



Array Processing Techniques for Direction of Arrival Estimation, Communications, and Localization in Vehicular and Wireless Sensor Networks

Marco Antonio Marques Marinho

Supervisors:
Alexey Vinel
João Paulo Carvalho Lustosa da Costa
Felix Antreich

Array Processing Techniques for Direction of Arrival Estimation,
Communications, and Localization in Vehicular and Wireless Sensor Networks
© Marco Antonio Marques Marinho
Halmstad University Dissertations no. 43
ISBN 978-91-87045-88-2 (printed)
ISBN 978-91-87045-89-9 (pdf)
Publisher: Halmstad University Press, 2018 | www.hh.se/hup

Aos meus pais:

Marco Antonio e Marcelita, tudo que há de bom em mim nasceu de vocês. Aprendi muito ao longo desses anos de doutorado longe de vocês, mas a maior lição que aprendi é vocês são as pessoas mais importantes da minha vida. Amo vocês acima de tudo.

Ao meu irmão:

Murilo, companheiro ao longo dessa jornada que escolhemos. Sem nossas conversas e discussões sobre a vida de um doutorando eu provavelmente teria perdido a batalha mais importante dessa guerra, a de mater-me são e sensato. Obrigado por tudo.

A minha noiva:

Stephanie, minha parceira e amiga. A distância ao longo desses anos não foi capaz de diminuir seu amor, companheirismo e carinho. Obrigado por sua dedicação e paciência nos momentos mais difíceis. Sua força em frente as dificuldades foi uma grande fonte de inspiração. Te amo.

Marco Antonio Marques Marinho

Abstract

Array signal processing in wireless communication has been a topic of interest in research for over three decades. In the fourth generation (4G) of the wireless communication systems, also known as Long Term Evolution (LTE), multi antenna systems have been adopted according to the Release 9 of the 3rd Generation Partnership Project (3GPP). For the fifth generation (5G) of the wireless communication systems, hundreds of antennas should be incorporated to the devices in a massive multi-user Multiple Input Multiple Output (MIMO) architecture. The presence of multiple antennas provides array gain, diversity gain, spatial gain, and interference reduction. Furthermore, arrays enable spatial filtering and parameter estimation, which can be used to help solve problems that could not previously be addressed from a signal processing perspective. The aim of this thesis is to bridge some gaps between signal processing theory and real world applications. Array processing techniques traditionally assume an ideal array. Therefore, in order to exploit such techniques, a robust set of methods for array interpolation are fundamental and are developed in this work. In this dissertation, novel methods for array interpolation are presented and their performance in real world scenarios is evaluated. Problems in the field of wireless sensor networks and vehicular networks are also addressed from an array signal processing perspective. Signal processing concepts are implemented in the context of a wireless sensor network. These concepts provide a level of synchronization sufficient for distributed multi antenna communication to be applied, resulting in improved lifetime and improved overall network behaviour. Array signal processing methods are proposed to solve the problem of radio based localization in vehicular network scenarios with applications in road safety and pedestrian protection.

Acknowledgments

To Prof. Felix Antreich, my friend, mentor, and co-supervisor. For helping me out through the rough patches along the way, not just academically, but also when got the short end of the stick after physical encounters with motor vehicles. Without his support and guidance I would not have arrived at this destination. It is a great honor to work and research alongside him.

To my supervisor Prof. João Paulo Carvalho Lustosa da Costa, to whom I own so much, for the endless support throughout my entire academic career. I am thankful for his patience, trust and for the endless opportunities he has provided me during my academic career. It is thanks to him I have gotten to experience so much of the world and had the opportunity to know so many interesting fields of research.

To my supervisor Prof. Alexey Vinel, who deposited so much trust in me by accepting me as a student in Halmstad. He always did his best and dedicated so much time to ensure I could apply my knowledge to interesting and relevant topics during my stay in Sweden. I appreciate all his effort to make my Ph.D. experience productive and stimulating. Without his help many interesting results would never have come to be.

To Prof. Edison Pignaton de Freitas for his support and incentive across my academic career. He was always able to keep a good humour even when I had lost mine. His seamlessly endless ability to moderate conflict and connect people have help me throughout this journey and serve as a great inspiration.

To Prof. Fredrik Tufvesson for the time dedicated to following my work and for all the suggestions and guidance offered.

Contents

Introduction	1
1 Array Interpolation	5
1.1 Overview and Contribution	5
1.2 Motivation	7
1.3 Data Model	11
1.4 Preliminaries	13
1.4.1 Forward Backward Averaging (FBA)	13
1.4.2 Spatial Smoothing (SPS)	14
1.4.3 Model Order Selection	14
1.4.4 Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT)	15
1.4.5 Vandermonde Invariance Transformation (VIT)	17
1.5 Array Interpolation	18
1.6 Classical Interpolation	19
1.7 Sector Selection and Discretization	23
1.7.1 UT Discretization	26
1.7.2 Principal Component Discretization	28
1.8 Linear Adaptive Array Interpolation	29
1.8.1 Data Transformation and Model Order Selection	32
1.8.2 Linear UT Array Interpolation	34
1.9 Multidimensional Linear Interpolation	35
1.9.1 Tensor Algebra Concepts	35
1.9.2 Multidimensional Data Model	36
1.9.3 Multidimensional Interpolation	37
1.10 Nonlinear Array Interpolation	38
1.10.1 MARS based interpolation	39
1.10.2 GRNN based interpolation	40
1.11 Numerical Simulation Results	41
1.11.1 Multidimensional Linear Performance Results	41
1.11.2 Nonlinear Performance Results	44
1.12 Summary	49

2 Cooperative MIMO for Wireless Sensor Networks	51
2.1 Overview and Contribution	51
2.2 Motivation	52
2.3 Wireless Sensor Networks Organization	54
2.4 Cooperative MIMO	55
2.5 Energy Analysis	58
2.5.1 Conventional Techniques	59
2.5.2 Cooperative MIMO	61
2.6 Synchronization	64
2.6.1 Effects of Synchronization Error on Cooperative MIMO	64
2.6.2 Proposed Coarse Synchronization Scheme	65
2.6.3 Proposed Fine Synchronization Schemes	66
2.6.4 Synchronization error propagation	70
2.7 Adaptive C-MIMO Clustering	73
2.7.1 Adaptive C-MIMO clustering	74
2.7.2 Numerical Simulations	78
2.8 Summary	85
3 Array Processing Localization for Vehicular Networks	87
3.1 Overview and Contribution	87
3.2 Motivation	88
3.3 Data Model	90
3.4 Space-Alternating Generalized Expectation Maximization (SAGE) Algorithm	91
3.5 Scenario Description	92
3.6 Array Processing Localization	93
3.6.1 Flip-Flop Estimation	93
3.6.2 Joint Direct Position Estimation	96
3.6.3 DOA only estimation	97
3.6.4 Applicability of DOA Estimation for Positioning	98
3.7 Three Dimensional DOA Based Estimation	99
3.7.1 Scenario Description	99
3.7.2 Definition of the Attitude Angles	99
3.7.3 Direction vector computation	100
3.7.4 Position estimation	100
3.8 Simulation Results	101
3.8.1 Results for Simulated Data	101
3.8.2 Results for Real Data	104
3.9 Summary	107
Conclusion	109
Author's Publications	113
References	117

Appendices	131
A Left Centro-Hermitian Matrix	133
B Forward Backward Averaging and Spatial Smoothing	135
C Proof of Theorem 1	139
D Proof of Theorem 2	141
E Multivariate Adaptive Regression Splines (MARS)	143
F Generalized Regression Neural Networks (GRNNs)	145
G Sensor Node Localization	147
H TRIAD Algorithm	151

List of Figures

1.1	Signal with DOA θ impinging on a uniform linear array (ULA), whose the antenna space is Δ	8
1.2	On the left side there is real data from an antenna array, while on the right side the data transformed by the transformation matrix \mathbf{B}	9
1.3	Graphical representation of a linear array	12
1.4	Example of SPS subarrays	14
1.5	Eigenvalue profile: eigenvalue index versus eingenvale	15
1.6	$P(\theta, \phi)$	24
1.7	Selected sectors and example of sector bounds	25
1.8	Selected sectors and respective bounds for one-dimensional case	26
1.9	UT of the approximated $l_c(\mathbf{X}; \boldsymbol{\theta}, \phi)$ for two sources	28
1.10	Example of transformed regions	30
1.11	Transformation error with respect to combined sector size	31
1.12	Example of estimated model order versus number of subarrays	33
1.13	Tensor $\mathcal{B} \in \mathbb{R}^{2 \times 2 \times 2}$	35
1.14	Results for a standard deviation of $\frac{\pi}{20}$ radians	42
1.15	Results for a standard deviation of $\frac{\pi}{10}$ radians	43
1.16	Results for a standard deviation of $\frac{\pi}{5}$ radians	43
1.17	6x1 antenna array: ANSYS HFSS model (left), prototype (right).	45
1.18	DOA estimation performance with 0.4λ element separation	46
1.19	DOA estimation performance with 0.2λ element separation	46
1.20	DOA estimation performance for model built with simulated data with 0.4λ element separation	47
1.21	DOA estimation performance for model built with simulated data with 0.2λ element separation	48
1.22	TLS-ESPRIT failure rate for 0.2λ element separation	48
2.1	Examples of peer-to-peer and cluster networks	55
2.2	Example of cooperative MIMO and SISO communications in a wireless sensor network	56
2.3	C-MIMO communication steps	56

2.4	Performance comparison between standard SISO systems and 2×2 MIMO systems using ZF, MMSE and ML equalization	58
2.5	Examples of single-hop and multi-hop communication	59
2.6	Performance comparison of equalization methods in the presence of synchronization errors in a 2×2 MIMO system	65
2.7	Performance comparison of equalization methods in the presence of synchronization errors in a 4×4 MIMO system	66
2.8	Performance comparison of equalization methods in the presence of synchronization errors in a 7×7 MIMO system	67
2.9	Performance of ML synchronization, $f = 20$ MHz, $B = 22$ MHz	69
2.10	Performance of proposed DFT based synchronization, $f = 20$ MHz, $B = 22$ MHz	70
2.11	Histogram of synchronization error using the ML method (2.22,2.24)	71
2.12	Histogram of synchronization error using the DFT method (2.25)	71
2.13	Comparison of individual node energy depletion using C-MIMO or multihop only	80
2.14	Comparison of individual node energy depletion using multihop only after 2500 seconds	82
2.15	Comparison of individual node energy depletion using C-MIMO after 2500 seconds	82
2.16	Comparison of individual node energy depletion using multihop only after 4500 seconds	82
2.17	Comparison of individual node energy depletion using C-MIMO after 4500 seconds	82
2.18	Comparison of individual node energy depletion using multihop only after 5500 seconds	82
2.19	Comparison of individual node energy depletion using C-MIMO after 5500 seconds	82
2.20	Example of WSN areas with different energetic demands	83
2.21	Mean available energy at the center region of the network	83
2.22	Mean available energy at the edges of the network	84
2.23	Standard deviation of available energy at the center region of the network	84
2.24	Standard deviation of available energy at the edges of the network	85
3.1	Scenario description for two vehicles	93
3.2	Block diagram for the proposed flip-flop method	95
3.3	System model composed of a URA at the base station and an antenna array at the UAV.	99
3.4	Definitions of pitch α , yaw β and roll γ	100
3.5	Position estimation error vs SNR	102
3.6	Position estimation error vs distance from source	103
3.7	Position estimation error vs number of NLOS components	104
3.8	Position estimation error vs K-factor	104

3.9	Position estimation error over the x- and y-axes	105
3.10	Antenna array and an overview of its placement	105
3.11	Patio where measurements were performed and map of user distribution	106
3.12	Position estimation error for measured data using the proposed joint estimation	107
3.13	Position estimation error for measured data using the proposed flip-flop estimation	107
F.1	GRNN as a parallel neural network	146
G.1	Polarization ellipse	147

List of Tables

1.1	Summary of DOA estimation methods, dividing into three families: maximum likelihood (ML), beamformers, and subspace methods	9
2.1	Energy consumption for the Mica2 platform	60
2.2	Transmission energy consumption	64
2.3	Total network energy consumption	79
2.4	Undeliverable packets	79

Introduction

This work aims to make possible the application of signal processing techniques in scenarios that were previously unsuitable for it. To this end, the topic of array interpolation is addressed, allowing imperfect arrays to be used for array processing. Furthermore, in this dissertation, array signal processing techniques are proposed for wireless sensor networks and vehicular networks.

The following research questions are addressed in this work:

1. How to estimate the directions of arrival of correlated sources that are impinging on an antenna array that does not possess the necessary characteristics (geometry or electromagnetic response) for doing so?
2. Can array processing techniques be used to provide improved lifetime and communication performance in wireless sensor network?
3. Can array processing techniques be used to reliably estimate position of vehicles and pedestrians in vehicular scenarios based on their radio transmissions?

While the overall goal of this thesis is to apply array processing to practical scenarios, to allow the reader more freedom each research question is discussed on an individual chapter. Each chapter is presented as self-contained as possible. Therefore, the reader may read Chapter 1, Chapter 2, Chapter 3 independently. At the beginning of each chapter an overview, listing the contents of the chapter to allow the reader to skip to sections of interest. Together with the overview, a detailed list of the contributions is presented. Furthermore, each chapter is given its own section providing a brief summary of the discussion contained within, presented in Section 1.12, Section 2.8, and Section 3.9, respectively.

The first question is addressed in Chapter 1, where robust array interpolation methods are proposed. By applying array interpolation, data with arbitrary geometries or responses can be transformed into a tractable mathematical model. The problem of interpolation is addressed from the ground up, and problems, such as the array discretization, are tackled. Linear and nonlinear interpolation approaches are shown, providing a wide spectrum of alternatives that can be used according to the desired accuracy and computational complexity of the application. By using numerical simulations and

measurements, this work shows that the proposed methods outperform the state-of-the-art approaches. Section 1.2 presents a detailed background, and motivation on array interpolation. The six major contributions presented in this chapter are:

1. A signal adaptive approach for sector detection. Avoiding the pitfalls of having to deal with multiple sectors or the bias introduced by transforming the entire field of view of the array.
2. Two well defined methods for sector discretization aiming to preserve the statistical integrity of the transformed regions while minimizing the dimensionality of the data.
3. A linear approach for interpolation that aims to maximize the transformed signal regions while minimizing the transformation error, minimizing the parameter estimation bias.
4. A nonlinear extension for the linear approach, allowing the application of the proposed linear method to arrays with an arbitrary number of dimensions.
5. Two nonlinear array interpolation methods. The nonlinear approaches are suitable for dealing with arrays with highly distorted responses or improper geometries.
6. The performance of the proposed methods is tested using real antenna measurements, validating the proposed methods under real scenarios.

Chapter 2 deals with the second research question. The developed array interpolation tools can be employed to distributed antenna systems, such as wireless sensor networks (WSNs). While interpolated clusters can perform beamforming and detect radio signal sources, a more critical problem in WSNs is the energy efficiency, specially with respect to communication. Another antenna array technique, the multiple-input multiple-output (MIMO) communication technique can be used to achieve improved energy efficiency. However, to this end, it is necessary to achieve a very high degree of synchronization. The problem of synchronization is tackled, and its effects on the performance of MIMO systems are assessed in detail. The energy performance of distributed MIMO systems is analysed, and a threshold for its application is derived. The energetic performance of a network employing cooperative MIMO is shown considerably outperforms state-of-the-art single-hop and multi-hop methods. A detailed introduction on the topic of cooperative MIMO in WSNs is presented in Section 2.2. The five major contributions presented in this chapter are:

1. A study on the energy efficiency of C-MIMO with respect to multi and single-hop communications, defining a threshold of selection between the multiple communication techniques.

2. An analysis on the impact of synchronization errors in distributed MIMO system with respect to the symbol period and the size of the communication clusters involved.
3. A study on the propagation of synchronization errors inside a WSN, and from that the derivation of a closed form expression regarding the resynchronization interval in distributed MIMO systems that takes into account the symbol duration and the clock drift of the nodes involved.
4. Two methods for synchronization for WSNs with different trade-offs between performance and computational complexity.
5. An adaptive method for selecting the cluster size of C-MIMO transmissions capable of maximizing energy efficiency and improve the homogeneity of the energy reserves across the network.

The third question is studied in Chapter 3, where the problem of localization in the context of vehicular networks is tackled. A broadband data model considering the spherical curvature of the received wavefront is presented. This data model is leveraged to allow the estimation of the position of a transmitter using two separate antenna arrays. The problem is parametrized to allow the direct estimation of the coordinates of a transmitter by jointly processing data from both subarrays. Alternative localization approaches with reduced computational load based on estimating the DOA and range of a transmitter are also proposed. By using numerical simulations and measurements, this work shows that the proposed methods outperform the state-of-the-art approaches, achieving a sub-meter accuracy under certain conditions. A detailed introduction to the problem of localization in vehicular scenarios can be found in Section 3.2, while Section 3.1 presents an overview of the chapter and its research contributions. The four major contributions presented in this chapter are:

1. A DOA only localization method capable of localizing transmitters near to a vehicle at a low computational cost.
2. A localization method based on range and DOA estimations performed at independent subarrays. The geometrical relationship between the individual estimates is used to speed up the rate of convergence of the SAGE algorithm on the individual subarrays.
3. A novel parametrization for the received waveform data model allowing joint localization estimation to be performed at both subarrays simultaneously.
4. Localization methods that do not require the transmission of specific location data that can be used to mitigate spoofing attacks in VANETs.

Chapter 1

Array Interpolation

In this chapter the following research question is addressed:

How to estimate the directions of arrival of correlated sources that are impinging on an antenna array that does not possess the necessary characteristics (geometry or electromagnetic response) for doing so?

1.1 Overview and Contribution

The remainder of this section is organized as follows:

- **Motivation:** The problem of direction of arrival estimation is discussed and a detailed introduction on the subject of array interpolation presented along with the challenges this work aims to tackle.
- **Data Model:** The spatial data model of a signal received at an antenna array is defined.
- **Preliminaries:** Here important concepts (spatial smoothing, forward backward average, Vandermonde invariant transformation, model order selection, and Estimation of Signal Parameters via Rotational Invariance Techniques) used throughout the remainder of this chapter are presented in detail.
- **Array Interpolation:** The problem of array interpolation is defined and a distinction between linear and nonlinear interpolation is made.
- **Classical Interpolation:** Here the prior methods available in the literature for array interpolation are briefly detailed.
- **Sector Detection and Discretization:** A signal adaptive approach that allows sectors of interest to be detected based on the received signal power is presented. Furthermore, the problem of discretizing these

sectors is addressed with two different approaches introduced, employing the unscented transform (UT) and principal component analysis (PCA) respectively.

- **Linear Interpolation:** This section presents a linear interpolation approach that aims to minimize the transformation bias while transforming as much as possible of the received signal. A calculation of a transform matrix employing the UTs parameters is discussed. The problem of selecting the proper smoothing length for proper signal decorrelation is discussed.
- **Multidimensional Interpolation:** Here, a multidimensional extension of the linear approach is presented. To derive this extension a multidimensional data model is presented and important tensor algebra concepts are detailed.
- **Nonlinear Interpolation:** This section presents two approaches for performing array interpolation in a nonlinear fashion. A general regression neural network (GRNN) and a multivariate adaptive regression splines (MARS) based approach are detailed.
- **Simulation Results:** Finally, the performance of the proposed methods is shown by means of numerical simulations. A set of such simulations is performed using a real measured response from an antenna array in order to validate the proposed methods using real data.

The research contributions presented in this chapter are:

1. A signal adaptive approach for sector detection. Avoiding the pitfalls of having to deal with multiple sectors or the bias introduced by transforming the entire field of view of the array.
2. Two well defined methods for sector discretization aiming to preserve the statistical integrity of the transformed regions while minimizing the dimensionality of the data.
3. A linear approach for interpolation that aims to maximize the transformed signal regions while minimizing the transformation error, minimizing the parameter estimation bias.
4. A nonlinear extension for the linear approach, allowing the application of the proposed linear method to arrays with an arbitrary number of dimensions.
5. Two nonlinear array interpolation methods. The nonlinear approaches are suitable for dealing with arrays with highly distorted responses or improper geometries.
6. The performance of the proposed methods is tested using real antenna measurements, validating the proposed methods under real scenarios.

1.2 Motivation

Sensor arrays have been applied in various fields of modern electrical engineering [59]. From up and coming applications such as massive Multiple Input Multiple Output (MIMO) [62] to microphone arrays capable, for instance, of detecting and separating multiple sound sources or playing individual songs depending on the seat of a car [105]. The mathematical tools provided by array signal processing together with the evolution of electronic circuitry have made such applications a reality. The problem of parameter estimation is of major interest in the field of array signal processing. Among the parameters that can be estimated, the direction of arrival (DOA) of an electromagnetic wave received at the array has experienced significant attention. The parameters to be estimated are those related to the received wave signals, spatial parameters such as direction of arrival (DOA) and direction of departure (DOD), temporal parameters such as time delay of arrival (TDOA). Moreover, frequency parameters such as Doppler shift can be estimated by fusing together the data received at various antennas and using a preexisting knowledge such as the geometry in which the antennas are organized and their responses to signals arriving with different center frequencies or from different directions.

Among the parameters that can be estimated, the DOA has received particular interest. Once one has knowledge of the DOAs of the received signal the received signals may be spatially filtered, i.e., the desired signal with a known DOA can be effectively separated from signals with other direction by using beamforming techniques. This spatial filtering is analogous to bandpass filtering in time domain. The DOAs are also one of the main parameters of radio detection and ranging (RADAR) systems as the one proposed as a precise radio altimeter in [53] and are specially crucial in electronic warfare.

The DOA estimation problem with antenna arrays can be seen in Figure 1.1. Note that the Figure presents a planar wavefront. Thus, it assumes that the transmitter is located in the far-field and that the dimensions of the array are sufficiently small for the spherical surface of the propagating signal to be approximated as a plane. Furthermore, it also assumes that the received wave is narrowband, usually represented by a single tonal wave in the literature. The array shown in this figure is known as a uniform linear array (ULA) and has certain properties assumed for it given the far-field and narrowband assumption. For a signal with a given DOA there is an exact and unique phase difference between the signals received by two consecutive antennas, i.e., there is no ambiguity with respect to the DOA. Thus, for this kind of arrays, if there is only one signal, one can estimate the DOA by looking only at this phase difference.

Although precise DOA estimation is possible in arrays of arbitrary response by employing the very computationally demanding classical maximum likelihood (ML) method [86, 87] or its extensions such as the expectation maximization (EM) [32, 79, 81, 84] or the space alternating generalized expectation maximization (SAGE) [33, 38], these extensions cannot be guaranteed to con-

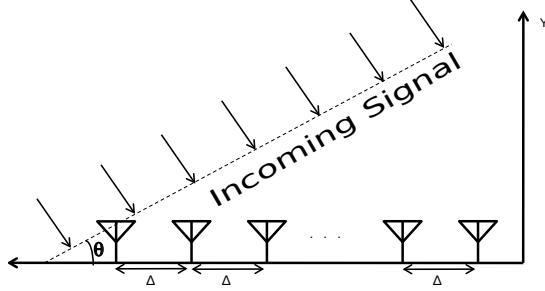


Figure 1.1: Signal with DOA θ impinging on a uniform linear array (ULA), whose the antenna space is Δ

verge to its local maxima, requiring a good initialization of the parameters, and can be very computationally demanding as the number of wavefronts with parameters to be estimated can be quite high [19]. Another alternative for DOA estimation in arbitrary array geometries is the conventional beamformer [4], this method requires a peak search and cannot properly separate closely spaced wavefronts unless a very large number of antennas is present at the array. An improvement over the conventional beamformer is the Capon minimum variance distortionless response (Capon-MVDR) [16]. This method offers increased resolution when compared with the traditional beamformer but suffers in the presence of highly correlated wavefronts scenarios and requires a peak search. The Multiple SIgnal Classification (MUSIC) [95] algorithm is a subspace based algorithm that can be applied with arrays of arbitrary response and also requires a peak search.

In contrast to ML and peak search based methods, there are also, in the literature, a vast number of algorithms that are either closed-form or require very few iterations. Examples of such techniques are the Iterative Quadratic Maximum Likelihood (IQML) [10], Root Weighted Subspace Fitting (Root-WSF) [101] and Root-MUSIC [3] methods. However, all of these methods rely on a Vandermonde or left centro-hermitian array response. Left centro-hermitian matrices are defined in Appendix A. Spatial Smoothing (SPS) [37] and Forward Backward Averaging (FBA) [90] also require Vandermonde and centro-hermitian array responses, respectively. These techniques enable the application of precise closed form DOA estimation methods and precise model order estimation (estimation of the number of impinging wavefronts) in the presence of highly correlated or even coherent signals. Another important DOA estimation technique is the Estimation of Signal Parameters via Rotational Invariance (ESPRIT) algorithm [93], this technique requires a shift-invariant array response. Shift-invariance is a less demanding requirement of the array response compared to a Vandermonde or left centro-hermitian response. This is less demanding when compared to a Vandermonde or left centro-hermitian

Family	Methods
Maximum Likelihood	Classical ML [86, 87] EM [32, 79, 81, 84] SAGE [33, 38] IQML [10]
Beamformers	Classical beamformer [4] CAPON-MVDR [16]
Subspace Methods	MUSIC [95] Root-MUSIC [3] ESPRIT [93] Root-WSF [101]

Table 1.1: Summary of DOA estimation methods, dividing into three families: maximum likelihood (ML), beamformers, and subspace methods

response. Table 1.1 presents a summary of the DOA estimation methods mentioned.

Arrays that have Vandermonde and left centro-hermitian responses are very hard to achieve in real antenna array implementations. Mutual coupling of the antennas, positioning errors in antennas, material tolerances, hardware biases, and the surrounding environment of the array affect the response of individual antennas and, thus, of the antenna array. To deal with such imperfections, a mapping between the real and desired virtual array response can be created, and the received data can then be interpolated into this virtual array. Array interpolation can be performed in either a linear or nonlinear fashion. This mapping is computed by obtaining a mapping/transformation capable of transforming the array as depicted in Figure 1.2.

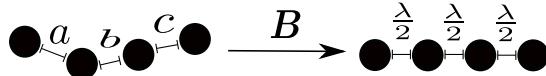


Figure 1.2: On the left side there is real data from an antenna array, while on the right side the data transformed by the transformation matrix \mathbf{B}

In general, when the response of the real array is relatively similar to the desired response, linear interpolation is capable of obtaining good results. However, if the real response strongly differs from the desired virtual response, a nonlinear interpolation method usually becomes necessary. Another application of the array interpolation technique can be seen in [61] where the Vandermonde Invariance Transformation (VIT) was developed. The VIT does not try to address the physical imperfections of the array response but instead transforms the response of an array with a uniform Vandermonde response into one with a nonuniform phase response. The VIT provides a noise shaping

effect by lowering the noise power over a desired angular region and allowing a more precise DOA estimation at the cost of increased computational load.

Bronez [11] first presented a solution for mapping real imperfect array responses into precise and desired responses, and he has called such mapping array interpolation. The presented schemes rely on dividing the field of view of the array response into smaller angular regions, called sectors. For each of these sectors, a transformation matrix is calculated using the empirical knowledge of the true array response. This is called sector-by-sector array interpolation. The larger the region transformed the larger the bias introduced becomes due to transformation imprecision. Friedlander and Weiss [40] extended sector-by-sector array interpolation to include FBA, SPS, and DOA estimation algorithms such as Root-MUSIC [39]. Bühren et al. [13] presented a design metric for the virtual array seeking to preserve the directional characteristics of the true array. Bühren et al. [12] and Weiss and Gavish [95] presented methods for array interpolation that allow the application of ESPRIT for DOA estimation. All of these interpolation methods are performed on by sector-by-sector basis, dividing the entire field of view of the array and performing multiple DOA estimations, one for each sector.

When performing array interpolation for a sector there is no guarantee with respect to what happens with signals received from outside the angular region for which the transformation matrix was derived (out-of-sector signals). If the out-of-sector signal is correlated with any possible in-sector signal, a large bias in the DOA estimation is introduced. Pesavento et al. [89] proposed a method for filtering out of sector signals using cone programming. A similar approach was used by Lau et al. [63, 64] where correlated out-of-sector signals were addressed by filtering out the out-of-sector signals.

More recently, array interpolation has been applied to more specific scenarios. Liu et al. [68] extended the concept of linear array interpolation to coprime arrays allowing for better identifiability of received signals. Hosseini and Sebt [52] applied the concept of linear array interpolation to sparse arrays by selecting virtual arrays as small as possible while retaining the aperture of the original sparse array.

This work proposes an adaptive single sector selection method aiming to circumvent the problem of out-of-sector signals by interpolating all regions of the array where power is received using a single transformation. This work presents an alternative for applying ESPRIT with FBA and SPS, relying on a signal adaptive sector construction and discretization and extended the method to the multidimensional case. The proposed method prevents out-of-sector problems by building a single sector using a signal adaptive approach while also bounding the transformation error. This, however, can lead to large sectors and, in the case of a small number of antennas, this will lead to larger transformation errors and DOA estimation biases. To deal with these limitations, the work at hand also presents a novel pre-processing step used for sector discretization based on PCA and derives a novel concept from the UT [74].

Furthermore, this work extends the concepts of linear interpolation and sector selection to the multidimensional case, allowing the application of the proposed methods to arrays and data models with an arbitrary number of dimensions.

This work proposes a novel nonlinear interpolation approach using MARS. The proposed approach is capable of interpolating arrays with a limited number of antennas and very distorted responses while keeping the DOA estimation bias low, heavily outperforming array linear interpolation methods. This performance improvement comes at the cost of increased computational cost. This work extends the MARS with advanced sector discretization methods that lead to a better performance and lower complexity. Furthermore, a more detailed mathematical description of the MARS method is presented. Expanding on the topic of nonlinear interpolation, the work at hand additionally presents a novel interpolation approach based on GRNNs. The proposed method is capable of achieving a performance similar to that of MARS in certain SNR regimes at a lower computational cost compared to the MARS based method.

In summary, this work extends the concepts of array interpolation presented in previous research to explore the concepts of statistical significant sector discretization and nonlinear interpolation. Previous research focused on concepts for sector-by-sector processing, whereas this work aims to use a single unified sector discretized in a statistically significant manner. Moreover, this work develops novel nonlinear methods for array interpolation and presents a method suitable for real-time implementations capable of providing better performance than linear interpolation methods previously derived in the literature.

The results of the proposed discretization and interpolation methods are assessed via a set of studies considering measured responses obtained from a real physical system. The results show that all the proposed methods significantly improve DOA estimation considering a physical system and its inherent imperfections. Furthermore, this work analyses the performance of the proposed interpolation methods when measurements of the true array response are not available and only simulated responses for building the interpolation models are available.

1.3 Data Model

This model considers a set of L wavefronts impinging onto a linear antenna array composed of M antennas. Figure 1.3 shows a graphical representation of the linear array. Considering the antenna elements are placed along the y-axis, the array geometry can be used to estimate the azimuth angle θ . For arrays composed of isotropic antennas, the elevation angle ϕ would have no influence on the antenna response, however, this work considers arrays with imperfect responses that have varying amplitude and phase responses of the different elements of the array with respect to the elevation angle of the received signals.

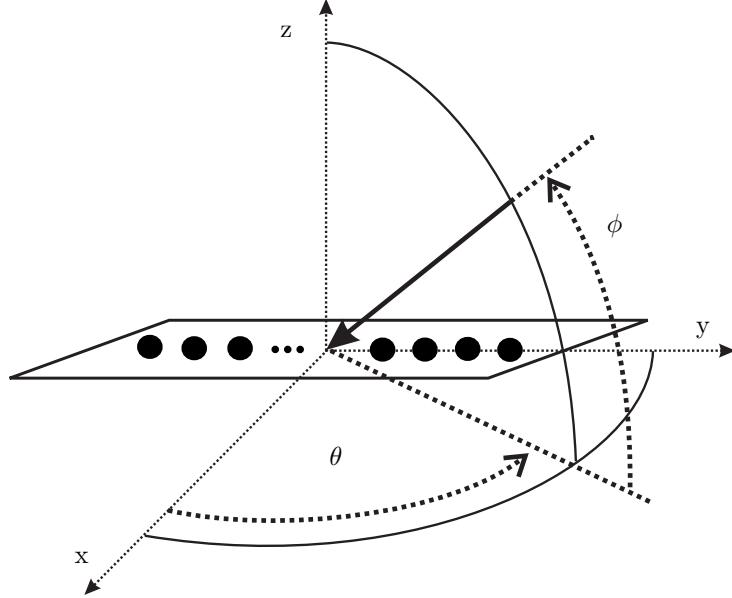


Figure 1.3: Graphical representation of a linear array

The received baseband signal at the array output can be expressed in matrix form as

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \in \mathbb{C}^{M \times N}, \quad (1.1)$$

where $\mathbf{S} \in \mathbb{C}^{L \times N}$ is the matrix containing the N symbols transmitted by each of the L sources, $\mathbf{N} \in \mathbb{C}^{M \times N}$ is the noise matrix with its entries drawn from $\mathcal{CN}(0, \sigma_n^2)$, and

$$\mathbf{A} = [\mathbf{a}(\theta_1, \phi_1), \mathbf{a}(\theta_2, \phi_2), \dots, \mathbf{a}(\theta_L, \phi_L)] \in \mathbb{C}^{M \times L}, \quad (1.2)$$

where θ_i and ϕ_i are the azimuth and elevation angles of the i -th signal, and $\mathbf{a}(\theta_i, \phi_i) \in \mathbb{C}^{M \times 1}$ is the antenna array response.

The received signal covariance matrix $\mathbf{R}_{\mathbf{XX}} \in \mathbb{C}^{M \times M}$ is given by

$$\mathbf{R}_{\mathbf{XX}} = \mathbb{E}\{\mathbf{XX}^H\} = \mathbf{A}\mathbf{R}_{\mathbf{SS}}\mathbf{A}^H + \mathbf{R}_{\mathbf{NN}}, \quad (1.3)$$

where $(\cdot)^H$ stands for the conjugate transpose, and

$$\mathbf{R}_{\mathbf{SS}} = \begin{bmatrix} \sigma_1^2 & \gamma_{1,2}\sigma_1\sigma_2 & \cdots & \gamma_{1,L}\sigma_1\sigma_L \\ \gamma_{1,2}^*\sigma_1\sigma_2 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & \\ \gamma_{1,L}^*\sigma_1\sigma_L & \gamma_{2,L}^*\sigma_2\sigma_L & \cdots & \sigma_L^2 \end{bmatrix}, \quad (1.4)$$

where σ_i^2 is the power of the i -th signal and $\gamma_{i,i'} \in \mathbb{C}$, $|\gamma_{i,i'}| \leq 1$ is the cross-correlation coefficient between signals i and i' with $i \neq i'$. $\mathbf{R}_{\text{NN}} \in \mathbb{C}^{M \times M}$ the spatial covariance matrix of the noise. In case the entries of the noise matrix are drawn from $\mathcal{CN}(0, \sigma_n^2)$, $\mathbf{R}_{\text{NN}} = \mathbf{I}_M \sigma_n^2$, where \mathbf{I}_M denotes an $M \times M$ identity matrix.

An estimate of the received signal covariance matrix can be obtained by

$$\hat{\mathbf{R}}_{\mathbf{XX}} = \frac{\mathbf{XX}^H}{N} \in \mathbb{C}^{M \times M}. \quad (1.5)$$

1.4 Preliminaries

In this section some important concepts used throughout the chapter are explained. Subsections 1.4.1 and 1.4.2 present a brief overview of the FBA and SPS algorithms. The problem of model order selection is detailed in subsection 1.4.3. The ESPRIT parameter estimation algorithm is shown in subsection 1.4.4. Finally, subsection 1.4.5 explains the vandermonde invariance transformation (VIT).

1.4.1 Forward Backward Averaging (FBA)

In the case where only two highly correlated or coherent sources are present the FBA [90] is capable of decorrelating the signals and allowing the application of subspace based methods. For an uniform linear array (ULA) or uniform rectangular array (URA) the steering vectors remain invariant, up to scaling, if their elements are reversed and complex conjugated. Let $\mathbf{Q} \in \mathbb{Z}^{M \times M}$ by an exchange matrix with ones on its anti-diagonal and zeros elsewhere as defined in (A.1). Then, for a M element ULA it holds that

$$\mathbf{Q}\mathbf{A}^* = \mathbf{A}. \quad (1.6)$$

The forward backward averaged covariance matrix can be obtained as

$$\hat{\mathbf{R}}_{\mathbf{XX}_{FBA}} = \frac{1}{2}(\hat{\mathbf{R}}_{\mathbf{XX}} + \mathbf{Q}\hat{\mathbf{R}}_{\mathbf{XX}}^*\mathbf{Q}). \quad (1.7)$$

By looking into the structure of the covariance matrix shown in (1.3), (1.7) can be rewritten as

$$\hat{\mathbf{R}}_{\mathbf{XX}_{FBA}} = \mathbf{A} \frac{1}{2}(\mathbf{R}_{\text{SS}} + \mathbf{R}_{\text{SS}}^*)\mathbf{A}^H + \frac{1}{2}(\mathbf{R}_{\text{NN}} + \mathbf{Q}\mathbf{R}_{\text{NN}}^*\mathbf{Q}). \quad (1.8)$$

The results is that the matrix $\mathbf{R}_{\text{SS}} + \mathbf{R}_{\text{SS}}^* = 2\Re(\mathbf{R}_{\text{SS}})$ is composed of purely real numbers, that is, if the correlation coefficient between the signals is given by a purely imaginary number then the FBA offers maximum decorrelation, if, however, the correlation coefficient is given only by a real number, then the FBA offers no improvement regarding signal decorrelation. FBA can also

be used with the sole purpose of improving the variance of the estimates by virtually doubling the number of available samples without resulting in noise correlation.

1.4.2 Spatial Smoothing (SPS)

In scenarios where more than two highly correlated sources are present the FBA alone cannot restore the full rank of the signal covariance matrix by itself and thus another solution is necessary. SPS was first presented as a heuristic solution in [110] and extended to signal processing in [37]. The idea behind Spatial Smoothing is to split a uniform array into multiple overlapping subarrays as shown in Figure 1.4, the steering vectors of the subarrays are again assumed to be identical up to scaling. Therefore, the covariance matrices of each subarray can be averaged.

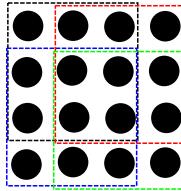


Figure 1.4: Example of SPS subarrays

Spatial Smoothing induces a random phase modulation that tends to decorrelate the signals that cause the rank deficiency [59]. The SPS covariance matrix can be written as

$$\hat{\mathbf{R}}_{\mathbf{XX}_{SPS}} = \frac{1}{V} \sum_{v=1}^V \mathbf{J}_v \hat{\mathbf{R}}_{\mathbf{XX}} \mathbf{J}_v^T, \quad (1.9)$$

where \mathbf{J}_v is an appropriate selection matrix to select the signals of the respective subarray.

The rank of the SPS covariance matrix can be shown to increase by 1 with probability 1 [21] for each additional subarray used until it reaches the maximum rank of L . SPS however comes at the cost of reducing array aperture, since the subarrays are composed of less antennas than the original full array.

More details on FBA and SPS can be found in Appendix B.

1.4.3 Model Order Selection

Model order selection is selecting the optimal trade-off between model resolution and its statistical reliability. In the specific case of this work, model order selection is mostly employed to select the eigenvectors of the signal covariance matrix that account for most of its power, each of these eigenvectors, in turn,

represents the statistics of a received signal. Therefore, in this work, model order selection is mostly used to estimate the number of signals received at the antenna array. This is done by analyzing the profile of the eigenvalues of the signal covariance matrix and looking for a big gap that should separate the eigenvalues related to the signal from the ones related to the noise subspace. If the signals are highly correlated a single eigenvalue can be related to two or more signals, leading to a biased estimation. For this reason, FBA and SPS must be applied in such cases.

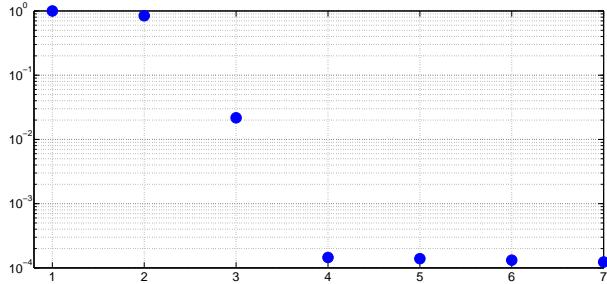


Figure 1.5: Eigenvalue profile: eigenvalue index versus eingenvale

Figure 1.5 presents the example of an eigenvalue profile of two incoming signals after FBA and SPS have been applied. The SNR in this case is 30 dB and eight antennas are used. Note, that visually identifying the model order is very difficult. To address this problem, techniques that were originally developed for model fitting of statistical models were extended to detecting the number of signals impinging over an antenna array. Model order selection schemes such as the Akaike information criterion (AIC) and the minimum description length (MDL) methods [106] and more recently the RADOI [91] are capable of properly detecting the number of received signals in the eigenvalue profile as shown in Figure 1.5. However, this is only possible after FBA and SPS. For multidimensional problems more accurate methods such as [24], [27] can be used.

1.4.4 Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT)

Once the number of signals has been estimated and with the FBA-SPS covariance matrix at hand a joint estimation of the DOAs of all the incoming signals can be performed. This work opts to use the estimation of parameters via rotational invariance techniques (ESPRIT) method [93] since it is a closed form algorithm that can be very easily extended to multidimensional scenarios [47]. It is important to highlight that the current state of array interpolation in the literature [12] states that ESPRIT cannot be employed with a transformation matrix calculated as shown previously in this work.

The ESPRIT parameter estimation technique is based on subspace decomposition. Matrix subspace decomposition is usually done by applying the Singular Value Decomposition (SVD). The SVD of the matrix $\mathbf{X} \in \mathbb{C}^{M \times N}$ is given by

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^H, \quad (1.10)$$

where $\mathbf{U} \in \mathbb{C}^{M \times M}$ and $\mathbf{V}^{N \times N}$ are unitary matrices called the left-singular vectors and right-singular vectors of \mathbf{X} and $\Sigma \in \mathbb{C}^{M \times N}$ is pseudo diagonal matrix containing the singular values of \mathbf{X} . The signal subspace $\mathbf{E}_S \in \mathbb{C}^{M \times \hat{L}}$ of \mathbf{X} can be constructed by selecting only the singular vectors related to the \hat{L} largest singular values, the remaining singular vectors form the noise subspace $\mathbf{E}_N \in \mathbb{C}^{M \times M - \hat{L}}$ of \mathbf{X} .

Equivalently eigenvalue decomposition can be applied on the auto correlation matrix $\hat{\mathbf{R}}_{\mathbf{XX}}$ of \mathbf{X} spanning the same subspace

$$\hat{\mathbf{R}}_{\mathbf{XX}} = \mathbf{E}\Lambda\mathbf{E}^{-1}, \quad (1.11)$$

where $\mathbf{E} \in \mathbb{C}^{M \times M}$ and $\Lambda \in \mathbb{C}^{M \times M}$ contains the eigenvectors and eigenvalues of $\hat{\mathbf{R}}_{\mathbf{XX}}$. The eigenvectors related to the \hat{L} largest eigenvalues span the same signal subspace \mathbf{E}_S of the single value decomposition. The same holds for the noise subspace of the EVD and left singular vectors of the SVD, \mathbf{E}_N .

This classic eigendecomposition is suitable when the noise received at the antenna array is spatially white, since, in the case at hand, a transformation is applied to the data, even if the received noise was originally white it turns in colored noise. To deal with colored noise the generalized eigenvalue decomposition (GEVD) can be used to take the noise correlation into account, the GEVD of the matrix pair $\hat{\mathbf{R}}_{\mathbf{XX}}, \hat{\mathbf{R}}_{\mathbf{NN}}$ is given by

$$\bar{\mathbf{R}}_{\mathbf{XX}}\bar{\Gamma} = \bar{\mathbf{R}}_{\mathbf{NN}}\bar{\Gamma}\Lambda, \quad (1.12)$$

where $\mathcal{E} \in \mathbb{C}^{M \times M}$ is a matrix containing the generalized eigenvectors and $\Lambda \in \mathbb{R}^{M \times M}$ is a matrix containing the generalized eigenvalues in its diagonal. Notice that this decomposition is the same as the EVD (1.11) for $\bar{\mathbf{R}}_{\mathbf{NN}} = \mathbf{I}$. The subspace $\bar{\Gamma}_S \in \mathbb{C}^{M \times \hat{L}}$ is formed selecting the generalized eigenvectors related to the \hat{L} largest generalized eigenvalues. This subspace, however, does not span the same column subspace as the original steering matrix, and needs to be reprojected onto the original manifold subspace or dewhitened. This can be done by

$$\mathbf{\Gamma}_s = \bar{\mathbf{R}}_{\mathbf{NN}}\bar{\Gamma}_s. \quad (1.13)$$

With this subspace estimate at hand the Total Least Squares (TLS) ESPRIT [93] is applied. Two subsets of the signal subspace that are related trough the shift invariance property need to be selected.

Let $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ represent the subspace subsets selected as previously mentioned. A matrix $\mathbf{\Gamma}_{1,2}$ is constructed as

$$\mathbf{\Gamma}_{1,2} = \begin{bmatrix} \mathbf{\Gamma}_1^H \\ \mathbf{\Gamma}_2^H \end{bmatrix} [\mathbf{\Gamma}_1 \mathbf{\Gamma}_2]. \quad (1.14)$$

Performing an eigendecomposition of $\mathbf{\Gamma}_{1,2}$ and ordering its eigenvalues in the decreasing order and its eigenvectors accordingly the eigenvector matrix \mathbf{V} can be divided into blocks as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{1,1} & \mathbf{V}_{1,2} \\ \mathbf{V}_{2,1} & \mathbf{V}_{2,2} \end{bmatrix}. \quad (1.15)$$

Finally, the parameters can be obtained by finding the eigenvalues of

$$\Phi = \text{eig}\left(-\frac{\mathbf{V}_{1,2}}{\mathbf{V}_{2,2}}\right). \quad (1.16)$$

The parameters in Φ can represent a phase delay respective to a DOA if DOAs are being estimated, or can represent the ratio of the strength at which a signal appears in the different polarizations.

For multidimensional arrays another option is to employ methods based on the parallel factor analysis (PARAFAC) decomposition such as [29], [28] instead of the ESPRIT.

1.4.5 Vandermonde Invariance Transformation (VIT)

Finally, once the first set of estimates has been obtained the VIT can be applied, this transformation aims to shape the noise away from the regions where the signal is arriving in order to obtain improved estimates [61]. The VIT is an array interpolation approach that transforms a vandermonde system with a response linear with respect to the angle of arrival into an also vandermonde system with a highly nonlinear response with respect to the angle of arrival. The VIT promotes a nonlinear transformation with respect to the selected spatial frequency $\mu(\theta)$. Let $\mathbf{u} = [1, e^{j\mu(\theta)}, \dots, e^{j\mu(\theta)(M-1)}]$, the VIT performs the following transformation

$$\bar{\mathbf{u}}^{(VIT)} = \mathbf{T}(\theta)\mathbf{u} = \left(\frac{e^{j\mu(\theta)} - r}{1 - r}\right) \begin{pmatrix} 1 \\ e^{j\nu(\theta)} \\ \vdots \\ e^{j\nu(\theta)(M-1)} \end{pmatrix}, \quad (1.17)$$

r and ν are design parameters that can be chosen considering a compromise between the level of noise suppression desired around μ and the linearity of the output.

The VIT can be used to apply a phase attenuation to the dataset, which in turn shapes the noise, reducing the power of the noise in the region near $\mu(\theta)$ and increasing it over its vicinity. Thus, the VIT needs to be applied angle wise, i.e, a set of initial estimates of the angles θ is used to calculate a VIT centered over the given angles, and a second estimate is performed. This second estimation yields an offset θ_{offset} with respect to the original θ , giving the final estimation $\theta_{\text{VIT}} = \theta + \theta_{\text{offset}}$. Due to the mentioned noise

shaping, this final estimation offers increased precision, but comes at the cost of transforming the dataset and applying the chosen DOA estimation method \hat{L} times.

The VIT can be interpreted as a zoom, similar to an optical zoom, with the first estimates a zoom can be used on the regions of the manifold where signal has been estimated to arrive. The region can be inspected with the zoom effect to detect any imprecisions from the first estimate. The increased performance comes at the cost of \hat{L} extra DOA estimations.

1.5 Array Interpolation

Arrays can be interpolated by means of a one-to-one mapping given by

$$\mathbf{f} : \mathbf{A}_{\mathcal{S}} \rightarrow \bar{\mathbf{A}}_{\mathcal{S}}, \quad (1.18)$$

where \mathcal{S} is a sector. It is worth noting that even if an array has the required geometry for the usage of a certain array processing technique, its response may not adhere to an underlying assumed mathematical model. Therefore, in such cases, geometrically well behaved arrays may need to be interpolated to allow the application of the desired technique.

Definition 1.1. A sector \mathcal{S} is a finite and countable set of 2-tuples (pairs) of angles (θ, ϕ) containing all the combinations of azimuth and elevation angles representing a region of the field of view of the array. The sector \mathcal{S} defines the column space of $\mathbf{A}_{\mathcal{S}}$ and $\bar{\mathbf{A}}_{\mathcal{S}}$.

$\mathbf{A}_{\mathcal{S}} \in \mathbb{C}^{M \times |\mathcal{S}|}$ is the array response matrix formed out of the array response vectors $\mathbf{a}(\theta, \phi) \in \mathbb{C}^{M \times 1}$ of the angles given by the elements of \mathcal{S} . $\mathbf{A}_{\mathcal{S}}$ contains the true array response of the physical system, which may not possess important properties such as being left centro-hermitian or Vandermonde. $\bar{\mathbf{A}}_{\mathcal{S}} \in \mathbb{C}^{M' \times |\mathcal{S}|}$ is the interpolated version of $\mathbf{A}_{\mathcal{S}}$ with columns $\bar{\mathbf{a}}(\theta, \phi) \in \mathbb{C}^{M' \times 1}$ being the array response of the so-called virtual or desired array, having all the properties necessary for posterior processing. $|\mathcal{S}|$ is the cardinality of the set \mathcal{S} , i.e. the number of elements in the set.

Definition 1.2. The mapping \mathbf{f} is said to be array size preserving if $M = M'$.

Definition 1.3. The mapping \mathbf{f} is said to be geometry preserving if it is size preserving and if the underlying array geometry for the true and virtual array is equivalent.

This work limits itself to mappings of linear planar arrays that are size and geometry preserving, as given in Definitions 1.2 and 1.3, respectively.

Linear array interpolation is usually done using a least squares approach. The problem is set up as finding a transformation matrix \mathbf{B} that is given by

$$\mathbf{B}\mathbf{A}_{\mathcal{S}} = \bar{\mathbf{A}}_{\mathcal{S}}. \quad (1.19)$$

The snapshot matrix \mathbf{X} can be transformed by multiplying it from the left-hand side with the transform matrix \mathbf{B} , which is equivalent to applying a linear model for each of the outputs of the virtual antenna array. Therefore, in linear array interpolation \mathbf{f} is given by the transformation matrix \mathbf{B} .

This model is usually not capable of transforming the response perfectly across the entire field of view except for the case where a large number of antenna elements is present or a very small sector is used. Large transformation errors often result in a large bias on the final DOA estimation, thus, usually, the response region is divided into a set of regions called sectors, and a different transform matrix is set up for each sector (sector-by-sector processing).

Nonlinear interpolation is an alternative to the linear approach capable of providing better accuracy under more challenging scenarios at the cost of increased complexity. In nonlinear interpolation, the mapping \mathbf{f} is given by a nonlinear function, such that

$$\mathbf{f}(v + \beta q) \neq \mathbf{f}(v) + \beta \mathbf{f}(q) \quad (1.20)$$

for some v, q and β .

1.6 Classical Interpolation

The SPS and FBA methods shown for decorrelating signals are limited to arrays being Vandermode and/or left centro-hermitian. As mentioned previously such geometries are hard to achieve in real implementations due to problems such as space limitation or the non linearity of the antenna lobes with respect to the DOA of received waves. Also, important subspace DOA estimation methods such as the Root-MUSIC [3] rely on the ULA geometry and although the ESPRIT [93] requires only the shift invariance property, in the literature most authors apply ESPRIT in either a ULA or URA. Thus, to allow the application of such methods to generic arrays some sort of data transformation needs to be used first. This type of transformation of known in the literature are array interpolation or array mapping. All the current approaches in the literature use this sector-by-sector processing, the works in [39], [89] and [13] differ only by the way the matrix \mathbf{A} is set up. The linear approach to array interpolation shown in (1.19) is equivalent to applying a linear model for each of the outputs of the virtual antenna array. This linear model can be given

$$[\mathbf{y}]_m = [\mathbf{B}]_{1,m} [\mathbf{y}]_m + [\mathbf{B}]_{2,m} [\mathbf{y}]_m + \dots + [\mathbf{B}]_{M,m} [\mathbf{y}]_m, \quad (1.21)$$

where $[\cdot]_{i,j}$ is the element of a matrix indexed by i and j , and $m \in \{1, \dots, M\}$. In classical array interpolation, since the receiver has no prior information about the DOA of the received signals \mathbf{A} and $\bar{\mathbf{A}}$ are constructed by dividing the field of view of the array into K continuous regions, called sectors, with upper bound u_k and lower bound l_k . This is the sector-by-sector approach. The region $[l_k, u_k]$ is then discretized according to

$$\mathcal{S}_k = [l_k, l_k + \Delta, \dots, u_k - \Delta, u_k], \quad (1.22)$$

where Δ is the angular resolution of the transformation. These angles are used to generate the respective set of steering vectors and construct \mathbf{A} and $\bar{\mathbf{A}}$ according to

$$\mathbf{A}_{\mathcal{S}_k} = [\mathbf{a}(l_k), \mathbf{a}(l_k + \Delta), \dots, \mathbf{a}(u_k - \Delta), \mathbf{a}(u_k)] \in \mathbb{C}^{M \times \frac{u_k - l_k}{\Delta}},$$

$$\bar{\mathbf{A}}_{\mathcal{S}_k} = [\bar{\mathbf{a}}(l_k), \bar{\mathbf{a}}(l_k + \Delta), \dots, \bar{\mathbf{a}}(u_k - \Delta), \bar{\mathbf{a}}(u_k)] \in \mathbb{C}^{M \times \frac{u_k - l_k}{\Delta}},$$

where M is the number of antennas, including their different polarizations, present at the array. The transformation is usually not perfect since the transformation matrix \mathbf{B} does not have enough degrees of freedom to transform the entire discrete sector. \mathbf{B} is obtained as the best fit between the transformed response $\mathbf{B}\mathbf{A}_{\mathcal{S}_k}$ and the desired response $\bar{\mathbf{A}}_{\mathcal{S}_k}$, this is done by applying a least squares fit to the overdetermined systems yielding

$$\mathbf{B} = \bar{\mathbf{A}}_{\mathcal{S}_k} \mathbf{A}_{\mathcal{S}_k}^\dagger \in \mathbb{C}^{M \times M}, \quad (1.23)$$

where $(\cdot)^\dagger$ is the pseudo inverse of the matrix. To access the precision of the transformation the Frobenius norm of the errors matrix $\mathbf{B}\mathbf{A}_{\mathcal{S}_k} - \bar{\mathbf{A}}_{\mathcal{S}_k}$ is compared with the Frobenius norm of the desired response steering matrix $\bar{\mathbf{A}}_{\mathcal{S}_k}$. The error of the transform is defined as

$$\epsilon(\mathcal{S}_k) = \frac{\|\bar{\mathbf{A}}_{\mathcal{S}_k} - \mathbf{B}\mathbf{A}_{\mathcal{S}_k}\|_F}{\|\bar{\mathbf{A}}_{\mathcal{S}_k}\|_F} \in \mathbb{R}^+. \quad (1.24)$$

Large transformation errors will results in a large bias in the final DOA estimates. The transformation error can, however, be kept as low as desired by making the sectors \mathcal{S}_k smaller. It is worth noting that for the sector-by-sector approach the transformation matrices can be calculated off-line, since they do not rely on the received data, and just reused when necessary.

With \mathbf{B} at hand the data can be transformed by

$$\bar{\mathbf{X}} = \mathbf{B}\mathbf{X}. \quad (1.25)$$

The transformed covariance is then equivalent to

$$\bar{\mathbf{R}}_{\mathbf{XX}} = \frac{\mathbf{B}\mathbf{X}(\mathbf{B}\mathbf{X})^H}{N} = \frac{\mathbf{B}\mathbf{X}\mathbf{X}^H\mathbf{B}^H}{N} = \mathbf{B}\hat{\mathbf{R}}_{\mathbf{XX}}\mathbf{B}^H. \quad (1.26)$$

From (1.26) it is possible to note that the transformation matrix can instead be applied directly to the covariance matrix. Plugging (1.4) into (1.26) results in

$$\begin{aligned} \bar{\mathbf{R}}_{\mathbf{XX}} &= \mathbf{B}\mathbf{A}\mathbf{R}_{\mathbf{SS}}\mathbf{A}^H\mathbf{B}^H + \mathbf{B}\mathbf{R}_{\mathbf{NN}}\mathbf{B}^H \\ &= \bar{\mathbf{A}}\mathbf{R}_{\mathbf{SS}}\bar{\mathbf{A}}^H + \mathbf{B}\mathbf{R}_{\mathbf{NN}}\mathbf{B}^H. \end{aligned} \quad (1.27)$$

Thus, although the transformation transforms \mathbf{A} into $\bar{\mathbf{A}}$ as desired it changes the characteristics of $\mathbf{R}_{\mathbf{NN}}$. If the noise was previously white it becomes colored

or, if the noise was already colored, it changes its color. Since most of the methods in the literature assume that $\mathbf{R}_{NN} = \sigma_n^2 \mathbf{I}$, i.e, white noise, the next step used in classical interpolation is a noise whitening step to restore the diagonal characteristic of \mathbf{R}_{NN}

$$\bar{\mathbf{R}}_{XX} = \bar{\mathbf{R}}_{NN}^{-\frac{1}{2}} \bar{\mathbf{R}}_{XX} \bar{\mathbf{R}}_{NN}^{-\frac{H}{2}}, \quad (1.28)$$

where $\bar{\mathbf{R}}_{NN} = \mathbf{B} \mathbf{R}_{NN} \mathbf{B}^H$. Different methods for array interpolation have been specially developed in order to apply the ESPRIT algorithm [12, 108]. In these methods a different transformation matrix is calculated for each of the shift invariant subarrays used in ESPRIT, in this way the noise between each subarray is not correlated and ESPRIT can be directly used, these methods, however, do not allow the application of FBA-SPS and thus are incapable of being used in the presence of highly correlated signals.

The current state-of-the-art for applying array interpolation in highly correlated signals is presented in [63, 64]. Normally, if the signals that fall out-of-sector are not correlated with the in sector signals the influence of the out-of-sector signals in the processing posterior to the transformation is not high. On the other hand, if a signal located out-of-sector is correlated with an in sector signal, the posterior estimations can be gravely affected. These works try to keep the out-of-sector response under control by taking into account the response over the entire array manifold. The problem can be formulated as a weighted least squares problem [64]

$$\min_{\mathbf{T}} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} w(\theta, \phi) \| \mathbf{T}\mathbf{a}(\theta, \phi) - v(\theta, \phi) \bar{\mathbf{a}}(\theta, \phi) \|^2 d\theta d\phi, \quad (1.29)$$

where $w(\theta, \phi)$ is an arbitrary weighting function and $v(\theta, \phi)$ defines the DOA dependent gain that shapes the array response, it should control the response of the array for out-of-sector signals while leaving the response unaltered for signals within the sector. An optimal solution would be to set $v(\theta, \phi)$ as a rectangular function, setting its gain to zero to any DOA outside of the sector and to 1 for the entire sector. This, however, would result in extremely large transformation errors, specially due to the discontinuities found in the edges of the sector. Another alternative is to shape $v(\theta, \phi)$ as a root raised cosine function in θ and ϕ , allowing the gain to roll off along the edges of the sector. By defining the sectors $\Delta\theta = [-\theta_0, \theta_0]$ and $\Delta\phi = [-\phi_0, \phi_0]$ the gain is given by (1.30).

This approach is capable of dealing with closely spaced and highly correlated out-of-sector signals since it addresses the out-of-sector response. However, since this approach transforms the entire field of view for each of the sectors used it introduces a large transformation bias in the final estimates.

$$v(\theta, \phi) = \begin{cases} 1, & |\theta| \leq \theta_0 \text{ and } |\phi| \leq \phi_0 \\ \frac{1}{2} + \frac{1}{4} \cos\left(\frac{\pi(|\theta| - \theta_0)}{2\theta_0 - \pi}\right) + \frac{1}{4} \cos\left(\frac{\pi(|\phi| - \phi_0)}{2\phi_0 - \pi}\right), & \theta_0 < |\theta| \leq (\pi - \theta_0) \text{ or } \phi_0 < |\phi| \leq (\pi - \phi_0) \\ 0, & (\pi - \theta_0) < |\theta| \leq \pi \text{ or } (\pi - \phi_0) < |\phi| \leq \pi \end{cases}, \quad (1.30)$$

1.7 Sector Selection and Discretization

Whatever array interpolation method is used, linear or nonlinear, the array response can only be interpolated after sector detection/selection and sector discretization. The following section presents a discussion on detection and selection of sectors of the field of view of the antenna array and how to discretize the detected sectors.

Traditionally, the field of view of the array is divided into sectors that are small enough to keep the interpolation bias bounded. These sectors can be made larger or smaller depending on the accuracy desired for the interpolation. After the field of view is divided the sectors are discretized uniformly using a fine grid and sector-by-sector interpolation is performed.

To avoid performing sector-by-sector interpolation this work uses a single combined sector containing all regions of the field of view of the array where significant power is received. Considering that the response of the true array needs to be known to construct a transformation of the interpolation of the received signal, this response can be used to detect angular regions where significant power is received in order to detect sectors adaptively. To this end, the conventional beamformer [4] can be applied to provide an estimate of angular regions where significant power is received. The beamformer yields the normalized power response

$$P(\theta, \phi) = \frac{\mathbf{a}^H(\theta, \phi) \hat{\mathbf{R}}_{\mathbf{XX}} \mathbf{a}(\theta, \phi)}{\mathbf{a}^H(\theta, \phi) \mathbf{a}(\theta, \phi)} \in \mathbb{R}. \quad (1.31)$$

Since the beamformer is only a delay and sum approach it is not affected by signal correlation and can provide low resolution estimates even in completely correlated signal environments. Although one could argue for using weighted beamformers such as the CAPON-MVDR [16] since this method offers increased resolution when compared the traditional beamformer it suffers in the presence of highly correlated wavefronts and requires a matrix inversion, leading to a higher computational load. Therefore, this work employs the conventional beamformer in this first step due to its robustness and simplicity.

Figure 1.6 shows an example of the real data output of (1.31) for a signal received at the six element linear physical antenna array shown in Figure 1.17 with $\theta = 0^\circ$, $\phi = 20^\circ$, SNR = 30 dB, and $N = 10$.

In physical systems, the result of (1.31) is discrete in θ and ϕ , and can be written as

$$P[z, v] = P(-90^\circ + (z \cdot \Delta_\theta), -90^\circ + (v \cdot \Delta_\phi)) = P(\theta, \phi), \quad (1.32)$$

with $z \in \mathbb{N}_0$, $v \in \mathbb{N}_0$, $\theta \in \mathcal{D}_{\Delta_\theta}$, and $\phi \in \mathcal{D}_{\Delta_\phi}$ where

$$\mathcal{D}_{\Delta_\theta} = \{-90^\circ, -90^\circ + \Delta_\theta, \dots, 90^\circ - \Delta_\theta, 90^\circ\}, \quad (1.33)$$

$$\mathcal{D}_{\Delta_\phi} = \{-90^\circ, -90^\circ + \Delta_\phi, \dots, 90^\circ - \Delta_\phi, 90^\circ\}. \quad (1.34)$$

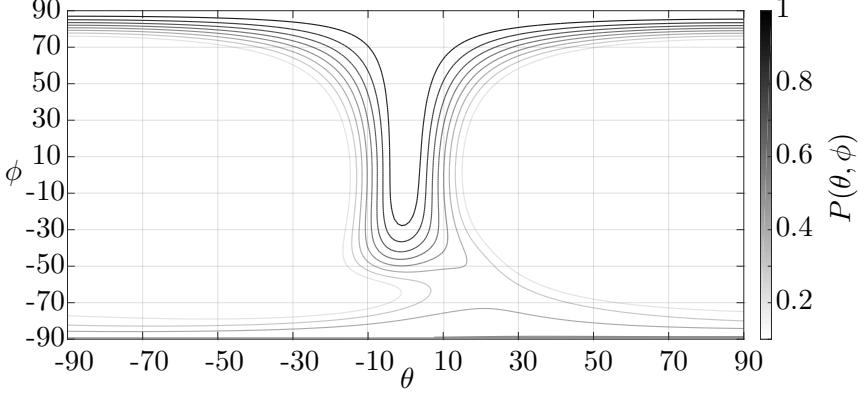


Figure 1.6: $P(\theta, \phi)$

Here, $\Delta_\theta \in \mathbb{R}^+$ and $\Delta_\phi \in \mathbb{R}^+$ are the resolution of the azimuth and elevation angles of the power response (1.31), respectively.

Figure 1.7 presents an example of the real data output of (1.32) when two signals are transmitted from $(-5^\circ, 0^\circ)$ and $(30^\circ, -10^\circ)$, with SNR = 30 dB, and $N = 10$ at the physical array shown in Figure 1.17. This normalized power response is scanned for sectors \mathcal{S}_k , and for each sector, the respective lower bounds $\theta_k^l \in \mathcal{D}_{\Delta_\theta}$ and $\phi_k^l \in \mathcal{D}_{\Delta_\phi}$, as well as the upper bounds $\theta_k^u \in \mathcal{D}_{\Delta_\theta}$ and $\phi_k^u \in \mathcal{D}_{\Delta_\phi}$ are defined as shown in Figure 1.7. A threshold $\alpha\sigma_n^2$, with $\alpha \in \mathbb{R}^+$, defines a sector \mathcal{S}_k considering the criterion $P(\theta, \phi) \geq \alpha\sigma_n^2$ and $k = 1, \dots, K$. A sector can be defined as

$$\mathcal{S}_k = \Theta_k \times \Phi_k = \{(\theta, \phi) | \theta \in \Theta_k \wedge \phi \in \Phi_k\}, \quad (1.35)$$

where

$$\Theta_k = \{\theta | \theta_k^l \leq \theta \leq \theta_k^u\} \subseteq \mathcal{D}_{\Delta_\theta}, \quad (1.36)$$

$$\Phi_k = \{\phi | \phi_k^l \leq \phi \leq \phi_k^u\} \subseteq \mathcal{D}_{\Delta_\phi}, \quad (1.37)$$

and

$$\forall_{k, k' = 1}^K \quad \mathcal{S}_k \cap \mathcal{S}_{k'} = \emptyset. \quad (1.38)$$

A combined sector can be defined as

$$\mathcal{S} = \mathcal{S}_1 \cup \dots \cup \mathcal{S}_k \cup \dots \cup \mathcal{S}_K \subseteq \mathcal{D}_{\Delta_\theta} \times \mathcal{D}_{\Delta_\phi}. \quad (1.39)$$

Most array signal processing techniques assume the presence of isotropic elements. However, in reality, the responses of antennas are never isotropic.

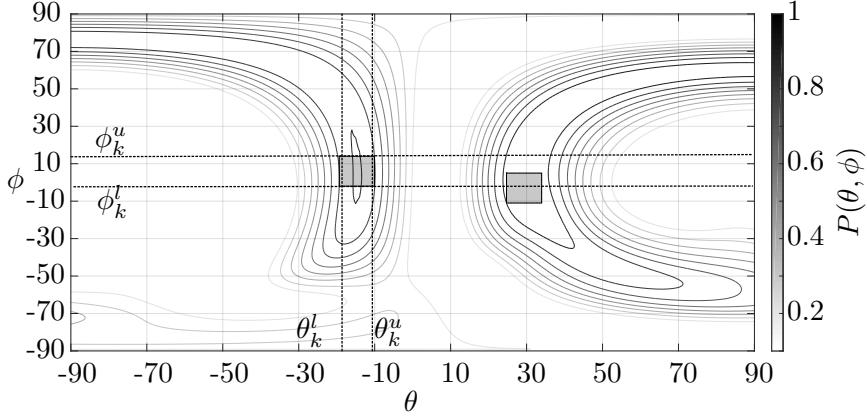


Figure 1.7: Selected sectors and example of sector bounds

Therefore, even for linear arrays, the sectors have to be built taking into account the elevation angle in order to obtain an accurate interpolation. However, for most linear antenna arrays with non-isotropic element responses, the resolution with which the elevation angle ϕ can be estimated is expected to be very low (assuming isotropic elements it is not possible at all). Furthermore, for elevation angles close to zenith, i.e. $\phi \approx 90^\circ$, the angular resolution in azimuth is decreased. This is expected even for arrays composed of perfectly isotropic antennas as there will be no phase difference for different azimuth angles when elevation angles are close to zenith. This is shown in Figure 1.6. The fact that there is a low angular resolution in elevation implies that the response of the array is not heavily affected by changes in elevation angle. Therefore, choosing a fixed elevation angle for the interpolation of a sector is often sufficient if the chosen elevation angle is sufficiently close to the elevation angle of the received signal.

The proposed sector detection and selection technique can be directly applied to two-dimensional arrays such as URAs. For such arrays, the elevation estimation will have a greatly improved resolution. Moreover, for such arrays, it will often be beneficial to interpolate the response in elevation to allow high-resolution techniques to be applied.

Within each of the detected sectors S_k several nearly coherent and closely-spaced signals could be received. Since these signals are represented by similar array responses, an interpolation considering this sector is efficient for all received signals. After interpolation, such signals can be separated with techniques such as SPS and FBA, allowing the application of a high-resolution DOA estimation method to jointly estimate the parameters of all the detected signals. Thus, the separation of closely-spaced sources is only limited by the

high-resolution DOA estimation method and by the number of elements of the antenna array, but not by the array interpolation itself.

To address cases where the noise floor is high a large α can be used. However, this means that only large sectors are detected, at the cost of discarding smaller sectors that are related to a signal component. On the other hand, smaller α means that smaller sectors are detected but at the cost of allowing noise to be mistakenly detected as a sector. Selecting noise regions as sectors will result in a smaller estimation bias than leaving regions with signal power out of the transform. Therefore, for cases where the noise floor is high compared to the signal strength, an α that sets the cutoff to be close to the noise floor is recommended.

For arrays where the response does not vary strongly with elevation, or for setups where signals are received by the array at a given narrow set of elevation angles, ϕ can be fixed. Thus, the output of (1.31) becomes one-dimensional, and the sector selection shown in Figure 1.7 can be simplified to a one-dimensional problem as shown in Figure 1.8.

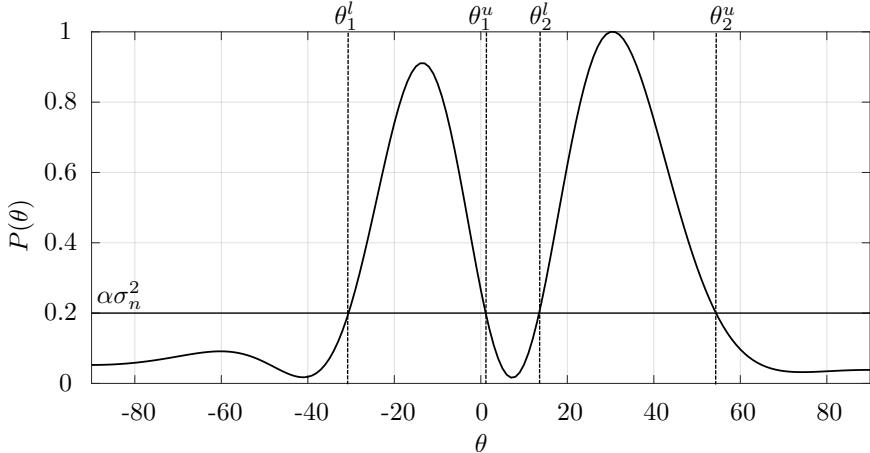


Figure 1.8: Selected sectors and respective bounds for one-dimensional case

In this case, the sectors \mathcal{S}_k are defined only by their lower and upper bounds θ_k^l and, θ_k^u respectively. Previous works on array interpolation created a discrete uniform set of angles within each sector [11, 89]. Instead of using uniform sector discretization, the next subsections present two different methods for sector discretization aiming to preserve the statistical properties of the considered sectors \mathcal{S}_k .

1.7.1 UT Discretization

The UT [54] is a transformation used to transform a continuous probability density function (PDF) into a discrete probability mass function (PMF). The

UT can be applied in array interpolation to discretize the detected sectors in such a way that the statistical properties, the moments, of the detected sectors are preserved in its discrete form. Thus, no statistical information is lost, since a PDF can be fully represented by its moments as stated in the following theorem.

Theorem 1. *A probability density function can be entirely described by its moments.*

Proof. A proof can be found in Appendix C. \square

The UT discretization can be applied by solving the nonlinear set of equations

$$\sum_{j=1}^{k-1} w_j S_j^k = w_1 S_1^k + \dots + w_{k-1} S_{k-1}^k = E\{r^k\}. \quad (1.40)$$

where S_j are known as sigma points, w_j are its respective weights, and r is the value assumed by a continuous random variable \tilde{r} . From (1.40) it is clear that in order to preserve the characteristics of \tilde{r} up to the k -th moment, it is necessary to calculate $k - 1$ sigma points and its weights by solving a nonlinear system of equations. Thus, there is a trade-off between simplicity in the calculation and the accuracy of the representation of higher order moments.

The UT can be applied to an approximation of the concentrated loss function $l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ and thus the log-likelihood $l(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ of the DOA estimation problem, which is derived in the following theorem. The concentrated loss function $l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ and the log-likelihood $l(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ of the DOA estimation problem at hand are defined in (D.1) and (D.4), and

$$\boldsymbol{\theta} = [\theta_1, \dots, \theta_d]^T, \quad (1.41)$$

$$\boldsymbol{\phi} = [\phi_1, \dots, \phi_d]^T. \quad (1.42)$$

Theorem 2. *The concentrated loss function $l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ for a DOA estimation problem can be approximated for each detected sector \mathcal{S}_k by the normalized power response of the conventional beamformer (1.31).*

Proof. A proof can be found in Appendix D. \square

Following the approximation derived in Theorem 2 the concentrated loss function and thus the log-likelihood with respect to each sector \mathcal{S}_k can be considered as a separate PDF each when applying the UT. In order to preserve the first and the second moment of the approximated log-likelihood, the number of points are chosen to transform each detected sector \mathcal{S}_k must be larger or equal than 3. Under this constraint and under the assumption that the input noise follows $\mathcal{CN}(0, \sigma_n^2)$, the respective PDF characteristics for each sector \mathcal{S}_k are mainly preserved. Figure 1.9 presents how the concentrated loss

function $l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ can be represented using the UT for a linear array assuming uniform antenna responses of the different antennas in elevation. A total transformation sector by concatenating all the sigma points given by the UT for the detected sectors can be written as

$$\mathbf{A}_S = [\mathbf{a}(S_1), \mathbf{a}(S_2), \dots, \mathbf{a}(S_T)] \in \mathbb{C}^{M \times T}, \quad (1.43)$$

where T is the total number of points used to discretize all the detected sectors.

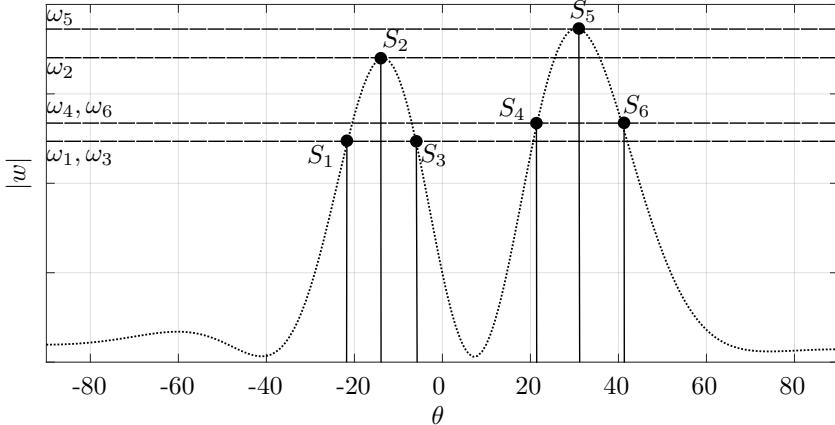


Figure 1.9: UT of the approximated $l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi})$ for two sources

The proposed UT-based discretization is well suited for linear arrays that have a uniform response in elevation. The calculation of the sigma points S_t and weights ω_t with $t = 1, \dots, T$ of the UT transformation is computationally expensive. Therefore, calculating multiple sets of points across multiple elevations may become infeasible depending on the application and system at hand. The UT can be very useful also when applying linear interpolation, as it can be used to ensure that the least squares calculation of the transformation matrix is a determined problem instead of an overdetermined one.

1.7.2 Principal Component Discretization

For most interpolation methods presented in the literature, the interpolated sectors are discretized using a simple uniform discretization [11, 89]. The discretization is usually done such that the discrete angles are closely spaced. Therefore, the steering vectors of such closely spaced angles are highly correlated. This work proposes the application of principal component analysis (PCA) to avoid problems caused by the correlation in predictors as well as to minimize the computational load of the transformation.

Ideally, \mathbf{A}_S would be set up such that $\mathbf{A}_S \mathbf{A}_S^H = \mathbf{I}_M$. However, as stated previously, this is often not the case with $\mathbf{A}_S \mathbf{A}_S^H = \mathbf{R}_{\mathbf{A}_S \mathbf{A}_S} \neq \mathbf{I}_M$. PCA aims to find a matrix \mathbf{P} such that

$$\mathbf{P} \mathbf{A}_S = \mathbf{B}_S, \quad (1.44)$$

where $\mathbf{B}_S \mathbf{B}_S^H = \mathbf{I}_M$. \mathbf{P} can be obtained by performing the eigendecomposition of $\mathbf{R}_{\mathbf{A}_S \mathbf{A}_S}$

$$\mathbf{R}_{\mathbf{A}_S \mathbf{A}_S} = \mathbf{E} \mathbf{\Sigma} \mathbf{E}^H, \quad (1.45)$$

and setting $\mathbf{P} = \mathbf{E}$. The proposed PCA pre-processing can be used to create a better set of predictor variables for the interpolation and to reduce the dimensionality of the problem by excluding any eigenvectors associated to very small eigenvalues.

1.8 Linear Adaptive Array Interpolation

In the case of linear interpolation, the transformation bias is closely tied to the total size of the transformed sector. While the UT and PCA discretization can minimize this problem, they cannot fully mitigate it. Therefore, this thesis presents an iterative approach that seeks to find out the maximum region of the field of view that can be transformed while keeping the transformation error bounded by a value. It also distributes the calculated region with respect to the power of each sector detected in the second step.

The first part of the next two steps is to calculate a weighting factor based on the power of each detected sector. This weighting factor will be used to optimally distribute the transformable region with respect to the received power, focusing the transformation on regions with more power and aiming to leave the least significant power outside of the total transformed region. Considering a one dimensional beamformer output, a weighting factor for each sector is calculated as

$$\xi_w = \frac{\sum_{z=\theta_w^l}^{\theta_w^u} P[z]}{\sum_{c=1}^W \sum_{z=\theta_c^l}^{\theta_c^u} P[z]} \in \mathbb{R}. \quad (1.46)$$

As mentioned in Section 1.6, classical array interpolation methods in the literature divide the array response into various sectors (sector-by-sector processing) in order to keep the error (1.24) small. However, in this thesis a single transform matrix is used based on a combination of the sectors detected in (1.31). In order to bound the error $\epsilon(\mathcal{S})$ a signal adaptive method for calculating the maximum transform size is used based on the weights calculated in (1.46). For a sector centered at μ_k , where

$$\mu_k = \left\lceil \frac{\theta_k^u - \theta_k^l}{2} \right\rceil_{\mathcal{D}_\Delta},$$

the discrete and countable set of angles used to transform this sector is given by

$$\mathcal{S}_k = \begin{cases} \left[\left\lceil \mu_k - \frac{\Xi \xi_k}{2} \right\rceil_{\mathcal{D}_\Delta}, \left\lceil \mu_k - \frac{\Xi \xi_k}{2} + \Delta \right\rceil_{\mathcal{D}_\Delta}, \dots, \left\lceil \mu_k + \frac{\Xi \xi_k}{2} \right\rceil_{\mathcal{D}_\Delta} \right] \\ , \Xi \xi_k < |\theta_k^u - \theta_k^l| \\ [\theta_k^l, \theta_k^l + \Delta, \dots, \theta_k^u], \Xi \xi_k \geq |\theta_k^u - \theta_k^l| \end{cases} \quad (1.47)$$

and

$$\mathcal{S}_k \cap \mathcal{S}_{\bar{k}} = \emptyset \quad \forall k, \bar{k} \in [1, \dots, K] \text{ and } k \neq \bar{k}, \quad (1.48)$$

where $\Xi \in \mathbb{R}^+$ is the total transform size in degrees of all sectors. The combined sector is given by

$$\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \dots \cup \mathcal{S}_K. \quad (1.49)$$

Thus, \mathcal{S} has a better resolution for the sectors \mathcal{S}_k where more power is present (weighted by ξ_k), i.e the transformation of the combined sector \mathcal{S} is weighted towards the sectors \mathcal{S}_k that include more signal power as shown in Figure 1.10.

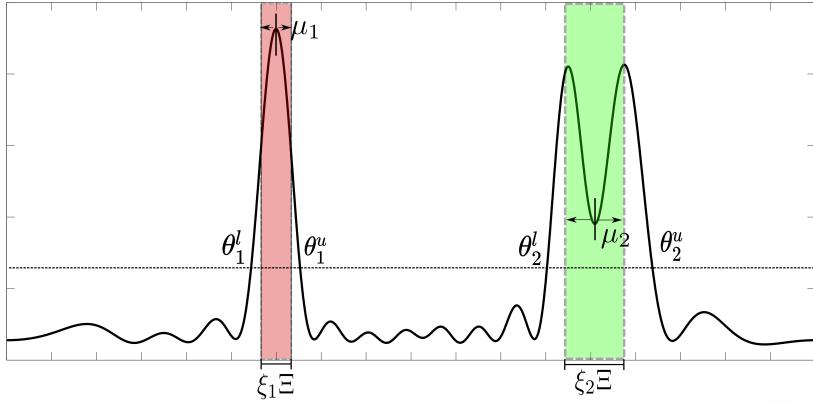


Figure 1.10: Example of transformed regions

The problem of obtaining the transform matrix \mathbf{B} is equivalent to solving a highly overdetermined system, thus

$$|\mathcal{S}| \rightarrow \infty \iff \epsilon(\mathcal{S}) \rightarrow \infty. \quad (1.50)$$

Thus, transforming the entire detected sectors may result in a very high transformation error introducing a very large bias into the final DOA estimates. To address this problem an upper bound to the transform error ϵ_{\max} needs to be defined and a search can be performed to find the maximum transform size covering the detected sectors that is still within the error upper bound. The

problem of finding the maximum Ξ with respect to ϵ_{\max} can be written as the optimization problem

$$\max_{\Xi} \epsilon(\mathcal{S}) \quad (1.51)$$

$$\text{subject to } \epsilon(\mathcal{S}) \leq \epsilon_{\max} \quad (1.52)$$

$$\Xi \leq \Xi_{\max} = \sum_{k=1}^K |\theta_k^u - \theta_k^l| \quad (1.53)$$

$$\Xi \geq \Xi_{\min} = M\Delta. \quad (1.54)$$

The problem in (1.51), (1.52), (1.53), (1.54) can efficiently be solved using a bisection search method, since once all μ_w and ξ_w have been defined, the error function increases monotonically for $\Xi > \Xi_{\min}$, as illustrated in Figure 1.11. $\epsilon(\mathcal{S})$ is greatly affected if the calculation of \mathbf{B} is either a heavily over-determined or an underdetermined system. Therefore, Ξ_{\min} is defined to ensure monotonicity of the problem given in (1.51).

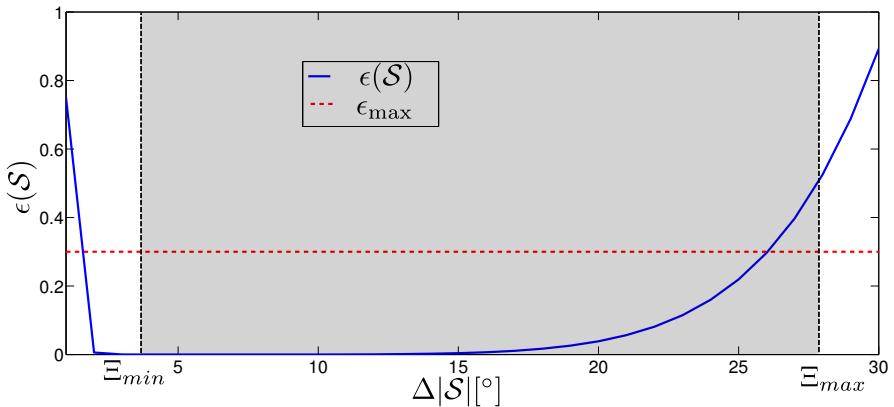


Figure 1.11: Transformation error with respect to combined sector size

In summation these two steps consist of calculating a set of transform matrices according to the weights given by (1.46) and looking for the maximum transform region that still yields a transformation errors bounded by ϵ_{\max} .

1.8.1 Data Transformation and Model Order Selection

Once all μ_k and ξ_k as well as Ξ have been defined the final transformation matrix can be calculated. In order to obtain further insight into the calculation of the transformation matrix (1.23) can be analysed in detail

$$\begin{aligned}\mathbf{B} &= \bar{\mathbf{A}}_S \mathbf{A}_S^\dagger \\ &= \bar{\mathbf{A}}_S \mathbf{A}_S^H (\mathbf{A}_S \mathbf{A}_S^H)^{-1} \\ &= \mathbf{R}_{\bar{\mathbf{A}}_S \mathbf{A}_S} \mathbf{R}_{\mathbf{A}_S \mathbf{A}_S}^{-1}.\end{aligned}\quad (1.55)$$

From (1.55) further insight into the calculation of the transformation matrix is available. The transformation matrix calculation can be seen as the transformation of the cross covariance between the real and desired array responses into the covariance of the desired array response if no errors are assumed. However, the desired array response covariance will be corrupted by the noise present in the measurements, thus, taking such errors into account (1.55) can be rewritten as

$$\mathbf{B} = \mathbf{R}_{\bar{\mathbf{A}}_S \mathbf{A}_S} (\mathbf{R}_{\mathbf{A}_S \mathbf{A}_S} + \mathbf{B} \mathbf{R}_{\mathbf{N} \mathbf{N}} \mathbf{B}^H)^{-1}. \quad (1.56)$$

The noise present after transformation is colored by the transformation itself, therefore, the calculation of \mathbf{B} will depend on \mathbf{B} itself. To solve this problem an iterative approach can be used, \mathbf{B} can be initialized using (1.23) and a new \mathbf{B} can be obtained using this estimate, after that a new iteration of \mathbf{B} is obtained using the previous value until convergence.

$$\begin{aligned}\mathbf{B}_0 &= \bar{\mathbf{A}}_S \mathbf{A}_S^\dagger \\ \text{for } i &\geq 1 \\ \mathbf{B}_i &= \mathbf{R}_{\bar{\mathbf{A}}_S \mathbf{A}_S} (\mathbf{R}_{\mathbf{A}_S \mathbf{A}_S} + \mathbf{B}_{i-1} \mathbf{R}_{\mathbf{N} \mathbf{N}} \mathbf{B}_{i-1}^H)^{-1} \\ \text{repeat until convergence}\end{aligned}\quad (1.57)$$

Since \mathbf{A}_S is formed of closely spaced angular responses its column vectors are highly correlated, this high correlation leads to a suboptimal transformation. Thus, one can obtain a more statistically significant \mathbf{B} by performing a reduced rank (RR) projection [92] that takes into account only the principal components of \mathbf{A}_S . Performing the singular value decomposition (SVD)

$$\mathbf{B}(\mathbf{A}_S + \boldsymbol{\Upsilon}) = \mathbf{U} \mathbf{D} \mathbf{V}^H, \quad (1.58)$$

the singularvectors of \mathbf{U} related to the r largest singularvalues are selected forming $\mathbf{U}_r \in \mathbb{C}^{M \times r}$ and the reduced rank \mathbf{B}_r is given by the orthogonal projection

$$\mathbf{B}_r = (\mathbf{U}_r \mathbf{U}_r^H) \mathbf{B} \in \mathbb{C}^{M \times M}. \quad (1.59)$$

r can be selected using a model order selection scheme such as the RADOI [91].

Once \mathbf{B} has been obtained, the transformed covariance including FBA and SPS is obtained by [40]

$$\bar{\mathbf{R}}_{\mathbf{XX}} = \frac{1}{2W} \sum_{w=1}^W \mathbf{J}_w^T (\mathbf{B} \hat{\mathbf{R}}_{\mathbf{XX}} \mathbf{B}^H + \mathbf{Q} \mathbf{B} \hat{\mathbf{R}}_{\mathbf{XX}}^* \mathbf{B}^H \mathbf{Q}) \mathbf{J}_w, \quad (1.60)$$

where $(\cdot)^*$ stands for the complex conjugation. While W can be chosen *a priori* it can also be adaptively chosen as to minimize the loss of effective array aperture while achieving a good estimate of L . The model order estimation method is given by $\text{MOE}(\bar{\mathbf{R}}_{\mathbf{XX}}(W)) = \hat{L}$. Therefore the following problem must be solved

$$(W, \hat{L}) = \arg \min_W \max_{\hat{L}} \{ \text{MOE}(\bar{\mathbf{R}}_{\mathbf{XX}}(W)) \}. \quad (1.61)$$

Figure 1.12 shows an example of selecting the optimal number of subarrays. After the maximum model order has been detected, increasing the number of subarrays offers no additional separation, since no extra signal was detected, and comes at the cost of sacrificing antenna aperture.

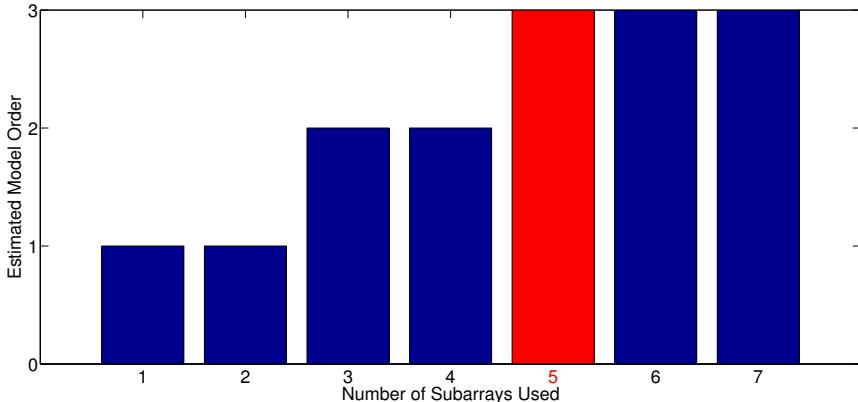


Figure 1.12: Example of estimated model order versus number of subarrays

It is important to notice that the estimated number of impinging signals \hat{L} can be different from the number of sectors detected in (1.31). Each of the detected sectors \mathcal{S}_k can be formed by a set of nearly coherent signals that now can be efficiently separated with (1.61) allowing the application of a high resolution DOA estimation method to jointly estimate the parameters of all the detected signals.

It is important to highlight now that if a ML method [86, 87] or its extensions such as the EM [32, 79, 81, 84] or SAGE [33, 38] would be chosen for estimating the direction of arrival of the received signals the model order would have to be estimated first, and to do so, in highly correlated signal environments,

FBA and SPS need to be applied since model order selection schemes rely on the analysis of the profile of the eigen or singular values. In practice, this means that even if a parametric ML method was to be used the steps taken so far into the algorithm would still need to be taken. Therefore, no claims of increased complexity can be made, up to this point, when the proposed DOA estimation algorithm is compared to a ML algorithm and the presented steps of the proposed algorithm can be viewed as a separate part that can be applied to use ML methods in highly correlated signals, arbitrary geometry array environments if one has no prior knowledge of the number of received signals.

1.8.2 Linear UT Array Interpolation

A linear transformation that takes into account the weights found by the UT for the sigma points can be found by extending (1.23), resulting in

$$\begin{aligned} \mathbf{B}_{\text{UT}} &= \bar{\mathbf{A}}_{\mathcal{S}_{\text{UT}}} \sqrt{\text{diag}\{[w_1, w_2, \dots, w_M]\}} \mathbf{A}_{\mathcal{S}_{\text{UT}}}^H \\ &\quad \left(\mathbf{A}_{\mathcal{S}_{\text{UT}}} \sqrt{\text{diag}\{[w_1, w_2, \dots, w_M]\}} \mathbf{A}_{\mathcal{S}_{\text{UT}}}^H \right)^{-1} \in \mathbb{C}^{M \times M}, \end{aligned} \quad (1.62)$$

where $\mathbf{A}_{\mathcal{S}_{\text{UT}}}$ stands for the steering matrix of sector \mathcal{S} after it has been discretized using the UT.

Following the proposed approach the transformation error shown in (1.24) is zero. However, the proposed approach requires the calculation of multiple sigma points through solving nonlinear systems of equations, and, for setting up such systems, it is necessary to obtain the moments of the likelihoods truncated from the beamformer output. To simplify the application and allow the proposed method to be used in real-time array interpolation for systems with limited processing power, this work proposes calculating sigma points for common sector likelihoods, that is, defining a set of sigma points for common expected beamformer waveforms that a real-time system may use. Using this set of points the interpolation matrix can be calculated and stored in a look-up-table as signals are detected by the system as the interpolation can then be applied.

The transformation matrix calculation can be seen as the transformation of the weighted cross-covariance between the real and desired array responses into the covariance of the desired array response. However, the desired array response covariance is corrupted by the noise present in the measurements. The noise present after transformation is colored by the transformation itself, therefore, the calculation of \mathbf{B}_{UT} will depend on \mathbf{B}_{UT} itself. To solve this problem, an iterative approach can be used,

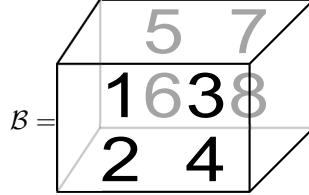


Figure 1.13: Tensor $\mathcal{B} \in \mathbb{R}^{2 \times 2 \times 2}$

$$\begin{aligned}
\mathbf{B}_{\text{UT}}^{(p)} &= \bar{\mathbf{A}}_{\mathcal{S}_{\text{UT}}} \sqrt{\text{diag}\{[w_1, w_2, \dots, w_M]\}} \mathbf{A}_{\mathcal{S}_{\text{UT}}}^H \\
&\left(\left(\mathbf{A}_{\mathcal{S}_{\text{UT}}} \sqrt{\text{diag}\{[w_1, w_2, \dots, w_M]\}} \mathbf{A}_{\mathcal{S}_{\text{UT}}}^H \right) + \mathbf{B}_{\text{UT}}^{(p-1)} \mathbf{R}_{\text{NN}} \mathbf{B}_{\text{UT}}^{(p-1)H} \right)^{-1} \\
&\in \mathbb{C}^{M \times M},
\end{aligned} \tag{1.63}$$

where $\mathbf{B}_{\text{UT}}^{(0)}$ can be initialized using a diagonal \mathbf{R}_{NN} with σ_n^2 filling its diagonal. After that a new iteration of $\mathbf{B}_{\text{UT}}^{(p)}$ is obtained using the previous value until convergence.

1.9 Multidimensional Linear Interpolation

This section extends the concepts of linear interpolation to arrays of two or more dimensions.

1.9.1 Tensor Algebra Concepts

This section introduces the tensor algebra concepts of tensor unfolding and r -mode product.

Tensor Unfolding

In order to operate with tensors using well-known linear algebra tools via the application of tensor unfoldings. This operation transforms a tensor of dimension $R \geq 2$ into a matrix. The r -mode unfolding of tensor \mathcal{A} is denoted by

$$[\mathcal{A}]_{(r)} \in \mathbb{C}^{M_r \times (M_1 \cdot M_2 \cdots M_{r-1} \cdot M_{r+1} \cdots M_R)}. \tag{1.64}$$

$[\mathcal{A}]_{(r)}$ is a matrix containing the r -mode vectors or r fibers of tensor \mathcal{A} along its rows. A r -mode vector can be obtained by fixing the index of all dimensions other than r and varying the index of the r -th dimension along its range.

The unfoldings of the tensor presented in Figure 1.13 are given by

$$[\mathcal{B}]_{(1)} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{bmatrix}, \tag{1.65}$$

$$[\mathcal{B}]_{(2)} = \begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{bmatrix}, \quad (1.66)$$

$$[\mathcal{B}]_{(3)} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}. \quad (1.67)$$

r-Mode Product

The r -mode product of a tensor $\mathcal{A} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_r}$ and a matrix $\mathbf{D} \in \mathbb{C}^{L \times M_r}$ is denoted by

$$\mathcal{C} = \mathcal{A} \times_r \mathbf{D} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times L \times \dots \times M_r}, \quad (1.68)$$

which is equivalent to

$$[\mathcal{A} \times_r \mathbf{D}] = \mathbf{D} \cdot [\mathcal{A}]_{(r)}. \quad (1.69)$$

Thus, the matrix \mathbf{D} must have the same number of columns as the numbers of elements in the r -th dimension of \mathcal{A} .

1.9.2 Multidimensional Data Model

The data model considers d planar wavefronts impinging on an antenna array containing $M_1 \times M_2 \times \dots \times M_R$ elements. The signal received at the m_1, m_2, \dots, m_R antenna at a given time snapshot t is given by

$$\begin{aligned} x_{m_1, m_2, \dots, m_R, t} = & \sum_{i=1}^d \gamma_{m_1, m_2, \dots, m_R, i} s_i[n] \prod_{r=1}^R e^{j \cdot (m_r - 1) \cdot \mu_i^{(r)}} \\ & + n_{m_1, m_2, \dots, m_R, n}, \end{aligned} \quad (1.70)$$

where $s_i[n]$ is the complex symbol transmitted by the i -th source at time snapshot n , $n_{m_1, m_2, \dots, m_R, t}$ is the zero mean circularly symmetric (ZMCS) white complex Gaussian noise present at antenna m_1, m_2, \dots, m_R at time snapshot n . $\gamma_{m_1, m_2, \dots, m_R}$ is an imperfection present at the antenna m_1, m_2, \dots, m_R . $\mu_i^{(r)}$ represents the spatial frequency of the signal transmitted by the i -th source over the r -th dimension, note that the spatial frequency can be mapped into different parameters depending on the dimension. In case of spatial dimension, it is mapped into a direction of arrival, while for frequency and time dimensions it can be mapped into time delay of arrival and Doppler shift, respectively.

For the received signal to have the PARAFAC structure its steering array must have a Khatri-Rao structure as follows. Let M_r be the size of the r -th dimension, and $M = \prod_{r=1}^R M_r$. A steering vector $\mathbf{a}_i^{(r)}$ containing the spatial frequencies relate to the i -th source over the r -th dimension can be defined as

$$\mathbf{a}_i^{(r)} = \left[1 \ e^{j \cdot \mu_i^{(r)}} \ \dots \ e^{(M_r - 1) \cdot j \cdot \mu_i^{(r)}} \right]^T. \quad (1.71)$$

Therefore, an array steering vector for the i -th source can be written as

$$\mathbf{a}_i = \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \dots \otimes \mathbf{a}_i^{(r)} \in \mathbb{C}^{M \times 1}, \quad (1.72)$$

while the array steering matrix for the r -th dimension can be constructed as

$$\mathbf{A}^{(r)} = [\mathbf{a}_1^{(r)}, \mathbf{a}_2^{(r)}, \dots, \mathbf{a}_d^{(r)}] \in \mathbb{C}^{M \times d}. \quad (1.73)$$

Finally, the total steering matrix of the d sources can be constructed in terms of Khatri-Rao products of all array steering matrices of all dimensions.

$$\mathbf{A} = \mathbf{A}^{(1)} \diamond \mathbf{A}^{(2)} \diamond \dots \diamond \mathbf{A}^{(r)} \in \mathbb{C}^{M \times d}. \quad (1.74)$$

For the tensor notation, a PARAFAC steering tensor for each of the d signal sources can be obtained by

$$\mathcal{A}_i = \mathbf{a}_i^{(1)} \circ \mathbf{a}_i^{(2)} \circ \dots \circ \mathbf{a}_i^{(r)} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R}, \quad (1.75)$$

where \circ is the outer product. A total array steering tensor can be constructed by concatenating the individual steering tensor for each of the signal sources over the $r+1$ dimension, as it follows

$$\mathcal{A} = [\mathcal{A}_1 |_{R+1} \mathcal{A}_2 |_{R+1} \dots |_{R+1} \mathcal{A}_d] \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times d}. \quad (1.76)$$

The $|_r$ operator represents the concatenation operation over the r -th dimension.

A non PARAFAC steering tensor can be written as

$$\mathcal{X} = (\mathcal{A} \odot \mathcal{G}) \times_R \mathbf{S}^T + \mathcal{N}, \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times N}, \quad (1.77)$$

here $\mathcal{N} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times N}$ is the noise tensor and contains the noise samples interfering with the measurements, and $\mathcal{G} \in \mathbb{C}^{M_1 \times M_2 \times \dots \times M_R \times d}$ is a tensor containing the antenna imperfections.

1.9.3 Multidimensional Interpolation

Array interpolation can be extended to the multidimensional case by considering the multidimensional data model and rewriting (1.19) as

$$\mathcal{I} \times_1 \bar{\mathbf{A}}_1 \times_2 \bar{\mathbf{A}}_2 \dots \times_R \bar{\mathbf{A}}_R = \mathcal{A} \times_1 \mathbf{B}_1 \times_2 \mathbf{B}_2 \dots \times_R \mathbf{B}_R, \quad (1.78)$$

where $\bar{\mathbf{A}}_r$ is the desired array response for the r -th dimension and \mathbf{B}_r is the transformation matrix for the r -th dimension given by

$$\mathbf{B}_r = [\mathcal{A}]_{(r)} \bar{\mathbf{A}}_r^\dagger, \quad (1.79)$$

where $(\cdot)^\dagger$ stands for the Moore–Penrose pseudo-inverse. Note that in this case the original array \mathcal{A} has no PARAFAC structure, while the array after the interpolation has the PARAFAC structure.

Similar to what is done in the matrix case, the power response is used to the detect sectors for interpolation. The power response for dimension r is given by

$$P(\mu^{(r)}) = \left| \frac{\mathbf{w}^{(r)H}(\mu^{(r)})\hat{\mathbf{R}}_{\mathbf{XX}}^{(r)}\mathbf{w}^{(r)}(\mu^{(r)})}{\mathbf{w}^{(r)H}(\mu^{(r)})\mathbf{w}^{(r)}(\mu^{(r)})} \right| \in \mathbb{R}, \quad (1.80)$$

where $(\cdot)^H$ is the conjugate hermitian and $\hat{\mathbf{R}}_{\mathbf{XX}}^{(r)}$ is the spatial covariance matrix for dimension (r) . $\hat{\mathbf{R}}_{\mathbf{XX}}^{(r)}$ is given by

$$\hat{\mathbf{R}}_{\mathbf{XX}}^{(r)} = \frac{M_r}{N \prod_{i=1}^R M_i} [\mathcal{X}]_{(r)} [\mathcal{X}]_{(r)}^H \in \mathbb{C}^{M_r \times M_r}. \quad (1.81)$$

Following the matrix approach, the output of (1.80) is scanned for K sectors, and for each sector the respective lower bound $l_k^{(r)} \in \mathcal{D}_\Delta$ and upper bound $u_k^{(r)} \in \mathcal{D}_\Delta$, where $k = 1, \dots, K$, are defined done in the matrix case. If K sectors are detected, a detected sector with bounds $[l_k^{(r)}, u_k^{(r)}]$ is said to be centred at

$$c_k = \left\lfloor \frac{|u_k^{(r)} - l_k^{(r)}|}{2} \right\rfloor_{\mathcal{D}_\Delta} \in \mathcal{D}_\Delta, \quad (1.82)$$

After the sectors are detected for each of the dimensions, they can be discretized and a transformation can be calculated following the same steps presented for the one dimensional case. After \mathbf{B}_r has been obtained for all the dimensions to be interpolated the received data can be transformed by

$$\mathcal{X} \times_1 \mathbf{B}_1 \times_2 \mathbf{B}_2 \dots \times_R \mathbf{B}_R. \quad (1.83)$$

The transformation also affects the noise component, leading to coloured noise at the output. This requires some sort of prewhitening to be applied prior to some DOA estimation algorithms. For prewhitening schemes, the reader may refer to [26] for the matrix case and [25] for the tensor case.

Depending on the structure chosen for the interpolated array many different DOA estimation methods can be applied. Multidimensional ESPRIT approaches such as [48] or methods based on the PARAFAC decomposition such as [29], [28] can be used.

1.10 Nonlinear Array Interpolation

In this section two nonlinear alternatives to array interpolation, allowing the application of FBA and SPS in highly correlated signal scenarios, are presented. Subsection 1.10.1 presents an interpolation approach based on the MARS regression method, allowing a model to be created expressing nonlinear relationships between the true and the desired array response. Subsection 1.10.2 shows an interpolation approach based on GRNNs, this approach

can be implemented in a parallel manner, making it a good candidate for array interpolation under real-time constraints. For the nonlinear approaches sector selection and discretization can be performed identically to when linear approaches are employed.

1.10.1 MARS based interpolation

This work proposes a nonlinear interpolation method based on the MARS technique. The MARS regression approach is detailed in Appendix E. The proposed interpolation approach builds a mapping that splits the real and imaginary parts of the response of each of the antennas of the array, that is, a total of $2M$ MARS models are built to interpolate the array response with respect to the combined sector \mathcal{S} . Based on (E.4) this mapping can be written as

$$\begin{aligned} \forall (\theta, \phi) \in \mathcal{S} : & \begin{bmatrix} \Re\{\bar{\mathbf{a}}(\theta, \phi)\} \\ \Im\{\bar{\mathbf{a}}(\theta, \phi)\} \end{bmatrix} = \mathbf{f}(\mathbf{a}(\theta, \phi)) \\ & = \begin{bmatrix} \sum_{\ell=1}^{L_m^R} c_{\ell,1}^R F_{\ell,1}^R(\Re\{\mathbf{a}(\theta, \phi)\}) \\ \vdots \\ \sum_{\ell=1}^{L_m^R} c_{\ell,M}^R F_{\ell,M}^R(\Re\{\mathbf{a}(\theta, \phi)\}) \\ \sum_{\ell=1}^{L_m^I} c_{\ell,1}^I F_{\ell,1}^I(\Im\{\mathbf{a}(\theta, \phi)\}) \\ \vdots \\ \sum_{\ell=1}^{L_m^I} c_{\ell,M}^I F_{\ell,M}^I(\Im\{\mathbf{a}(\theta, \phi)\}) \end{bmatrix}, \end{aligned} \quad (1.84)$$

where $c_{l_m}^R$ and $c_{l_m}^I$ are the weighting constants, $F_{l_m}^R$ and $F_{l_m}^I$ are the basis functions and L_m^R and L_m^I are the number of functions and weighting constants of each model. Furthermore, $\Re\{\bar{\mathbf{a}}(\theta, \phi)\}$ and $\Im\{\bar{\mathbf{a}}(\theta, \phi)\}$ denote the real and imaginary part of vector $\bar{\mathbf{a}}(\theta, \phi)$, respectively. Based on (E.6) the generalized cross-validation (GCV) for the backward pass (step used to reduce possible over-fitting of the model) can be written as

$$\Re\{\text{GCV}\} = \frac{\frac{1}{|\mathcal{S}|} \sum_{(\theta, \phi) \in \mathcal{S}} (\Re\{\bar{\mathbf{a}}_{[m]}(\theta, \phi)\} - \mathbf{f}_m^R(\mathbf{a}(\theta, \phi)))^2}{(1 - \frac{L_m^R + p \frac{L_m^R - 1}{2}}{|\mathcal{S}|})^2}, \quad (1.85)$$

$$\Im\{\text{GCV}\} = \frac{\frac{1}{|\mathcal{S}|} \sum_{(\theta, \phi) \in \mathcal{S}} (\Im\{\bar{\mathbf{a}}_{[m]}(\theta, \phi)\} - \mathbf{f}_m^I(\mathbf{a}(\theta, \phi)))^2}{(1 - \frac{L_m^I + p \frac{L_m^I - 1}{2}}{|\mathcal{S}|})^2}, \quad (1.86)$$

where \mathbf{f}_m^R and \mathbf{f}_m^I are the real and imaginary part of the model for the m -th antenna, respectively. Both $\Re\{\text{GCV}\}$ and $\Im\{\text{GCV}\}$ are used to avoid overfitting, that is, to reduce the complexity of the model created in order to avoid fitting of the noise or of measurement errors of the array response. $\bar{\mathbf{a}}_{[m]}(\theta, \phi)$

denotes the m -th element of the virtual array response. To use this approach in an real-time system the models can be built in an initialization step. Sectors of interest for the application of the MARS interpolation can be selected and a model can be built and stored as a look-up table. The data received can be interpolated snapshot-wise with

$$\begin{bmatrix} \Re\{\bar{\mathbf{x}}[n]\} \\ \Im\{\bar{\mathbf{x}}[n]\} \end{bmatrix} = \mathbf{f}(\mathbf{x}[n]), \quad (1.87)$$

where $\mathbf{x}[n] \in \mathbb{C}^{M \times 1}$ is the n -th column of matrix \mathbf{X} with $n = 1, \dots, N$.

In some cases, the antenna array will have a response which is similar enough to the desired response in some of its field of view so that a linear approach yields good results. Thus, the MARS approach can be used only on portions of the field of view where the real response differs strongly from the desired one.

The MARS approach is especially useful for arrays with a small number of elements. For linear approaches, the degrees of freedom of the linear approach are limited by the number of antennas in the virtual array. It is possible to increase the degrees of freedom by creating a virtual array with a larger number of antennas, however, this will lead to a transformation matrix that is ill-conditioned and results in a large bias in the DOA estimates. The MARS approach, on the other hand, can build as many hinge functions and relationships between them as necessary.

1.10.2 GRNN based interpolation

GRNNs are a type of neural network used for general regression problems that are extended in this work for the problem of array interpolation. An overview of GRNNs is presented in Appendix F.

Classic interpolation relies on transforming an array response over the discrete set of angles \mathcal{S} . In linear array interpolation, the choice of points that belong to \mathcal{S} is usually an arbitrary one, based solely on the predefined sector bounds and the chosen angular resolution. When applying a linear approach, angular resolution is usually chosen to be as high as possible, thus, the set of points belonging to \mathcal{S} is usually closely spaced. This is not necessary when applying the GRNN since it is capable of interpolating between multiple sets of training data. In fact, a dense set of training points will result in an increased computational load while providing very little benefit to the overall accuracy of the GRNN based array interpolation.

The neural network representing the real and imaginary parts of the array can be built by extending (F.1) and (F.2) to the application at hand, resulting in the mapping

$$\begin{bmatrix} \Re\{\bar{\mathbf{x}}[n]\} \\ \Im\{\bar{\mathbf{x}}[n]\} \end{bmatrix} = \mathbf{f}(\mathbf{x}[n], \bar{\mathbf{a}}(\theta, \phi), \mathbf{a}(\theta, \phi)), \quad (1.88)$$

and

$$\Re\{\bar{\mathbf{x}}_{[m]}[n]\} = \frac{\sum_{(\theta,\phi) \in \mathcal{S}} \Re\{\bar{\mathbf{a}}_{[m]}(\theta,\phi)\} e^{-\frac{\xi(\theta,\phi)}{2\sigma^2}}}{\sum_{(\theta,\phi) \in \mathcal{S}} e^{-\frac{\xi(\theta,\phi)}{2\sigma^2}}}, \quad (1.89)$$

$$\Im\{\bar{\mathbf{x}}_{[m]}[n]\} = \frac{\sum_{(\theta,\phi) \in \mathcal{S}} \Im\{\bar{\mathbf{a}}_{[m]}(\theta,\phi)\} e^{-\frac{\xi(\theta,\phi)}{2\sigma^2}}}{\sum_{(\theta,\phi) \in \mathcal{S}} e^{-\frac{\xi(\theta,\phi)}{2\sigma^2}}}, \quad (1.90)$$

where σ^2 is the smoothness parameter, and

$$\xi(\theta,\phi) = \|\mathbf{x}[n] - \mathbf{a}(\theta,\phi)\|_2^2. \quad (1.91)$$

Once the data has been interpolated to the desired array structure, array signal processing schemes such as FBA, SPS or ESPRIT can be applied. Differently, from other array interpolation methods, there is no noise color introduced by the interpolation, therefore pre-whitening is not necessary.

One of the advantages of applying GRNNs to array interpolation is that the sectors considered for transformation can be made very large while having a small effect on the accuracy of the interpolation. As shown in (1.89, 1.90), in case the input resembles the data used for the training in a certain region of the transformed sector, this data will have a larger weight. On the other hand, if the input is distant from any region of the transformed sector, this region will have very little influence on the output of the interpolation.

Therefore, with a GRNN it is possible to transform very large regions of the field of view of the array without suffering from the same problems as the ones present in linear array interpolation, where the transformation error grows as the transformed region grows. On the other hand, transforming a large region using the GRNN will increase the computational cost of the interpolation. As shown in Figure F.1 it is possible to implement the proposed GRNN array interpolation in a parallel manner, leading to a faster interpolation. Thus, the proposed GRNN array interpolation is a possible alternative for precise array interpolation in real-time applications.

1.11 Numerical Simulation Results

This section presents a performance assessment of the proposed multidimensional linear approach in Subsection 1.11.1 and of the proposed adaptive linear and nonlinear approaches in Subsection 1.11.2.

1.11.1 Multidimensional Linear Performance Results

In order to study the performance of the proposed multidimensional method a set of numerical simulations is performed. Three signals impinge on a 8×8 planar array. The spatial frequencies are set to $(-\frac{\pi}{3}, \frac{\pi}{3})$, $(\frac{\pi}{5}, \frac{\pi}{5})$ and $(\frac{\pi}{3}, -\frac{\pi}{3})$.

All three signals have a correlation coefficient of 0.8 with each other. As shown in (8) and in (11), the array response is corrupted by inducing a random phase delay at each sensor element. In this section, a normal distribution is assumed for these phase delays and the goal is to compensate these phase delays by applying the proposed multidimensional interpolation scheme. Additionally, an extra unknown phase delay with standard deviation equal to $\frac{1}{10}$ of the standard deviation of the known errors is introduced. The proposed technique is compared to the results of an array with no errors introduced (perfect array), to the interpolation method from Section 1.8 using a matrix approach extending the optimization problem to two dimensions, and to the corrupted array when no interpolation method is applied. For the matrix method, the transformation matrix is calculated according to (1.55) taking into account $[\mathcal{A}]_{(R+1)}$ and $\bar{\mathbf{A}}$, that is, the total steering matrices are used. The DOAs are estimated using the PARAFAC decomposition to estimate the factors and the translation invariance for DOA estimation. Estimations are performed using $N = 100$ and the final RMSE is the average with respect to 1000 Monte Carlo runs.

Figure 1.14 presents the results for phase errors following a normal distribution with standard deviation of $\frac{\pi}{20}$ radians. The results show that, for low SNRs where the noise is the main source of bias, the proposed technique is able to provide results similar to those obtained with a perfect array. However, the non interpolated array is subject to a constant bias due to the phase delay present at the sensor elements.

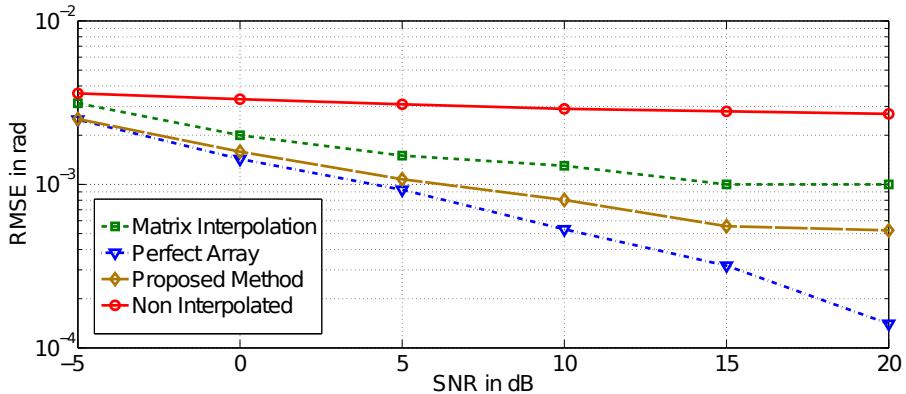


Figure 1.14: Results for a standard deviation of $\frac{\pi}{20}$ radians

Figure 1.15 presents the results for phase errors following a normal distribution with standard deviation of $\frac{\pi}{10}$ radians. The performance of the proposed interpolation method with comparison to the perfect array worsens as the errors get larger. Also, it is noticeable that the proposed method introduces a

bias that saturates the accuracy of the estimation approximately at the 5 dB mark. This bias is caused by the transformation error.

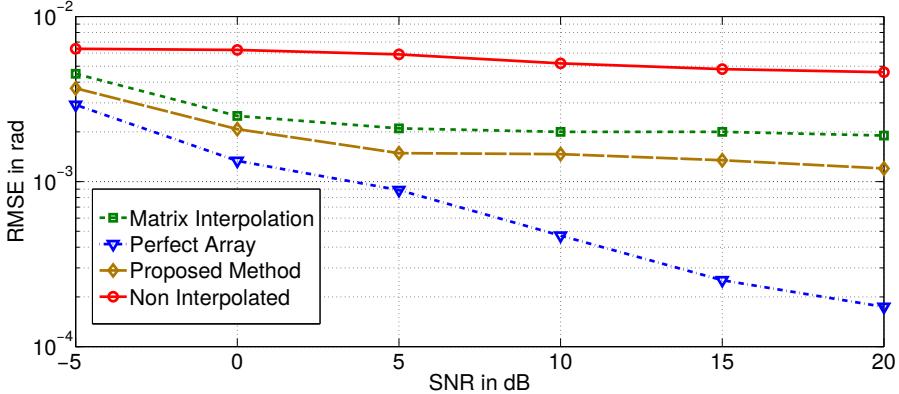


Figure 1.15: Results for a standard deviation of $\frac{\pi}{10}$ radians

Finally, Figure 1.16 presents the results for phase errors following a normal distribution with standard deviation of $\frac{\pi}{5}$ radians. The performance of the proposed method worsens considerably, as it becomes hard to compensate for the large phase errors introduced by each sensor elements. However, the performance of the proposed method is still superior to the estimation done without employing array interpolation.

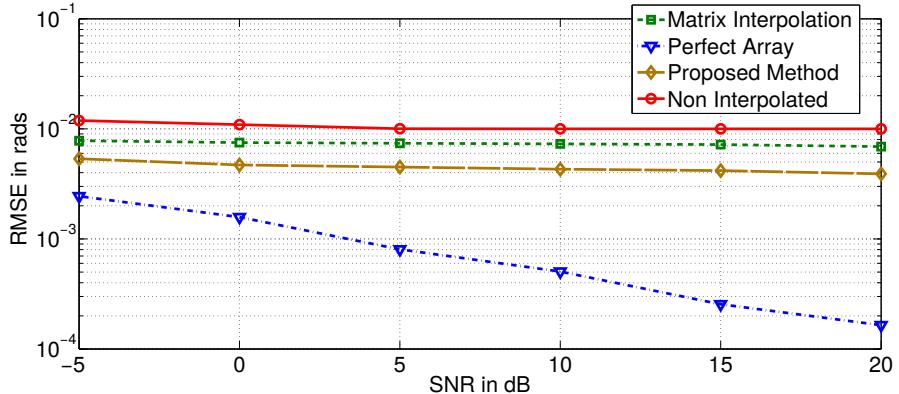


Figure 1.16: Results for a standard deviation of $\frac{\pi}{5}$ radians

In all simulations the proposed method outperformed its matrix counterpart. In the proposed method, the interpolation matrices are calculated for each dimension in a independent way. In the matrix approach, however, the dimensions are not decoupled and a total transformation matrix must be cal-

culated taking into account all the permutations possible between angular responses from each dimension. Thus, the calculation of this transformation matrix is highly overdetermined resulting in a larger transformation error. Another advantage of the proposed method is that the optimization problems are one dimensional, resulting in a smaller computational load.

1.11.2 Nonlinear Performance Results

The performance of the proposed algorithms is tested considering a manufactured 6×1 linear antenna array shown in Figure 1.17 [15]. The single antennas are circularly polarized dielectric resonator antennas (DRA) [14], exhibiting a strong miniaturization of the aperture. The antennas are tuned to work at GPS L1 frequency band centered around 1575.42 MHz in standalone configuration. Once in the array, no attempt has been made to re-tune them, in order to emphasize the mutual coupling effects and eventual degradation. This 6×1 linear array was especially designed for research on mutual coupling with adjustable inter-element spacing of the 6 antennas. The small dimensions of the single antennas allow to place them quite close to each other, up to a minimum distance of 30 mm ($\approx 0.16\lambda$ at L1 central frequency). The following simulations consider array responses of the manufactured array measured in an anechoic chamber as well as array responses simulated using a 3D electromagnetic simulator (ANSYS HFSS). Thus, realistic assessment of the proposed algorithms can be achieved considering measured and properly simulated array responses.

The simulations are performed considering the receiver has no previous knowledge of the received signal. The combined sector is detected and discretized according to the methods presented in Section 1.7. After discretization the combined sector is interpolated using the linear and nonlinear methods proposed in Sections 1.7 and 1.10. The model order of the interpolated data then is estimated by applying FBA and SPS adaptively and using RADOI [91] as the model order estimation method following the procedure proposed in [72]. Once model order is estimated the DOA estimates are obtained using TLS-ESPRIT [93].

For obtaining $\hat{\mathbf{R}}_{\mathbf{XX}}$ $N = 100$ snapshots are used and the Root Mean Squared Error (RMSE) is calculated with respect to 1000 Monte Carlo simulations. Two signals from $\theta_1 = 45^\circ$ and $\theta_2 = 15^\circ$ with $\sigma_1^2 = \sigma_2^2 = 1$ and $\gamma_{1,2} = 1$ according to (1.4) are impinging on the array. The given RMSE is

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{k=1}^K \left((\hat{\theta}_1^k - \theta_1)^2 + (\hat{\theta}_2^k - \theta_2)^2 \right)}, \quad (1.92)$$

where $\hat{\theta}_i^k$ is the estimate of θ_i at the k -th Monte Carlo run.

The Signal-to-Noise-Ratio (SNR) is defined as

$$\text{SNR} = \frac{\sigma_1^2}{\sigma_n^2} = \frac{\sigma_2^2}{\sigma_n^2} = \dots = \frac{\sigma_d^2}{\sigma_n^2}. \quad (1.93)$$

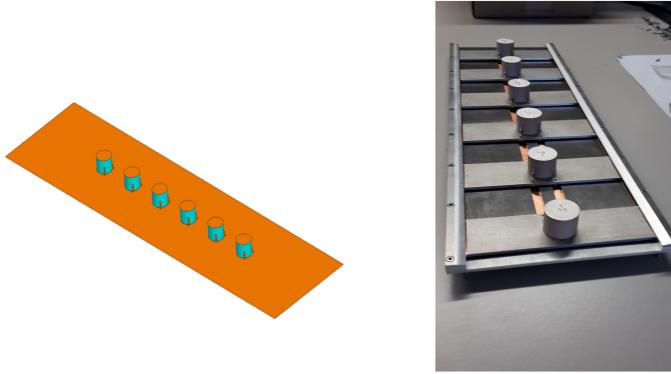


Figure 1.17: 6x1 antenna array: ANSYS HFSS model (left), prototype (right).

Performance of Nonlinear Methods With Measured Array Response Knowledge

The first set of simulations studies the performance of the proposed methods using the measured array response to build the interpolation models. Figure 1.18 presents the performance of the discretization and interpolation methods proposed in this work when the array is measured with 0.4λ inter-element spacing. As a benchmark, the results achieved with no interpolation, denoted as raw data in the figures, with the linear interpolation approach from [72], and the Cramér–Rao lower bound (CRLB) [87] are presented. The results show that, if no interpolation is applied, a constant bias is present for the DOA estimates even at high SNRs. The linear approach is capable of providing a DOA estimation accuracy smaller than one degree at positive SNRs. The proposed nonlinear approaches offer significantly improved performance when compared to the linear approach. The MARS approach is shown to benefit significantly from the UT discretization. This is expected as the UT discretization reduces the overall complexity and possible over-fitting of the MARS model, as the mapping needs to consider only the statistically significant regions of the sector. The proposed GRNN approach is shown to have a very similar performance when either PCA or the UT are used as a discretization method. As the computational complexity of the PCA-based discretization method is significantly smaller than that of the UT-based method, PCA becomes more attractive when the GRNN method is used for interpolation.

In the second set of simulations, presented in Figure 1.19, the inter-element spacing between is set to 0.2λ . In this configuration the mutual coupling between antenna elements is very strong and the radiation patterns are highly distorted when compared to isotropic patterns. The difference in performance

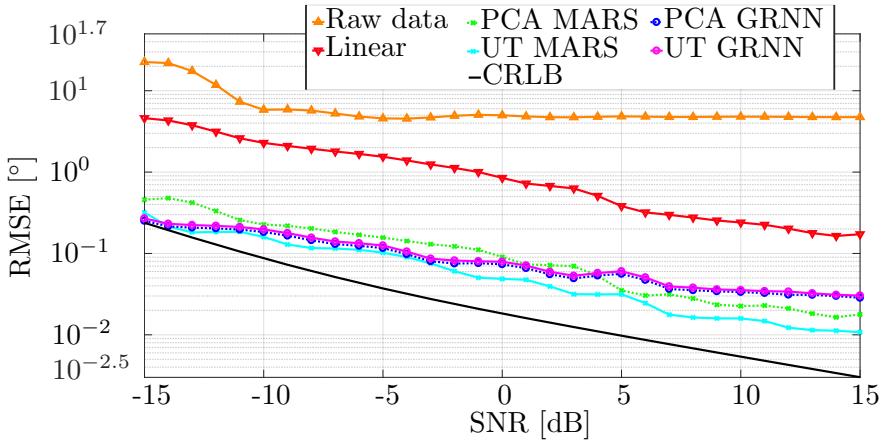


Figure 1.18: DOA estimation performance with 0.4λ element separation

between the proposed nonlinear methods and the bench mark linear method becomes more significant when compared to the 0.4λ inter-element spacing. This challenging scenario also highlights the performance difference between the proposed discretization and interpolation methods. The performance of the MARS based method with the UT pre-processing significantly outperforms its PCA variant.

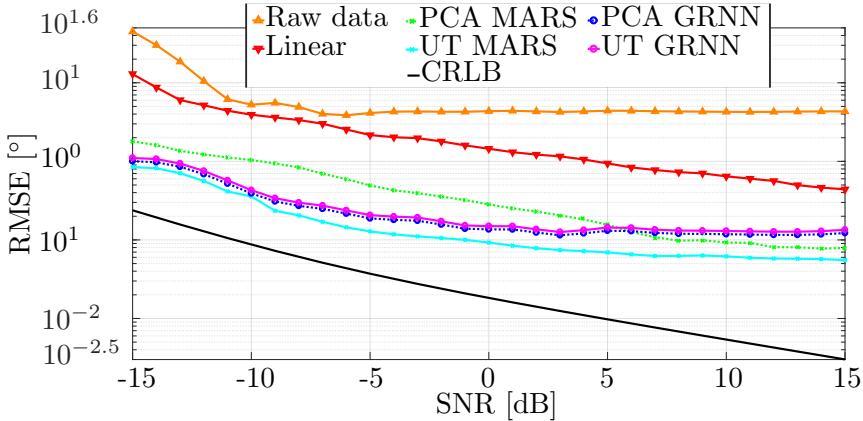


Figure 1.19: DOA estimation performance with 0.2λ element separation

Performance of Nonlinear Methods With Simulated Array Response Knowledge

Obtaining the response of an antenna over its entire field of view with a fine enough resolution for performing array interpolation requires a well equipped anechoic chamber. In case such a chamber is not available, the interpolation can be done using a properly simulated array response. With the next set of simulations the performance of the proposed methods if the array response that is considered by the algorithms is not the true (measured) array response but is derived from simulations using a 3D electromagnetic simulator (ANSYS HFSS) is studied. The results shown in Figure 1.20 were obtained by setting the inter-element spacing of the antennas to 0.4λ . The simulation results of the array interpolation methods show that the performance of all methods is degraded by the imperfect knowledge of the array response (simulated array response). Despite the decrease in performance, the proposed methods are capable of achieving an accuracy of a tenth of a degree for moderate SNRs.

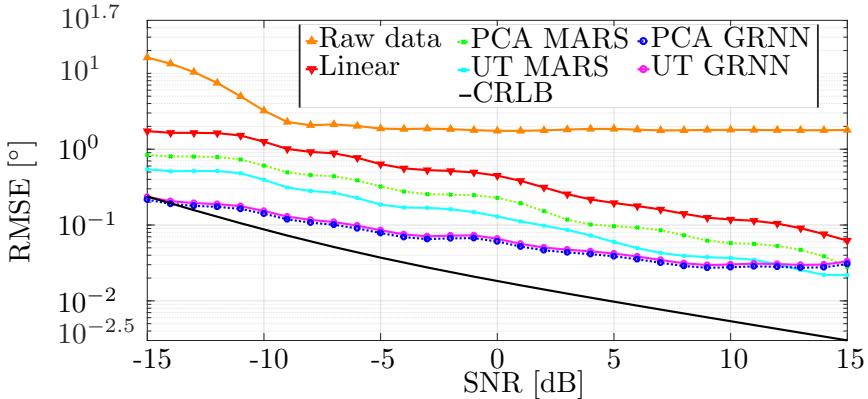


Figure 1.20: DOA estimation performance for model built with simulated data with 0.4λ element separation

Figure 1.21 presents the results of the simulations when the inter-element is 0.2λ and the model is built using a simulated array response. In this case the proposed PCA approaches suffer heavy degradation. This is due to the fact that the principal components of the simulated array response do not represent the principal components of the true array response (measured array response) well enough for building a proper model. In this scenario the linear approach is not capable of providing an accuracy below one degree.

Failure Rate Performance after Nonlinear Interpolation

In the last set of simulations the failure rate of the TLS-ESPRIT used for DOA estimation is studied. For this simulation a single signal is received by the array

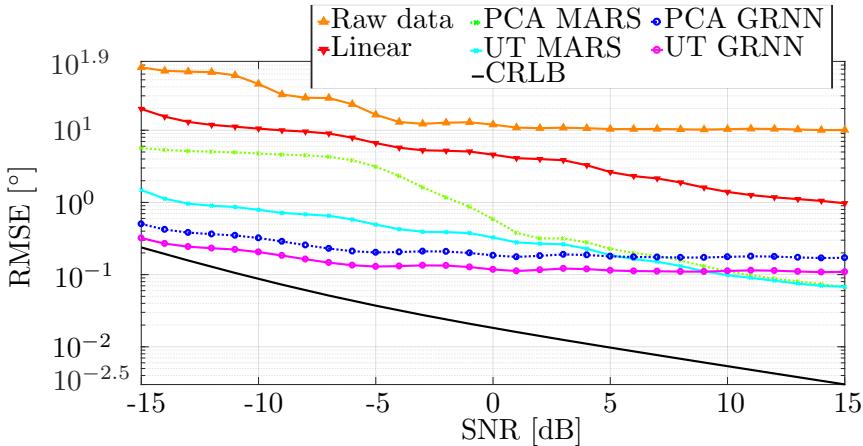


Figure 1.21: DOA estimation performance for model built with simulated data with 0.2λ element separation

with $\theta = 80^\circ$ and the array has an inter-element spacing of 0.2λ . Figure 1.22 compares the failure rates of non-interpolated data with interpolated data using the UT-based discretization and the MARS method. For these simulations TLS-ESPRIT is considered to fail if the eigendecomposition of the solution for the shift invariance equation of TLS-ESPRIT yields complex eigenvalues. The simulations show that the proposed interpolation method is capable of significantly reducing the failure rate of subspace based methods for regions near end-fire of the array.

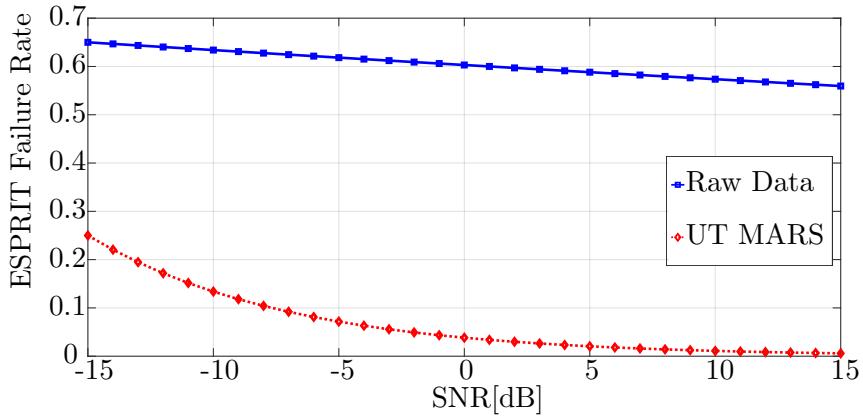


Figure 1.22: TLS-ESPRIT failure rate for 0.2λ element separation

1.12 Summary

This chapter presented multiple array interpolation algorithms with the goal of allowing low bias DOA estimation to be performed.

The first step is to detect sectors of interest, to this end a method for signal adaptive sector selection was presented. The proposed method scans the field of view of the array to detect and determine the sectors of interest for interpolation. Once the sectors have been detected they must be discretized before being transformed, in this chapter two alternatives for sector discretization, relying on the UT and the PCA were presented. These discretization methods are capable of discretizing the sector while maintaining the statistical integrity of the transformed data.

An adaptive linear approach for array interpolation was presented, aiming at minimizing the transformation bias with respect to the received signal. This approach was extended to multilinear arrays of arbitrary dimensions.

Nonlinear approaches have also been presented for the cases where the real array response deviates too heavily from the desired array response. A nonlinear interpolation method based on the MARS regression was presented capable of achieving good performance at moderate to high SNRs. A second nonlinear interpolation approach applying GRNNs was presented, with a reduced computational complexity when compared to the MARS method, the GRNN method was shown to provide the best performance under low SNR scenarios.

The efficiency of the proposed methods was tested by means of numerical simulations using a linear array with responses measured in an anechoic chamber as well as responses simulated in a 3D electromagnetic simulator (ANSYS HFSS), taking into account all realistic effects, such as mutual coupling between the elements, and non-perfect radiation patterns. Furthermore, simulations were performed using real measured array patterns, validating the performance of the proposed methods under real life conditions.

The methods were capable of accurate DOA estimation and reduced the estimation bias when compared to a noninterpolated array. Linear estimation, while offering the worst performance among the methods presented, has a very low complexity. Nonlinear estimation methods can offer improved performance when needed, at the cost of greatly increased computational complexity.

Chapter 2

Cooperative MIMO for Wireless Sensor Networks

In this chapter the following research question is addressed:

Can array processing techniques be used to provide improved lifetime and communication performance in wireless sensor network?

2.1 Overview and Contribution

The remainder of this section is organized as follows:

- **Motivation:** An introduction to wireless sensor networks and how the challenges they face can be addressed using array processing is presented in this section.
- **Wireless Sensor Networks Organization:** This section presents a brief description of the usual organization inside WSNs
- **Cooperative MIMO:** Here, the concept of Cooperative MIMO (C-MIMO) is detailed. The mathematical model behind MIMO is addressed and the steps involved in performing MIMO over a distributed set of devices are presented.
- **Energy Analysis:** In this section the energy consumption of single-hop, multi-hop and C-MIMO communications is studied. The threshold for when to switch between different communication techniques is derived to maximize the energy efficiency with respect to the distance involved in the communication.
- **Synchronization:** Here, the critical problem of synchronization for a distributed MIMO system is studied. The effects of synchronization errors in the performance of MIMO transmissions is analysed. Two different

methods for synchronization are presented as detailed, allowing for different trade-offs between precision and computational complexity. The error propagation involved in the synchronization of nodes over multiple hops is also addressed. Finally, a closed formula for the minimum resynchronization interval with respect to clock drift and symbol duration is derived.

- **Adaptive C-MIMO Clustering:** This section presents an approach for adaptively forming C-MIMO communication clusters based on the distance involved in the communication and on the energy reserves of neighbouring nodes. This approach seeks to maximize the energy efficiency of communications in the network while keeping the energy distribution between the nodes as even as possible.

The research contributions presented in this chapter are:

1. A study on the energy efficiency of C-MIMO with respect to multi and single-hop communications, defining a threshold of selection between the multiple communication techniques.
2. An analysis on the impact of synchronization errors in distributed MIMO system with respect to the symbol period and the size of the communication clusters involved.
3. A study on the propagation of synchronization errors inside a WSN, and from that the derivation of a closed form expression regarding the resynchronization interval in distributed MIMO systems that takes into account the symbol duration and the clock drift of the nodes involved.
4. Two methods for synchronization for WSNs with different trade-offs between performance and computational complexity.
5. An adaptive method for selecting the cluster size of C-MIMO transmissions capable of maximizing energy efficiency and improve the homogeneity of the energy reserves across the network.

2.2 Motivation

Recent advances in electronics and miniaturization have allowed the development of small, low-powered and low-cost devices capable of small range communication, a limited amount of processing and of collecting data about its surroundings. Wireless sensor networks (WSNs) are a collection of such devices, called nodes, organized in a cooperative network environment. They possess various advantages over traditional sensor applications. Due to their low cost and small size they can be deployed densely and randomly to monitor a certain phenomenon. On the other hand, traditional sensors usually require

extensive planning in their deployment and, depending on the nature of the phenomenon to be sensed, cannot be placed near it due to their high cost. WSNs are usually self-contained and capable of organizing the network on the fly. Due to this characteristics, these networks are suitable for deployment on harsh environments.

WSN usage ranges from military applications, such as battlefield surveillance and targeting, to healthcare applications, such as portable automated drug infusion systems in hospitals [1]. However, the extensive use of WSNs is still being hindered by the limited energy budget available in the nodes that constitute the networks. Furthermore, since WSNs are deployed even in harsh and hard-to-reach environments, replacing individual nodes, or their batteries, is not always practical. Therefore, extensive research is being performed to maximize the energy efficiency of WSNs [34].

Solutions for minimizing energy consumption in WSNs have been proposed through different layers by analyzing energy efficiency for all the tasks involved in WSNs. Given that communication is the most energy consuming task, diminishing the need for communication and increasing the efficiency of communication are the promising directions to achieve WSNs with longer lifetime [82]. While many proposals focus on enhancing energy efficiency in the network layer using energy efficient routing protocols [44, 50], some focus on enhancing the communication efficiency by proposing energy efficient protocols for medium access control (e.g., [18, 66]). Other proposals consider energy aware processing approaches for communication among sensor nodes in WSNs, as presented in [107], as well as the application of alternative solutions within the physical layer [97, 98].

Multi-hop communication is a widely used technique to obtain improved energy efficiency and to maximize network lifetime by spreading energy consumption over different nodes [17]. Multi-hop techniques take advantage of the cooperative nature of WSNs in splitting the distance involved in communication by employing intermediary nodes to forward data packets. Since free space loss is not linear, splitting the distance by using intermediary nodes results in reduced power demand, thus minimizing transmission energy consumption.

Another proposal made by taking advantage of the cooperative nature of WSNs is the formation of cooperative Multiple-Input Multiple-Output (C-MIMO) clusters [23]. The faster data transmission rates achievable with C-MIMO system allow for interaction between fast-moving mobile nodes and traditional static ones [76, 77]. C-MIMO also allows the application of antenna array techniques such as beamforming or DOA estimation. The usage of C-MIMO results in reduced hardware complexity on single nodes as it is not necessary to employ multiple antennas and radios at a single node. Thus, it is possible to use C-MIMO with no modifications to the hardware of existing sensor nodes. The complexity is transferred to the software, which is responsible for managing the communication involving a large number of nodes.

Keeping in view the characteristics and advantages of C-MIMO, a number of problems that come in the way of effectively implementing C-MIMO in WSNs are studied in this thesis and solutions are proposed. The proposed solutions deal with problems that may arise when implementing C-MIMO in WSNs, such as time synchronizing of sensor nodes, routing for C-MIMO communication, adaptively forming C-MIMO clusters and analyzing the energy efficiency of C-MIMO. This work aims at fostering implementation of C-MIMO for different applications and presenting a performance analysis of C-MIMO for WSNs, considering relevant technical aspects. Furthermore, the energy efficiency of a C-MIMO network, operating on top of the proposed framework, is studied based on realistic node specifications.

2.3 Wireless Sensor Networks Organization

WSNs are at their very principle decentralized, ad hoc and self contained systems. Decentralization, however, does not imply a lack of organization in WSNs, it refers only to the capacity of the network to function properly in the event that any of its components cease to function, i.e, the network does not depend on a single specific element to work. The ad hoc capability implies that a new element can join the network and operate on it without any prior configuration. Being self-contained means that wireless sensor networks do not rely on any external system to operate.

With these three characteristics in mind, an analysis can be made on the usual methods of organizing a WSN. These methods are usually divided into two categories, a peer-to-peer or planar architecture and a cluster based or hierarchical architecture.

In peer-to-peer networks there exists no distinction or separation between members, they often possess exactly the same capabilities and resources. This very loose environment distances itself from usual network organizations and provides a more homogenous and locally autonomous framework. Peer-to-peer network members interact with each other to share more than information but also all available resources such as processing power and storage capacity. However, the lack of any "leader" element also means that the task of making decisions as simple as allocating IP addresses to new members can be quite complex.

Cluster based networks, on the other hand, are closer to the conventional network architectures. In cluster based networks the elements are usually divided into two categories: Full Function Devices (FFD) and Reduced Function Devices (RFD). RFDs are elements with limited resources, usually only possessing a sensing element, a radio for wireless communication and the processing power necessary only to perform the said sensing and forwarding of this captured information. FFDs are more powerful elements, they are responsible for organizing clusters of RFDs, routing and repeating information across the network as well as performing in any network processing necessary, the sen-

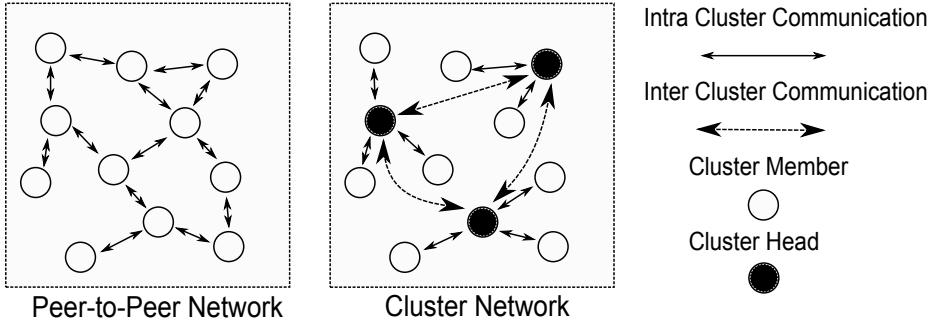


Figure 2.1: Examples of peer-to-peer and cluster networks

sors performing this tasks are known as cluster heads. Although cluster based networks require a more strict type of organization than their peer-to-peer counterpart, they offer a simple way to deal with problems that were arduous on the peer-to-peer networks, as decision making and resource allocating tasks are performed by the FFDs. A cluster network might be comprised entirely of FFDs, which nodes assuming cluster head function depending on various metrics, such as available resources or a pre-configured priority order. In the event that a cluster head ceases to work another FFD will assume its functions.

Regardless of the network organization WSNs are highly cooperative environments with elements working towards a common goal. Due to its highly strict energy budget it is specially important that sensors cooperate when communicating to minimize energy consumption and increase network lifetime.

2.4 Cooperative MIMO

C-MIMO differs from usual MIMO implementations due to the fact that the antenna arrays are present at a single device, but rather made out of the individual antennas present on separate devices. By introducing the concept of a MIMO cluster, a virtual MIMO device for implementing MIMO communications is created. This concept is illustrated in Figure 2.2.

In C-MIMO, a transmit node recruits nearby nodes to aid in data transmission, whereas the destination node recruits nearby nodes to aid in the reception and decoding of data. Reception can also be performed by a single node, resulting in a SIMO configuration. SIMO can be used to achieve longer distances in communication, or to split the energy necessary in long range communication between various nodes. On the other hand, a single node might transmit to a receiving cluster to minimize the bit error rate (BER) at long range, resulting in a MISO configuration.

When employing C-MIMO, the transmitted data can come from a single node or various nodes of the transmitting cluster. Figure 2.3 presents the steps necessary for a C-MIMO communication. The first step, represented by ① in

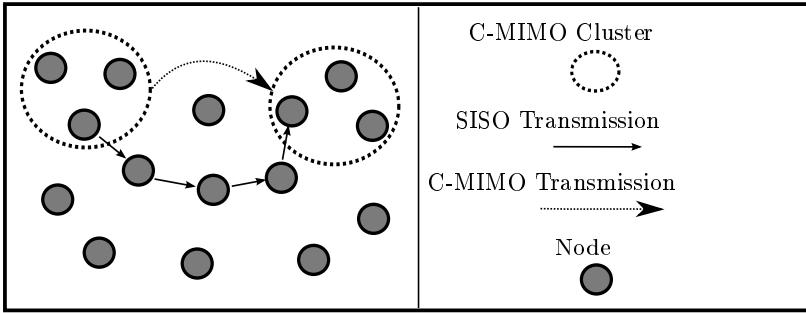


Figure 2.2: Example of cooperative MIMO and SISO communications in a wireless sensor network

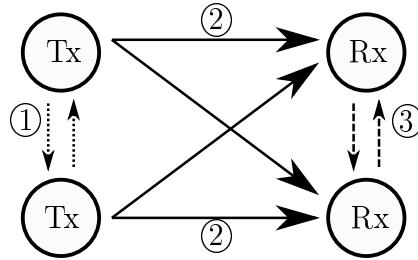


Figure 2.3: C-MIMO communication steps

Figure 2.3, is the intra-cluster communication at the transmit cluster (intra-tx) and involves synchronizing and exchanging data that needs to be transmitted among the transmitting nodes. Symbols are received and decoded before being forwarded in the next step, i.e, a decode and forward approach is used. If both sensors need to transmit data, this exchange is not necessary, as each sensor can transmit its own data. Nodes are assumed to be equipped with omnidirectional antennas. Assuming that the nodes remain static, synchronization is achieved according to the method proposed in Section 2.6.

In step ② in Figure 2.3, in inter-cluster communications (inter), both sensors transmit different symbols, during the same time slot, using a MIMO scheme. This work proposes the use of the so-called V-BLAST transmission, as described in [111], for MIMO communications to exploit spatial diversity. In normal transmissions, a single symbol would be transmitted over the channel during each time slot, whereas in the case of V-BLAST, the symbols are grouped into frames of the size of the receive array. In case of a C-MIMO cluster with M_t transmit nodes, frames of the size M_t are transmitted in the same time slot.

The received signal at the i -th receive node $x_i \in \mathbb{C}$, during a given time slot can be modeled as

$$x_i = \sum_{k=1}^{M_t} h_{i,k} \cdot s_k + n_i \in \mathbb{C}, \quad (2.1)$$

where $h_{i,k}$ represents the complex-valued impulse response of channel between transmit node k and receive node i , and s_k is the symbol transmitted by the k -th node. $n_i \in \mathbb{C}$ is the noise present at the i -th receive node during sampling. Equivalently, a matrix representation for the signals received at all receiving nodes can be written as

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{n}, \quad (2.2)$$

where $\mathbf{x} \in \mathbb{C}^{M_r \times 1}$ contains the signal received at each node of the M_r receive nodes, $\mathbf{s} \in \mathbb{C}^{M_t \times 1}$ is the vector of the symbols transmitted, and $\mathbf{n} \in \mathbb{C}^{M_r \times 1}$ contains the noise samples at each of the receive nodes, and

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,M_t} \\ \vdots & \ddots & \vdots \\ h_{M_r,1} & \cdots & h_{M_r,M_t} \end{bmatrix} \in \mathbb{C}^{M_r \times M_t} \quad (2.3)$$

is the channel matrix. To estimate the transmitted symbols, \mathbf{H} must be estimated. Traditionally, an estimate $\hat{\mathbf{H}}$, obtained by transmitting a set of pilot symbol vectors $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_U] \in \mathbb{C}^{M_t \times U}$, where $\mathbf{p}_i \in \mathbb{C}^{M_t \times 1}$, and assuming $U > M_r$ is given by

$$\hat{\mathbf{H}} = \mathbf{X}\mathbf{P}^\dagger, \quad (2.4)$$

where $(\cdot)^\dagger$ represents the pseudoinverse of a matrix. In this work, no precoding is used during step ② in Figure 2.3. Note that although precoding helps in improving the efficiency of C-MIMO, it entails higher computational complexity. For a more detailed discussion on trade-offs and optimal pilot symbol selection for MIMO channel estimation, the reader may refer to [7].

Finally in step ③ in Figure 2.3, at the receive cluster (intra-rx), the receive nodes exchange the received information to decode the received symbol. The binary output of the analog digital converter of the node receiver is sent to the node or nodes responsible for equalizing the received signal. Once the channel matrix estimate $\hat{\mathbf{H}}$ is obtained, the receiving cluster needs to equalize the received symbols to obtain an estimate of the transmitted symbols \mathbf{s} . Various methods exist for performing this equalization. Among them, this work evaluates the performance of Zero Forcing (ZF) [113] and Minimum Mean Square Error (MMSE) [56] with their respective successive interference cancellation (SIC) extensions [96, 114], and Maximum Likelihood (ML) [115] methods. If the data is destined to only one sensor of the receiving cluster, this exchange becomes unidirectional. Another option is to exchange only a portion of the

received information so that every sensor is responsible for decoding, thus alleviating the computational burden of a single node.

Figure 2.4 depicts a comparison between standard SISO systems and a 2×2 MIMO configuration for different equalization methods. For the simulations in Figure 2.4, 300 symbols are transmitted using a binary phase shift keying (BPSK) constellation, 30 of which are used to perform channel estimation. The communication channels are generated using independent Rayleigh distributions for each transmitter-receiver pair, which are considered to be constant during a transmission interval. The ML equalization method is the most efficient one in terms of minimizing the BER, followed by the MMSE method, which employs the SIC. The ZF presents worse results, with a BER being inferior to that of the SISO case. This is due to the noise enhancement that follows ZF equalization. In the best case scenario, with perfect knowledge of a completely uncorrelated channel matrix, the ZF BER curves superpose the SISO curves [104]. The SISO performance shown in Figure 2.4 is obtained by the theoretical SISO BER for a transmission over a Rayleigh channel. With MIMO, the same BER can be reached at lower SNRs; therefore, less power is required to achieve the same BER at the receiver. For the remainder of this work, the MMSE with SIC is considered as the equalization method, since its performance is similar to that of the ML for intermediary SNR regimes and, in general, less complex.

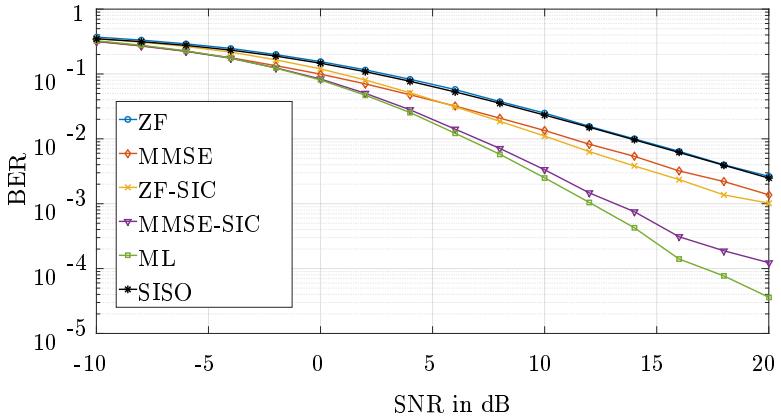


Figure 2.4: Performance comparison between standard SISO systems and 2×2 MIMO systems using ZF, MMSE and ML equalization

2.5 Energy Analysis

This section studies the energetic performance of communication using conventional and C-MIMO techniques.

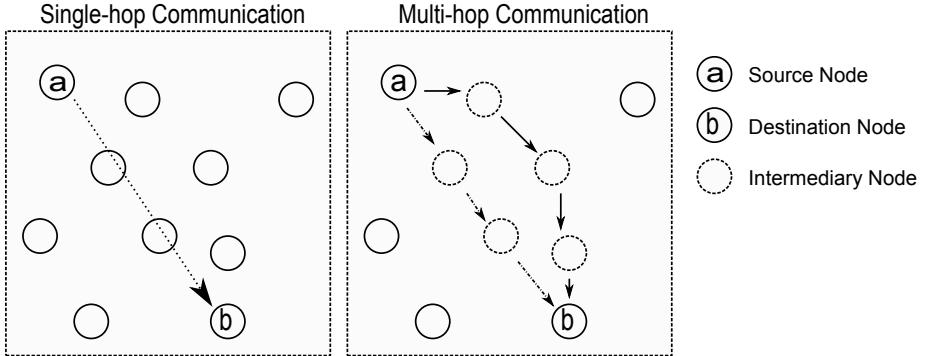


Figure 2.5: Examples of single-hop and multi-hop communication

2.5.1 Conventional Techniques

In standard WSNs communication is usually done using either single-hop transmissions or multi-hop transmissions. Single-hop transmissions consist of end-to-end communications without aid of intermediary nodes while multi-hop transmissions consist of using multiple intermediary nodes as routers to convey the necessary data to the destination node.

Figure 2.5 shows an example of single and multi-hop configurations. Note that the multi-hop configuration presents multiple available paths for signal transmission, optimal path selection is a problem related to the routing protocol, the reader may refer to [20, 69, 112].

The increased efficiency provided by multi-hop communication is due to the fact that the attenuation suffered by a wireless signal increases exponentially with the distance. However, an analysis must be made to determine the point where multi-hop ceases to outperform single-hop. According to the energy consumption model proposed in [17]

$$E_t = \alpha, \quad (2.5)$$

$$E_r = \beta, \quad (2.6)$$

where E_t is the energy necessary for transmitting a single symbol over a certain distance and E_r is the energy necessary for receiving and decoding the given symbol. The parameter α is directly dependent on the distance between the transmitting and receiving nodes and can be written as

$$\alpha = \begin{cases} a + b \cdot d^\gamma & ; d_{max} \geq d > d_{min} \\ a + b \cdot d_{min}^\gamma & ; d \leq d_{min} \end{cases} \quad (2.7)$$

here d_{min} defines the maximum distance that can be reached by setting the transmit power of the transmitting node radio to its minimum configurable value, d_{max} is the maximum reachable distance by setting the transmit power

Mode	Current
Rx	$7.03mA$
Tx (-18 dBm)	$3.72mA$
Tx (-13 dBm)	$5.21mA$
Tx (-10 dBm)	$5.37mA$
Tx (-6 dBm)	$7.05mA$
Tx (-2 dBm)	$8.47mA$
Tx (0 dBm)	$11.57mA$
Tx (+3 dBm)	$13.77mA$
Tx (+4 dBm)	$17.37mA$
Tx (+5 dBm)	$21.48mA$
Tx (+10 dBm)	$27.01 mA$

Table 2.1: Energy consumption for the Mica2 platform

to its highest configurable value, d is the distance between the transmitting and receiving nodes and γ is the path loss coefficient. To evaluate the energy efficiency of both techniques a comparison is made a transmission over d_{max} using both techniques. Let $d_{max} = k \cdot d_{min}$ and the path loss coefficient be equal to the free space loss coefficient $\gamma = 2$ and the power necessary for receiving a signal being equivalent to the power necessary for minimal transmission $\beta = a$. In fact receiving is usually more energy demanding than transmitting at low power as seen on table 2.1.

The total energy consumed in a single-hop transmission can be described by using (2.5) and (2.6) as

$$E_r + E_t(d_{max}) = 2a + b \cdot d_{max}^2 = 2a + b \cdot (k \cdot d_{min})^2, \quad (2.8)$$

equivalently, the energy consumed by the multi-hop transmission over k symmetric hops can be written as

$$k \cdot E_r + k \cdot E_t(d_{min}) = k \cdot a + k \cdot (a + b \cdot d_{min}^2) = 2 \cdot k \cdot a + b \cdot k \cdot d_{min}^2. \quad (2.9)$$

From (2.8) and (2.9) it is possible to derive the condition that makes single-hop more energy efficient than multi-hop

$$k \leq \frac{2a}{b \cdot d_{min}^2}. \quad (2.10)$$

According to [6] [80], a condition necessary for minimizing energy consumption in multi-hop is that the hop distance is the same for all hops. For n intermediary nodes placed between two nodes separated by a distance D the hop distance is

$$d_{hop} = \frac{D}{n}, \quad (2.11)$$

replacing d_{min} with d_{hop} in (2.9) and taking its derivative with respect to n , the number of hops that minimizes the energy consumption in multi-hop communications can be found

$$n_{opt} = \sqrt{\frac{b}{2a}} \cdot D. \quad (2.12)$$

By replacing n by n_{opt} in (2.11) the optimum hop distance in terms of energy consumption can be written as

$$d_{char} = \sqrt{\frac{2a}{b}}, \quad (2.13)$$

where d_{char} is known as the characteristic distance. Note that d_{char} depends only upon the values of a and b , thus it is a parameter intrinsic to the sensor in question.

Care must be taken when employing the multi-hop approach to avoid reducing energy efficiency by using an unnecessary number of hops. When adequately employed the multi-hop approach can lead to significant energy saving in WSNs. However, multi-hopping suffers from some serious disadvantages. Data forwarding is usually done on a best effort delivery basis. That means that the transmitting node has no guarantee of the transmitted data reaching its destination, or that it will be delivered within a given time frame. Multi-hop networking can lead to data congestion on nodes that are located between node clusters that generate heavy traffic. This heavy traffic will also result in a high drain of energy resources in the nodes responsible for forwarding the data. This will result in a high delivery delay and will eventually lead to depletion of energy in these midway nodes, resulting in a disconnected network. Data relaying is usually done in a decode and forward fashion, this can result in a high delay even when there is no network congestion present.

Single-hop transmissions are not affected by network congestion as they are end-to-end communications, but they require a very high signal power when employed over large distances. If a single sensor is responsible for producing a large amount of data that needs to be transmitted over a large distance, this will lead to a very fast depletion of its energy resources. This uneven depletion is highly prejudicial to WSN operation, since replacing individual nodes might be as costly as replacing the entire network. Furthermore, single-hop transmissions might be unattainable over large distances due to the limited power at which sensor radios usually operate.

2.5.2 Cooperative MIMO

By introducing the C-MIMO approach, the drawbacks of the multi-hop approach, such as larger delay and possible uneven energy consumption, can be avoided while still achieving energy efficient communications. With the fundamentals of C-MIMO reviewed in Section 2.4, a reliable, low BER and energy

efficient communication can be implemented by taking advantage of the cooperative nature of WSNs.

With the steps of Figure 2.3 in mind, it is possible to derive an analytical expression for the energy consumption of C-MIMO. Defining E_{intra} as the energy necessary to transmit an entire data packet to another sensor located within the C-MIMO cluster frontier, shown as ① and ③ in Figure 2.3, the energy spent on spreading the packet for transmission and for consolidating it for equalization and decoding is given by

$$E_{\text{tx}_{\text{intra}_{\text{tx}}}} = \frac{M_t - 1}{M_t} E_{\text{intra}}, \quad (2.14)$$

$$E_{\text{tx}_{\text{intra}_{\text{rx}}}} = \frac{M_r - 1}{M_r} E_{\text{intra}}, \quad (2.15)$$

respectively. For an $M_t \times M_r$ C-MIMO transmission, a central node that spreads its packet to neighboring nodes needs to perform $M_t - 1$ transmissions with a cost of $\frac{E_{\text{intra}}}{M_t}$, as the packet is split into M_t smaller packets for C-MIMO transmission. The same applies to consolidating the packet on the receiving C-MIMO cluster, where $M_r - 1$ transmissions with a cost of $\frac{E_{\text{intra}}}{M_r}$, are sent to a central node for equalization and decoding.

Dissemination and consolidation of packets on the C-MIMO clusters also consume significant energy on the radios listening to the transmissions. Therefore, $M_t - 1$ receptions with a cost of $\frac{E_{\text{rx}}}{M_t}$, where E_{rx} is the energy cost for receiving an entire data packet, are necessary at the transmitting cluster. At the receiving cluster, $M_r - 1$ receptions with a cost of $\frac{E_{\text{rx}}}{M_r}$ are necessary when the received symbols are transmitted to the central node for decoding. Thus, the energy spent with receptions at the transmitting and receiver clusters is given by

$$E_{\text{rx}_{\text{intra}_{\text{tx}}}} = \frac{(M_t - 1)}{M_t} E_{\text{rx}}, \quad (2.16)$$

$$E_{\text{rx}_{\text{intra}_{\text{rx}}}} = \frac{(M_r - 1)}{M_r} E_{\text{rx}}, \quad (2.17)$$

respectively. Additionally, as the packet is transmitting over long distances using C-MIMO, the M_r receiving nodes must listen. Therefore, another M_r receptions with $\frac{E_{\text{rx}}}{M_r}$ cost are necessary.

Finally, $E_{\text{tx}_{\text{inter}}}$ is the energy spent on transmitting a data packet, using higher power for the long distance transmission step ② on Figure 2.3. M_t nodes transmit $\frac{1}{M_t}$ of the data packet.

The total cost for transmitting a packet, using cooperative MIMO, is then given by

$$\begin{aligned} E_{\text{C-MIMO}} = & E_{\text{tx}_{\text{inter}}} + E_{\text{tx}_{\text{intra}_{\text{tx}}}} + E_{\text{tx}_{\text{intra}_{\text{rx}}}} \\ & + E_{\text{rx}_{\text{intra}_{\text{tx}}}} + E_{\text{rx}_{\text{intra}_{\text{rx}}}} + E_{\text{rx}}. \end{aligned} \quad (2.18)$$

C-MIMO is capable of reaching large distances without demanding much power from a single node. In addition, since multiple copies of the same signal are received, the BER is considerably smaller at the same SNR. This makes the C-MIMO capable of reaching long distances, using much less power than SISO configurations do. C-MIMO configurations can lead to even lower BER ratios, which allow even less power to be used for long range transmissions.

Next, assuming a multihop communication where the nodes involved are uniformly spaced and the energy necessary to transmit a data packet to the next node is equivalent to the energy necessary to transmit inside a C-MIMO cluster, E_{intra} , the energy used for a multihop transmission is given by

$$E_{\text{mhop}} = k(E_{\text{intra}} + E_{\text{rx}}), \quad (2.19)$$

where k is the number of hops necessary to reach the destination node. From (2.18) and (2.19), the condition under which C-MIMO becomes more efficient than multihop, i.e. $E_{\text{C-MIMO}} < E_{\text{mhop}}$, is

$$\begin{aligned} E_{\text{inter}} &< \left(k - \left(\frac{M_t - 1}{M_t} + \frac{M_r - 1}{M_r} \right) \right) E_{\text{intra}} \\ &\quad + \left(k - \left(\frac{M_t - 1}{M_t} + \frac{M_r - 1}{M_r} + 1 \right) \right) E_{\text{rx}}. \end{aligned} \quad (2.20)$$

From (2.20), it is clear that C-MIMO is more efficient than multihop only when the nodes are separated by a large number of hops.

Figure 2.4 shows that a BER of 10^{-3} can be reached using MIMO communication with SNR 8 dB lower than standard SISO communication. From the specifications of the Harvard Mica2 platform [99], switching from 4 dBm transmission to -15 dBm transmission means that the radio consumes approximately half of the energy to achieve the same BER.

Table 2.2 presents a comparison of the amounts of energy consumed by the different communication methods, namely, single-hop, multi-hop, and 2×2 C-MIMO. When communicating with a node located 210 m away from the originating node, C-MIMO performs only slightly better than the single-hop and worse than the multi-hop approach. Even though the difference in performance between MIMO and single-hop is marginal, the energy consumption is spread out among various nodes when using the C-MIMO. Such a characteristic is highly desirable in WSNs. For longer distances, when using multi-hop, as the increasing number of receptions starts becoming highly energy-demanding, C-MIMO outperforms multi-hop. From Table 2.2 the increased efficiency provided by the C-MIMO over large distances can be observed.

Note, that the computational complexity required in C-MIMO is higher than that required in a standard SISO communication. However, the energy consumed due to additional computational complexity is significantly smaller than the energy saved by efficient communication. As a rule of thumb, in a standard communication, the energy spent in processing is 10^3 to 10^4 times less than the energy spent in communication [2].

Distance \ Method	Single-hop	Multi-hop	MIMO
210 m	100%	95%	99%
420 m	Not feasible	100%	83%

Table 2.2: Transmission energy consumption

2.6 Synchronization

Synchronization is fundamental to optimal performance of MIMO systems, as poor synchronization can lead to a very large BER or even ruin any used precoding. The subject of network-wide synchronization in WSNs has been extensively studied [42, 43, 70, 109], and several algorithms have been proposed to keep a common clock across the entire network. Most solutions suggest maintaining a relationship between clocks across the network, instead of trying to forcefully synchronize clocks across all nodes. For networks relying on GPS synchronization, it has been shown that very precise synchronization is possible by keeping the variations as small as 200 ns [71]. However, relying on GPS receivers leads to increased energy demand in a WSN that is no longer self-contained. Broadcast synchronization schemes capable of achieving 1 μ s of accuracy have been proposed in [36].

Subsection 2.6.1 presents an analysis of the effects of synchronization errors on WSNs operating with C-MIMO and V-BLAST. To mitigate such effects, two synchronization schemes are proposed for WSNs operating with C-MIMO. A method for coarse synchronization is presented in Subsection 2.6.2. Since WSNs can also be composed of nodes with limited processing capabilities, a method for accurate synchronization that requires low computational effort is proposed. Similarly, for networks with higher processing power, a method that provides highly accurate synchronization is proposed. These methods are detailed in Subsection 2.6.3. Finally, Subsection 2.6.4 presents an analysis of the propagation of synchronization errors in WSNs, besides a study on how to choose the appropriate symbol duration and re-synchronization interval of the network to allow the use of C-MIMO.

2.6.1 Effects of Synchronization Error on Cooperative MIMO

For networks operating at 256 ksps rate, the resulting symbol duration is approximately 4 μ s. In this case, a synchronization error of 1 μ s represents 25 % of symbol duration. In practice, synchronization errors result in a lower SNR, implying a lower BER performance for a fixed noise power at the receiver. Figures 2.6, 2.7, and 2.8 compare the performances of the equalization methods discussed in this work in the presence of a synchronization errors, expressed as fractions of the symbol period T_s , for MIMO systems with different numbers

of elements. The results show that, for smaller configurations, errors as small as $0.2 T_s$ can have a large impact on the BER performance of MIMO communications, specially for SIC variants due to the error propagation problem caused when a symbol is decoded incorrectly. While larger configurations are shown to be more robust to synchronization errors, achieving such configuration in a C-MIMO scheme may not be always possible, since it depends, for instance, on the node density of the network or the overall processing power of the node responsible for equalization and decoding. Therefore, another method for synchronizing transmitting nodes and receiving nodes is necessary to take full advantage of MIMO communications in WSNs.

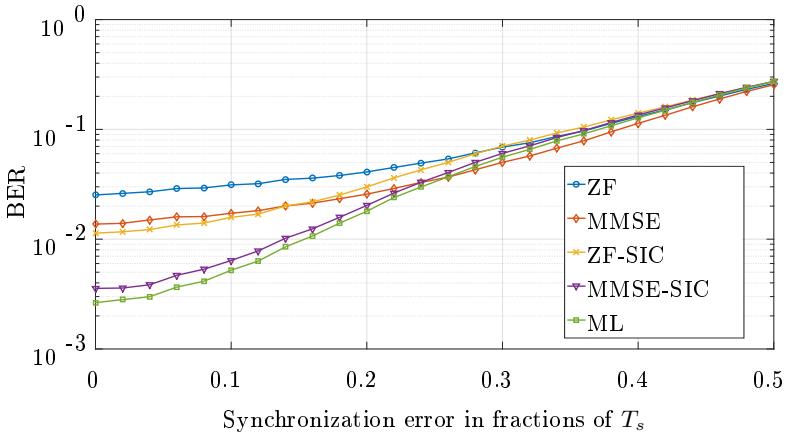


Figure 2.6: Performance comparison of equalization methods in the presence of synchronization errors in a 2x2 MIMO system

2.6.2 Proposed Coarse Synchronization Scheme

To achieve a degree of synchronization compatible with the operation of C-MIMO, this work proposes scheduling a tonal transmission between a pair of nodes or tonal broadcast to a group of nodes. The sampling on the receiving nodes starts at the scheduled time, and the clock error can be compensated by comparing the received tonal wave with a reference wave kept internally. The comparison can be done using ML or discrete Fourier transform based methods. However, owing to the periodic nature of the tonal waves, an element of ambiguity creeps in for delays higher than a wave period. Therefore, the receive node needs to be synchronized first with an accuracy of a period of the transmitted tonal wave.

To avoid problems created by sampling with a difference of more than a period of the tonal wave, the time length of the tonal transmission needs to be known to the receiving nodes. If a signal with less than the expected length is

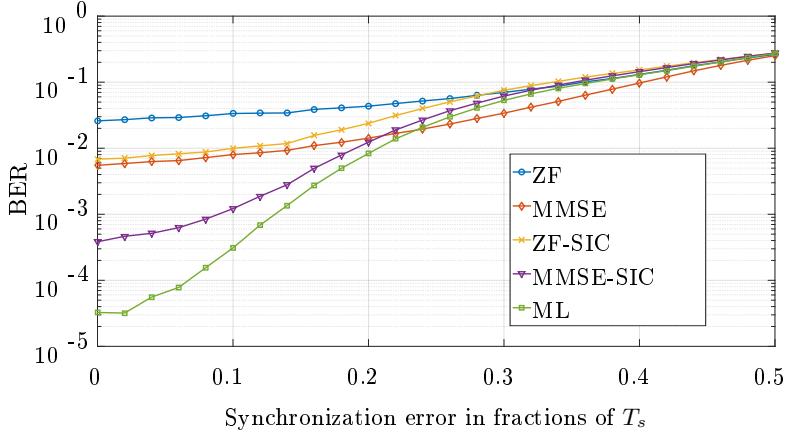


Figure 2.7: Performance comparison of equalization methods in the presence of synchronization errors in a 4x4 MIMO system

received, the receiver can compensate it by starting sampling earlier or later and adjusting its internal clock accordingly. An estimate $\hat{\delta}_p$ of the sampling synchronization error δ_p can be compensated prior to applying the sliding correlator to achieve precise synchronization by finding

$$\hat{\delta} = \operatorname{argmax}_{\delta_p} \int_{\delta_p}^{T_p + \delta_p} \operatorname{rect}\left(\frac{t - \frac{T_p}{2} + \delta_p}{T_p}\right) \sin(2\pi f t) dt, \quad (2.21)$$

where T_p is the time length of the tonal wave used for synchronization, f is the frequency of the synchronization signal and t is the time. That is, the sampling is synchronized by finding the delay that maximizes the received energy during a period of T_p .

2.6.3 Proposed Fine Synchronization Schemes

Once the nodes are synchronized with a precision of a wave period, more precise synchronization can be achieved. A possible solution is the application of the ML synchronization method

$$\hat{\delta} = \operatorname{argmin}_{\delta} \sum_{z=0}^{T_p} (r(q) - \sin(2\pi(q + \delta)f))^2, \quad (2.22)$$

where q represents discrete time instants, $\hat{\delta}$ is the estimate of the clock offset δ between two nodes synchronizing, and

$$r(q) = \sin(2\pi(q + \delta)f). \quad (2.23)$$

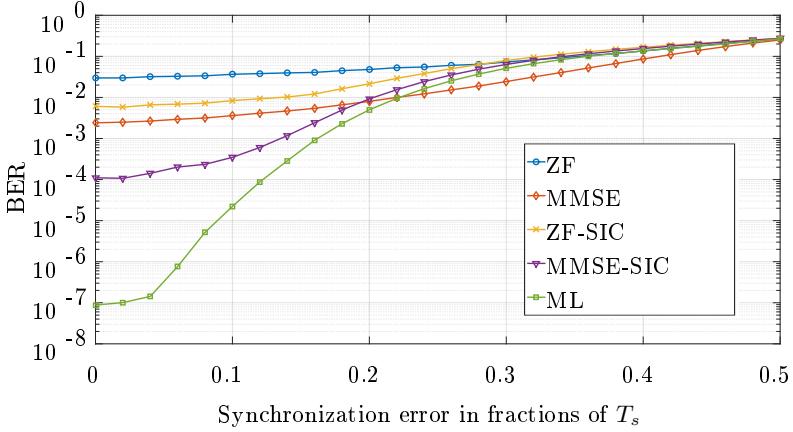


Figure 2.8: Performance comparison of equalization methods in the presence of synchronization errors in a 7x7 MIMO system

In the discrete domain, the estimator can be written as

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}} \| \mathbf{r} - \mathbf{l}(\delta) \|_2^2, \quad (2.24)$$

where $\mathbf{r} \in \mathbb{R}^{T_p F_s}$ is the sampled version of the synchronization waveform arriving at the receiver with a delay, relative to its clock, of δ and $\mathbf{l}(\hat{\delta}) \in \mathbb{R}^{T_p F_s}$ is the sampled version of a local copy of the waveform with a delay of $\hat{\delta}$. Using a high resolution, i.e., very small values of $\hat{\delta}$ as candidate values, this estimation method can achieve very precise synchronization, but entails very high computational load.

An alternative of lesser complexity is to use the discrete Fourier transforms (DFT) of the sampled signals to calculate the cross correlation between the received waveform and a local copy. Using this method, an estimation of the cross correlation is obtained as

$$\mathbf{c} = \mathbf{F}^{-1} (\mathbf{F} \mathbf{r}_0 \odot (\mathbf{F} \mathbf{l}_0)^*) \in \mathbb{C}^{2T_p F_s + 1}, \quad (2.25)$$

where $\mathbf{F} \in \mathbb{R}^{2T_p F_s + 1 \times 2T_p F_s + 1}$ is the DFT matrix, $(\cdot)^*$ is the complex conjugation, \odot is the Hadamard product, \mathbf{l} is a local copy of the waveform following the receiver's clock, and $\mathbf{r}_0 \in \mathbb{R}^{2T_p F_s + 1}$ and $\mathbf{l}_0 \in \mathbb{R}^{2T_p F_s + 1}$ are the zero padded versions of \mathbf{r} and \mathbf{l} . By finding the index l , where $|[\mathbf{c}]_l|$ is maximum, an estimate $\hat{\delta}$ of the delay δ can be obtained by

$$\hat{\delta} = \frac{l - T_p F_s + 1}{F_s}. \quad (2.26)$$

The proposed DFT based method offers reduced computational complexity compared as to the complexity of the ML method obtained from (2.22) and

(2.24), since the estimation is performed in a closed form. The ML computational complexity is a function of the resolution Δ . Therefore, by defining the ML search interval as $(-\frac{\pi}{2}, \frac{\pi}{2})$, the computational complexity of the ML method is given by

$$O(3\frac{\pi}{\Delta}T_pF_s), \quad (2.27)$$

while that of the DFT based method is given by

$$O(4T_pF_s \log(2T_pF_s)). \quad (2.28)$$

Note that the DFT based method requires one FFT and an inverse FFT, as the FFT of the reference signal can be stored and does not need to be calculated. Therefore, the DFT based method is more efficient than the ML method unless a very low resolution is chosen for the ML approach. However, the DFT method has a resolution that is limited by the sampling frequency F_s as shown in (2.26). Therefore, high synchronization accuracy demands a high sampling frequency or time interpolation to increase the number of time samples.

The effects of the filtering used for data transmission should also be taken into account in the synchronization process. The input filter alters the autocorrelation of the white noise leading to a colored noise. The noise at the output of the filter is given by

$$n^c(t) = w(t) * n(t), \quad (2.29)$$

where $n(t)$ is the uncorrelated noise before filtering, $w(t)$ is the filter response, $n^c(t)$ is the noise at the output of the filter, and $*$ is the convolution operator. The convolution can be represented in the frequency domain as

$$N^c(\omega) = W(\omega)N(\omega), \quad (2.30)$$

where $N^c(\omega)$, $w(\omega)$, and $N(\omega)$ are the Fourier transforms respectively of $n^c(t)$, $h(t)$, and $n(t)$. Therefore, the autocorrelation of the noise at the output is given by

$$\begin{aligned} E\left\{|N^c(\omega)|^2\right\} &= E\left\{|W(\omega)N(\omega)|^2\right\} \\ &= E\left\{|W(\omega)|^2\right\}E\left\{|N(\omega)|^2\right\} \\ &= \sigma_n^2 E\left\{|W(\omega)|^2\right\}, \end{aligned} \quad (2.31)$$

where σ_n^2 is the variance of the uncorrelated noise before filtering, and $E\{\cdot\}$ is the expected value operator. Assuming an ideal brick-wall filter

$$W(\omega) = \text{rect}\left(\frac{\omega}{2B}\right), \quad (2.32)$$

$$w(t) = \mathcal{F}^{-1}\{W(\omega)\} = 2B \text{sinc}(2Bt), \quad (2.33)$$

where B is the bandwidth of the input filter. Therefore, the noise samples at the output of the filter are uncorrelated only if they are sampled with $2B$ sampling rate. A possible solution to this would be to increase the bandwidth of the filter, which would, however, lead to a decrease in the effective SNR. Another problem of employing a very high sampling frequency is that it leads to increased energetic demands at the receiver, and, therefore, defeats the purpose of C-MIMO by leading to a decrease in energy efficiency.

Figure 2.9 highlights the effects of noise correlation arising from oversampling during synchronization when using a 20 MHz tonal wave for the ML method. For the simulations, the sampling lasts for 2 periods of the tonal wave. The front end filter is an ideal low pass brick wall filter with bandwidth $B = 22$ MHz. The sampling frequency for simulations varies between two and ten times the bandwidth of the front end filter. The results show that, due to the effects of noise correlation, the higher number of samples, obtained from oversampling, results in a very small increase in performance.

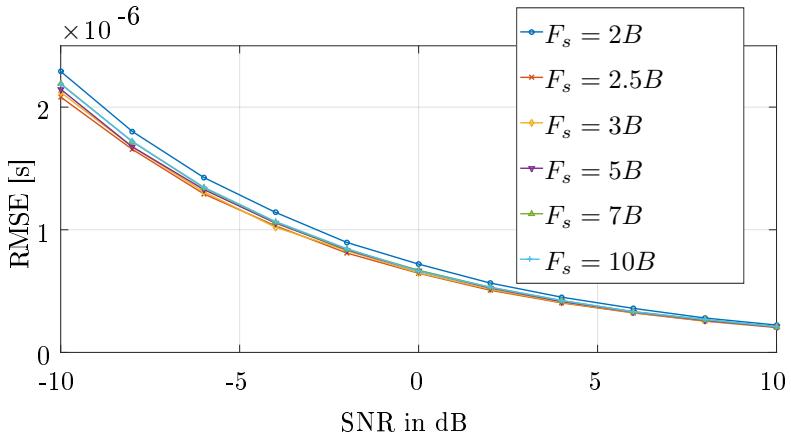


Figure 2.9: Performance of ML synchronization,
 $f = 20$ MHz, $B = 22$ MHz

Figure 2.10 presents the synchronization performance, obtained by using the proposed DFT based method. In the proposed method, the performance is degraded by noise only in very low SNR regimes, while in high SNR regimes, the major source of error comes from limited resolution. In this case, increasing the sampling rate led to significant performance improvement, despite introducing noise correlation. Noise being not the dominant source of error, increasing the sampling rate results in higher resolution, and, therefore, in reducing the errors caused by low discrete resolution. An alternative, when using a low sampling frequency, is to interpolate the received samples of the waveform in time, thus leading to a higher sample rate. However, interpolation is beneficial only for low sampling frequencies, where the noise between samples is approx-

imately uncorrelated. For high sampling frequencies, interpolation provides no gain. In Figure 2.10, the samples taken at $F_s = 2B$ and $F_s = 2.5B$ are interpolated to have four times the original number of samples. The curves for the synchronization results, using the interpolated signals are shown in the legend with the marker (i).

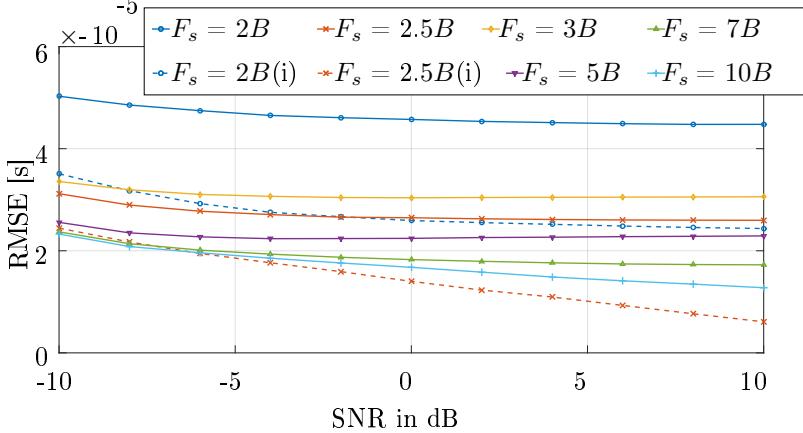


Figure 2.10: Performance of proposed DFT based synchronization, $f = 20$ MHz, $B = 22$ MHz

2.6.4 Synchronization error propagation

According to Section 2.5.2, C-MIMO is efficient only when used in long distance communication. However, direct synchronization of pairs of sensors over long distances is not feasible due to the large energy demanded for long range SISO transmissions. A possible solution would be to employ C-MIMO and a beamforming algorithm to reach distant sensors. This alternative requires the network to possess knowledge of the position of the sensors, and such knowledge is not assumed for this work. Therefore, another alternative is to propagate synchronization across network, using a multihop approach. This subsection overlooks the effects of error propagation in synchronization when a multihop approach is used, and how to define the synchronization interval and symbol period that allow the effective use of C-MIMO. The effects of synchronization errors caused by propagation time, clock drift, and errors in synchronization methods are analysed and consolidated into a single expression.

Using the ML estimator according to (2.22) and (2.24) in the presence of white noise, the distribution of the synchronization error between two sensors is expected to be normal. However, for high sampling frequencies, the distribution needs to be analysed. Figure 2.11 shows the histogram of the synchronization error for ML estimator. The result is obtained considering the

same scenario as that one used in obtaining the results shown in Figure 2.9. Furthermore, the SNR is fixed at 10 dB and a sampling frequency of $F_s = 7B$ is considered. Despite the discrete nature of the DFT method in (2.25), this error also approaches, under a high enough resolution, a normal distribution as shown in Figure 2.12.

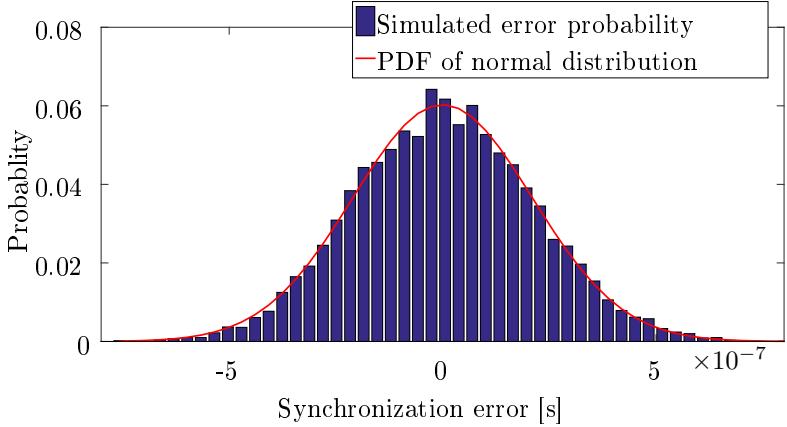


Figure 2.11: Histogram of synchronization error using the ML method (2.22,2.24)

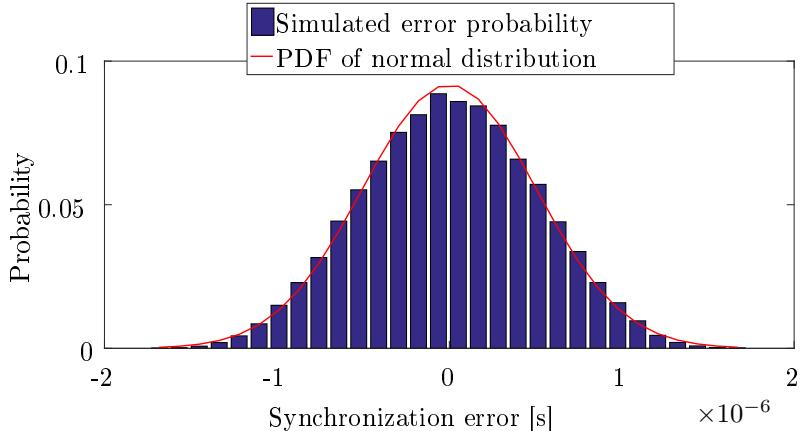


Figure 2.12: Histogram of synchronization error using the DFT method (2.25)

Figures 2.11 and 2.12 show that the PDF of the error can be approached by a normal distribution with a given variance $\sigma_{\epsilon_{\text{sync}}}^2$, despite the presence

of correlated noise. A random variable representing the final synchronization error ϵ_f , can then be defined as

$$\epsilon_f = \sum_{i=1}^k \epsilon_i, \quad (2.34)$$

where ϵ_i is the synchronization error of the i -th node on the synchronization chain, composed of k nodes. The variance $\sigma_{\epsilon_f}^2$ of ϵ_f is given by

$$\sigma_{\epsilon_f}^2 = \sum_{i=1}^k \sum_{j=1}^k [\mathbf{R}_{\epsilon\epsilon}]_{i,j}, \quad (2.35)$$

where $\mathbf{R}_{\epsilon\epsilon}$ is the covariance matrix of the synchronization errors of the nodes on the synchronization chain, and $[\cdot]_{i,j}$ is the element of a matrix indexed by i and j . Assuming that the errors are uncorrelated between all the nodes involved, $\sigma_{\epsilon_f}^2$ is then given by

$$\sigma_{\epsilon_f}^2 = \sum_{i=1}^k \sigma_{\epsilon_i}^2. \quad (2.36)$$

Therefore, under the assumption that all the nodes have a synchronization error that follows a normal distribution with the same variance, the synchronization error between nodes, separated by k hops, can be approximated as a zero mean normal distribution with variance

$$\sigma_{\epsilon_f}^2 = k\sigma_{\epsilon_{\text{sync}}}^2. \quad (2.37)$$

The synchronization does not take into account the propagation time of the transmitted wave. This correction can be performed by measuring the approximate distance between a pair of sensors, using the received signal strength indicator (RSSI). Methods such as RSSI are capable of estimating with 5 to 10 % error [30]. Another alternative is to apply a sensor localization method, one such method employing array processing concepts is presented in Appendix G. In addition, the nodes may also, if fixed, know their positions. Considering that the localization error between the sensors involved in the synchronization chain is a function of the distance between these sensors, the total synchronization error δ_d due to the propagation can be written as

$$\delta_d = \frac{\varphi d}{c}, \quad (2.38)$$

where φ is the relative distance estimation error, d is the total distance involved in synchronization and c is the speed of light.

Another important source of synchronization imprecision is the clock drift between two sensors. Given that the nodes of a network operate on clocks

with a given maximum drift rate ρ , they can drift apart at a maximum rate of 2ρ . A function ζ , defining the synchronization error between two sensors, which are k synchronization hops apart at time instant τ_1 and which have last synchronized at time instant τ_0 , is given by

$$\zeta = \mathcal{N}(0, k\sigma_{\epsilon_{\text{sync}}}^2) + \delta_d + 2\rho\Delta_\tau, \quad (2.39)$$

where $\Delta_\tau = \tau_1 - \tau_0$. As shown in Figures 2.6, 2.7, and 2.8, the synchronization error needs to be, at the most, 20 % of the symbol period to guarantee a good BER performance. Therefore, to use the benefits of C-MIMO, the following relationship must hold

$$\begin{aligned} |\mathcal{N}(0, k\sigma_{\epsilon_{\text{sync}}}^2)| + \delta_d + 2\rho\Delta_\tau &\leq 0.2T_s \\ |\mathcal{N}(0, k\sigma_{\epsilon_{\text{sync}}}^2)| &\leq 0.2T_s - \delta_d - 2\rho\Delta_\tau, \end{aligned} \quad (2.40)$$

where T_s is the symbol duration and

$$0.2T_s > \delta_d + 2\rho(\tau_1 - \tau_0). \quad (2.41)$$

The probability of the synchronization error being within the acceptable bound is then given by

$$\text{erf}\left(\frac{0.2T_s - (\delta_d + 2\rho\Delta_\tau)}{\sqrt{2k\sigma_{\epsilon_{\text{sync}}}^2}}\right). \quad (2.42)$$

Therefore, when designing a WSN for operation with C-MIMO, the parameters T_s and Δ_τ need to be defined to ensure that the probability of the synchronization error is within the acceptable bound. Also, care must be taken to ensure that the energetic efficiency of C-MIMO is not offset by a large energy demand from constant synchronization with a small Δ_τ and with very large symbol periods T_s .

2.7 Adaptive C-MIMO Clustering

After achieving synchronization, C-MIMO can be used to improve energy efficiency of WSNs. However, prior to transmitting the data, a selection must be made between using multi-hop or C-MIMO, since C-MIMO becomes beneficial only for long distance transmission, as explained in Section 2.4. Also, as WSNs have to be deployed even in harsh environments, the network must be capable of dealing with node outages (temporary or permanent), due to extreme heat or cold conditions, temporary link outages due to eventual interference in a given area or even energy depletion and consequent permanent failure of nodes. Routing algorithms developed for WSNs must be capable of choosing alternative routes, if a given route becomes unavailable. Depending on the frequency of such failures, different routing algorithms can be employed. For networks with fairly stable links, a routing table can be maintained at each node, with

either the full path to a given destination or only the next hop, depending on the memory available with the nodes or the limitations in maximum packet overhead. If the communication links are not reliable, up-to-date maintenance of a large routing table in every node of the network might not be possible. In such cases, routing algorithms such as Ad hoc On-Demand Distance Vector (AODV), are generally used, in which paths are discovered only on demand, thus avoiding the need for maintaining routing tables on the nodes.

This Section is divided into two subsections. Subsection 2.7.1 details the proposed method for adaptively forming C-MIMO clusters according to the energy efficiency of different cluster sizes and according to the impact of different clusters sizes on the distribution of available energy over neighboring nodes. Subsection 2.7.2 presents the numerical simulations for the proposed method, highlighting positive effects on network lifetime and energy distribution.

2.7.1 Adaptive C-MIMO clustering

Possible C-MIMO formations for a given node can be represented as new edges with different values on a graph. Therefore, the existence of C-MIMO can be transparent to the chosen routing algorithm. Moreover, since C-MIMO can be used with different cluster sizes, it is important to choose the best possible cluster size for reaching a given node in the network and representing it on the routing graph as a vertex with the lowest possible cost. Another important characteristic of C-MIMO is its ability to distribute energy consumption among various nodes, and, depending on the number of neighbors of a given node, the node may choose the neighbors that cooperate on a C-MIMO transmission to maximize the life time of the network.

Choosing between C-MIMO, multi-hop, and single-hop is not enough. Since C-MIMO can be used with different cluster sizes, it is important to choose the best possible cluster size for reaching a given node in the network, representing it on the routing graph as a vertex with the lowest possible cost. However, minimizing energy consumption is, by itself, not enough to maximize the lifetime of a WSN. If energy consumption is low but the energy distribution is uneven, some nodes will fail before others. This distribution can be area dependent, for instance, regions with a high traffic of information will tend to be depleted ahead of the rest of the network. Node failures in such areas might lead to a disconnected network, reducing the effective lifetime of the network. Generally, it is desirable that the nodes in a WSN have their energy reserves evenly depleted, so that the network can stay fully connected as long as it is operational. Therefore, by being aware of the energy reserves of its neighbours, a node can, when using C-MIMO, choose the configuration and neighbours that minimize the difference between the energy reserves of its neighbours.

This work proposes calculating the cost of a transmission as a function of transmission power, cluster size and effects on the energy reserves of neighbours. The metric shown in (2.43) is used to define the cost of a transmission between a pair of nodes indexed by $n = 1, \dots, L$ and $m = 1, \dots, L$. In

$$J(n, m, w) = \begin{cases} \alpha(\sigma_0(n, w) - \sigma_1(n, w)) + E_{\text{C-MIMO}}(w), & \text{if } n \neq m \text{ and } w \neq 1 \text{ and } m \in \mathcal{N}_{n,w} \\ E_{\text{intra}} + E_{\text{rx}}, & \text{if } n \neq m \text{ and } w = 1 \text{ and } m \in \mathcal{N}_{n,1} \\ 0, & \text{if } n = m \\ \infty, & \text{if } n \neq m \text{ and } m \notin \mathcal{N}_{n,w} \end{cases} \quad (2.43)$$

this expression L is the number of nodes that compose the network, $w = 1, \dots, W$ indexes the possible sizes for the C-MIMO clusters up to the limit W , $E_{\text{C-MIMO}}(w)$ is the energy cost for a C-MIMO transmission using clusters composed of w elements, $\sigma_0(n, m, w)$ and $\sigma_1(n, m, w)$ are the standard deviations of the energy available in the node's neighbours before and after the possible C-MIMO transmission. $\mathcal{N}_{n,w}$ is the set of nodes that can be reached by node n with a C-MIMO cluster size w . For a fixed cluster size, the nodes n and m always choose the neighbours with the highest amount of energy to maximize the network lifetime. However, it is possible that some formations, even if the total energy cost is lower, cause the variance of the energy available on the neighbours to increase too much. This means that the distribution of the available energy between the nodes is less uniform leading to some nodes running out of energy before others, and resulting in a disconnected network. Finally, α is a weighting factor that dictates the importance of keeping a uniform energy reserve versus the importance of minimizing overall communication energy consumption. If the transmit node n has knowledge about the state of the energy reserves of the receive cluster it may also take into account its changes on the calculation of the cost.

Each node n stores the cost relative to reaching a given neighbor in a cost matrix

$$[\mathbf{R}_n]_{m,w} = J(n, m, w) \in \mathbb{R}. \quad (2.44)$$

A graph used for routing can be set up by grouping all of the cost matrices from the nodes that form the network into a tensor structure

$$\mathcal{R} = [\mathbf{R}_1 | \mathbf{R}_2 | \dots | \mathbf{R}_L] \in \mathbb{R}^{L \times L \times W}, \quad (2.45)$$

where $|$ denotes a stacking operation onto the first dimension. However, constructing a graph based on this structure will result in vertices that are connected by multiple edges, since nodes can communicate using different C-MIMO cluster sizes. To avoid this problem, some edges must be filtered from the graph. This decision is simple, and can be based solely on the cost of the edges, leaving only the edges with the smallest cost between two vertices. Therefore, the decision on what cluster size $s_{n,m}$ to use to communicate between a pair of nodes is given by

$$s_{n,m} = \underset{w}{\operatorname{argmin}} [\mathcal{R}]_{n,m,w} \in \mathbb{N}^+. \quad (2.46)$$

Algorithm 1 presents the proposed solution for choosing the optimal cluster formation for reaching a given node. A list of neighbors is present at each node and associated to it. Moreover, the list includes information about the number of nodes involved in the C-MIMO clusters for reaching the intended destination.

The cost can be a function of variables such as transmission power, cluster size and effects on the energy reserves of neighbours. This work uses the following metric to define the cost of C-MIMO transmission

$$J(\sigma_0^2, \sigma_1^2, E_{\text{C-MIMO}}) = \alpha(\sigma_0^2 - \sigma_1^2) + E_{\text{C-MIMO}}, \quad (2.47)$$

Algorithm 1 Lower Cost Hop

```
1: load neighborList
2: if routePacketReceived then
3:   nodes  $\leftarrow$  routePacket.getOriginNodes()
4:   size  $\leftarrow$  routePacket.clusterSize()
5:   cost  $\leftarrow$  calculateCost(routePacket)
6:   for node : nodes do
7:     if neighborList.contains(node) then
8:       if node.cost() > cost then
9:         neighborList.remove(node)
10:        neighborList.add(node, cost, size)
11:      end if
12:    else
13:      neighborList.add(node, cost, size)
14:    end if
15:  end for
16: end if
```

where $E_{c\text{-mimo}}$ is the energy cost for a C-MIMO transmission as computed in (2.18), σ_o^2 and σ_1^2 are the variances of the energy available in the neighbouring nodes, before and after the possible C-MIMO transmission. For a fixed cluster size, the node always chooses the neighbours with the highest amount of energy to maximize the network lifetime. However, it is possible that some formations, even if the total energy cost is low, significantly increase the difference between the energy levels of the neighbours. Thus, the distribution of the available energy between the nodes becomes less uniform. This will lead to some nodes running out of energy before others do, which may ultimately result in a disconnected network. Finally, α is the weighting factor that dictates need for maintaining a uniform energy reserve versus the need for minimizing overall communication energy consumption. If the transmit node has knowledge about the state of the energy reserves of the receive cluster it may also take into account its changes on the calculation of the cost.

Other parameters such as delay, packet loss or available bandwidth may also be included in the metrics of cost function. If one of the nodes transmitting the packet is not entered in the list, this entry is added on the neighbour's list. If an entry exists, the cost of the received packet is compared to the cost given in the list (Lines 7 - 8 of Algorithm 1); if the cost of the received packet is lower than the cost given in the list, this entry is replaced (Lines 9 – 10 of Algorithm 1). This optimal entry can be then sent to the transmitting node or cluster, assuming that all connection characteristics are reflective, or the transmitting node or cluster itself may use the response to calculate the cost associated with the reverse path. Using this algorithm, a node can decide when a packet needs to be forwarded, and the optimal number of nodes that need to be involved in a C-MIMO transmission to reach a certain neighbour on the list. The graph

structure is not modified; therefore, classic graph routing algorithms can still be employed.

The proposed algorithm can take advantage of the C-MIMO technique and of the traditional multi-hop technique to achieve the lowest possible energy consumption in the network. The same algorithm is also used to select the optimal number of cluster members that must be present in a C-MIMO transmission. Thus, C-MIMO transmissions can be adaptively employed by modifying the cluster configurations to suit the changes in the network topology, but without disturbing the operation of standard transmission techniques.

2.7.2 Numerical Simulations

For numerical simulations, this work assumes that the sensors follow the communication characteristics of Harvard Mica2 platform [99]. An area of $100 \text{ m} \times 100 \text{ m}$ is filled with 300 nodes placed with coordinates following independent uniform distribution. These nodes are capable of transmitting with a power ranging from -20 dBm to 10 dBm , with respective current drains at the radio ranging from 3.7 mA to 21.5 mA . The current drain at the radios during reception is 7 mA , which is considerably higher than the drain on low power transmissions.

The noise power density at the radios during reception is considered to be -174 dBm/Hz and the radios are assumed to have a 5 dB noise figure. The communication is assumed to take effect by using 2.4 GHz center frequency and a 20 MHz channel, treated with a 22 MHz filter. The resulting noise power at the receiver was -96 dBm .

Transmissions follow the free-space path loss model, and the antennas have no gain. Therefore, the attenuation in dB of a signal, received d kilometres away from its receiver, is given by

$$\text{Loss}(d) = 20 \log(d) + 100.05. \quad (2.48)$$

Following the results shown in Figure 2.18, the desired system BER is set to 10^{-2} . The SNR necessary to achieve the same SNR varies with the size of the C-MIMO cluster; larger cluster formation requires lower SNR to reach a given BER.

According to [5], the probability that a given node has at least n neighbours located at most r distance units away is given by

$$P(n) = 1 - \sum_{i=0}^{n-1} \frac{(\rho \pi r^2)^i}{i!} e^{-\rho \pi r^2}, \quad (2.49)$$

where ρ is the node density of the network. For the simulations $r = 7.9 \text{ m}$, which is the maximum distance reachable with an SNR of 10 dB using a transmission power of -8 dBm . This shows that the probabilities of a node having at least 2, 3 or 4 neighbours are 98 %, 93.2 %, and 83.8 % respectively.

$\lambda \backslash M$	1	2	3	4
0	100 %	88.3 %	73 %	70.1 %
$\frac{2}{3}$	100 %	90 %	75.5 %	72.9 %
2	100 %	91.8 %	78.6 %	76.6 %
4	100 %	92.3 %	81 %	78.7 %
6	100 %	92.9 %	82.4 %	80 %

Table 2.3: Total network energy consumption

$\lambda \backslash M$	1	2	3	4
$\frac{2}{3}$	450	202	147	44
2	551	252	163	117
4	652	261	183	174
6	746	320	194	187

Table 2.4: Undeliverable packets

Node failures are also included in the simulations. The node failure rate followed an exponential distribution with parameter λ varying from 0 to 6 failures/s. The average duration of each failure is 30 s. Packets are generated on the network at a rate of $\frac{2}{3}$ packets/s, with one random node serving as the origin of the other as the destination. A packet is considered undeliverable if, at the time of generation, it could not be routed to its destination, because either the destination node is in a failure state or the destination was unroutable.

Each simulation could run for a total of 24000 simulated seconds, and the results show an average of 1000 Monte Carlo simulations. Tables 2.3 and 2.4 show the simulation results obtained in terms of energy consumption and undeliverable packets with a varying number of maximum allowed C-MIMO cluster sizes.

For Table 2.3, the reference is always the network in which no C-MIMO is present, i.e $M = 1$, since it always presents the largest energy consumption. The results show that allowing larger C-MIMO cluster reduces the energy consumption of the network. Since the simulated area is not very large, the benefit of allowing the larger C-MIMO clusters, composed of 4 sensors, is not very high as compared to the benefit available from a cluster composed of only 3 sensors. If the network area would have been larger, larger C-MIMO formations would have brought increased benefits, since long distance transmissions become more common on such network. Therefore, when designing networks or choosing the application of C-MIMO for existing networks, the total network area is an important parameter. Networks with small areas have no use

for large C-MIMO clusters, because most communications happen over small distances.

Table 2.3 also highlights the fact that the decrease in the effectiveness of C-MIMO, due to higher failure rates, is more than that in standard multi-hop only network. As nodes fail, the effective density of the network decreases. Consequently, each node has fewer operational neighbours, reducing the probability that larger C-MIMO configurations can be used, according to (2.49).

Table 2.4 shows the results for the number of packets considered undeliverable. For a network, not using C-MIMO, when the failure rate of the nodes is 6, almost half of the packets produced are not delivered. Increasing the maximum C-MIMO cluster size to 2 can significantly reduce the number of dropped packets, because larger distances can be reached by using C-MIMO. Just as energy consumption, networks with a large failure rate can benefit very little from allowing large C-MIMO cluster sizes, since those can rarely form in the sparser network.

Figure 2.13 presents the result of a simulation comparing the lifetime of a network using C-MIMO with a maximum cluster size of 3 and that of the same network when no C-MIMO is used. For this simulation, the packet generation rate is $\frac{2}{3}$ packets/s, and each node has enough energy to perform 120 transmissions with -8 dBm power. For each configuration, the simulation runs until the energy of 200 nodes has been depleted. The plot shows that, when C-MIMO

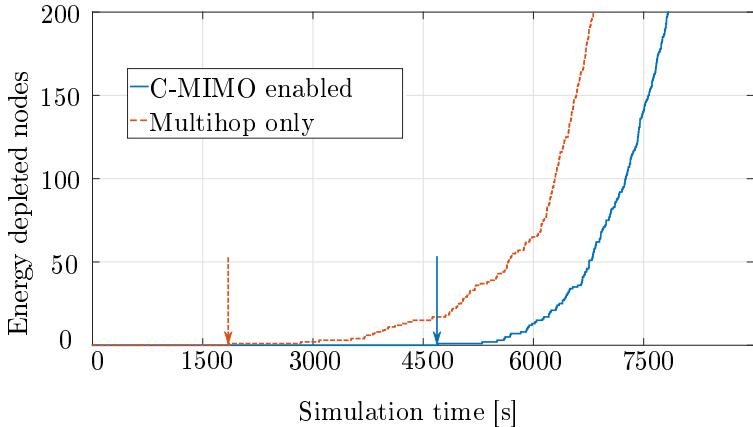


Figure 2.13: Comparison of individual node energy depletion using C-MIMO or multihop only

is enabled, not only does C-MIMO prolong the overall lifetime of the network significantly, it also delays the depletion of individual nodes considerably. This is a very desirable outcome for networks in which a certain parameter, such as temperature or humidity, must be sensed. In such networks, the energy depletion of individual nodes leads to smaller resolution in the sensed parameter.

This result comes not only from the reduced energy cost of C-MIMO for long distances, but also from its ability to split the energy cost between a group of nodes.

The definition of system lifetime varies widely, and there is as yet no consensus on this issue. However, some understandings that frequency appear in literature can be regarded as acceptable definitions for WSN lifetime, such as the time of the first failure, the moment when the network becomes disconnected and turns off for good, or the moment from which the network can no longer provide any services for some reason [31]. Figure 2.13 illustrates that, from the moment the first node failed (indicated by an arrow), there is a large gain when C-MIMO is used for almost double the time used until the first failure. Therefore, considering that the lifetime of a network ends with the first failure, the lifetime of the network under evaluation is doubled.

Figure 2.20 presents an example of a WSN with areas highlighted based on their expected communication energy consumption. In this example, the center area, shown with a checkered background, is expected to expend more energy than the rest of the network due to communication. The nodes that are located in the center are responsible for forwarding a large number of packets from nodes that are located closer to the edges of the network. Therefore, the nodes located in this area reach the end of their energy reserves faster than the rest of the network. On the other hand, nodes that are located on the stripped area have longer lifetimes as these nodes are rarely responsible for routing information for other nodes.

Figures 2.21 and 2.22 highlight the effects of the proposed C-MIMO cluster formation on the available energy reserve of nodes located at the center and near the edges of the network. The results shown in Figures 2.21 are expected, since the nodes that are located in the center have to perform less re-transmissions of data from other nodes on the network. Since C-MIMO can reach longer distances the nodes that generate the data can communicate with more distant nodes without requiring a large number of intermediary hops. On the other hand, while the results in Figure 2.22 show a decrease in the mean available energy for nodes located near the edges of the network. This, however, is desirable. The decrease in available energy indicates that the nodes located on the edges that, without C-MIMO, would generate packets and hand them over for other nodes to transmit across the network are now able to transmit over larger distances and spend their own energy reserves. This alleviates the energetic strain on sensors located at the center of the network. The results also show that, if a fixed C-MIMO cluster size is used the center nodes have their energy reserves depleted faster since fixing the cluster size prevents minimizing the energy consumption with respect to the distance between nodes.

Figures 2.23 and 2.24 present the effects of the proposed C-MIMO cluster formation on the distribution of energy on the network. Figure 2.23 shows that when no C-MIMO is used the energy on the center is poorly distributed, illustrated by the large standard deviation σ between the energy reserves of

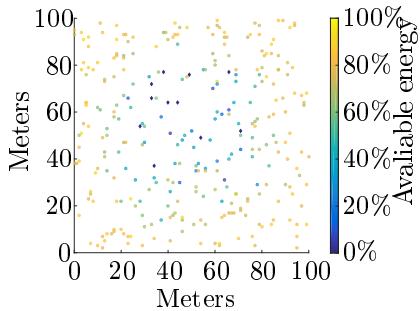


Figure 2.14: Comparison of individual node energy depletion using multihop only after 2500 seconds

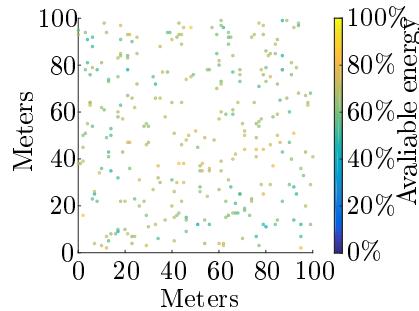


Figure 2.15: Comparison of individual node energy depletion using C-MIMO after 2500 seconds

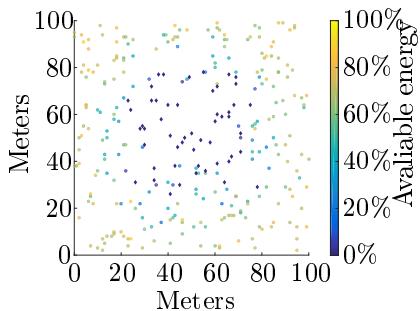


Figure 2.16: Comparison of individual node energy depletion using multihop only after 4500 seconds

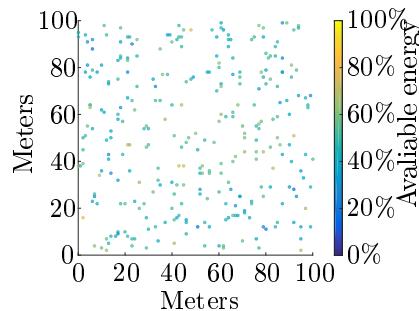


Figure 2.17: Comparison of individual node energy depletion using C-MIMO after 4500 seconds

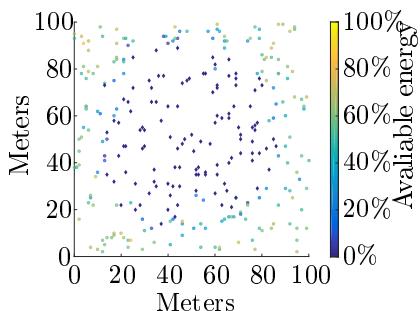


Figure 2.18: Comparison of individual node energy depletion using multihop only after 5500 seconds

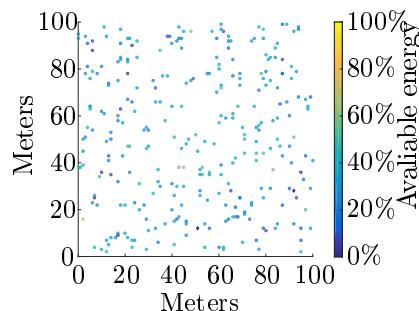


Figure 2.19: Comparison of individual node energy depletion using C-MIMO after 5500 seconds

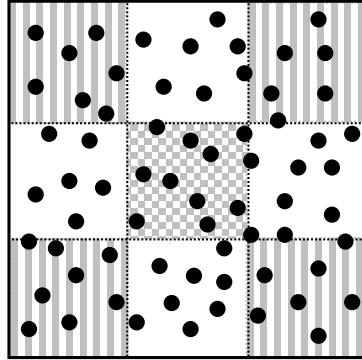


Figure 2.20: Example of WSN areas with different energetic demands

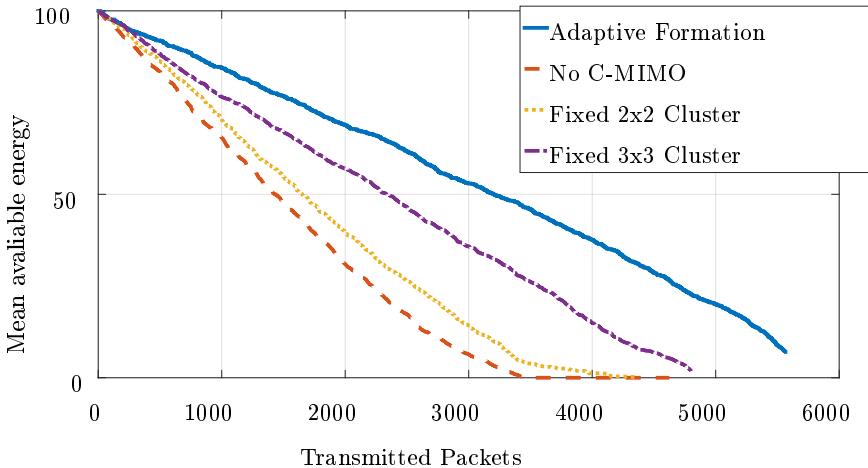


Figure 2.21: Mean available energy at the center region of the network

the nodes. Although the distribution seems to improve after a certain instant, this is only true due to the fact that the nodes on the center are running out of energy, resulting in a smaller σ as the number of nodes with very small energy reserves grows. Figure 2.24 shows that also for the edges of the network the energy distribution is more uniform when the proposed algorithm is used.

The results of comparing energy dispersion can be further analysed by plotting the individual energy reserves of nodes across the network. Figures 2.14, 2.16, and 2.18 show the status of the energy available, in percentage, in the network nodes after 2500, 4500 and 5500 seconds of simulation without enabling C-MIMO. Figures 2.15, 2.17, and 2.19 present the results in respect of the same parameters for a network after enabling the C-MIMO with a maximum cluster size of 3. Nodes with enough energy for at least one transmission

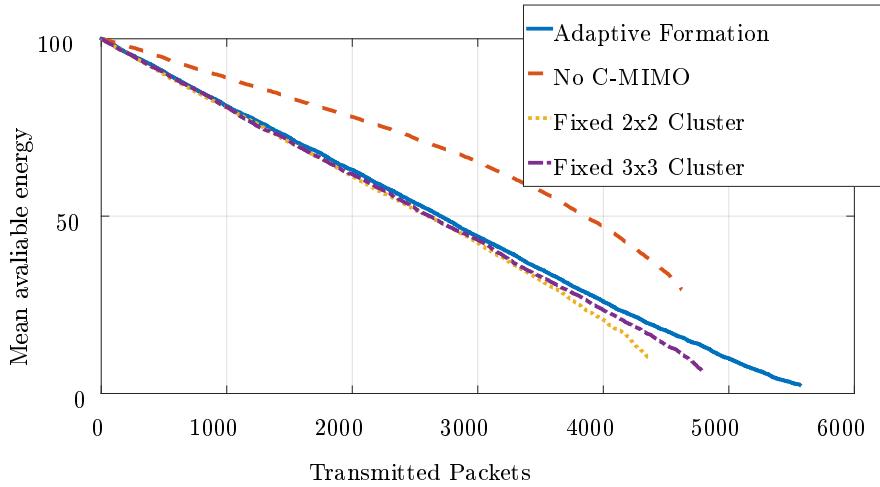


Figure 2.22: Mean available energy at the edges of the network

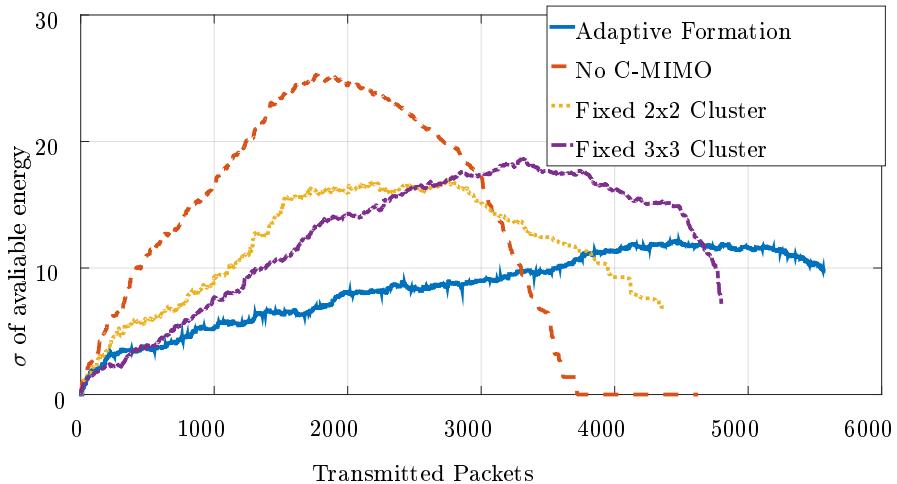


Figure 2.23: Standard deviation of available energy at the center region of the network

are shown as circles and the depleted as diamond symbols. When no C-MIMO is used, the nodes on the center of the network are depleted of their energy much faster than the nodes located along the edges of the network. These nodes are responsible for forwarding the information of the nodes placed along the edges. This results in a disconnected network and thus no information can be collected from areas that has no node coverage and where no sensors are available. However, when C-MIMO is enabled, the energy is more evenly

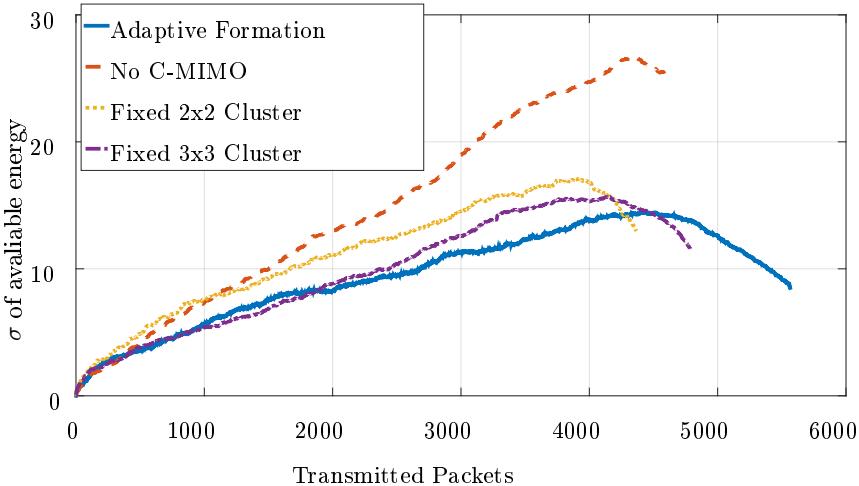


Figure 2.24: Standard deviation of available energy at the edges of the network

spread across the network. The nodes at the center are no longer depleted ahead of the nodes along the edges, since they are now capable of long range communication using C-MIMO. The network remains fully connected for a longer period with active nodes to cover a larger area for a longer period.

2.8 Summary

In this chapter the problems of communication energy consumption, advanced synchronization, and improved cluster formation for cooperative MIMO in WSNs are addressed.

An analytical study regarding the energy efficiency of C-MIMO in relation to that of multi-hop communications is presented. This analysis relates the number of hops necessary to reach a certain node with the cost involved in C-MIMO communications to clearly define the conditions under which C-MIMO is a better choice than multihop.

With respect to synchronization in WSNs, a method enabling the various nodes involved in C-MIMO to adequately schedule transmission and reception is presented. A coarse step for achieving an initial rough estimate, based on a simple sliding correlator, is also presented. For more precise synchronization ML and a DFT based methods are proposed. It is shown that the ML method is capable of achieving a higher degree of accuracy, whereas the DFT based method is computationally less complex.

Furthermore, an analysis of the effect of synchronization errors on C-MIMO schemes and of how the synchronization errors spread over the network is presented. Based on this analysis, a relation between symbol periods, clock

drift, and synchronization error is derived. This can be used to ensure that an acceptable degree of synchronization is achieved for a given data rate by using C-MIMO.

Finally, a method that allows the formation of efficient C-MIMO cluster is proposed, availing the benefits provided by this communication scheme, such as reduced energy cost and better distribution of energy demand. The proposed method is shown to extend network lifetime by reducing individual node failures and to achieve networks with a more homogeneous distribution of energy between the nodes, resulting in networks that remain fully connected for longer periods.

Chapter 3

Array Processing Localization for Vehicular Networks

In this chapter the following research question is addressed:

Can array processing techniques be used to reliably estimate position of vehicles and pedestrians in vehicular scenarios based on their radio transmissions?

3.1 Overview and Contribution

The remainder of this section is organized as follows:

- **Motivation:** The concept of vehicular networks and smart transportation systems is introduced. The problem of localizing vehicles and road users in such environment is shown and a solution from an array processing perspective is introduced.
- **Data Model:** The broadband spherical wave data model of a signal received at an antenna array is defined.
- **Space-Alternating Generalized Expectation Maximization (SAGE) Algorithm:** Here, the space-alternating generalized expectation-maximization (SAGE) algorithm for parameter estimation is briefly explained.
- **Scenario Description:** The assumed scenario for the localization in vehicular networks is shown and some constraints are discussed.
- **Array Processing Localization:** In this section three different approaches for localization in vehicular scenarios are presented. These approaches vary in performance and complexity and can be applied based on the specific performance and output rate requirements.

- **Three Dimensional DOA Based Estimation:** Here, the DOA based method is extended to the three dimensional case considering the estimation of position and attitude of UAVs as an application.
- **Simulation Results:** Finally, the performance of the proposed methods is shown by means of numerical simulations. A set of such simulations is performed using a real measured response from an antenna array to validate the proposed methods using real data.

The research contributions presented in this chapter are:

1. A DOA only localization method capable of localizing transmitters near to a vehicle at a low computational cost.
2. A localization method based on range and DOA estimations performed at independent subarrays. The geometrical relationship between the individual estimates is used to speed up the rate of convergence of the SAGE algorithm on the individual subarrays.
3. A novel parametrization for the received waveform data model allowing joint localization estimation to be performed at both subarrays simultaneously.
4. Localization methods that do not require the transmission of specific location data that can be used to mitigate spoofing attacks in VANETs.

3.2 Motivation

Vehicular ad hoc networks (VANETs) are a promising technology with applications such as road safety, traffic control, automated vehicle control, and platooning [60]. Many of such applications require the localization of all vehicles that compose a VANET to be known or estimated with a high degree of accuracy. Furthermore, not only must the accuracy be sufficient the estimate must also be reliable. Road safety, automated vehicles, and road control require the vehicles to be able to process and to estimate the position of vulnerable road users (VRUs), as well as the position of the other vehicles in the vicinity, so that they can operate adequately.

For the detection of VRUs, a variety of methods have been proposed. These methods must be precise and fast enough to meet different safety requirements regarding a multitude of possible situations. A set of requirements for VRU detection systems is presented in [49]. Most of these methods for VRU detection in the literature rely on computer vision [45]. These methods can provide an accuracy of up to 75% [35, 46], which might be insufficient for safety-critical systems such as collision detection and prevention/avoidance.

The localization of other vehicles is also of interest and can be obtained by using a variety of methods. One of the main source of location information in

a VANET is the usage of Global Navigation Satellite Systems (GNSS) information to obtain, for each vehicle, an estimate of its location. GNSS such as the Global Positioning System (GPS) can provide location information with a mean error of fifteen meters in urban environments [83]. However, this accuracy is not sufficient for emerging applications in Intelligent transportation systems (ITS) such as platooning [102] or for safety-critical applications. Furthermore, GNSS requires that a set of at least four line of sight signals are received from different satellites, this is not always possible in dense urban environments, leading to a possible large number of outages.

Other means of vehicle positioning such as using the received signal strength indicator (RSSI) of the received data signals [88] can also be used and are capable of complementing GNSS information and making safety applications possible. Other vehicle positioning approaches such as dead reckoning [57] can be used to provide position estimation during outages of GNSS. However, these approaches assume that the information that is provided by neighboring vehicles can be trusted.

Even in scenarios where the position provided by some vehicles can be assumed to be reliable and precise enough for the application at hand problems such as spoofing can still emerge. In a spoofing attack it is possible to falsify GNSS information such that a vehicle will obtain a position estimation that has been altered by the attacker. Furthermore, since GNSS information is usually spread using a data dissemination method [85], the position information can be falsified by the transmitting vehicle or by intermediary vehicles of the network. Position verification approaches [65] for VANETs can mitigate but not fully solve the problem of spoofing attacks.

Approaches that leverage the presence of inertial sensors in smartphones inside a vehicle, using them to provide robust positioning for vehicles in a network have been proposed [22], providing an accuracy of approximately 7 meters. Methods for localization in vehicular networks based on building a social model of the interactions of vehicles in the networks can provide position estimates with an accuracy of up to 10 meters [67]. While these methods provide an improvement over a GNSS only positioning solution, the accuracy provided is not sufficient for the safety-critical applications in vehicular networks. Approaches that rely on a dedicated localization infrastructure can provide exact localization methods with an accuracy superior to one meter [94] at the cost of an increased system complexity and may not be available outside of urban and busy roads.

A promising technology that can be applied to VANETs is the Multiple-input multiple-output (MIMO) communication scheme. MIMO has been used in modern wireless communications standards to allow for better spectral efficiency, faster data rates, and more robust communication. The usage of MIMO in vehicular network scenarios has been proposed to improve the network performance [103], suppress possible jamming attacks [58], and increase network capacity [55]. The antennas used for MIMO applications can be leveraged

to provide accurate positioning of vehicles on the network, without requiring a dedicated infrastructure. Since the usage of MIMO requires the presence of multiple antennas at both receiver and transmitter, this antennas can be setup as, for instance, linear antenna arrays to allow the application of array signal processing techniques to provide an extra source of localization information on the network.

This work proposes the usage of array signal processing tools for direction of arrival (DOA) estimation to passively estimate the position of vehicles transmitting a signal. The proposed method allows the position of a vehicle to be estimated by all vehicles within its communication range passively. This estimation does not require the transmission of localization specific messages, thus, it can be used to alleviate the network load. The proposed estimation method is also robust to data falsification since it relies on parameters that are estimated on the physical layer and cannot be altered or falsified easily, requiring a very complex system to do so.

3.3 Data Model

The assumed model considers a set of L wavefronts impinging onto an ULA composed of M antenna elements. The signal is assumed to be transmitted using an orthogonal frequency-division multiplexing (OFDM) scheme composed of K subcarriers. This model takes into account the curvature of the spherical wavefront. Using the Fresnel approximation, after removing the cyclic prefix and taking the discrete Fourier transform (DFT) of the received signal the space-frequency response of the k -th subcarrier received at antenna m during time snapshot t can be written as

$$x_{m,k,t} = \sum_{l=1}^L s_{l,k,t} \alpha_{m,l,k} e^{\omega_l(m-1) + \psi_l(m-1)^2} \cdot e^{j2\pi k \Delta_f(\tau_l)} + n_{m,k,t}, \quad (3.1)$$

where,

$$\omega_l = -\frac{2\pi\Delta_m \sin(\theta_l)}{\lambda}, \quad (3.2)$$

$$\psi_l = \frac{\pi\Delta_m^2 \cos^2(\theta_l)}{\lambda r_l}, \quad (3.3)$$

$s_{l,k,t}$ is the symbol transmitted by the l -th source at time instant t , $\alpha_{m,l,k}$ is the complex channel gain coefficient, Δ_m is the separation between antenna elements of the array, θ_l is the DOA of the l -th signal, r_l is the range of the l -th source, λ is the wavelength of the carrier frequency, Δ_f is the frequency separation between the subcarriers of the OFDM signal, τ_l is the propagation delay of the l -th received signal, and $n_{m,k,t}$ is additive complex white Gaussian noise.

The model can be rewritten in matrix form as

$$\mathbf{X} = \mathbf{D}(\mathbf{A} \diamond \mathbf{Z})\mathbf{S} + \mathbf{N} \in \mathbb{C}^{MK \times T}, \quad (3.4)$$

where \diamond is the Khatri-Rao product, T is the total number of snapshots available, \mathbf{D} is a diagonal matrix containing the channel gain coefficients, $\mathbf{S} \in \mathbb{C}^{L \times T}$ contains the symbols transmitted by the L sources for all T snapshots, $\mathbf{N} \in \mathbb{C}^{MK \times T}$ is the noise matrix with its entries drawn from $\mathcal{CN}(0, \sigma_n^2)$, and

$$\mathbf{A} = [\mathbf{a}(\theta_1, r_1), \mathbf{a}(\theta_2, r_2), \dots, \mathbf{a}(\theta_L, r_L)] \in \mathbb{C}^{M \times L}, \quad (3.5)$$

$$\mathbf{Z} = [\mathbf{z}(\tau_1), \mathbf{z}(\tau_2), \dots, \mathbf{z}(\tau_L)] \in \mathbb{C}^{K \times L}, \quad (3.6)$$

with

$$\mathbf{a}(\theta_l, r_l) = \left[1, e^{\omega_l + \psi_l}, \dots, e^{\omega_l(M-1) + \psi_l(M-1)^2} \right]^T \in \mathbb{C}^{M \times 1}, \quad (3.7)$$

$$\mathbf{z}(\tau_l) = \left[e^{j2\pi 1\Delta_f(\tau_l)}, e^{j2\pi 2\Delta_f(\tau_l)}, \dots, e^{j2\pi K\Delta_f(\tau_l)} \right]^T \in \mathbb{C}^{K \times 1}, \quad (3.8)$$

being the space and frequency steering vectors of the l -th received signal respectively.

3.4 Space-Alternating Generalized Expectation Maximization (SAGE) Algorithm

The SAGE algorithm is a generalization of the EM algorithm, both are algorithms used to derive maximum likelihood estimates in an iterative fashion. The idea behind these algorithms is to split an L -dimensional multi-dimensional problem that might be too computationally demanding to perform within an expected time constrain into L separate one-dimensional problems that can be solved within a shorter time. SAGE and EM utilize the concept of hidden data space to avoid dealing with the complete observed data set. In the case of a signal received at the array, the complete data can be written as a function of the hidden data space

$$\mathbf{Y} = \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_L, \quad (3.9)$$

$$\mathbf{X}_l = \mathbf{h}(\mathbf{p}_l)\mathbf{s}_l + \mathbf{n}_l, \quad (3.10)$$

where \mathbf{h} is a steering vector constructed from the parameter vector $\mathbf{p}_l = [p_l^1, p_l^2, \dots, p_l^P]$, \mathbf{s}_l is the symbol vector and \mathbf{n} is the noise component and is complex normally distributed with zero mean and variance σ_l^2 .

An iteration of SAGE is split into two different steps, the expectation and the maximization steps. The expectation step of the c -th cycle of the i -th iteration of SAGE is given by

$$\begin{aligned}\hat{\mathbf{X}}_l(\mathbf{p}_l^{[i,c-1]}, \mathbf{s}_l^{[i,c-1]}) &= E[\mathbf{X}_l | \mathbf{Y}, \mathbf{p}_l^{[i,c-1]}, \mathbf{s}_l^{[i,c-1]}] \\ &= \mathbf{h}(\mathbf{p}_l^{[i,c-1]}) \mathbf{s}_l^{[i,c-1]} \\ &+ \frac{1}{L} (\mathbf{Y} - \sum_{\substack{j=1 \\ j \neq i}}^L \mathbf{h}(\mathbf{p}_j^{[i,c-1]}) \mathbf{s}_j^{[i,c-1]})\end{aligned}\quad (3.11)$$

$$\hat{\mathbf{R}}_{\mathbf{X}_l} = \hat{\mathbf{X}}_l(\mathbf{p}_l^{[i,c-1]}, \mathbf{s}_l^{[i,c-1]}) \hat{\mathbf{X}}_l(\mathbf{p}_l^{[i,c-1]}, \mathbf{s}_l^{[i,c-1]})^H \quad (3.12)$$

The maximization step can then be performed, to estimate the parameter p_l^j the remaining parameters of $\mathbf{p}_l^{[i,c-1]}$ are kept fixed and the following problem is solved

$$p_l^j = \underset{p_l^j}{\operatorname{argmax}} \mathbf{h}(\mathbf{p}_l^{[i,c-1]})^H \hat{\mathbf{R}}_{\mathbf{X}_l}(\mathbf{p}_l^{[i,c-1]}, \mathbf{s}_l^{[i,c-1]}) \mathbf{h}(\mathbf{p}_l^{[i,c-1]}). \quad (3.13)$$

The transmitted symbols are estimated following

$$\mathbf{s}_l^{[i,c]} = \mathbf{h}(\mathbf{p}_l^{[i,c]})^H \hat{\mathbf{X}}_l(\mathbf{p}_l^{[i,c-1]}, \mathbf{s}_l^{[i,c-1]}). \quad (3.14)$$

The SAGE algorithm iterates over the parameters of all received signals until a convergence threshold is met. For the remainder of this work the superscripts for iteration and cycle will be omitted to simplify the notation.

3.5 Scenario Description

This work assumes that the vehicles connected to the VANET are equipped with two linear antenna arrays at two distinct locations on their frames. Figure 3.1 depicts an example of a vehicle with antenna arrays equipped at the wing mirror position triangulating the position of a VRU.

To avoid ambiguities in DOA estimation the antenna elements of the array must be placed no further than half a wave length apart. Assuming the communication occurs at the 5.8 GHz frequency band and that DOA estimation is to be done at carrier frequency the antenna elements can be placed as far as 2.5 cm apart. Antenna elements can be installed close together to allow more antennas to be installed on the same area or to reduce the total size of the antenna array, however this comes at the cost of increasing mutual coupling between antenna elements. This work assumes that the antenna arrays are installed at the wing mirror position of the vehicles composing the network.

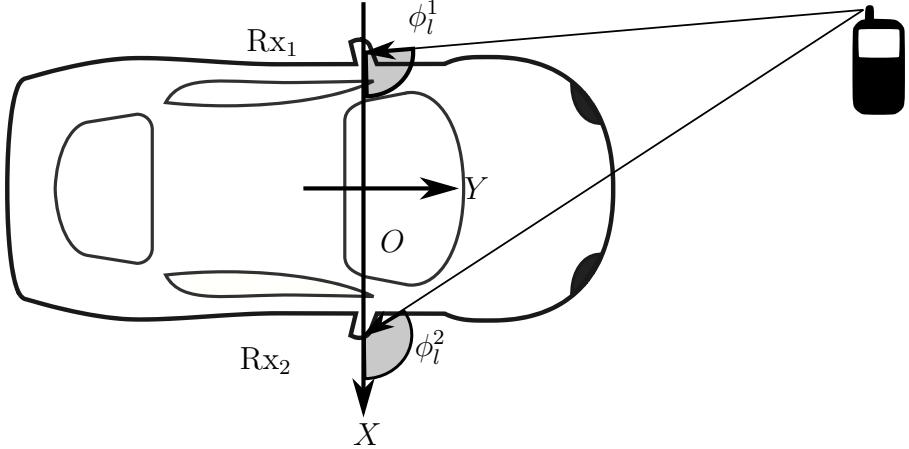


Figure 3.1: Scenario description for two vehicles

3.6 Array Processing Localization

This section presents three methods for estimating the position of a transmitter in a vehicular network context, namely the flip-flop, joint, and DOA only methods.

3.6.1 Flip-Flop Estimation

The first proposed method, addressed in this work as the flip-flop method, consists of iteratively performing individual estimations at each subarray and using the estimates of each subarray to update the SAGE algorithm, that is used to derive ML parameter estimates based on the data of each antenna array, of the remaining one.

The SAGE algorithm consists of the so-called expectation and maximization steps. The expectation step consists of obtaining an estimate for the signal received from a given source (hidden data space). Given that the data available is the set of superimposed signals received at the antenna, this estimate can be obtained as

$$\hat{\mathbf{X}}_i^k = \mathbf{X}^k - \sum_{\substack{l=1 \\ l \neq i}}^L \left(\hat{\mathbf{h}}(\mathbf{p}_l^k) \hat{\mathbf{s}}_l \right), \quad (3.15)$$

where $\hat{\mathbf{X}}_i^k$ is the estimate for the signal received from the i -th source at the k -th subarray, with $k \in [1, 2]$, and \mathbf{p} is the parameter vector

$$\mathbf{p}_l^k = \left[\hat{\theta}_l^k, \hat{r}_l^k, \hat{\tau}_l^k \right], \quad (3.16)$$

where $\hat{\theta}_l^k$, \hat{r}_l^k , and \hat{r}_l^k are the azimuth, delay, and range estimates for the l -th source at the k -th subarray, respectively. $\hat{\mathbf{s}}_l$ is a vector with the estimates of the symbols transmitted by the l -th source. Furthermore, $\hat{\mathbf{h}}$ is the current estimate of \mathbf{h} , which is defined as

$$\mathbf{h}(\mathbf{p}_l^k) = \left[\mathbf{a}(\hat{\theta}_l^k, \hat{r}_l^k) \otimes \mathbf{z}(\hat{r}_l^k) \right] \in \mathbb{C}^{MK \times 1}. \quad (3.17)$$

Following the expectation step, the maximization step transforms a D -dimensional estimation problem into D one-dimensional estimation problems by fixing all except one of the parameters during the estimation. Let the cost function d be defined as

$$d(\mathbf{p}_l^k, \mathbf{X}^k) = \left| \mathbf{h}(\mathbf{p}_l^k)^H \mathbf{X}^k \right|^2. \quad (3.18)$$

To obtain a complete estimate of \mathbf{p} the algorithm will solve the set of D optimization problems. These optimization problems can be written as

$$\hat{\theta}_l^k = \operatorname{argmax}_{\theta_l^k} d\left(\mathbf{p}_l^k, \hat{\mathbf{X}}_l^k\right), \quad (3.19)$$

$$\hat{r}_l^k = \operatorname{argmax}_{r_l^k} d\left(\mathbf{p}_l^k, \hat{\mathbf{X}}_l^k\right), \quad (3.20)$$

$$\hat{\tau}_l^k = \operatorname{argmax}_{\tau_l^k} d\left(\mathbf{p}_l^k, \hat{\mathbf{X}}_l^k\right). \quad (3.21)$$

To estimate each parameter, a different set of parameters is fixed, and a one-dimensional search is performed over the parameter that is currently being estimated. This process is iteratively performed until the estimates of all parameters converges.

Once the range and DOA parameter has been estimated for one of the subarrays, an estimate of the position of the transmitter can be obtained. Assuming that the center of a line crossing the car and both wing mirrors to be the origin of the reference coordinate system, denoted as O , it is necessary to obtain the angle ϕ_l^k , which, as shown in Figure 3.1, is a complementary angle to θ_l^k . Thus, the relationship between ϕ_l^k and θ_l^k is given by

$$\phi_l^k = -\frac{1}{\hat{\theta}_l^k + 2\pi}. \quad (3.22)$$

With this parameter at hand and following the coordinate system shown in Figure 3.1, an estimate of the position of the l -th transmitter with respect to the signal received at the k -th subarray Rx _{k} is given by

$$x_l^k = \hat{r}_l^k \cos(\phi_l^k) + x_{\text{Rx}_k}, \quad (3.23)$$

$$y_l^k = \hat{r}_l^k \sin(\phi_l^k). \quad (3.24)$$

where x_l^k and y_l^k are the estimated coordinates for the position of the l -th transmitter, where x_{Rx_l} is the position of the center of the l -th antenna arrays over the X axis.

A position estimated with respect one of the subarrays can be then mapped into a DOA and range estimation to the remaining subarray. Thus, once the full set of the parameters from one the subarrays have been estimated after a SAGE iteration, the position estimate extracted from such parameters can be used to update the current estimates for the remaining subarray before its next SAGE iteration. This can improve the rate of convergence for the SAGE algorithms as well as prevent the search of one of the SAGE algorithms from running into a local maximum and converging to an imprecise estimate. Figure 3.2 presents a block diagram illustrating the flow of the proposed method.

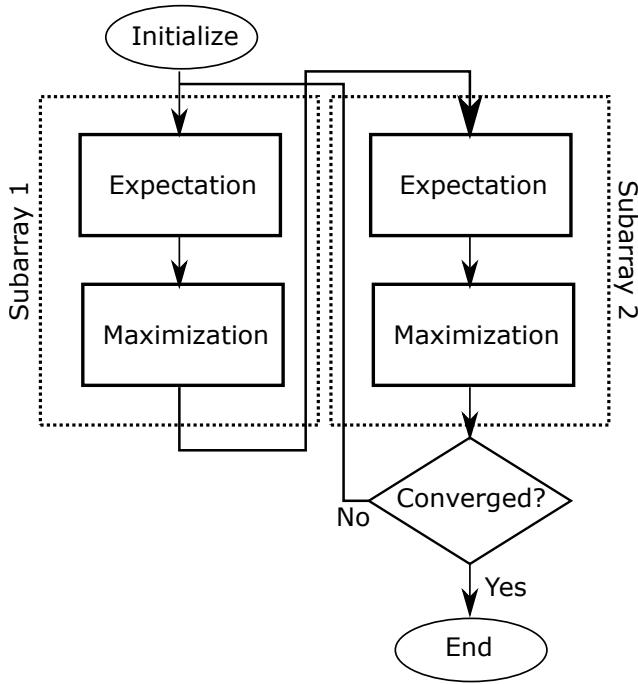


Figure 3.2: Block diagram for the proposed flip-flop method

Ideally, after the SAGE algorithm has converged at all subarrays, the transmitter would be detected at the same point concerning both subarrays, that is, $\hat{x}_l^1 = \hat{x}_l^2$ and $\hat{y}_l^1 = \hat{y}_l^2$. However, due to the noise present at the antenna measurements and possible numerical errors induced during the estimation of the position of the transmitter with respect to the two subarrays, they will be

different, thus $\hat{x}_l^1 \neq \hat{x}_l^2$ and $\hat{y}_l^1 \neq \hat{y}_l^2$. To solve this, the final estimation can be given as a function of the estimates for each of the subarrays as:

$$\hat{x}_l = \frac{\gamma \hat{x}_l^1 + \nu \hat{x}_l^2}{\gamma + \nu}, \quad (3.25)$$

$$\hat{y}_l = \frac{\gamma \hat{y}_l^1 + \nu \hat{y}_l^2}{\gamma + \nu}, \quad (3.26)$$

where γ and ν are weighting coefficients that represent how reliable are the position estimates of each subarray. These weights can be set, for instance, as a function of the received signal strength at each of the subarrays. In this case γ and ν are given by

$$\gamma = \frac{\left(\mathbf{a}(\hat{\theta}_l^1) \otimes \mathbf{z}(\hat{\tau}_l^1) \right)^H \mathbf{X}^1 \mathbf{X}^{1H} \left(\mathbf{a}(\hat{\theta}_l^1) \otimes \mathbf{z}(\hat{\tau}_l^1) \right)}{\left(\mathbf{a}(\hat{\theta}_l^1) \otimes \mathbf{z}(\hat{\tau}_l^1) \right)^H \left(\mathbf{a}(\hat{\theta}_l^1) \otimes \mathbf{z}(\hat{\tau}_l^1) \right) \text{tr} \left(\mathbf{X}^1 \mathbf{X}^{1H} \right)}, \quad (3.27)$$

$$\nu = \frac{\left(\mathbf{a}(\hat{\theta}_l^2) \otimes \mathbf{z}(\hat{\tau}_l^2) \right)^H \mathbf{X}^2 \mathbf{X}^{2H} \left(\mathbf{a}(\hat{\theta}_l^2) \otimes \mathbf{z}(\hat{\tau}_l^2) \right)}{\left(\mathbf{a}(\hat{\theta}_l^2) \otimes \mathbf{z}(\hat{\tau}_l^2) \right)^H \left(\mathbf{a}(\hat{\theta}_l^2) \otimes \mathbf{z}(\hat{\tau}_l^2) \right) \text{tr} \left(\mathbf{X}^2 \mathbf{X}^{2H} \right)}. \quad (3.28)$$

3.6.2 Joint Direct Position Estimation

To avoid performing two independent estimations, one at each subarray, and consequently having merge both results, the estimation problem can be rewritten to a new parametrization. A set of coordinates for a transmitter will have a unique mapping to a given set of DOA and range at each subarray. Thus, the problem can be directly written with respect to these parameters, the coordinates of a transmitter, which are common to both subarrays.

The range parameter present in (3.1) can be written as

$$r_l^k = \sqrt{(x_{\text{Rx}_k} - x_l)^2 + (y_l)^2}, \quad (3.29)$$

and the DOA parameter θ_l^k can be written as

$$\theta_l^k = \arctan \left(\frac{y_l}{x_{\text{Rx}_k} - x_l} \right), \quad (3.30)$$

where $x_l = x_l^1 = x_l^2$ and $y_l = y_l^1 = y_l^2$.

Notice that after the re-parametrization, x_l and y_l are the only unknowns, and are common for both subarrays. Thus, to jointly estimate these parameters an augmented received signal matrix containing the inputs of subarrays Rx₁ and Rx₂ can be built as

$$\bar{\mathbf{X}} = \begin{bmatrix} \mathbf{X}^1 \\ \mathbf{X}^2 \end{bmatrix} \in \mathbb{R}^{2MK \times T}. \quad (3.31)$$

SAGE can then be applied to the dataset $\bar{\mathbf{X}}$ only. The expectation step for this parametrization is given by

$$\hat{\mathbf{X}}_i = \bar{\mathbf{X}}^k - \sum_{\substack{l=1 \\ l \neq i}}^L \left(\hat{\mathbf{h}}(\mathbf{p}_l) \hat{\mathbf{s}}_l \right), \quad (3.32)$$

where \mathbf{p}_l is a unique parameter vector for both subarrays given by

$$\mathbf{p}_l = [x_l, y_l]. \quad (3.33)$$

The cost function can then be written as

$$d(\mathbf{p}_l, \bar{\mathbf{X}}) = \left| \mathbf{h}(\mathbf{p}_l)^H \bar{\mathbf{X}} \right|^2. \quad (3.34)$$

and the optimization problems can be solved directly for the transmitter coordinates x_l and y_l . Since the dataset of both subarrays is jointly used to perform the estimation, the distance between both subarrays is a contributing factor for the accuracy of the estimates obtained, as it improves the identifiability of the range parameter.

This method is, however, not without its drawbacks. The computational load when employing the proposed joint estimation is higher than the computational load involved in the flip-flop method. The search space for the flip-flop method is limited to $[-\frac{p_i}{2}, \frac{p_i}{2}]$ in the DOA estimation, and to $[0, r_{\max}]$ in the range estimation, where r_{\max} is the maximum assumed communication distance. For the joint estimation method, the search space is, in theory, $[0, r_{\max}]$ for both x_l and y_l , assuming that a transmitter can be located anywhere within a semicircle with radius r_{\max} . While the search space may be reduced by assuming certain constraints on a transmitting vehicle, it might lead to large biases in the presence of non-line-of-sight (NLOS) components that are not arriving from a position contained within this reduced search space. Due to the restricted search space, such components will not be properly estimated, and, therefore, their contribution to the received signal will not be properly separated from the remaining components, leading to possibly large estimation biases.

3.6.3 DOA only estimation

Alternatively, if a planar wavefront model is assumed, a position estimation for the transmitter can be obtained by performing individual parameter estimations at each of the subarrays independently. In this case, the DOA estimation of an incoming signal at each of the subarrays can be used to triangulate the transmitter, thus a narrowband model is sufficient. Notice that in this case, the phase delay factor $\psi_l(m-1)^2$ does not need to be considered in (3.1), therefore it is not possible to obtain a direct ranging estimate. With DOA

estimates at hand and following the coordinate system shown in Figure 3.1 the lines representing the signal received at Rx_1 and Rx_2 can be written as

$$y_l^1 = \tan(\phi_l^1)x_l^1 - \tan(\phi_l^1)x_{Rx_1}, \quad (3.35)$$

$$y_l^2 = \tan(\phi_l^2)x_l^2 - \tan(\phi_l^2)x_{Rx_2}. \quad (3.36)$$

The position of the transmitter array can be obtained by calculating the point where (3.35) and (3.36) intersect, this can be done by solving

$$x_l = \frac{\tan(\phi_l^1)x_{Rx_1} - \tan(\phi_l^2)x_{Rx_2}}{\tan(\phi_l^1) - \tan(\phi_l^2)}, \quad (3.37)$$

$$y_l = \frac{\tan(\phi_l^1)\tan(\phi_l^2)(x_{Rx_2} - x_{Rx_1})}{\tan(\phi_l^2) - \tan(\phi_l^1)}. \quad (3.38)$$

While this method requires only DOA estimates at each of the subarrays to provide a position estimate for the transmitter, it is highly sensitive to imprecisions in the DOA estimates as it is a highly nonlinear function of these estimates.

3.6.4 Applicability of DOA Estimation for Positioning

Another possible solution for estimating the position of a transmitter is to use the DOAs and TDOAs that have been estimated. Using the estimates of the time delay of arrival (TDOA) $\hat{\tau}_l^k$, an estimation of the range between the transmitting and receiving array can be obtained as

$$\hat{r}_l^k = c\hat{\tau}_l^k, \quad (3.39)$$

where c is the speed of light. However, estimating the range by means of the TDOA is less robust than the one obtained through the SAGE estimation using the spherical wave model. This is due to the fact that the TDOA can only be estimated with respect to an internal reference at the receiver. That is, the TDOA estimation will be influenced by the synchronization between the transmitter and the receiver. Whatever error is present in synchronization will be present in the TDOA estimation, as the internal reference of the receiver will differ from the one of the transmitter. Therefore, unless a very reliable source of synchronization is present, this method of positioning is not suitable for safety-critical applications, as an error in TDOA estimation of just one nano second will already result in a ranging error of approximately 30 cm. If the synchronization between receiver and transmitter is to be relied on, the position of the transmitter can be estimated following (3.23) and (3.24) and substituting the range estimate for the one shown in (3.39).

3.7 Three Dimensional DOA Based Estimation

This section presents an extension of the proposed DOA only based method to the three dimensional case. The assumed scenario is presented in subsection 3.7.1. Subsection 3.7.2 defines the attitude angles while subsection 3.7.3 details the computation of the direction vectors. Finally, the proposed estimation method is presented in 3.7.4.

3.7.1 Scenario Description

This method assumes a MIMO system composed of a URA and an antenna array at the UAV as shown in Figure 3.3. The URA is the center of the 3-D space formed by x , y and z . The UAV is endowed with T transmitting antennas and the URA at the base station in Figure 3.3 has size M_v by M_v , where $M_v \geq T$.

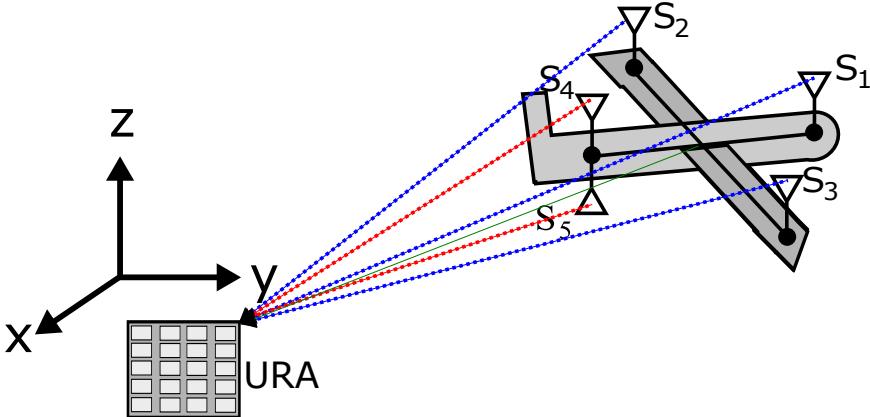


Figure 3.3: System model composed of a URA at the base station and an antenna array at the UAV.

The origin of a three dimensional coordinate system is placed center position of the URA and the far-field and narrowband assumptions are considered applicable. The antennas present at the UAV are considered to be points in space, and shall be denoted as $P_{s_1} \dots P_{s_d} \in \mathbb{R}^3$. The distance between any two antennas i and j , $\|\overrightarrow{P_{s_i}P_{s_j}}\|$, is considered to be known.

3.7.2 Definition of the Attitude Angles

Figure 3.4 illustrates a graphical representation of the definitions of pitch, yaw and roll for this work. The pitch axis is defined as the line that passes through the nose and tail antennas, the roll axis as the line that passes through

both wing antennas and the yaw axis as a line perpendicular to these two and passing through the center of the UAV. The attitude angles are defined as the rotation of the plane along these axes in relation to a given inertial frame of reference.

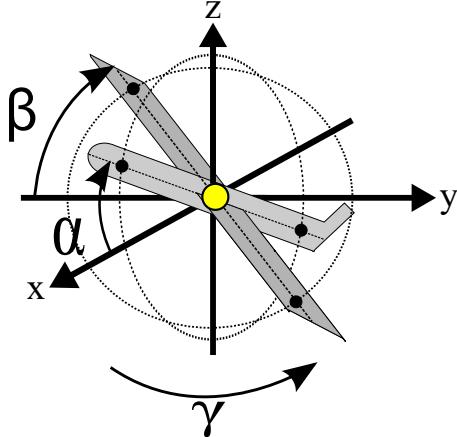


Figure 3.4: Definitions of pitch α , yaw β and roll γ

3.7.3 Direction vector computation

The first step is to estimate the DOAs of the signals transmitted by each of the T antennas of the UAV. Once the DOAs have been estimated, the next step is to generate a direction vector for the signals. The center of the URA is considered to be the origin of our coordinate system and shall be denoted as O . Antennas are considered points in space. Since the URA is the origin of the coordinate system, all signals pass through the origin, and thus an equation representing the coordinates x_{P_i} , y_{P_i} and z_{P_i} of any point P_i within the line representing the i -th signal can be written as

$$\begin{aligned} x_{P_i} &= \|m_{P_i}\| \cdot \sin(\phi_i) \cos(\theta_i), \\ y_{P_i} &= \|m_{P_i}\| \cdot \sin(\phi_i) \sin(\theta_i), \\ z_{P_i} &= \|m_{P_i}\| \cdot \cos(\phi_i), \end{aligned} \quad (3.40)$$

where θ_i and ϕ_i are the azimuth and elevation of the i -th signal respectively and $\|m_{P_i}\| \in \mathbb{R}$ is the magnitude of the vector $\overrightarrow{OP_i}$.

3.7.4 Position estimation

After obtaining the direction vector for all signals, an estimate of the positions of the antennas on the coordinate system can be obtained by solving a

system of equations based on the distance between each pair of antennas. The Euclidean distance between two points $A \in \mathbb{R}^3$ and $B \in \mathbb{R}^3$ is given by

$$m_{\overrightarrow{AB}} = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2}. \quad (3.41)$$

To find the estimate of the point where antennas i and j are located x_A , x_B , y_A , y_B , z_A and z_B can be replaced by x_{P_i} , x_{P_j} , y_{P_i} , y_{P_j} , z_{P_i} and z_{P_j} respectively. Since the distances between each pair of antennas and the DOAs of each signal are known, $\|m_{P_i}\|$ and $\|m_{P_j}\|$ can be obtained, yielding an estimation of the positions of antennas i and j with respect to the center of the URA. Note that there are $n(n - 1)/2$ available equations to choose from, where n is the number of antennas present at the UAV. Once the position of the antennas in the UAV has been calculated, the TRIAD algorithm can be used to determine its attitude. The TRIAD algorithm is detailed in Appendix H.

3.8 Simulation Results

This section presents a performance assessment of the proposed flip-flop and joint positioning methods. Subsection 3.8.1 presents the results obtained for simulated data while subsection 3.8.2 presents the results obtained using measured data.

3.8.1 Results for Simulated Data

The performance of the proposed flip-flop and joint positioning algorithms is assessed by a set of numerical simulations. The simulation assumes two antenna subarrays at the wing mirrors; these are considered to be ULAs composed of $M = 5$ antennas with inner element spacing of $\Delta_m = \frac{\lambda}{2}$. For the simulations, this work assumes the transmitter is using the LTE standard, with the maximum fast Fourier transform size of 2048, of which 1200 are effective subcarriers, and 15 kHz subcarrier spacing. Estimations are done assuming $T = 100$ OFDM symbols have been sampled with a normal cyclic prefix. Frequency correlation is introduced by following the model proposed in [51]. To assess the performance of the proposed methods with respect to the bandwidth of the transmitted signal, a bandwidth ranging from a tonal signal (narrowband) to 20 MHz (largest possible LTE channel bandwidth) is considered. The ranging estimation is done only using the range parameter present in the spherical wave model, since the TDOA estimate reliability depends on the synchronization between transmitter and receiver. The antenna arrays are assumed to be placed 1.80 m apart from each other, the average car width. Unless stated otherwise, the simulations assume the presence of 3 non line of sight (NLOS) components with respective scatterers randomly placed between the transmitter and the receiving array. The number of components is not assumed to be known for the purpose of estimating the parameters

using SAGE. Unless stated otherwise, the Rician K factor for the simulations is kept fixed at 3 dB. The SAGE algorithm for both the flip-flop and joint estimation methods is initialized using a position estimates obtained using the DOA only method. The root mean squared error (RMSE) is derived based on 1000 Monte Carlo simulations and is calculated as

$$\text{RMSE} = \sqrt{(\hat{x} - x)^2 + (\hat{y} - y)^2}. \quad (3.42)$$

The first set of simulations assesses the performance of the proposed method for different SNR. For this set of simulations the distance between the receiving array and the transmitter is kept fixed at 20 m and the SNR varies from -5 to 25 dB. Figure 3.5 show that, for the joint estimation method, the extra frequency samples provided by a broader bandwidth have a large impact on performance. When a narrowband case is considered, SAGE does not converge to an accurate estimate under low to moderate SNR conditions. However, once more frequency samples are introduced, the joint estimation method is capable of sub-centimetre accuracy even for low SNRs. The flip-flop method performance is shown to be more stable in the narrowband case, however, it is still outperformed by the joint method for high SNRs.

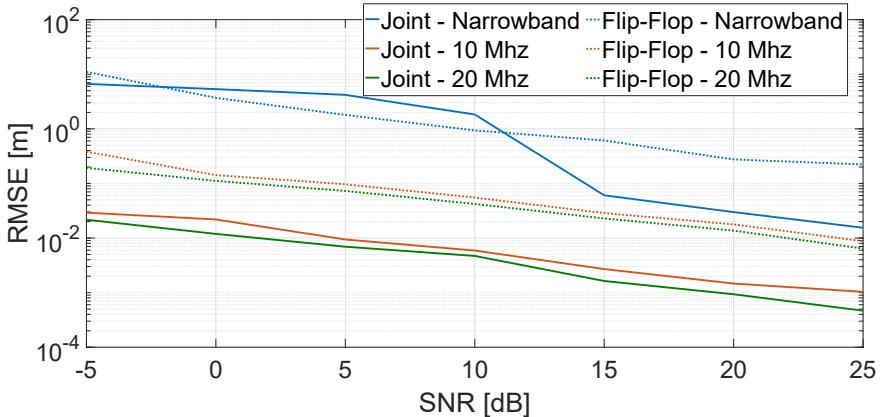


Figure 3.5: Position estimation error vs SNR

The second set of simulations studies the performance of the proposed methods dependence on the distance from the transmitter. For this set of simulations, the SNR is kept fixed at 15 dB, and the distance between the receiving array and transmitter varies from 5 to 50 meters. The results in Figure 3.6 show that, for both methods, the accuracy is degraded as the distance from the transmitter increases. Despite having the worst performance the flip-flop method is still capable of sub-meter performance for distances up to the 35 meters from the transmitter in the narrowband case.

The next set of simulations presented in Figure 3.7 studies the performance of the proposed algorithms as the number of NLOS components increases

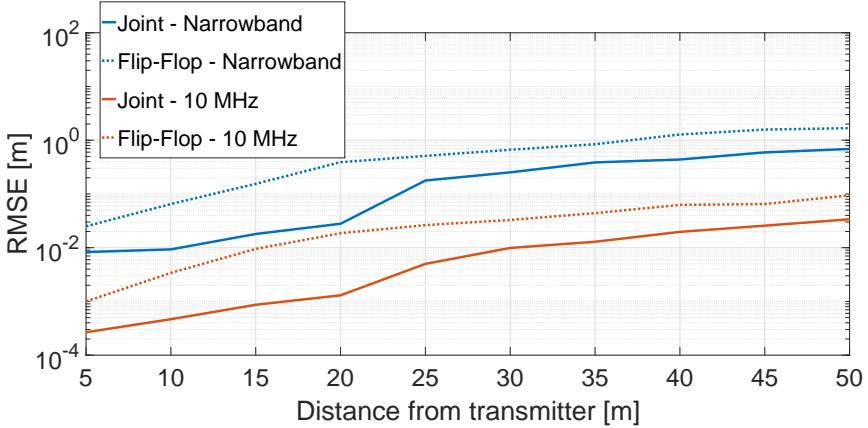


Figure 3.6: Position estimation error vs distance from source

while the K-factor is kept fixed at 3 dB, except for the case where no NLOS component is present, where the K-factor tends to infinity. For this simulation, the SNR is kept fixed at 15 dB, and the distance from the transmitter to the receiver is fixed at 20 m. The results show that the number of NLOS components has only a moderate impact on the performance of the proposed method if the K-factor is reasonable. As the number of NLOS components increases, the probability of closely spaced sources increases and the probability of NLOS paths with similar TDOAs increases, leading to a possible higher spatial correlation and frequency correlation. The increased spatial correlation is especially harmful as it can make a NLOS component non separable from the LOS component by SAGE. The results also show that, in the unlikely scenario where only the LOS component is present, the performance of the proposed methods is greatly improved.

Next, in order to assess the effect of the K-factor on the proposed methods, the number of NLOS components is kept fixed at 6 and the K-factor is varied. Figure 3.8 shows that, unlike increasing the number of NLOS components, the performance effects of large K-factor on the proposed methods is substantial.

The fifth set of simulations, with results presented in Figure 3.9, shows the estimation error over the x- and y-axis of the position estimation. For this set of simulations the distance from the transmitter is fixed at 20 meters and the performance presented is for the joint estimation method using a 10 MHz bandwidth. The results show that most of the estimation error is a result of the error over the y-axis. The range parameters is present only in the ψ component of the model. This component has a highly nonlinear relationship with the range parameter, thus, even small errors in the identification and estimation of this component will lead to rather large errors in the range estimate. However, even for low SNRs, the error over the y-axis component is kept below one centimetre.

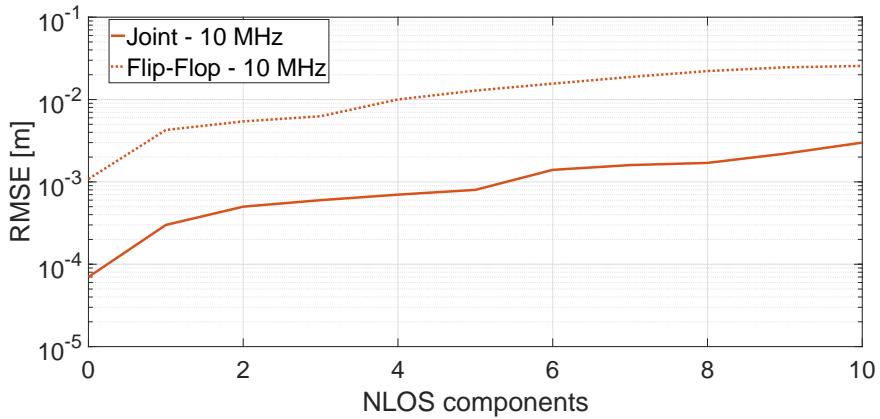


Figure 3.7: Position estimation error vs number of NLOS components

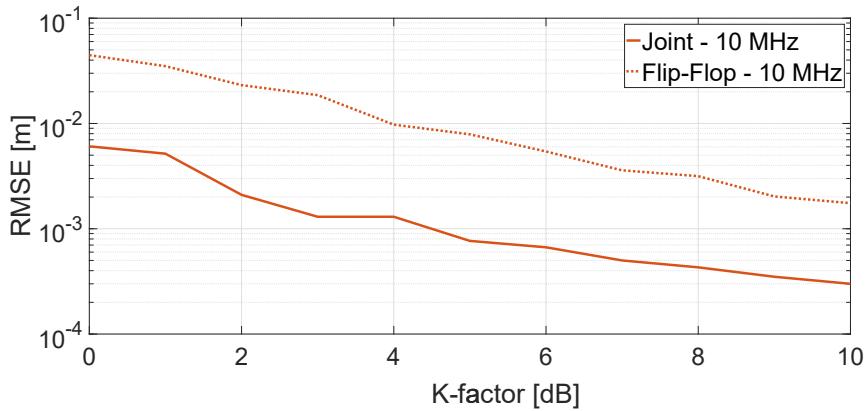


Figure 3.8: Position estimation error vs K-factor

3.8.2 Results for Real Data

To assess the performance of the proposed method in real world scenarios, a measured data set is used. The measurements were performed using a center frequency of 2.59 GHz with 50 MHz of bandwidth. The antenna array is a ULA composed of 128 elements with inner element spacing of $\frac{\lambda}{2} = 5.8\text{cm}$ and spanning 8 meters. Measurements were performed at the Lund University Department of Computer Science. The antenna array was placed on the roof of the entrance and the measurements were performed on a patio overlooked by the array. Figures 3.10a and 3.10b present an overview of the placement of the antenna and a picture of the antenna array that was used for the mea-

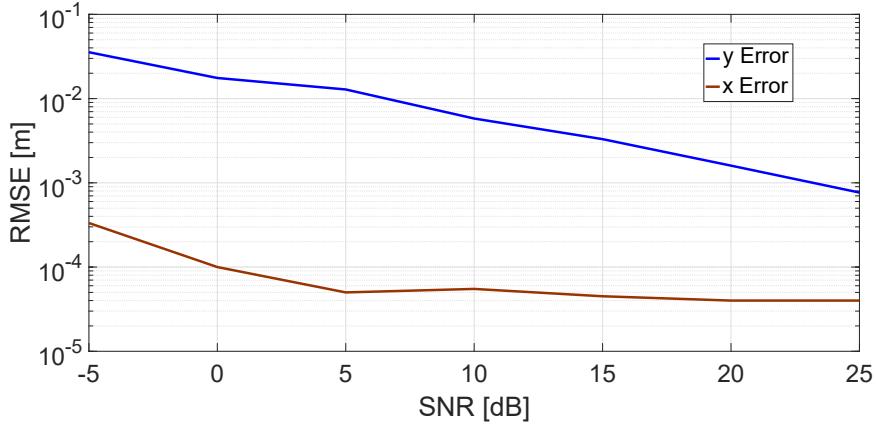


Figure 3.9: Position estimation error over the x- and y-axes

surements. Figure 3.11b presents a map of the location of users on the patio shown in Figure 3.11a with respect to the position of the antenna array.

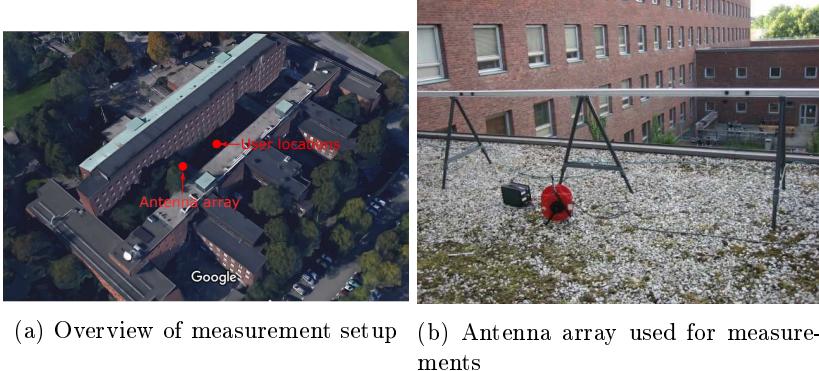


Figure 3.10: Antenna array and an overview of its placement

To assess the performance of the proposed algorithms using real data the positions of users 5, 10, 15, 19, and 26 are chosen to be estimated, as depicted in Figure 3.11b. One of the shortcomings of the measured data set is that the elevation of the antenna array is not the horizon, as assumed on the model, which might introduce some bias on the final position estimates. To better approach the vehicular scenario, two subarrays composed of 5 elements are selected from the 128 available elements. These subarrays are chosen as to be approximately 1.8 meters apart and equidistant from the center of the array.

Figures 3.12 and 3.13 present the performance of the proposed position estimation methods when applied to the measured data set. Synthetic noise is



(a) Patio where measurements were performed

(b) User distribution for measurements

Figure 3.11: Patio where measurements were performed and map of user distribution

added to the measurements in order to assess the performance of the proposed method with respect to SNR in real scenarios. The results show that, for user 5, there is a constant bias present in the estimation. This is due to the fact that user 5 has a very low elevation angle with respect to the antenna array, this will lead to the measurements not fitting the assumed model. Furthermore, the antenna responses for such a low elevation angle can be very different from the assumed isometric patterns. For the remaining users the performance behaves similarly to the results obtained with the simulated data set, users that are located further from the array will have a smaller localization accuracy than those closer to the array.

The overall difference in performance from the real data set to the simulated data set is most likely due to noise already present in the measurements and which is introduced by instruments such as power amplifiers and converters used in the set-up. Despite these imprecisions, the proposed joint estimation method was capable of achieving sub meter performance for an added SNR of up to 10 dB for all except user 5, and presented centimetre-level accuracy for a SNR of 20 dB. The flip-flop method was also able to achieve sub-meter accuracy for SNRs larger than 10 dB with a faster convergence rate when compared to the joint method.

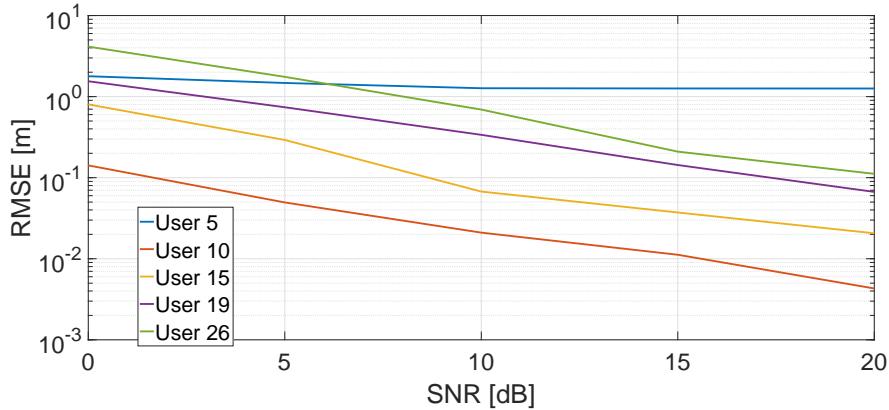


Figure 3.12: Position estimation error for measured data using the proposed joint estimation

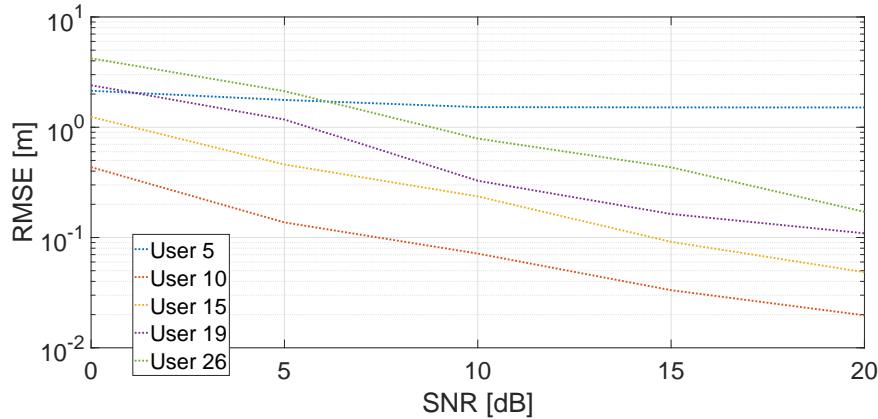


Figure 3.13: Position estimation error for measured data using the proposed flip-flop estimation

3.9 Summary

This chapter presented novel methods for position estimation for vehicular network systems. The proposed methods rely on using a broadband and spherical wave model.

The SAGE algorithm is used to provide position estimates in two distinct ways. The presented flip-flop method estimates the DOA and range of received signals at two subarrays separately, and uses the results of one array to update

the problem on the remaining one. This method has moderate accuracy and a fast convergence time.

The proposed joint estimation method is used to obtain a direct estimate of the coordinates of the transmitter by parametrizing the problem directly over such coordinates. While this method has a slower convergence time, it is capable of providing superior performance compared to the flip-flop method.

The performance of the proposed methods is studied using simulated and real measured data. Using the simulated data set this work shows that the proposed methods are capable of sub centimetre level accuracy when the only error source is additive white Gaussian noise. In case of using the real data set this work proves that the proposed method is capable of achieving centimetre-level accuracy for moderate distances and large SNRs, and will provide sub meter accuracy for moderate SNRs when other sources of error are introduced by the measurement chain.

The proposed methods can be used as a stand-alone localization estimation methods as well as for spoofing mitigation, as they rely solely on estimating parameters from the physical layer. Such parameters are extremely hard to fake with hardware available today. Furthermore, they do not require any specific data to be transmitted and can be applied when any data exchange is happening on the network. Therefore, the proposed methods do not imply an increased network load.

Conclusion

In this work the following research questions were tackled:

1. How to estimate the directions of arrival of correlated sources that are impinging on an antenna array that does not possess the necessary characteristics (geometry or electromagnetic response) for doing so?
2. Can array processing techniques be used to provide improved lifetime and communication performance in wireless sensor network?
3. Can array processing techniques be used to reliably estimate position of vehicles and pedestrians in vehicular scenarios based on their radio transmissions?

The first research question was addressed in Chapter 1. This chapter discussed two approaches for sector detection and discretization, one approach is based on the UT while the other is based on the PCA. The presented approaches aim to preserve the statistical characteristics of the signal to be transformed while minimizing the transformation bias introduced in the process.

Linear array interpolation was discussed and an extension of linear interpolation to the multidimensional case was presented. The approach generalizes the problem of linear interpolation to an arbitrary number of dimensions and allows the application of linear interpolation to arrays other than linear.

Two different alternatives for nonlinear array interpolation were also presented in Chapter 1. A way of performing array interpolation using the MARS method was presented. MARS is capable of expressing non-linear relationships between the real and the desired response. MARS is shown to achieve good performance for larger SNRs and can provide performances close to the CRLB. Furthermore, a novel approach for array interpolation extending GRNNs to the problem of array interpolation was also presented. Using GRNNs allows the interpolation to take into account larger sectors of the field of view of the antenna array and to achieve DOA estimation results with smaller estimation error compared to standard linear array interpolation. The proposed method is specially beneficial for arrays with a limited number of elements or in the case where a larger number of signals is present at the array, since in these cases the transformation error for the linear approach will lead to considerable bias

in the DOA estimation. As it is possible to implement the proposed GRNN array interpolation in a parallel manner, this leads to a fast interpolation especially suitable for real time applications. While nonlinear methods are capable of providing an improvement of the linear interpolation, it is computationally expensive and does not always provide a large improvement.

The performance of the proposed nonlinear interpolation methods was assessed using the response obtained from a real array. For such array, the proposed nonlinear methods were capable of vastly outperforming their linear counterpart.

The second research question was addressed in Chapter 2, which presents a framework for tackling practical problems for the implementation of C-MIMO in WSNs. The energy efficiency of C-MIMO in relation to that of multihop communications were studied and expressions comparing the consumption of multi-hop and C-MIMO are derived. This allows a routing mechanism to properly decide when to employ multi-hop or C-MIMO, achieving the minimal possible energy consumption for data transmission.

A method allowing the various nodes involved in C-MIMO to properly schedule transmission and reception was presented. A simple sliding correlator synchronization method was also presented as means of achieving coarse synchronization. More precise synchronization methods based on ML and a DFT were presented. It was shown that the ML method is capable of achieving a higher degree of accuracy, whereas the DFT based method is computationally less complex. Furthermore, an analysis of the effect of synchronization errors on C-MIMO schemes and of how the synchronization errors spread over the network was presented. Based on this analysis, a relation between symbol periods, clock drift, and synchronization error was derived. This can be used to ensure that an acceptable degree of synchronization is achieved for a given data rate by using C-MIMO.

Also in Chapter 2, a method that allows the formation of efficient C-MIMO cluster was proposed, availing the benefits provided by this communication scheme, such as reduced energy cost and better distribution of energy demand. The proposed method was shown to extend network lifetime by reducing individual node failures and to achieve networks with a more homogeneous distribution of energy between the nodes, resulting in networks that remain fully connected for longer periods.

The last research question was addressed in Chapter 3, where two different methods for position estimation in vehicular networks were presented. These methods rely on the presence of separate antenna arrays on the frame of the vehicle and use a spherical wave model in order to estimate the position of a transmitter.

A fast converging flip-flop method that performs estimations of DOA and range at each subarray individually was presented. This method uses the output of the estimation at each subarray to update the estimation problem on the remaining subarray. A joint estimation method, that uses the data of both

subarrays simultaneously to estimate the coordinates of the transmitter directly was also presented. This method is capable of superior accuracy when compared to the flip-flop method at the cost of increased convergence time.

Both methods were tested using simulated and a real measured data set. The simulated data shows that the proposed methods are capable of centimetre level accuracy under ideal conditions. The real data set shows that the proposed methods are still capable of achieving sub metre accuracy when other sources of error, such as the ones introduced by conversion and amplification, are present.

In summary, this work aimed to bring array processing closer to real life problems. Array interpolation allows the application, in real systems, of signal processing tools that were previously confined to theory. Cooperative MIMO can be of great use to WSNs to extend their lifetime and, with that, their usability. Finally, array processing can be used to help solve the problem of localization in the context of vehicular networks, solving the problem of reliability that is present when GPS or RSSI information is used. This can enable important technologies such as automated driving and platooning to become a reality.

Author's Publications

- [ath1] S. Caizzone, W. Elmarissi, M. A. M. Marinho, and F. Antreich. Direction of arrival estimation performance for compact antenna arrays with adjustable size. *The IEEE International Microwave Symposium (IMS2017)*, 2017.
- [ath2] E. P. de Freitas, R. K. Miranda, M. A. M. Marinho, J. P. C. L. da Costa, and C. E. Pereira. A practical implementation of a cooperative antenna array for wireless sensor networks. In *Conference on Smart Spaces*, pages 309–318. Springer, 2015.
- [ath3] E. P. Freitas, J. P. C. L. Costa, A. L. F. Almeida, and M. A. M. Marinho. Applying MIMO techniques to minimize energy consumption for long distances communications in wireless sensor networks. In Springer-Verlag, editor, *NEW2AN/ruSMART 2012*, pages 379–390, 2012.
- [ath4] J. Milanezi Jr, J. P. C. L. da Costa, R. S. Ferreira Junior, M. A. M. Marinho, R. A. Shayani, and R. T. de Sousa Júnior. Energy harvesting photovoltaic system to charge a cell phone in indoor environments. In *International Conference on Composite Materials & Renewable Energy Applications 2014*, 2014.
- [ath5] R. S. Ferreira Jr., M. A. M. Marinho, K. Liu, J. P. C. L. da Costa, A. V. Amaral, and H. C. So. Improved landing radio altimeter for unmanned aerial vehicles based on an antenna array. 2012.
- [ath6] J. P. A. Maranhão, J. P. C. L. da Costa, E. P. de Freitas, M. A. M. Marinho, and G. Del Galdo. Multi-hop cooperative XIXO transmission scheme for delay tolerant wireless sensor networks. In *Smart Antennas (WSA 2016); Proceedings of the 20th International ITG Workshop on*, pages 1–5. VDE, 2016.
- [ath7] M. A. M. Marinho, F. Antreich, S. Caizzone, J. P. C. L. da Costa, A. Vinel, and E. P. de Freitas. Robust nonlinear array interpolation for direction of arrival estimation of highly correlated signals. *Signal Processing*, 144:19–28, 2018.

- [ath8] M. A. M. Marinho, F. Antreich, and J. P. C. L. da Costa. Improved array interpolation for reduced bias in DOA estimation for gnss. In *The Institute of Navigation, International Technical Meeting (ION ITM) 2014*, 2014.
- [ath9] M. A. M. Marinho, F. Antreich, J. P. C. L. da Costa, and J. A. Nossek. Reduced rank tls array interpolation for DOA estimation. In *18th International ITG Workshop on Smart Antennas (WSA 2014)*, 2014.
- [ath10] M. A. M. Marinho, F. Antreich, J. P. C. L. da Costa, and J. A. Nossek. A signal adaptive array interpolation approach with reduced transformation bias for DOA estimation of highly correlated signals. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '14). IEEE International Conference on*, 2014.
- [ath11] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, A. L. F. de Almeida, G. Del Galdo, E. P. de Freitas, and A. Vinel. Array interpolation based on multivariate adaptive regression splines. In *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016 IEEE*, pages 1–5. IEEE, 2016.
- [ath12] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, E. P. de Freitas, and A. Vinel. Adaptive communication and cooperative MIMO cluster formation for improved lifetime in wireless sensor networks. In *2016 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE), Aachen, Germany, September 26-29, 2016*, 2016.
- [ath13] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, and L. R. A. X. de Menezes. Unscented transformation based array interpolation. In *Proc. 40th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, Australia, April 2015.
- [ath14] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, and L. R. A. X. de Menezes. Multidimensional array interpolation applied to direction of arrival estimation. In *19th International ITG Workshop on Smart Antennas (WSA 2015)*, 2015.
- [ath15] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, A. L. F. de Almeida, and R. T. de Sousa Jr. Using cooperative MIMO techniques and UAV relay networks to support connectivity in sparse wireless sensor networks. In *Computing, Management and Telecommunications (ComManTel), 2013 International Conference on*, 2013.
- [ath16] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, A. L. F. de Almeida, and R. T. de Sousa Jr. Using MIMO techniques to enhance communication among static and mobile nodes in wireless sensor networks. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, 2013.

- [ath17] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, and R. T. de Sousa Jr. Applying cooperative MIMO technique in an adaptive routing mechanism for wireless sensor networks. In *2013 IEEE Conference on Wireless Sensors (ICWiSe2013)*, 2013.
- [ath18] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, and R. T. de Sousa Jr. Synchronization for cooperative MIMO in wireless sensor networks. In *NEW2AN/ruSMART 2013*, 2013.
- [ath19] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, and R. T. de Sousa Jr. Sensor localization via diversely polarized antennas. In *International Workshop on Performance Control in Wireless Sensor Networks (PWSN)*, 2014.
- [ath20] M. A. M. Marinho, R. S. Ferreira Jr., J. P. C. L. da Costa, E. P. de Freitas, K. Liu, H. C. So, and F. Antreich. Antenna array based positioning scheme for unmanned aerial vehicles. In *17th International ITG Workshop on Smart Antennas (WSA 2013)*, 2013.
- [ath21] M. A. M. Marinho, A. Vinel, F. Antreich, J. P. C. L. da Costa, and E. P. de Freitas. Antenna array based localization scheme for vehicular networks. In *Computer and Information Technology (CIT), 2017 IEEE International Conference on*, pages 142–146. IEEE, 2017.
- [ath22] R. K. Miranda, J. P. C. L. da Costa, M. A. M. Marinho, E. P. de Freitas, and R. T. de Sousa Jr. Evaluation of space-time-frequency (stf)-coded MIMO-OFDM systems in realistic channel models. In *28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014)*, 2014.

References

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: A survey. *Computer Networks*, 38:393–422, 2002.
(Cited on pages 53 and 1)
- [2] K. P. Ang, L. and Seng, L. W. Chew, L. S. Yeong, and W. C. Chia. *Wireless Multimedia Sensor Networks on Reconfigurable Hardware: Information Reduction Techniques*. Springer Science & Business Media, 2013.
(Cited on pages 63 and 58)
- [3] A. J. Barabell. Improving the resolution performance of eigenstructured based direction-finding algorithms. In *Proceedings of ICASSP 83*, 1983.
(Cited on pages 8, 9, 19, 11, and 22)
- [4] M. S. Bartlett. Smoothing periodograms from time series with continuous spectra. *Nature*, 161:686–687, 1948.
(Cited on pages 8, 9, 23, 10, 25, and 32)
- [5] C. Bettstetter. On the minimum node degree and connectivity of a wireless multihop network. In *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing*, pages 80–91. ACM, 2002.
(Cited on pages 78 and 73)
- [6] M. Bhardwaj and A. P. Chandrakasan. Upper bounds on the lifetime of sensor networks. In *Proceedings of IEEE ICC01*, 2001.
(Cited on page 60)
- [7] M. Biguesh and A. B. Gershman. Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals. *IEEE Transactions on Signal Processing*, 54:884–893, 2006.
(Cited on pages 57 and 55)
- [8] H. Black. A passive system for determining the attitude of a satellite. *AIAA Journal*, 2:1350–1351, July 1964.
(Cited on page 151)

- [9] H. Black. Early developments of transit, the navy navigation satellite system. *Journal of Guidance, Control and Dynamics*, 13:577–585, July 1990.
(Cited on page 151)
- [10] Y. Bresler and A. Macovski. Exact maximum likelihood estimation of superimposed exponentials signals in noise. *IEEE ASSP Magazine*, 34:1081–189, 1986.
(Cited on pages 8, 9, and 11)
- [11] T. P. Bronez. Sector interpolation of non-uniform arrays for efficient high resolution bearing estimation. In *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, 1988.
(Cited on pages 10, 26, 28, and 12)
- [12] M. Buhren, M. Pesavento, and J. F. Bohme. A new approach to array interpolation by generation of artificial shift invariances: Interpolated ESPRIT. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, 2003.
(Cited on pages 10, 15, 21, 12, 18, and 24)
- [13] M. Buhren, M. Pesavento, and J. F. Böhme. Virtual array design for array interpolation using differential geometry. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, volume 2, 2004.
(Cited on pages 10, 19, 12, and 24)
- [14] S. Caizzone and A. Dreher. Miniaturized DRA array for GNSS applications. In *2015 9th European Conference on Antennas and Propagation (EuCAP)*, pages 1–2. IEEE, 2015.
(Cited on pages 44 and 45)
- [15] S. Caizzone, W. Elmarissi, M. A. M. Marinho, and F. Antreich. Direction of arrival estimation performance for compact antenna arrays with adjustable size. *The IEEE International Microwave Symposium (IMS2017)*, 2017.
(Cited on page 44)
- [16] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings IEEE*, 57:1408–1418, 1969.
(Cited on pages 8, 9, 23, 10, 25, and 32)
- [17] C. Chen, J. Ma, and K. Yu. Designing energy-efficient wireless sensor networks with mobile sinks. In *Sensys06*, 2006.
(Cited on pages 53, 59, and 1)
- [18] S. Cho and A. Chandrakasan. Energy-efficient protocols for low duty cycle wireless microsensor. In *Proceedings of the 33rd Annual Hawaii*

International Conference on System Sciences, 2000.

(Cited on pages 53 and 1)

- [19] P. J. Chung and J. F. Böhme. Comparative convergence analysis of EM and SAGE algorithms in DOA estimation. *IEEE Transactions on Signal Processing*, 49:2940–2949, 2001.
(Cited on pages 8 and 10)
- [20] D. S. J. D. Couto, D. Aguayo, J. Bicket, and R. Morris. A high-throughput path metric for multi-hop wireless routing. *Wirel. Netw.*, 11:419–435, 2005.
(Cited on page 59)
- [21] J. H. Cozzens and M. J. Sousa. Source enumeration in a correlated signal environment. *IEEE Transactions on Acoustics Speech and Signal Processing*, 42:304–317, 1994.
(Cited on pages 14 and 17)
- [22] S. B. Cruz, T. E. Abrudan, Z. Xiao, N. Trigoni, and J. Barros. Neighbor-aided localization in vehicular networks. *IEEE Transactions on Intelligent Transportation Systems*, 2017.
(Cited on page 89)
- [23] S. Cui and A. J. Goldsmith. Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks. *IEEE Journal on Selected Areas in Communications*, 2004.
(Cited on pages 53 and 2)
- [24] J. P. C. L. da Costa, M. Haardt, F. Romer, and G. Del Galdo. Enhanced model order estimation using higher-order arrays. *Conference Record of The Forty-First Asilomar Conference on Signals, Systems & Computers*, pages 412–416, 2007.
(Cited on pages 15 and 18)
- [25] J. P. C. L. da Costa, K. Liu, H. C. So, F. Roemer, M. Haardt, and S. Schwarz. Generalized multidimensional prewhitening for enhanced signal reconstruction and parameter estimation. In *Signal Processing, Elsevier publisher*, 2013.
(Cited on pages 38 and 42)
- [26] J. P. C. L. da Costa, F. Roemer, and M. Haardt. Deterministic prewhitening to improve subspace parameter estimation techniques in severely colored noise environments. In *Proc. 54th International Scientific Colloquium (IWK)*, 2009.
(Cited on pages 38 and 42)
- [27] J. P. C. L. da Costa, F. Roemer, M. Haardt, and R. T. de Sousa Jr. Multi-dimensional model order selection. *EURASIP Journal on Advances*

in Signal Processing, 26, 2011.

(Cited on pages 15 and 18)

- [28] J. P. C. L. da Costa, F. Roemer, M. Weis, and M. Haard. Robust R-D parameter estimation via closed-form PARAFAC. In *Proc. ITG Workshop on Smart Antennas (WSA '10)*, 2010.
(Cited on pages 17, 38, 20, and 42)
- [29] J. P. C. L. da Costa, D. Schulz, F. Roemer, M. Haardt, and J. A. Apolinario Jr. Robust R-D parameter estimation via closed-form PARAFAC in kronecker colored environments. In *Proc. 7-th International Symposium on Wireless Communications Systems (ISWCS 2010)*, 2010.
(Cited on pages 17, 38, 20, and 42)
- [30] V. Daiya, J. Ebenezer, S. A. V. S. Murty, and B. Raj. Experimental analysis of RSSI for distance and position estimation. In *Recent Trends in Information Technology (ICRTIT), 2011 International Conference on*, pages 1093–1098. IEEE, 2011.
(Cited on pages 72 and 67)
- [31] L. M. S. De Souza, H. Vogt, and M. Beigl. A survey on fault tolerance in wireless sensor networks. *Interner Bericht. Fakultät für Informatik, Universität Karlsruhe: Karlsruhe, Germany*, 2007.
(Cited on pages 81 and 74)
- [32] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Soc. B.*, 39(1), 1977.
(Cited on pages 7, 9, 33, 10, and 38)
- [33] F. A. Dietrich. A tutorial on channel estimation with SAGE. *Technical Report TUM-LNS-TR-06-03*, 2006.
(Cited on pages 7, 9, 33, 10, and 38)
- [34] I. Dietrich and F. Dressler. On the lifetime of wireless sensor networks. *ACM Transactions on Sensor Networks*, 5, 2009.
(Cited on pages 53 and 1)
- [35] P. Dollar, C. Wojek, B. Schiele, and P. Perona. Pedestrian detection: An evaluation of the state of the art. *IEEE transactions on pattern analysis and machine intelligence*, 34(4):743–761, 2012.
(Cited on page 88)
- [36] J. Elson, L. Girod, and D. Estrin. Fine-grained network time synchronization using reference broadcasts. In *Proceedings of the Fifth Symposium on Operating Systems Design and Implementation*, 2002.
(Cited on pages 64 and 59)

- [37] J. E. Evans, J. R. Johnson, and D. F. Sun. Application of advanced signal processing techniques to angle of arrival estimation in ATC navigation and surveillance system. Technical report, 1982.
(Cited on pages 8, 14, 11, and 16)
- [38] J. A. Fessler and A. O. Hero. Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Processing*, 42(10), October 1994.
(Cited on pages 7, 9, 33, 10, and 38)
- [39] B. Friedlander. The root-MUSIC algorithm for direction finding with interpolated arrays. *Signal Processing*, 30:15–29, 1993.
(Cited on pages 10, 19, 12, and 24)
- [40] B. Friedlander and A. J. Weiss. Direction finding using spatial smoothing with interpolated arrays. *Aerospace and Electronic Systems, IEEE Transactions on*, 28:574–587, 1992.
(Cited on pages 10, 33, 12, and 38)
- [41] J. H. Friedman. Multivariate adaptive regression splines. *The annals of statistics*, pages 1–67, 1991.
(Cited on pages 143, 144, 4, 113, and 114)
- [42] S. Ganeriwal, D. Ganesan, H. Shim, V. Tsiatsis, and B. Srivastava. Estimating clock uncertainty for efficient duty-cycling in sensor networks. In *Proceedings of the Third International Conference on Embedded Networked Sensor Systems (Sensys)*, 2005.
(Cited on pages 64 and 59)
- [43] S. Ganeriwal, R. Kumar, and M. B. Srivastava. Timing-sync protocol for sensor networks. In *Proceedings of the Sensys03*, 2003.
(Cited on pages 64 and 59)
- [44] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin. Highly-resilient, energy-efficient multipath routing in wireless sensor networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(4):1125, 2001.
(Cited on pages 53 and 1)
- [45] D. M. Gavrila, J. Giebel, and S. Munder. Vision-based pedestrian detection: The protector system. In *Intelligent Vehicles Symposium, 2004 IEEE*, pages 13–18. IEEE, 2004.
(Cited on page 88)
- [46] A. M. Geronimo, D. Lopez, A. D. Sappa, and T. Graf. Survey of pedestrian detection for advanced driver assistance systems. *IEEE transactions on pattern analysis and machine intelligence*, 32(7):1239–1258, 2010.
(Cited on page 88)

- [47] M. Haardt. Efficient one-, two- and multidimensional high-resolution array signal processing. In *3-8265-2220-6*, Aachen, Germany, 1996. Shaker Verlag.
 (Cited on page 15)
- [48] M. Haardt, F. Roemer, and G. Del Gado. Higher-order SVD based subspace estimation to improve parameter estimation accuracy in multi-dimensional harmonic retrieval problems. *IEEE Transactions on Signal Processing*, 56(7):3198–3213, July 2008.
 (Cited on pages 38 and 42)
- [49] A. Habibovic and J. Davidsson. Requirements of a system to reduce car-to-vulnerable road user crashes in urban intersections. *Accident Analysis & Prevention*, 43(4):1570–1580, 2011.
 (Cited on page 88)
- [50] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proceedings of the 33rd Hawaii International Conference on System Sciences*, 2000.
 (Cited on pages 53 and 1)
- [51] P. Hoher. A statistical discrete-time model for the wssus multipath channel. *IEEE Transactions on vehicular technology*, 41(4):461–468, 1992.
 (Cited on page 101)
- [52] S. M. R. Hosseini and M. A. Sebt. Array interpolation using covariance matrix completion of minimum-size virtual array. *IEEE Signal Processing Letters*, 24(7):1063, 2017.
 (Cited on pages 10 and 12)
- [53] R. S. Ferreira Jr., M. A. M. Marinho, K. Liu, J. P. C. L. da Costa, A. V. Amaral, and H. C. So. Improved landing radio altimeter for unmanned aerial vehicles based on an antenna array. 2012.
 (Cited on pages 7 and 10)
- [54] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *Int. symp. aerospace/defense sensing, simul. and controls*, volume 3, pages 3–2. Orlando, FL, 1997.
 (Cited on pages 26, 4, and 28)
- [55] M. Khurana, C. Ramakrishna, and S. N. Panda. Capacity enhancement using MU-MIMO in vehicular ad hoc network. *International Journal of Applied Engineering Research*, 12(16):5872–5883, 2017.
 (Cited on page 89)
- [56] N. Kim, Y. Lee, and H. Park. Performance analysis of MIMO system with linear MMSE receiver. *Wireless Communications, IEEE Transactions on*,

7(11):4474–4478, 2008.

(Cited on pages 57 and 55)

- [57] T. King, H. Füßler, M. Transier, and W. Effelsberg. Dead-reckoning for position-based forwarding on highways. In *Proc. of the 3rd International Workshop on Intelligent Transportation (WIT 2006)*, pages 199–204, 2006.
(Cited on pages 89 and 3)
- [58] D. Kosmanos, N. Prodromou, A. Argyriou, L. A. Maglaras, and H. Janicke. MIMO techniques for jamming threat suppression in vehicular networks. *Mobile Information Systems*, 2016, 2016.
(Cited on page 89)
- [59] H. Krim and M. Viberg. Two decades of array signal processing research: The parametric approach. *IEEE Signal Processing Magazine*, 13:67–94, 1996.
(Cited on pages 7, 14, and 17)
- [60] V. Kumar, S. Mishra, and N. Chand. Applications of VANETs: Present & future. *Communications and Network*, 5(01):12, 2013.
(Cited on pages 88 and 3)
- [61] T. Kurpjuhn, M. T. Ivrlac, and J. A. Nossek. Vandermonde invariance transformation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, 2001.
(Cited on pages 9, 17, 12, and 20)
- [62] Erik G Larsson, Ove Edfors, Fredrik Tufvesson, and Thomas L Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, 2014.
(Cited on page 7)
- [63] B. K. Lau, G. J. Cook, and Y. H. Leung. An improved array interpolation approach to DOA estimation in correlated signal environments. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, 2004.
(Cited on pages 10, 21, 12, and 24)
- [64] B. K. Lau, M. Viberg, and Y. H. Leung. Data-adaptive array interpolation for DOA estimation in correlated signal environments. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, 2005.
(Cited on pages 10, 21, 12, and 24)
- [65] T. Leinmuller, E. Schoch, and F. Kargl. Position verification approaches for vehicular ad hoc networks. *IEEE Wireless Communications*, 13(5):16–21, 2006.
(Cited on pages 89 and 3)

- [66] P. Lettieri and M. B. Srivastava. Adaptive frame length control for improving wireless link throughput, range and energy efficiency. In *INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, 1998.
(Cited on pages 53 and 1)
- [67] K. Lin, J. Luo, L. Hu, M. S. Hossain, and A. Ghoneim. Localization based on social big data analysis in the vehicular networks. *IEEE Transactions on Industrial Informatics*, 13(4):1932–1940, 2017.
(Cited on page 89)
- [68] C. Liu, P. P. Vaidyanathan, and P. Pal. Coprime coarray interpolation for DOA estimation via nuclear norm minimization. In *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*, pages 2639–2642. IEEE, 2016.
(Cited on pages 10 and 12)
- [69] W. Lou, W. Liu, and Y. Zhang. Performance optimization using multi-path routing in mobile ad hoc and wireless sensor networks. *Combinator. Optim. Commun. Netw.*, 2:117–146, 2006.
(Cited on page 59)
- [70] D. Lucarelli and I.-J. Wang. Decentralized synchronization protocols with nearest neighbor communication. In *Proceedings of the Sensys04*, 2004.
(Cited on pages 64 and 59)
- [71] J. Mannermaa, K. Kalliomaki, T. Mansten, and S. Turunen. Timing performance of varios GPS receivers. In *Proceedings of the 1999 Joint Meeting of the European Frequency and Time Forum and the IEEE International Frequency Control Symposium*, 1999.
(Cited on pages 64 and 59)
- [72] M. A. M. Marinho, F. Antreich, J. P. C. L. da Costa, and J. A. Nossek. A signal adaptive array interpolation approach with reduced transformation bias for DOA estimation of highly correlated signals. In *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '14). IEEE International Conference on*, 2014.
(Cited on pages 44, 45, 46, and 47)
- [73] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, A. L. F. de Almeida, G. Del Galdo, E. P. de Freitas, and A. Vinel. Array interpolation based on multivariate adaptive regression splines. In *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016 IEEE*, pages 1–5. IEEE, 2016.
- [74] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, and L. R. A. X. de Menezes. Unscented transformation based array interpolation. In *Proc.*

40th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brisbane, Australia, April 2015.
(Cited on page 10)

- [75] M. A. M. Marinho, J. P. C. L. da Costa, F. Antreich, and L. R. A. X. de Menezes. Multidimensional array interpolation applied to direction of arrival estimation. In *19th International ITG Workshop on Smart Antennas (WSA 2015)*, 2015.
- [76] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, A. L. F. de Almeida, and R. T. de Sousa Jr. Using cooperative MIMO techniques and UAV relay networks to support connectivity in sparse wireless sensor networks. In *Computing, Management and Telecommunications (Com-ManTel), 2013 International Conference on*, 2013.
(Cited on pages 53 and 2)
- [77] M. A. M. Marinho, E. P. de Freitas, J. P. C. L. da Costa, A. L. F. de Almeida, and R. T. de Sousa Jr. Using MIMO techniques to enhance communication among static and mobile nodes in wireless sensor networks. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, 2013.
(Cited on pages 53 and 2)
- [78] F. L. Markley. Attitude determination using vector observations and the singular value decomposition. In *AAS/AIAA Astrodynamics Specialist Conference*, 1987.
(Cited on page 153)
- [79] G. J. McLachlan and T. Krishnan. *The Em Algorithm and Extensions*. John Wiley & Sons, Inc., New York, 1997.
(Cited on pages 7, 9, 33, 10, and 38)
- [80] V. Mhatre and C. Rosenberg. Design guidelines for wireless sensor networks: Communication, clustering and aggregation. In *Elsevier Ad Hoc Networks*, 2004.
(Cited on page 60)
- [81] M. I. Miller and D. R. Fuhrmann. Maximum-likelihood narrow-band direction finding and the EM algorithm. *IEEE Transactions on Acoustics Speech and Signal Processing*, 38:1560–1577, 1990.
(Cited on pages 7, 9, 33, 10, and 38)
- [82] R. A. F. Mini and A. A. F. Loureiro. Energy in wireless sensor networks. In *Middleware for Network Eccentric and Mobile Applications*, pages 3–24. Springer, 2009.
(Cited on pages 53 and 1)

- [83] M. Modsching, R. Kramer, and K. ten Hagen. Field trial on GPS accuracy in a medium size city: The influence of built-up. In *3Rd workshop on positioning, navigation and communication*, pages 209–218, 2006.
(Cited on pages 89 and 3)
- [84] T. K. Moon. The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, November 1996.
(Cited on pages 7, 9, 33, 10, and 38)
- [85] T. Nadeem, P. Shankar, and L. Iftode. A comparative study of data dissemination models for VANETs. In *Mobile and Ubiquitous Systems-Workshops, 2006. 3rd Annual International Conference on*, pages 1–10. IEEE, 2006.
(Cited on pages 89 and 3)
- [86] B. Ottersten, M. Viberg, and T. Kailath. Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data. *IEEE Transactions on Signal Processing*, 40(3), March 1992.
(Cited on pages 7, 9, 33, 10, and 38)
- [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*, pages 99–151. Springer, 1993.
(Cited on pages 7, 9, 33, 45, and 47)
- [88] R. Parker and S. Valaee. Vehicular node localization using received-signal-strength indicator. *IEEE Transactions on Vehicular Technology*, 56(6):3371–3380, 2007.
(Cited on pages 89 and 3)
- [89] M. Pesavento, A. B. Gershman, and Z.-Q. Luo. Robust array interpolation using second-order cone programming. *Signal Processing Letters*, 9:8–11, 2002.
(Cited on pages 10, 19, 26, 28, 12, and 24)
- [90] S. U. Pillai and B. H. Kwon. Forward/backward spatial smoothing techniques for coherent signal identification. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37:8–9, January 1989.
(Cited on pages 8, 13, 11, and 16)
- [91] E. Radoi and A. Quinquis. A new method for estimating the number of harmonic components in noise with application in high resolution radar. *EURASIP Journal on Applied Signal Processing*, pages 1177–1188, 2004.
(Cited on pages 15, 32, 44, 18, 37, and 46)
- [92] G. C. Reinsel and R. P. Velu. *Multivariate Reduced-rank Regression*. Springer New York, 1998.
(Cited on pages 32 and 37)

- [93] R. Roy and T. Kailath. ESPRIT - estimation of signal parameters via rotation invariance techniques. *IEEE Transactions on Acoustics Speech and Signal Processing*, 17, 1989.
(Cited on pages 8, 9, 15, 16, 19, 44, 22, and 46)
- [94] F. A. Santos, A. T. Akabane, R. S. Yokoyama, A. A. F. Loureiro, and L. A. Villas. A roadside unit-based localization scheme to improve positioning for vehicular networks. In *Vehicular Technology Conference (VTC-Fall), 2016 IEEE 84th*, pages 1–5. IEEE, 2016.
(Cited on page 89)
- [95] R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3), March 1986.
(Cited on pages 8, 9, and 10)
- [96] C. Shen, Y. Zhu, S. Zhou, and J. Jiang. On the performance of V-BLAST with zero-forcing successive interference cancellation receiver. In *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, volume 5, pages 2818–2822. IEEE, 2004.
(Cited on pages 57 and 55)
- [97] E. Shih, B. H. Calhoun, S.H. Cho, and A. P. Chandrakasan. Energy-efficient link layer for wireless microsensor networks. In *VLSI, 2001. Proceedings. IEEE Computer Society Workshop on*, 2001.
(Cited on pages 53 and 1)
- [98] E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, and A. Chandrakasan A. Wang. Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks. In *Proceedings of ACM MobiCom01*, 2001.
(Cited on pages 53 and 1)
- [99] V. Shnayder, M. Hempstead, B. Chen, G. W. Allen, and M. Welsh. Simulating the power consumption of large-scale sensor network applications. In *Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 188–200. ACM, 2004.
(Cited on pages 63, 78, 58, and 72)
- [100] M. D. Shuster. Maximum likelihood estimation of spacecraft attitude. *The Journal of the Astronautical Sciences*, 108:2089–2117, 2001.
(Cited on page 153)
- [101] P. Stoica and A. Nehorai. A novel eigenanalysis method for direction estimation. In *Proceedings IEEE F.*, 1990.
(Cited on pages 8, 9, and 11)

- [102] D. V. A. H. G. Swaroop and J. K. Hedrick. Constant spacing strategies for platooning in automated highway systems. *Journal of dynamic systems, measurement, and control*, 121(3):462–470, 1999.
(Cited on pages 89 and 3)
- [103] A. Triwinarko, I. Dayoub, and P. Wikanta. Using MIMO and cross layer design for VANETs: A review. In *Signals and Systems (ICSigSys), 2017 International Conference on*, pages 13–18. IEEE, 2017.
(Cited on page 89)
- [104] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge university press, 2005.
(Cited on pages 58 and 56)
- [105] T. Virtanen. Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. *IEEE transactions on audio, speech, and language processing*, 15(3):1066–1074, 2007.
(Cited on page 7)
- [106] M. Wax and T. Kailath. Detection of signals by information by information theoretic criteria. *IEEE Transactions on Acoustics Speech and Signal Processing*, 33:387–392, 1985.
(Cited on pages 15 and 18)
- [107] M. Weiser, B. Welch, A. Demers, and S. Shenker. Scheduling for reduced CPU energy. In *Proceedings of 1st USENIX Symposium on Operating System Design and Implementation*, pages 13–23, November 1994.
(Cited on pages 53 and 1)
- [108] A. J. Weiss and M. Gavish. The interpolated ESPRIT algorithm for direction finding. In *Electrical and Electronics Engineers in Israel, 1991. Proceedings., 17th Convention of*, 1991.
(Cited on pages 21 and 24)
- [109] G. Wener-Allen, G. Tewari, A. Patel, M. Welsh, and R. Nagpal. Firefly inspired sensor network synchronicity with realistic radio effects. In *Proceedings of the Third International Conference on Embedded Networked Sensor Systems (Sensys)*, 2005.
(Cited on pages 64 and 59)
- [110] N. Wiener. Extrapolation, interpolation and smoothing of stationary time series. *MIT Press*, 1949.
(Cited on pages 14 and 16)
- [111] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela. V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. In *Signals, Systems and Electronics*, 1998.
(Cited on pages 56 and 54)

- [112] A. Woo, T. Tong, and D Culler. Taming the underlying challenges of reliable multihop routing in sensor networks. In *Proc. ACM SenSys 2003*, 2003.
(Cited on page 59)
- [113] R. Xu and F. C. M. Lau. Performance analysis for MIMO systems using zero forcing detector over fading channels. *IEE Proceedings-Communications*, 153(1):74–80, 2006.
(Cited on pages 57 and 55)
- [114] A. Zanella, M. Chiani, and M. Z. Win. MMSE reception and successive interference cancellation for MIMO systems with high spectral efficiency. *Wireless Communications, IEEE Transactions on*, 4(3):1244–1253, 2005.
(Cited on pages 57 and 55)
- [115] X. Zhu and R. D. Murch. Performance analysis of maximum likelihood detection in a MIMO antenna system. *Communications, IEEE Transactions on*, 50(2):187–191, 2002.
(Cited on pages 57 and 55)

Appendices

Appendix A

Left Centro-Hermitian Matrix

Definition A.1. A exchange matrix \mathbf{Q}_n is defined as a $n \times n$ matrix with one in its anti-diagonal and zeros elsewhere

$$\mathbf{Q}_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{C}^{n \times n}. \quad (\text{A.1})$$

Thus, left hand multiplication of a matrix by an exchange matrix reverses the order of its rows while a right hand multiplication will reverse the order of its columns.

Definition A.2. A matrix $\mathbf{W} \in \mathbb{C}^{i \times j}$ is called centro-Hermitian if

$$\mathbf{Q}_i \mathbf{W}^* \mathbf{Q}_j = \mathbf{W}, \quad (\text{A.2})$$

where $(\cdot)^*$ stands for complex conjugation.

Definition A.3. A matrix $\mathbf{W} \in \mathbb{C}^{i \times j}$ is called left centro-Hermitian if

$$\mathbf{Q}_i \mathbf{W}^* = \mathbf{W} \Leftrightarrow \mathbf{Q}_i \mathbf{W} = \mathbf{W}^* \quad (\text{A.3})$$

Appendix B

Forward Backward Averaging and Spatial Smoothing

One of the goals of this work is to enable the application of FBA and SPS in arrays with arbitrary responses. In this section these techniques are detailed and its effect in a possible rank deficient signal correlation matrix is shown. Both these techniques require a regular array geometry such as a ULA or URA as mentioned, this can be virtually obtained via array interpolation as shown in this work. The Forward Backward Averaged received signal matrix is constructed as

$$\mathbf{Z} = [\mathbf{X} | \mathbf{Q}_M \mathbf{X}^* \mathbf{Q}_N], \quad (\text{B.1})$$

where \mathbf{Q}_n is a $n \times n$ exchange matrix as given in (A.1). Let \mathbf{A} be a steering matrix for $d = 3$ and $M = 3$ in a centro-hermitian array

$$\mathbf{A} = \begin{bmatrix} e^{-j\mu_1} & e^{-j\mu_2} & e^{-j\mu_3} \\ 1 & 1 & 1 \\ e^{j\mu_1} & e^{j\mu_2} & e^{j\mu_3} \end{bmatrix}. \quad (\text{B.2})$$

Thus,

$$\mathbf{Q}_M \mathbf{A} = \begin{bmatrix} e^{j\mu_1} & e^{j\mu_2} & e^{j\mu_3} \\ 1 & 1 & 1 \\ e^{-j\mu_1} & e^{-j\mu_2} & e^{-j\mu_3} \end{bmatrix} \quad (\text{B.3})$$

and

$$\mathbf{Q}_M \mathbf{A}^* = \begin{bmatrix} e^{-j\mu_1} & e^{-j\mu_2} & e^{-j\mu_3} \\ 1 & 1 & 1 \\ e^{j\mu_1} & e^{j\mu_2} & e^{-j\mu_3} \end{bmatrix}. \quad (\text{B.4})$$

The signal matrix for $N = 3$ can be written as

$$\mathbf{S} = \begin{bmatrix} s_{1,1} & s_{1,2} & s_{1,3} \\ s_{2,1} & s_{2,2} & s_{2,3} \\ s_{3,1} & s_{3,2} & s_{1,3} \end{bmatrix}, \quad (\text{B.5})$$

and

$$\mathbf{SQ}_N^* = \begin{bmatrix} s_{1,1}^* & s_{1,2}^* & s_{1,3}^* \\ s_{2,3}^* & s_{2,2}^* & s_{2,1}^* \\ s_{3,3}^* & s_{3,2}^* & s_{3,1}^* \end{bmatrix}. \quad (\text{B.6})$$

The steering matrix is not affected by the transformation, i.e., the DOAs of the normal and Forward Backward Averaged version are the same, therefore \mathbf{X} and $\mathbf{Q}_M \mathbf{X}^* \mathbf{Q}_N$ can be concatenated. Also, \mathbf{SQ}_N^* virtually enhances the number of samples to a degree that depends on the correlation between the incoming signals, as shown latter in the appendix. When only noise is present

$$\mathbf{Z} = \begin{bmatrix} n_{1,1} & n_{1,2} & n_{1,3} & n_{1,3}^* & n_{1,2}^* & n_{1,1}^* \\ n_{2,1} & n_{2,2} & n_{2,3} & n_{2,3}^* & n_{2,2}^* & n_{2,1}^* \\ n_{3,1} & n_{3,2} & n_{3,3} & n_{3,3}^* & n_{3,2}^* & n_{3,1}^* \end{bmatrix}, \quad (\text{B.7})$$

since the entries $n_{i,j}$ are drawn from $\mathcal{CN}(0, \sigma_n^2)$ the noise samples are spatially uncorrelated.

A similar analysis can be made for Spatial Smoothing dividing the array into two subarrays of length $L = 2$. The steering matrices are

$$\mathbf{A}^{(1)} = \begin{bmatrix} e^{-j\mu_1} & e^{-j\mu_2} & e^{-j\mu_3} \\ 1 & 1 & 1 \end{bmatrix}, \quad (\text{B.8})$$

$$\mathbf{A}^{(2)} = \begin{bmatrix} 1 & 1 & 1 \\ e^{j\mu_1} & e^{j\mu_2} & e^{j\mu_3} \end{bmatrix}. \quad (\text{B.9})$$

Therefore, $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ introduce the same phase delay between the signal samples and can be used to obtain two independent DOA estimations from

$$\mathbf{A}^{(1)} \mathbf{S}, \quad (\text{B.10})$$

$$\mathbf{A}^{(2)} \mathbf{S}, \quad (\text{B.11})$$

at the price of sacrificing an antenna for each estimation. These estimations can be, for simplicity, averaged uniformly.

A deeper analysis can be made by looking into the effects on the covariance of the received signal. Let

$$\mathbf{D} = \text{diag}[e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_d}]. \quad (\text{B.12})$$

The covariance matrix of the signal received at the i -th subarray of length W can be written as

$$\mathbf{R}_{\mathbf{XX}}^{(i)} = \mathbf{A}_W \mathbf{D}^{\frac{W-M}{2} + (i-1)} \mathbf{R}_{\mathbf{SS}} [\mathbf{D}^{\frac{W-M}{2} + (i-1)}]^H \mathbf{A}_W^H + \sigma_n^2 \mathbf{I}. \quad (\text{B.13})$$

The backward covariance matrix of the i -th subarray is given by $\mathbf{R}_{\mathbf{XX}}^{(i)}_{WB} = \mathbf{Q} [\mathbf{R}_{\mathbf{XX}}^{(i)}]^* \mathbf{Q}$, where \mathbf{Q} is an exchange matrix as given in (A.1). Given that $\mathbf{Q} \mathbf{A}_W^* = \mathbf{A}_W$ the backward signal covariance matrix is given by

$$\mathbf{R}_{\mathbf{XX}}^{(i)}_{WB} = \mathbf{A}_W [\mathbf{D}^{\frac{W-M}{2} + (i-1)}]^* [\mathbf{R}_{\mathbf{SS}}]^* [\mathbf{D}^{\frac{W-M}{2} + (i-1)}]^{*H} \mathbf{A}_M^H + \sigma_n^2 \mathbf{I}. \quad (\text{B.14})$$

By performing a uniform weighting across the subarrays the total covariance matrix is given by

$$\mathbf{R}_{\mathbf{XX}SSFB} = \frac{1}{2L} \sum_{i=1}^L (\mathbf{R}_{\mathbf{XX}}^{(i)} + \mathbf{R}_{\mathbf{XX}}^{(i)}), \quad (\text{B.15})$$

where L is the total number of subarrays employed. Since $\mathbf{D}^* = \mathbf{D}^{-1}$ the following relationship holds

$$\begin{aligned} \mathbf{R}_{\mathbf{XX}SSFB} &= \mathbf{A}_W \left[\frac{1}{2L} \sum_{i=1}^L \mathbf{D}^{\frac{W-M}{2}+(i-1)} \mathbf{R}_{\mathbf{SS}} [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^H \right. \\ &\quad \left. + [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^* [\mathbf{R}_{\mathbf{SS}}]^* [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^{*H} \mathbf{A}_W^H + \sigma_n^2 \mathbf{I} \right]. \end{aligned} \quad (\text{B.16})$$

The smoothed signal covariance can be rewritten as

$$\begin{aligned} \mathbf{R}_{\mathbf{SS}SSFB} &= \frac{1}{2L} \sum_{i=1}^L \mathbf{D}^{\frac{W-M}{2}+(i-1)} \mathbf{R}_{\mathbf{SS}} [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^H \\ &\quad + [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^* [\mathbf{R}_{\mathbf{SS}}]^* [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^{*H} \\ &= \frac{1}{L} \sum_{i=1}^L \operatorname{Re}[\mathbf{D}^{\frac{W-M}{2}+(i-1)} \mathbf{R}_{\mathbf{SS}} [\mathbf{D}^{\frac{W-M}{2}+(i-1)}]^H]. \end{aligned} \quad (\text{B.17})$$

Thus, if the cross correlation between the signals is a purely imaginary number it is completely mitigated by the Forward Backward averaging alone. Considering that only spatial smoothing was applied, its elements are given by

$$[\mathbf{R}_{\mathbf{SS}}^{(L)}]_{ij} = [\mathbf{R}_{\mathbf{SS}}]_{ij} \frac{1}{L} \sum_{l=1}^L \mathbf{D}_{ii}^{\frac{W-M}{2}+(l-1)} [\mathbf{D}_{jj}^{\frac{W-M}{2}+(l-1)}]^*. \quad (\text{B.18})$$

When $i = j$

$$[\mathbf{R}_{\mathbf{SS}}^{(L)}]_{ii} = [\mathbf{R}_{\mathbf{SS}}]_{ii}, \quad (\text{B.19})$$

when $i \neq j$ then

$$\begin{aligned} \frac{1}{L} \sum_{l=1}^L \mathbf{D}_{ii}^{\frac{W-M}{2}+(l-1)} [\mathbf{D}_{jj}^{\frac{W-M}{2}+(l-1)}]^* &= \frac{1}{L} \sum_{l=1}^L e^{j \frac{W-M}{2}(\theta_i - \theta_j)} e^{-j(l-1)(\theta_i - \theta_j)} \\ &= e^{j \frac{W-M}{2} \Delta \theta_{ij}} \frac{1}{L} \sum_{m=1}^{L-1} e^{jm \Delta \theta_{ij}}. \end{aligned} \quad (\text{B.20})$$

Thus, the effectiveness of Spatial Smoothing depends directly on the difference between the $\Delta \theta_{ij}$ between the angles of arrival of the signals indexed by i and j , this is why Spatial Smoothing is applied mostly in a heuristic manner and why, in this work, a simple way of choosing the minimum number of subarrays that provides proper signal decorrelation was proposed.

Appendix C

Proof of Theorem 1

Proof. Let a random variable \tilde{r} have a probability distribution function $p_{\tilde{r}}(r)$, $r \in \mathbb{R}$. Assuming that a moment generating function $M_{\tilde{r}}(k)$ exists, it is equivalent to the two-sided Laplace transform of \tilde{r} . Hence,

$$M_{\tilde{r}}(-k) = \mathcal{L}\{p_{\tilde{r}}(r)\}(k) + \mathcal{L}\{p_{\tilde{r}}(-r)\}(-k), \quad (\text{C.1})$$

where $\mathcal{L}\{p_{\tilde{r}}(r)\}(k)$ is the Laplace transform of $p_{\tilde{r}}(r)$. Thus, $M_{\tilde{r}}(-k)$ is unique iff both sides of the Laplace transform are unique. Therefore, since the Laplace transform of a function is unique, so is $M_{\tilde{r}}(-k)$. Thus, a PDF can be uniquely described by its moments provided that all moments exist and that $p_{\tilde{r}}(r)$ is integrable on any finite subset over its support. \square

Appendix D

Proof of Theorem 2

Proof. The log-likelihood can be given as

$$l(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi}) = -MN \log(\pi\sigma_n^2) - \frac{1}{\sigma_n^2} \text{tr}\{(\mathbf{X} - \mathbf{AS}) - (\mathbf{X} - \mathbf{AS})^H\}, \quad (\text{D.1})$$

where $\text{tr}\{\cdot\}$ is the trace operator. Differentiating (D.1) with respect to σ_n^2 and equating to zero leads an estimate for the noise variance

$$\hat{\sigma}_n^2 = \frac{1}{MN} \text{tr}\{(\mathbf{X} - \mathbf{AS}) - (\mathbf{X} - \mathbf{AS})^H\}. \quad (\text{D.2})$$

Substituting (D.2) into (D.1), differentiating with respect to \mathbf{S} , and equating to zero gives an estimate for \mathbf{S}

$$\hat{\mathbf{S}} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{X}. \quad (\text{D.3})$$

Then, substituting (D.2) and (D.3) into (D.1) yields

$$\begin{aligned} (\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}) &= \arg \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \text{tr}\{\mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \hat{\mathbf{R}}_{\mathbf{XX}}\} \\ &= \arg \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi}). \end{aligned} \quad (\text{D.4})$$

Assuming that the received signals have low spatial correlation with

$$\forall_{i, i'}^d, \mathbf{a}^H(\theta_i, \phi_i) \mathbf{a}(\theta_{i'}, \phi_{i'}) \rightarrow 0, \quad (\text{D.5})$$

the concentrated loss function close to the global optimum can be approximated as

$$\begin{aligned}
& \arg \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi}) \\
& \approx \arg \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \sum_{i=1}^d \frac{\mathbf{a}^H(\theta_i, \phi_i) \hat{\mathbf{R}}_{\mathbf{xx}} \mathbf{a}(\theta_i, \phi_i)}{\mathbf{a}^H(\theta_i, \phi_i) \mathbf{a}(\theta_i, \phi_i)} \\
& \approx \arg \max_{\boldsymbol{\theta}, \boldsymbol{\phi}} \sum_{k=1}^K \frac{1}{\chi_k} P(\theta, \phi) |_{(\theta, \phi) \in \mathcal{S}_k}, \tag{D.6}
\end{aligned}$$

where with $\theta \in \mathcal{D}_{\Delta_\theta}$ and $\phi \in \mathcal{D}_{\Delta_\phi}$

$$\chi_k = \sum_{(\theta, \phi) \in \mathcal{S}_k} P(\theta, \phi) \Delta_\theta \Delta_\phi. \tag{D.7}$$

□

Remark 1. The assumption that signals have low spatial correlation can be found valid for many cases, especially, for arrays with $M \rightarrow \infty$. However, signals can still have high correlation in time with cross-correlation coefficient $\gamma_{i,i'} \rightarrow 1$.

Remark 2. The approximation introduced in (D.6) results to $\forall_{(\theta, \phi) \notin \mathcal{S}} l_c(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{\phi}) \triangleq 0$. The normalization χ_k with respect to each sector \mathcal{S}_k makes it possible to consider the concentrated loss function for each sector \mathcal{S}_k as a separate PDF when applying the UT.

Appendix E

Multivariate Adaptive Regression Splines (MARS)

The MARS [41] regression model is an extension of piece wise linear regression models. MARS uses a set of splines as basis functions, these splines in turn have the input variables of the model as their parameters. MARS is an attractive regression technique due to the fact that it is capable of representing nonlinear relationships between input and output variables through the multiplication of basis functions. MARS is also a non-parametric regression model, therefore it is capable of deriving the relationship between input and output variables purely based on the data and requires no presumptions.

The MARS model relies on basis functions known as hinge functions. Hinge functions take the form

$$h(k, z) = \begin{cases} \max(0, k - z) \\ \text{or} \\ \max(0, z - k) \end{cases} \quad (\text{E.1})$$

where

$$\max(a, b) = \begin{cases} a \text{ if } a > b \\ b \text{ if } a \leq b \end{cases}, \quad (\text{E.2})$$

and k is known as a knot. The set $\mathcal{C} = \{\max(0, k - z), \max(0, z - k)\}$ is called a reflected pair.

Let $\mathbf{z}_i \in \mathbb{R}^I$ be a vector containing the observations of the I input variables belonging to the set \mathcal{I} for a given training data observation. Let O be the total number of training data observations and \mathcal{O} be the set containing all the observed values for the input variable during training. The set of candidate basis functions \mathcal{H} for a given MARS model can be built using all the possible

combinations of the pairs $\max(0, k - z)$ and $\max(0, z - k)$ with $k \in \mathcal{O}$ and $z \in \mathcal{I}$. Therefore,

$$|\mathcal{H}| = 2IO, \quad (\text{E.3})$$

MARS performs two steps in order to create the model of the form

$$f(\mathbf{z}) = \beta_0 + \sum_{l=1}^L \beta_l h_l(\mathbf{z}), \quad (\text{E.4})$$

where $h_l(\mathbf{z}) \in \mathcal{H}$ and β_0, \dots, β_L are weighting coefficients calculated using the normal least squares, and \mathbf{z} is a vector containing the observations of the input variables. Note that despite having as an argument the vector with all input variables, each $h_l(\mathbf{z})$ is only a function of a single element $z \in \mathbf{z}$.

The process of building the MARS model is divided in two phases called the forward and the backwards pass. In the forward pass, basis functions are added to the model set \mathcal{M} in reflected pairs with $\mathcal{M} \subseteq \mathcal{H}$. The basis functions from the candidate set \mathcal{H} are added to the model set based on their influence on the total regression error. Only once the set \mathcal{M} has at least one basis function, products of candidate functions and functions already present in \mathcal{M} are also considered. Therefore, for $|\mathcal{M}| > 1$ the following candidate functions are assessed using

$$c_m h_c(\mathbf{z}), \quad (\text{E.5})$$

where $h_c(\mathbf{z}) \in \mathcal{H}$ and $c_m \in 1 \cup \mathcal{M}$. Therefore, the model is capable of creating nonlinear relationships by multiplying basis functions, but products of basis functions are only added to the model if their lower order part of the product is already present on the model. Basis functions are added until the model reaches a given maximum model complexity ν_{\max} and basis functions products are only considered up to a maximum order ϑ_{\max} .

After the model has reached its maximum allowed complexity, it is probably over-fitted. Thus, the second phase of the MARS model building, the backward pass, is used to reduce the model complexity. In this phase, the basis functions that are in the set \mathcal{M} and that when removed impact the squared error the least, are removed. To estimate the optimal model complexity, MARS uses the generalized cross-validation (GCV), which is defined as

$$\text{GCV}(\nu) = \frac{\sum_{o=1}^O (y_o - f_\nu(\mathbf{z}_o))^2}{\left(1 - \frac{\nu+p}{O}\right)^2}, \quad (\text{E.6})$$

where y_i is the i -th training output, f_ν is a model of complexity ν and p is a penalty factor. Different values for p are discussed in [41], where a value of 3 is suggested as a default value. The GCV penalizes a model based on its complexity and creates a compromise between accuracy and complexity. Maximizing the GCV reduces over-fitting and creates a more general model.

Appendix F

Generalized Regression Neural Networks (GRNNs)

GRNNs are probabilistic neural networks, thus the number of training samples required is significantly reduced. Since measurements of real antenna array responses can be limited in resolution due to the computational complexity for electromagnetic simulations or the hardware available at measurement anechoic chambers, the usage of a probabilistic neural network can be beneficial.

A GRNN relies on centering a normal distribution around the training samples according to

$$\hat{c}(\mathbf{b}) = \frac{\sum_{i=1}^I c_i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^I \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}, \quad (\text{F.1})$$

where

$$D_i^2 = (\mathbf{b} - \tilde{\mathbf{b}}_i)^T (\mathbf{b} - \tilde{\mathbf{b}}_i), \quad (\text{F.2})$$

$\hat{c}(\mathbf{b})$ is the estimated output given the input vector \mathbf{b} , $\tilde{\mathbf{b}}_i$ is the i -th training input vector, c_i is the i -th training output, σ^2 is known as the smoothness parameter and is usually chosen heuristically, and $i = 1, \dots, I$ where I is the number of training samples available.

In a GRNN the weight of a given training sample on the estimation of the output is given by (F.2) which is the euclidean distance between a training input and the system input. A GRNN allows a small number of training samples to be used due to the fact that it is capable of interpolating between training samples using the calculated weights, as shown in (F.1).

Figure F.1 depicts how a GRNN with input vector $\mathbf{b} \in \mathbb{R}^J$ can be built as a parallel neural network. The individual layers can have their computations performed in parallel making it possible to speed up the application of the neural network.

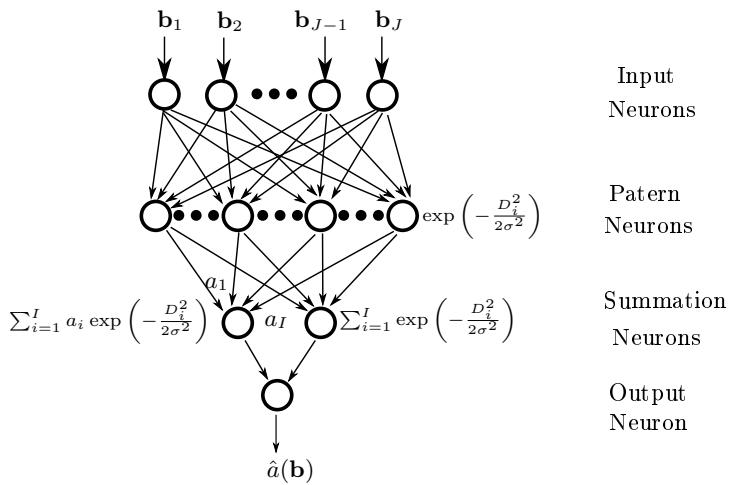


Figure F.1: GRNN as a parallel neural network

Appendix G

Sensor Node Localization

Assume a network formed consisting of K sensors randomly placed at points S_1, S_2, \dots, S_K where

$$S_1 = [x_1, y_1]. \quad (\text{G.1})$$

Furthermore, assuming the presence of a set of nodes S_i, \dots, S_j that posses prior information about their own localization in the network, it is necessary that

$$|i, \dots, j| \geq 3, \quad (\text{G.2})$$

and that the orientation of all sensors in know with respect to a common reference, this can be done by employing a beacon transmitter or assuming an internal compass is present.

The electric field of a propagating wave can be presented as

$$E = -E_x \mathbf{e}_x + E_y \mathbf{e}_y, \quad (\text{G.3})$$

where E_x and E_y are the horizontal and vertical components of the electric field. These components define a polarization ellipse shown in Figure G.1.

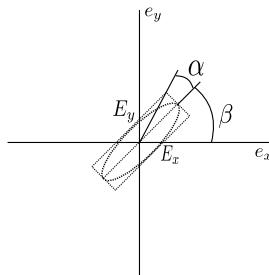


Figure G.1: Polarization ellipse

Note that the electric field components can be written now in terms of the electric angles α and β as

$$E_x = E \cos(\gamma) \quad (\text{G.4})$$

$$E_y = E \sin(\gamma) e^{j\eta} \quad (\text{G.5})$$

where

$$\cos(2\gamma) = \cos(2\alpha) \cos(2\beta) \quad (\text{G.6})$$

$$\tan(\eta) = \tan(2\alpha) \csc(2\beta). \quad (\text{G.7})$$

Considering a wavefront impinging over a crossed dipole with components parallel to the x - and y -axis of the polarization ellipse the output at such dipoles will be proportional to the components E_x and E_y respectively and be written as

$$\begin{aligned} E &= (-E_x)\mathbf{e}_x + (E_y \cos(\theta))\mathbf{e}_y \\ &= -\cos(\gamma)\mathbf{e}_x + \sin(\gamma) \cos(\theta) e^{j\eta} \mathbf{e}_y \end{aligned} \quad (\text{G.8})$$

where θ is the angle of arrival of the received wave with respect to the x -axis.

The received signal can now be expressed in matrix as

$$\mathbf{X} = \mathbf{us} + \mathbf{N} \quad (\text{G.9})$$

where $\mathbf{X} \in \mathbb{C}^{2 \times N}$ is the matrix containing the measured outputs at each of the dipoles, N is the number of measured snapshots, $\mathbf{s} \in \mathbb{C}^{1 \times N}$ is the vector contain the original transmitted signal transmitted, $\mathbf{N} \in \mathbb{C}^{2 \times N}$ is the matrix containing the Additive White Gaussian noise present at sampling and the vector $\mathbf{u} \in \mathbb{C}^{2 \times 1}$ is defined as

$$\mathbf{u} = \begin{bmatrix} -\cos(\gamma) \\ \sin(\gamma) \cos(\theta) e^{j\eta} \end{bmatrix}. \quad (\text{G.10})$$

This localization approach consists of analysing the ratio between the outputs of the crossed dipole in order to estimate the direction of arrival of the received signal. Given the ratio

$$r = \frac{-\cos(\gamma)}{\sin(\gamma) \cos(\theta) e^{j\eta}}, \quad (\text{G.11})$$

the angle θ can be obtained by

$$\theta = \cos^{-1} \left(\frac{-\cos(\gamma)}{r \sin(\gamma) e^{j\eta}} \right). \quad (\text{G.12})$$

One way to estimate the ratios is to simply average a large number of samples from both antennas in order to reduce the effects of the noise and obtain the ratio between the powers. It is important to highlight at this point

that although there is a need for a crossed dipole with two independent outputs to be present it is not necessary that both outputs are connect to individual receiver radios. A single symbol can be measured at both outputs by dividing the symbol duration, assuming the transmitted energy is constant over the entire symbol duration.

A more precise approach is to employ the ESPRIT algorithm in order to obtain the ratio between the received symbols. The advantage of the ESPRIT algorithm is that it is capable of dealing much more efficiently with noise since it relies on an eigendecomposition that separates noise and signal subspaces.

Although the ESPRIT method is capable of obtaining more precise estimates it generates a larger computational load. This load is, however, justified as a small error in the estimated ratio may lead to a large error in the estimation of the DOA. This computationally demanding step is also only necessary once every time the topology of the network changes, thus, the ratio at which sensors change location needs to be taken into account when choosing the estimation method.

Note that the DOA is given with respect to the reference of the x -axis and cannot distinguish from with direction, front or rear, the signal is arriving. Furthermore, θ ranges from $[-\pi, \pi]$ and thus it is also not possible to tell from with sector is the signal actually arriving since the only information that the system has is $\cos(\theta)$. The end result is that each sensor possesses two estimated lines in the ground plane where the transmitting node may be located. However, by acquiring a set of line estimates it is possible to obtain a single estimate of the transmitting sensor localization. The problem is reduced to the least squares problem of finding the point with minimum distance from any of the possible combination of line estimates.

By writing the representing the line estimates as line equations of the type $Ax + By + C = 0$ in the sensor coordinate system, the estimate of the sensor position can be found by solving

$$\min_p \frac{|A_{p_1}x_0 + B_{p_1}y_0 + C_{p_1}|}{\sqrt{A_{p_1}^2 + B_{p_1}^2}} + \frac{|A_{p_2}x_0 + B_{p_2}y_0 + C_{p_2}|}{\sqrt{A_{p_2}^2 + B_{p_2}^2}} + \dots, \quad (\text{G.13})$$

where p is an index set containing the possible combinations of estimated lines. While more than three sensors can be used to obtain increased accuracy it also results in a higher computational load involved in the calculation of the minima. Once the lines are choose the final location estimate $S_0 = [x_0, y_0]$ can be found by taking the derivative of the above with respect to x_0 and y_0 and finding the point where it is equal to zero.

Furthermore, this technique may be used in conjunction with other localization methods such as the RSSI. A set of candidate locations can be selected and the RSSI information can be used to choose the candidate that best fits the RSSI information.

Appendix H

TRIAD Algorithm

The TRIAD [8, 9] algorithm allowed the development of early satellite navigation systems and still represents the state of practice for space and aircraft instrument based attitude estimation. It involves only two linearly independent reference vectors and their respective measured directions.

By defining two linearly independent reference vectors

$$\begin{aligned}\vec{R}_1 &= [x_{\vec{R}_1}, y_{\vec{R}_1}, z_{\vec{R}_1}]^T, \\ \vec{R}_2 &= [x_{\vec{R}_2}, y_{\vec{R}_2}, z_{\vec{R}_2}]^T,\end{aligned}\tag{H.1}$$

and their respective measured direction vectors

$$\begin{aligned}\vec{r}_1 &= [x_{\vec{r}_1}, y_{\vec{r}_1}, z_{\vec{r}_1}]^T, \\ \vec{r}_2 &= [x_{\vec{r}_2}, y_{\vec{r}_2}, z_{\vec{r}_2}]^T.\end{aligned}\tag{H.2}$$

The TRIAD algorithm tries to find the rotation matrix \mathbf{W} that satisfies

$$\vec{r}_i = \mathbf{W} \vec{R}_i, \tag{H.3}$$

where $\mathbf{W}^T \mathbf{W} = \mathbf{I}$ and $\det(\mathbf{W}) = \pm 1$, i.e. \mathbf{W} is an orthogonal matrix and preserves the magnitude of any vector \vec{R}_i it operates on. Since \vec{R}_1 and \vec{R}_2 are linearly independent vectors, a linearly independent vector orthogonal to both can be obtained by

$$\vec{R}_3 = \vec{R}_1 \times \vec{R}_2, \tag{H.4}$$

where \times is the vector cross product operation, the same can be done for \vec{r}_1 and \vec{r}_2

$$\vec{r}_3 = \vec{r}_1 \times \vec{r}_2. \tag{H.5}$$

Since a rotation applied to \vec{R}_1 and \vec{R}_2 would also rotate \vec{R}_3 , a linear system can be written as

$$[\vec{r}_1, \vec{r}_2, \vec{r}_3] = \mathbf{R}[\vec{R}_1, \vec{R}_2, \vec{R}_3], \tag{H.6}$$

here the commas represent the concatenation of two column vectors. In the noise free case the present linear system holds exact and will yield an orthogonal matrix \mathbf{W} . However, in the presence of noise, the result might be a non orthogonal matrix. To address this problem the TRIAD algorithm replaces \vec{R}_1 , \vec{R}_2 and \vec{R}_3 by

$$\begin{aligned}\vec{S}_1 &= \frac{\vec{R}_1}{\|\vec{R}_1\|}, \\ \vec{S}_2 &= \frac{\vec{R}_1 \times \vec{R}_2}{\|\vec{R}_1 \times \vec{R}_2\|}, \\ \vec{S}_3 &= \frac{\vec{S}_1 \times \vec{S}_2}{\|\vec{S}_1 \times \vec{S}_2\|},\end{aligned}\tag{H.7}$$

respectively. \vec{r}_1 , \vec{r}_2 and \vec{r}_3 are also replaced by

$$\begin{aligned}\vec{s}_1 &= \frac{\vec{r}_1}{\|\vec{r}_1\|}, \\ \vec{s}_2 &= \frac{\vec{r}_1 \times \vec{r}_2}{\|\vec{r}_1 \times \vec{r}_2\|}, \\ \vec{s}_3 &= \frac{\vec{s}_1 \times \vec{s}_2}{\|\vec{s}_1 \times \vec{s}_2\|}.\end{aligned}\tag{H.8}$$

respectively. Note that by construction the matrices $[\vec{S}_1, \vec{S}_2, \vec{S}_3]$ and $[\vec{s}_1, \vec{s}_2, \vec{s}_3]$ are orthogonal matrices, since their columns are made orthogonal to each other. This avoids the computationally intensive task of calculating the matrix inverse, since $[\vec{S}_1, \vec{S}_2, \vec{S}_3]^{-1} = [\vec{S}_1, \vec{S}_2, \vec{S}_3]^T$. Thus, an estimate of \mathbf{W} can be found by

$$\hat{\mathbf{W}} = [\vec{s}_1, \vec{s}_2, \vec{s}_3] \cdot [\vec{S}_1, \vec{S}_2, \vec{S}_3]^T.\tag{H.9}$$

As the rotation matrix is directly dependent on the pitch, yaw and roll an estimation can be extracted from $\hat{\mathbf{W}}$. The yaw, pitch and roll rotation matrices are given by

$$\begin{aligned}\mathbf{R}(\alpha) &= \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \\ \mathbf{R}(\beta) &= \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \\ \mathbf{R}(\gamma) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & -\sin(\gamma) \\ 0 & \sin(\gamma) & \cos(\gamma) \end{bmatrix}.\end{aligned}\tag{H.10}$$

If the rotation follows the order of yaw, pitch and roll, then

$$\hat{\mathbf{W}} = \mathbf{R}(\alpha)\mathbf{R}(\beta)\mathbf{R}(\gamma)$$

$$\hat{\mathbf{W}} = \begin{bmatrix} \cos(\hat{\alpha})\cos(\hat{\beta}) & \cos(\hat{\alpha})\sin(\hat{\beta})\sin(\hat{\gamma}) - \sin(\hat{\alpha})\cos(\hat{\gamma}) \\ \sin(\hat{\alpha})\cos(\hat{\beta}) & \sin(\hat{\alpha})\sin(\hat{\beta})\sin(\hat{\gamma}) - \cos(\hat{\alpha})\cos(\hat{\gamma}) \\ -\sin(\hat{\beta}) & \cos(\hat{\beta})\sin(\hat{\gamma}) \\ \cos(\hat{\alpha})\sin(\hat{\beta})\cos(\hat{\gamma}) + \sin(\hat{\alpha})\sin(\hat{\gamma}) & \\ \sin(\hat{\alpha})\sin(\hat{\beta})\cos(\hat{\gamma}) + \cos(\hat{\alpha})\sin(\hat{\gamma}) & \\ \cos(\hat{\beta})\cos(\hat{\gamma}) & \end{bmatrix} \quad (\text{H.11})$$

where $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ are estimates of the yaw pitch and roll respectively. Note that changing the order of rotation will change the structure of $\hat{\mathbf{W}}$.

The TRIAD algorithm allows an estimate of the UAV attitude to be obtained by obtaining only two linearly independent position vectors, such as a wing and the nose position vectors. Other methods for estimating the attitude exist, such as the QUEST [100] and SVD based methods such as [78]. These methods are capable of taking into account a broad set of measurements and reference vectors resulting in a more accurate estimation. However, they are more computationally bulky than the TRIAD algorithm and, usually, slower, thus not practical for real-time estimation scenarios.