H\boldsymbol{W}_i\left(\boldsymbol{x}_i - \hat{\gamma}_{t,i}\boldsymbol{a}(\theta_t)\right) \\ &= constants + \sum_{i=1}^{L} \frac{\boldsymbol{a}(\theta_t)^H\boldsymbol{W}_i\boldsymbol{x}_i}{\boldsymbol{a}(\theta_t)^H\boldsymbol{W}_i\boldsymbol{a}(\theta_t)}\boldsymbol{x}_i^H\boldsymbol{W}_i\boldsymbol{a}(\theta_t). \end{aligned} \tag{A.22}$$

Maximizing the complete data log-likelihood function given in the above equation with respect to θ_t, results in the following estimator for SOI angle:

$$\hat{\theta}_t = \arg\max_{\theta} \sum_{i=1}^{L} \frac{\left|\boldsymbol{a}(\theta)^H\boldsymbol{W}_i\boldsymbol{x}_i\right|^2}{\boldsymbol{a}(\theta)^H\boldsymbol{W}_i\boldsymbol{a}(\theta)}. \tag{A.23}$$

SOI Amplitudes Staying Constant at All Samples :

If amplitudes of SOI component stay constant at every sample, complete data log-likelihood can be rewritten as

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) = constants - \sum_{i=1}^{L} \left(\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t)\right)^H \boldsymbol{W}_i\left(\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t)\right). \tag{A.24}$$

If we take derivative of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ with respect to γ_t^* and equating the result to zero, we obtain

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})}{\partial \gamma_t^*} &= \sum_{i=1}^{L} \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\left(\boldsymbol{x}_i - \gamma_t\boldsymbol{a}(\theta_t)\right) \\ &= \sum_{i=1}^{L} \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{x}_i - \gamma_t\sum_{i=1}^{L} \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{a}(\theta_t) = 0. \end{aligned} \tag{A.25}$$

Then, $\hat{\gamma}_t$ is found to be

$$\hat{\gamma}_t = \frac{\sum_{i=1}^{L} \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{x}_i}{\sum_{i=1}^{L} \boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{a}(\theta_t)} = \frac{\boldsymbol{a}^H(\theta_t)\sum_{i=1}^{L}\boldsymbol{W}_i\boldsymbol{x}_i}{\boldsymbol{a}^H(\theta_t)\left(\sum_{i=1}^{L}\boldsymbol{W}_i\right)\boldsymbol{a}(\theta_t)}. \tag{A.26}$$

If $\hat{\gamma}_t$ is put in Eq.(A.24), the following $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old})$ is obtained:

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{old}) &= constants - \sum_{i=1}^{L} \left(\boldsymbol{x}_i - \hat{\gamma}_t\boldsymbol{a}(\theta_t)\right)^H \boldsymbol{W}_i \left(\boldsymbol{x}_i - \hat{\gamma}_t\boldsymbol{a}(\theta_t)\right) \\ &= constants - \sum_{i=1}^{L} \boldsymbol{x}_i^H\boldsymbol{W}_i \left(\boldsymbol{x}_i - \hat{\gamma}_t\boldsymbol{a}(\theta_t)\right) \\ &= constants - \sum_{i=1}^{L} \boldsymbol{x}_i^H\boldsymbol{W}_i\boldsymbol{x}_i \\ &\qquad + \frac{\boldsymbol{a}^H(\theta_t)\sum_{i=1}^{L}\boldsymbol{W}_i\boldsymbol{x}_i}{\boldsymbol{a}^H(\theta_t)\left(\sum_{i=1}^{L}\boldsymbol{W}_i\right)\boldsymbol{a}(\theta_t)} \sum_{i=1}^{L} \boldsymbol{x}_i^H\boldsymbol{W}_i\boldsymbol{a}(\theta_t) \end{aligned} \tag{A.27}$$

Maximizing likelihood function given in Eq. (A.27) with respect to θ_t is equivalent to maximizing the last term, resulting for the following estimator for SOI angle:

$$\hat{\theta}_t = \arg\max_{\theta} \frac{\left|\boldsymbol{a}^H(\theta)\sum_{i=1}^{L}\boldsymbol{W}_i\boldsymbol{x}_i\right|^2}{\boldsymbol{a}^H(\theta)\left(\sum_{i=1}^{L}\boldsymbol{W}_i\right)\boldsymbol{a}(\theta)}. \tag{A.28}$$

This concludes the derivation.

Table A.1: Intermittent Jamming Method 2

Algorithm 6 Intermittent Jam. with Knowledge of Possible Jamming Power Levels

Initialization: Initialize $\boldsymbol{\theta} = \{\gamma_{t,1}, \cdots, \gamma_{t,L}, \theta_t, \alpha_1, \cdots, \alpha_M\}$

Expectation Phase:

Compute Eq.(A.9) :

$$M_{m,i}^{old} = \frac{\alpha_m^{old} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{old} \boldsymbol{a}(\theta_t^{old}), \boldsymbol{I} + s_m \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}{\sum_{k=1}^{M} \alpha_k^{old} \mathcal{CN}\left(\boldsymbol{x}_i; \gamma_{t,i}^{old} \boldsymbol{a}(\theta_t^{old}), \boldsymbol{I} + s_k \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right)}.$$

Compute Eq.(A.18) : $$\boldsymbol{W}_i = \sum_{m=1}^{M} M_{m,i}^{old} \left(\boldsymbol{I} - \frac{s_m}{1 + N s_m} \boldsymbol{a}(\theta_j)\boldsymbol{a}^H(\theta_j)\right).$$

Maximization Phase:

Compute Eq.(A.16) : $$\hat{\alpha}_m = \frac{1}{L} \sum_{i=1}^{L} M_{m,i}^{old}$$

if Signal model = non-coherent **then**

Compute Eq.(A.21) : $$\hat{\gamma}_{t,i} = \frac{\boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{x}_i}{\boldsymbol{a}^H(\theta_t)\boldsymbol{W}_i\boldsymbol{a}(\theta_t)}.$$

Compute Eq.(A.23) : $$\hat{\theta}_t = \arg\max_{\theta} \sum_{i=1}^{L} \frac{\left|\boldsymbol{a}(\theta)^H \boldsymbol{W}_i \boldsymbol{x}_i\right|^2}{\boldsymbol{a}(\theta)^H \boldsymbol{W}_i \boldsymbol{a}(\theta)}.$$

end if

if Signal model = coherent **then**

Compute Eq.(A.26) : $$\hat{\gamma}_t = \frac{\boldsymbol{a}^H(\theta_t) \sum_{i=1}^{L} \boldsymbol{W}_i \boldsymbol{x}_i}{\boldsymbol{a}^H(\theta_t) \left(\sum_{i=1}^{L} \boldsymbol{W}_i\right) \boldsymbol{a}(\theta_t)}.$$

Compute Eq.(A.28) : $$\hat{\theta}_t = \arg\max_{\theta} \frac{\left|\boldsymbol{a}^H(\theta) \sum_{i=1}^{L} \boldsymbol{W}_i \boldsymbol{x}_i\right|^2}{\boldsymbol{a}^H(\theta) \left(\sum_{i=1}^{L} \boldsymbol{W}_i\right) \boldsymbol{a}(\theta)}.$$

end if

Termination Check:

if the convergence criteria is satisfied **then**

return $\hat{\sigma}^2$, $\hat{\gamma}_{t,i}(\hat{\gamma}_t)$ and $\hat{\theta}_t$

else

Go to Expectation Phase

end if

REFERENCES

[1] S. M. Kay, *Fundamentals of Statistical Processing, Volume I: Estimation Theory*. Prentice Hall, 1993.

[2] F. Gini, R. Reggiannini, and U. Mengali, "The modified Cramer-Rao bound in vector parameter estimation," *IEEE Trans. Commun.*, vol. 46, pp. 52–60, Jan 1998.

[3] E. Aboutanios and B. Mulgrew, "Iterative frequency estimation by interpolation on Fourier coefficients," *IEEE Trans. Signal Process.*, vol. 53, pp. 1237–1242, Apr. 2005.

[4] Ç. Candan, "Fine resolution frequency estimation from three DFT samples: Case of windowed data," *Signal Processing*, vol. 114, pp. 245–250, Sept. 2015.

[5] R. A. Poisel, *Electronic Warfare Target Location Methods*. Canton Street, MA, USA: Artech House, 2 ed., 2012.

[6] R. Peterson, R. Ziemer, and D. Borth, *Introduction to Spread Spectrum Communications*. New Jersey, USA: Prentice-Hall, Inc., 2 ed., 1995.

[7] O. Besson, P. Stoica, and Y. Kamiya, "Direction finding in the presence of an intermittent interference," *IEEE Trans. Signal Process.*, vol. 50, pp. 1554–1564, July 2002.

[8] L. B. VanBrunt, *Applied ECM Volume 2*. USA: EW Engineering, Inc., 1 ed., 1982.

[9] M. A. Richards, *Fundamentals of Radar Signal Processing*. Two Penn Plaza, NY, USA: McGraw-Hill, 3 ed., 2022.

[10] Ş. Bilgi Akdemir and Ç. Candan, "Maximum-likelihood direction of arrival estimation under intermittent jamming," *Digital Signal Processing*, vol. 113, p. 103028, 2021.

[11] Ş. Bilgi Akdemir and Ç. Candan, "Study of angle of arrival estimation performance of MVDR and Capon methods under intermittent interference," in *Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, May 2018.

[12] E. Tuncer and B. Friedlander, *Classical and Modern Direction-of-Arrival Estimation*. 30 Corporate Drive, MA, USA: Academic Press, 1 ed., 2009.

[13] E. V. Stansfield, "Accuracy of an interferometer in noise," *Proc. IEE Radar Sonar Navig.*, vol. 143, pp. 217–226, Aug 1996.

[14] A. De Martino, *Introduction to Modern EW Systems*. Canton Street, MA, USA: Artech House, 2 ed., 2018.

[15] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, pp. 276–280, March 1986.

[16] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 984–995, July 1989.

[17] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 8, pp. 336–339, April 1983.

[18] B. D. Rao and K. V. S. Hari, "Performance analysis of Root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 1939–1949, Dec 1989.

[19] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 720–741, May 1989.

[20] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound: Further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 2140–2150, Dec 1990.

[21] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Trans. Signal Process.*, vol. 39, pp. 446–453, Feb 1991.

[22] B. Porat and B. Friedlander, "Analysis of the asymptotic relative efficiency of the MUSIC algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, pp. 532–544, April 1988.

[23] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, pp. 67–94, July 1996.

[24] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Prentice Hall, 2005.

[25] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 1081–1089, October 1986.

[26] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1132–1143, July 1990.

[27] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, pp. 1783–1795, Oct 1990.

[28] J. Sheinvald, M. Wax, and A. J. Weiss, "On maximum-likelihood localization of coherent signals," *IEEE Trans. Signal Process.*, vol. 44, pp. 2475–2482, Oct 1996.

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

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

[31] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. Spring Street, NY, USA: Springer Science+Business Media, LLC, 1 ed., 2008.

[32] J. Yang, Y. Yang, B. Lei, J. Lu, and L. Yang, "Nonuniform linear array DOA

estimation using EM criterion," *Digital Signal Processing*, vol. 86, pp. 36 – 41, 2019.

[33] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, 2005.

[34] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, 2006.

[35] M. M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4646–4655, 2010.

[36] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing, Volume 7* (R. Chellappa and S. Theodoridis, eds.), ch. 11, pp. 509–581, Berlin, Heidelberg: Academic Press, 2017.

[37] J. Zhang, T. Qiu, and S. Luan, "An efficient real-valued sparse Bayesian learning for non-circular signal's DOA estimation in the presence of impulsive noise," *Digital Signal Processing*, vol. 106, p. 102838, 2020.

[38] Q. Liu, H. C. So, and Y. Gu, "Off-grid DOA estimation with nonconvex regularization via joint sparse representation," *Signal Processing*, vol. 140, pp. 171 – 176, 2017.

[39] L. Hu, Z. Shi, J. Zhou, and Q. Fu, "Compressed sensing of complex sinusoids: An approach based on dictionary refinement," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3809–3822, 2012.

[40] Z. Liu, Z. Huang, and Y. Zhou, "An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 1–11, 2012.

[41] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[42] A. Das and T. J. Sejnowski, "Narrowband and wideband off-grid direction-of-arrival estimation via sparse Bayesian Learning," *IEEE J. Ocean. Eng.*, vol. 43, no. 1, pp. 108–118, 2018.

[43] M. Carlin, P. Rocca, G. Oliveri, F. Viani, and A. Massa, "Directions-of-arrival estimation through Bayesian compressive sensing strategies," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3828–3838, 2013.

[44] Z. Liu, Z. Huang, and Y. Zhou, "Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 1–12, 2013.

[45] Z.-M. Liu, L. Zheng, D.-W. Feng, and Z.-T. Huang, "Direction-of-arrival estimation for coherent sources via sparse Bayesian learning," *International Journal of Antennas and Propagation*, vol. 2014, pp. 1–8, 04 2014.

[46] A. Das, "Deterministic and Bayesian sparse signal processing algorithms for coherent multipath directions-of-arrival (DOAs) estimation," *IEEE J. Ocean. Eng.*, vol. 44, no. 4, pp. 1150–1164, 2019.

[47] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, 2013.

[48] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[49] A. Das, "Theoretical and experimental comparison of off-grid sparse Bayesian direction-of-arrival estimation algorithms," *IEEE Access*, vol. 5, pp. 18075–18087, 2017.

[50] P. Chen, Z. Cao, Z. Chen, and C. Yu, "Sparse off-grid DOA estimation method with unknown mutual coupling effect," *Digital Signal Processing*, vol. 90, pp. 1–9, 2019.

[51] A. Das, W. S. Hodgkiss, and P. Gerstoft, "Coherent multipath direction-of-arrival resolution using compressed sensing," *IEEE J. Ocean. Eng.*, vol. 42, no. 2, pp. 494–505, 2017.

[52] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[53] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, 2013.

[54] G. Tang, B. N. Bhaskar, and B. Recht, "Near minimax line spectral estimation," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 499–512, 2015.

[55] Z. Yang and L. Xie, "Exact joint sparse frequency recovery via optimization methods," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5145–5157, 2016.

[56] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, 2016.

[57] Y. Chi and M. Ferreira Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Process. Mag.*, vol. 37, no. 2, pp. 39–57, 2020.

[58] F. Sellone and A. Serra, "A Novel Online Mutual Coupling Compensation Algorithm for Uniform and Linear Arrays," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 560–573, 2007.

[59] Z. Ye and C. Liu, "On the resiliency of MUSIC direction finding against antenna sensor coupling," *IEEE Trans. Antennas Propag.*, vol. 56, no. 2, pp. 371–380, 2008.

[60] Youming Li and M. H. Er, "Theoretical analyses of gain and phase error calibration with optimal implementation for linear equispaced array," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 712–723, 2006.

[61] A. Liu, G. Liao, C. Zeng, Z. Yang, and Q. Xu, "An eigenstructure method for estimating DOA and sensor gain-phase errors," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5944–5956, 2011.

[62] Z. Liu and Y. Zhou, "A unified framework and sparse bayesian perspective for

direction-of-arrival estimation in the presence of array imperfections," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3786–3798, 2013.

[63] E. D. di Claudio and R. Parisi, "WAVES: weighted average of signal subspaces for robust wideband direction finding," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2179–2191, 2001.

[64] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 823–831, 1985.

[65] Yeo-Sun Yoon, L. M. Kaplan, and J. H. McClellan, "TOPS: New DOA estimator for wideband signals," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 1977–1989, 2006.

[66] Z. Liu, Z. Huang, and Y. Zhou, "Direction-of-arrival estimation of wideband signals via covariance matrix sparse representation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4256–4270, 2011.

[67] A. D. Seifer, "Monopulse-radar angle tracking in noise or noise jamming," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 622–638, 1992.

[68] W. D. Blair, G. A. Watson, T. Kirubarajan, and Y. Bar-Shalom, "Benchmark for radar allocation and tracking in ECM," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 4, pp. 1097–1114, 1998.

[69] Kai-Bor Yu and D. J. Murrow, "Adaptive digital beamforming for angle estimation in jamming," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, no. 2, pp. 508–523, 2001.

[70] D. J. Rabideau, "Clutter and jammer multipath cancellation in airborne adaptive radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 565–583, 2000.

[71] F. Bandiera, A. Farina, D. Orlando, and G. Ricci, "Detection algorithms to discriminate between radar targets and ECM signals," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 5984–5993, 2010.

[72] A. De Maio and D. Orlando, "Adaptive radar detection of a subspace signal embedded in subspace structured plus Gaussian interference via invariance," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2156–2167, 2016.

[73] M. Umehira, T. Okuda, X. Wang, S. Takeda, and H. Kuroda, "An adaptive interference detection and suppression scheme using iterative processing for automotive FMCW radars," in *2020 IEEE Radar Conference (RadarConf20)*, pp. 1–5, 2020.

[74] A. Gomaa and N. Al-Dhahir, "A sparsity-aware approach for NBI estimation in MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1854–1862, 2011.

[75] M. Spuhler, D. Giustiniano, V. Lenders, M. Wilhelm, and J. B. Schmitt, "Detection of reactive jamming in DSSS-based wireless communications," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1593–1603, 2014.

[76] S. Liu, F. Yang, J. Song, and Z. Han, "Block sparse Bayesian learning-based NB-IoT interference elimination in LTE-advanced systems," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4559–4571, 2017.

[77] S. Chandran, *Advances in Direction-of-Arrival Estimation*. Artech, 2005.

[78] S. E. Lipsky, *Microwave Passive Direction Finding*. Raleigh, NC 27613, USA: SciTech Publishing, Inc., 1 ed., 2004.

[79] M. Richards, J. Scheer, and W. Holm, *Principles of Modern Radar*. No. 3. c. in Principles of Modern Radar, SciTech Pub., 2010.

[80] B. J. Slocumb, "Surveillance radar range-bearing centroid processing," in *Signal and Data Processing of Small Targets 2001* (O. E. Drummond, ed.), vol. 4473, pp. 74 – 85, International Society for Optics and Photonics, SPIE, 2001.

[81] B. J. Slocumb and D. L. Macumber, "Surveillance radar range-bearing centroid processing, part II: merged measurements," in *Signal and Data Processing of Small Targets 2006* (O. E. Drummond, ed.), vol. 6236, p. 623604, International Society for Optics and Photonics, SPIE, 2006.

[82] P. Swerling, "Maximum angular accuracy of a pulsed search radar," *Proceedings of the IRE*, vol. 44, no. 9, pp. 1146–1155, 1956.

[83] M. A. Richards, *Introduction to Radar Systems*. Boston, USA: McGraw-Hill, 3 ed., 2006.

[84] D. Barton and S. Sherman, *Monopulse Principles and Techniques, Second Edition*. 2011.

[85] L. Yang, M. R. McKay, and R. Couillet, "High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models," *IEEE Transactions on Signal Processing*, vol. 66, no. 7, pp. 1933–1947, 2018.

[86] B. P. Gibbs, *Advanced Kalman Filtering, Least Squares and Modeling*. John Wiley and Sons, Ltd, 2011.

[87] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar Array Processing* (S. Haykin, J. Litva, and T. J. Shepherd, eds.), pp. 99–151, Berlin, Heidelberg: Springer, 1993.

[88] S. V. Schell and W. A. Gardner, "High-resolution direction finding," in *Signal Processing and its Applications*, vol. 10 of *Handbook of Statistics*, pp. 755–817, Elsevier, 1993.

[89] C. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1 ed., 2007.

[90] F. Gini and R. Reggiannini, "On the use of Cramer-Rao-like bounds in the presence of random nuisance parameters," *IEEE Transactions on Communications*, vol. 48, pp. 2120–2126, Oct. 2000.

[91] F. Gini, "A radar application of a modified Cramer-Rao bound: Parameter estimation in non-Gaussian clutter," *IEEE Trans. Signal Process.*, vol. 46, pp. 1945–1953, July 1998.

[92] Ç. Candan, "A method for fine resolution frequency estimation from three DFT samples," *IEEE Signal Process. Lett.*, vol. 18, pp. 351–354, June 2011.

[93] M. Morelli, M. Moretti, and A. A. D'Amico, "Single-tone frequency estimation by weighted least-squares interpolation of Fourier coefficients," *IEEE Trans. Commun.*, vol. 70, pp. 526–537, Jan. 2022.

[94] M. Wang, Q. Wang, K. Han, and X. Jia, "Velocity estimation by two sample dft interpolation and weighted least square fusion for dual-frequency radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–11, 2023.

[95] Ç. Candan, "Analysis and further improvement of fine resolution frequency estimation method from three DFT samples," *IEEE Signal Processing Letters*, vol. 20, no. 9, pp. 913–916, 2013.

[96] M. Macleod, "Fast nearly ML estimation of the parameters of real or complex single tones or resolved multiple tones," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 141–148, 1998.

[97] P. D. Viesti, A. Davoli, G. Guerzoni, and G. M. Vitetta, "Recursive algorithms for the estimation of multiple superimposed undamped tones and their application to radar systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 2, pp. 1834–1853, 2023.

[98] A. Serbes, "Fast and efficient estimation of frequencies," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 4054–4066, 2021.

[99] E. Aboutanios, "A modified dichotomous search frequency estimator," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 186–188, 2004.

[100] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inf. Theory*, vol. 15, pp. 457–462, July 1969.

[101] M. Hadidi and S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. Automat. Contr.*, vol. 24, no. 6, pp. 944–948, 1979.

[102] A. Logothetis and V. Krishnamurthy, "Expectation maximization algorithms for MAP estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2139–2156, 1999.

[103] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1453–1464, 2004.

[104] J. Kullaa, "Detection, identification, and quantification of sensor fault in a sensor network," *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 208–221, 2013.

[105] D. Barber and A. T. Cemgil, "Graphical models for time-series," *IEEE Signal Processing Magazine*, vol. 27, no. 6, pp. 18–28, 2010.

[106] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[107] V. Venkatesh and S. Kay, "Cramer-Rao lower bounds for complex parameters." https://www.ele.uri.edu/faculty/kay/New%20web/downloadable%20files/Nageha_complex%20CRLB.pdf.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Bilgi Akdemir, Şafak
Foreign Languages : English, German, Russian

EDUCATION

Degree	Institution	Year of Graduation
M.Sc.	Middle East Technical University	2010
B.Sc.	Middle East Technical University	2003
High School	İzmir Fen Lisesi	1999

WORK EXPERIENCE

Year	Company	Title
2006-Present	TÜBİTAK BİLGEM İLTAREN	Senior Chief Expert Researcher
2003-2006	MİKES	Digital Design Engineer

PUBLICATIONS

International Journal Publications

1. Ş. Bilgi Akdemir and Ç. Candan, "Maximum-likelihood direction of arrival estimation under intermittent jamming", Elsevier Digital Signal Processing, Volume 113, 2021, 103028, ISSN 1051-2004, https://doi.org/10.1016/j.dsp.2021.103028.

International Conference Publications

1. E. Öztürk and Ş. Bilgi Akdemir, “A network based test interface allowing access of hardware developers to test equipment”, IEEE Autotestcon, Oct. 2017, doi: 10.1109/AUTEST.2017.8080507

2. M.B. Kocamış, H. Abacı, Ş. Bilgi Akdemir, S. Varma and A. Yıldırım “Deinterleaving for radar warning receivers with missed pulse consideration”, Proc. IEEE 13. European Radar Conference, Oct. 2016

National Conference Publications

1. O. Güneş and Ş. Bilgi Akdemir, “CA-CFAR based target detection in FMCW radars”, IEEE 31th Signal Processing and Communications Applications Conference (SIU), 2023, doi: 10.1109/SIU59756.2023.10224049

2. Ş. Bilgi Akdemir and Ç. Candan, "Study of angle of arrival estimation performance of MVDR and Capon methods under intermittent interference," IEEE 26th Signal Processing and Communications Applications Conference (SIU), 2018, doi: 10.1109/SIU.2018.8404731

3. O. Torun, M. B. Kocamış, H. Abacı, Ş. Bilgi Akdemir, S. Varma and A. Yıldırım, “Deinterleaving of radar signals with stagger PRI and dwell-switch PRI types”, IEEE 25th Signal Processing and Communications Applications Conference (SIU), 2017, doi: 10.1109/SIU.2017.7960383

4. M. B. Kocamış, O. Torun, H. Abacı, Ş. Bilgi Akdemir, S. Varma and A. Yıldırım, “Deinterleaving for radars with jitter and agile pulse repetition interval”, IEEE 25th Signal Processing and Communications Applications Conference (SIU), 2017, doi: 10.1109/SIU.2017.7960412

5. M.B. Kocamis, H. Abacı, Ş. Bilgi Akdemir, S. Varma and A. Yıldırım “Deinterleaving for radar warning receivers”, IEEE 24th Signal Processing and Communications Applications Conference (SIU), 2016,
doi: 10.1109/SIU.2016.7496033

6. M.B. Kocamış, M.E. Mıhçıoğlu, Ş.Bilgi Akdemir, S. Varma and A. Yıldırım, "A clustering approach for radar warning receivers", IEEE 23th Signal Processing and Communications Applications Conference (SIU), 2015, doi: 10.1109/SIU.2015.7130396

7. Ş. Bilgi Akdemir, A. Orduyılmaz, A. C. Gürbüz and A.Yıldırım, "Analysis of frequency modulated continuous wave signals using time-frequency domain shape features", IEEE 22th Signal Processing and Communications Applications Conference (SIU), 2014, doi: 10.1109/SIU.2014.6830595

8. Ş. Bilgi Akdemir, Ç. Candan, "Mismatched filter design in MIMO radar", IEEE 20th Signal Processing and Communications Applications Conference (SIU), 2012, doi: 10.1109/SIU.2012.6204508

9. Ş. Bilgi Akdemir, Ç. Candan, "Moving target detection with MIMO radar", IEEE 19th Signal Processing and Communications Applications Conference (SIU), 2011, doi: 10.1109/SIU.2011.5929598