DISSERTATION

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
from TELECOM ParisTech

Specialization: Communication and Electronics

Konstantinos Papakonstantinou

Applications of Statistical Signal Processing in Mobile Terminal Localization

Defense scheduled on the 8th of July 2010 before a committee composed of:

Reviewers	Prof. Bernard Fleury, Aalborg University
	Prof. Bernard Uguen, University of Rennes 1
Examiners	Prof. Merouane Debbah, SUPELEC
	Prof. Eric Moulines, Telecom ParisTech (ENST)
	Prof. Nikos Sidiropoulos, Technical University of Crete
Thesis Supervisor	Prof. Dirk Slock, EURECOM

THESE

présentée pour obtenir le grade de

Docteur de TELECOM ParisTech

Spécialité: Communication et Electronique

Konstantinos Papakonstantinou

Les Applications du traitement du signal statistique à la localisation mobile

Soutenance de thèse prévue le 08 Juillet 2010 devant le jury composé de :

Rapporteurs	Prof. Bernard Fleury, Aalborg University
	Prof. Bernard Uguen, Université de Rennes 1
Examinateurs	Prof. Merouane Debbah, SUPELEC
	Prof. Eric Moulines, Télécom ParisTech (ENST)
	Prof. Nikos Sidiropoulos, Technical University of Crete
Directeur de thèse	Prof. Dirk Slock, EURECOM

Abstract

This thesis outlines our attempt to tackle the problem of estimating the location of the mobile terminal (MT), using the wireless network's infrastructure and assuming realistic propagation environments. It is essentially a collection of the following:

- Generic localization methods and methods designed specifically for MIMO-OFDM systems,
- implementation algorithms ranging from low-complexity (least squares) to high-complexity ones (maximum likelihood),
- derivations of performance bounds (Cramer-Rao bound) for the localization methods and
- theorems and conjectures on the identifiability and the accuracy of the estimates.

Apart from the introduction and the conclusions, the document contains five more chapters. The context of each one of these chapters is based on one or two publications, while in chapter 4 also a few preliminary unpublished results can be found.

The purpose of the long introductory chapter is two-fold: On one hand we introduce the fundamentals of localization methods to the unfamiliar readers. The fundamentals are complemented with the main sources of errors encountered in localization methods and some existing attempts to mitigate them. On the other hand, we present some material that is used throughout this document, like eg. geometrical and statistical channel models, useful expressions for a MIMO-OFDM system, the maximum likelihood (ML) location estimation solution and last but not least a discussion on identifiability and performance, which will be continued in some of the following chapters.

In general, localization methods are 2-step processes: In the 1^{st} step a set of location-dependent parameters need to be estimated. We therefore present in chapter 2 a high-resolution low-complexity algorithm (4D Unitary

ESPRIT) for estimating four different kinds of LDP from the received signal of a MIMO-OFDM system.

Chapter 3 contains the basic principles of the 2^{nd} step of the NLoS localization methods that are based on the single-bounce model (SBM). It starts with a description of the original method that inspired us to also utilize the SBM for localization purposes. This method is essentially a hybrid method that utilizes angles of the signal components on both sides of a communication link and lengths of the paths. Following the description of the original method, some extensions that we have proposed, can be found in this chapter. These extensions include a weighted least squares solution (WLS) and a thorough study on the impact of network geometry on performance.

In chapter 4, SBM-based localization methods are extended in order to become applicable to environments that change dynamically due to the movement of the MT. The so called DSBM-based methods, where the initial "D" is used for "dynamic", are introduced and studied in depth. Specifically two different DSBM-methods are presented. The first method utilizes angles of arrival (AoA), angles of departure (AoD), path lengths and Doppler shifts (DS) while the second one is more interesting since it does not utilize the AoA. We derive closed-form solutions (WLS) for both and compare them.

The 2 steps of localization methods are combined into a single step, in the method called direct location estimation (DLE) that is described in chapter 5. The concept of estimating the MT location directly from the received signal was originally developed for LoS environments. With the aid of the channel models described in chapter 1 we designed localization methods that have the same principles but are also suitable for multipath and NLoS environments.

Chapter 6 presents a tracking method, i.e. a method that estimates only the speed components, which is not based on the assumption of single-bouncing. Albeit more general, the method presented therein is also more complex. Furthermore, in contrast to the methods of the previous chapters, it is based on Bayesian and not ML estimation.

Finally in the last chapter we make a list of conclusions that we have reached while working on NLoS localization. We comment on the context of each chapter, provide arguments to justify our choices, advocate in favor of our methods. We further make a list of ideas on how to extend this work and describe how to generalize the concept of SBM-based methods, so that they become widely accepted.

Acknowledgements

Αφιερωμένο στους γονείς μου, Δημήτρη και Ρούλα.
Ελπίζω να σας έκανα, έστω και λίγο, περήφανους.

This thesis summarizes most of the results obtained during the past three and half years that I was working at Eurecom, in the South of France. Eurecom is an amazing place to work and the main reason for this is not its location as any visitor would expect. The main reason is the people that work here and essentially transform the place into a warm and friendly environment. A few of my ex- and current colleagues that have made this place truly special, are Marios, Ruben, Marco, Florian and Gillian, Sara, Randa, Zuleita, Antony, Ikbal, Ken, David, Petros and Antitza, Nesli and Ufuk and Mariam and Virginia who visited us for a few months. Had it not been for them, I doubt I would have been productive, since I believe that productivity depends on your motivation and your mood! I would therefore like to thank each and every one of them. Some extra special thanks goes to Ikbal, for always being there to help me when I needed an English to French translator and to Ken for being a personal advisor on top of being a friend. Well since I started mentioning friends, I might as well thank all my friends around the globe for contributing, on their own way, to this work. Specifically, Vasiliki, Aki and Theo thanks for pushing me to do a PhD. It was so worth it. Actually, Vasiliki thank you for being always next to me throughout this period of time. To my close friends in Greece, I have to say thanks for making me realize again and again that I am still young and certainly know how to have a good time. Last but not least, Sophiana, thanks for being yourself and being close to me during the last few months of my PhD.

Well, to obtain a PhD it takes more than being in a good mood and having a good motive. It also requires collaboration with other researchers

and -at least for the first few months- guidance from an expert in the field. I consider myself to be very lucky, cause I got the opportunity to be supervised by and work closely with two incredibly smart, highly skilled and well reputed professors, Merouane Debbah and Dirk Slock. Merouane was there when I started my PhD and helped me take my first steps into the world of research. I admit that I was totally lost back then and his assistance was more than valuable. Soon however, he had to leave and at that point my senior advisor, Dirk, came in the picture. Dirk has an amazing background and intuition and thus it was natural for him to point me to the right research directions and to answer any research-related question I could possibly come up with. There were a few times that we actually argued about a topic we were investigating. Needless to say -and it hurts to admit- he was always right. Many thanks to both of you guys, part of my success is because of you. Other people that I got the chance to collaborate with, are the researchers working for the industrial and academic partners in the European project WHERE and the French project Semafor. Many thanks to all of them for their valuable feedback on my work.

My parents, Dimitris and Roula and my sister, Maria, could not be absent from the list of people I should thank. I feel lucky having their unlimited love and support throughout my whole life. As a matter of fact, as indicated above in Greek, although it is negligible compared to what they have offered to me, I dedicate my PhD to my parents.

Contents

List of Figures

Acronyms

Here are the main acronyms used in this document. The meaning of an acronym is usually indicated once, when it first occurs in the text. The English acronyms are also used for the French summary.

AoA	Angle(s) of Arrival.
AoD	Angle(s) of Departure.
app.	Appendix.
AWGN	Additive White Gaussian Noise.
BC	Broadcast Channel.
BE	Bayesian Estimation.
BS	Base Station.
cdf	cumulative density function.
CIR	Channel Impulse Response.
CRB	Cramer-Rao Bound.
DDM	Double Directional Model.
DL	Downlink.
DLE	Direct Location Estimation.
DFS	Doppler Frequency Shift(s).
DPD	Direct Position Determination.
DS	Doppler Shift(s).
DSBM	Dynamic Single Bounce Model.
EKF	Extended Kalman Filter.
EM	Expectation-Maximization algorithm.
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Techniques.
eq.	Equation.
FCC	(U.S.) Federal Communications Commission.
FIM	Fisher Information Matrix.
FP	Fingerprinting.

i.i.d.	independent and identically distributed.
io	input-output.
KF	Kalman Filter.
LDP	Location Dependent Parameters.
LE	Location Estimation.
l.h.s.	left hand side.
LMDP	Location and Motion Dependent Parameters.
LoP	Lines of Position.
LoS	Line of Sight.
LS	Least Squares.
MAP	Maximum a Posteriori.
ME	Maximum Entropy.
MIMO	Multiple Input Multiple Output.
MISO	Multiple Input Single Output.
ML	Maximum Likelihood.
MMSE	Minimum Mean Square Error.
MPC	Multipath Components.
MSE	Mean Square Error.
MT	Mobile Terminal.
MUSIC	Multiple Signal Classification.
NLoS	Non Line of Sight.
pdf	probability density function.
PD	Positive Definite.
pmf	probability mass function.
PSD	Positive Semi-Definite.
r.h.s.	right hand side.
RMS	Root Mean Square.
RMSE	Root Mean Square Error.
RSS	Received Signal Strength.
Rx	Receiver(s).
SBM	Single-Bounce Model.
SIMO	Single Input Multiple Output.
SNR	Signal-to-Noise Ratio.
SISO	Single-Input Single-Output.
s.t.	such that.
SVD	Singular Value Decomposition.
TDoA	Time Differences of Arrival.
ToA	Time(s) of Arrival.
TLS	Total Least Squares.
Tx	Transmitter(s).

WLS	Weighted Least Squares.
w.r.t.	with respect to.
2D (3D, 4D)	two (three, four) dimensional.

Notation

Throughout this thesis, calligraphic upper-case letters denote sets. Upper case and lower case boldface symbols will represent matrices and column vectors respectively. For any defined vector $\boldsymbol{a}$, the corresponding capital symbol $\boldsymbol{A}$ will represent a diagonal matrix whose main diagonal is equal to $\boldsymbol{a}$, i.e. $\boldsymbol{A} = diag\{\boldsymbol{a}\}$ and vice versa. Extending this, a_i will denote the ith entry of $\boldsymbol{a}$ and the $\{i, i\}$ entry of $\boldsymbol{A}$. It therefore suffices to define any of the above (a vector, a diagonal matrix or just an entry), to define all 3.

$\mathrm{tr}\{\cdot\}$	Trace of the matrix in brackets.
$\det\{\cdot\}$	Determinant of the matrix in brackets.
$\lvert a\rvert$	Absolute value of a.
$\lVert\mathbf{a}\rVert$	Euclidean norm of vector $\mathbf{a}$.
$\lVert\mathbf{A}\rVert$	Frobenius norm of matrix $\mathbf{A}$.
$\lvert\mathbb{S}\rvert$	The cardinality of set $\mathbb{S}$.
$\lfloor a \rfloor$	Floor operation, rounds (the elements of) a to the nearest integers towards minus infinity.
$\lceil a \rceil$	Ceil operation, rounds (the elements of) a to the nearest integers towards infinity.
$\hat{(\cdot)}$	An estimate of the quantity in parentheses.
$\tilde{(\cdot)}$	The error in the estimate of the quantity in parentheses.
$\mathbf{A}^*$	The complex conjugate of matrix $\mathbf{A}$.
$\mathbf{A}^\dagger$	The complex conjugate transpose (Hermitian) of matrix $\mathbf{A}$.
$\mathbf{A}^t$	The transpose of matrix $\mathbf{A}$.
$\mathbf{A}^{-1}$	The inverse of matrix $\mathbf{A}$.
$[\mathbf{A}]_{(k:l,m:n)}$	A submatrix of $\mathbf{A}$ containing the common elements of rows k-l and columns m-n.
$\mathbf{A}^+$	The left pseudo-inverse of matrix $\mathbf{A}$, given by $\left(\mathbf{A}^\dagger\mathbf{A}\right)^{-1}\mathbf{A}^\dagger$.
$\mathbf{A} = \mathrm{diag}\,(\mathbf{a})$	The diagonal matrix with the entries of vector $\mathbf{A}$ along its main diagonal, if the latter is defined.

$\mathbf{a} = \text{diag}(\mathbf{A})$	The vector equal to the main diagonal of $\mathbf{A}$, if the latter is defined.
$\mathbf{A}^{1/2}$	Hermitian square root of the positive semidefinite matrix $\mathbf{A}$.
$\mathbf{A} \geq \mathbf{B}$	means that $\mathbf{A} - \mathbf{B}$ is non-negative definite.
$T(\cdot)$	Linear transformation of the vector in parenthesis.
$p(\cdot)$	pdf of the continuous random variable in parenthesis.
$P[\cdot]$	pmf of the discrete random variable in brackets.
$\mathcal{E}\{\cdot\}$	Expected value of the random variable in brackets.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Circularly symmetric complex Gaussian random vector of mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\mathcal{U}$	Uniform distribution
$sgn\{\cdot\}$	Sign of the variable in brackets.
$\max, \min$	Maximum and minimum.
$\text{argmax}\{\cdot\}$	Argument that maximizes the function(al) in brackets.
$\text{argmin}\{\cdot\}$	Argument that minimizes the function(al) in brackets.
$\sim$	Distributed according to.
$\odot$	Hadamard Product.
$\otimes$	Kronecker Product.
$\boxtimes$	Khatri-Rao (Column-wise Kronecker) Product.

Symbol Index

The definition of symbols that are repeated throughout this thesis are given in the tables below. There will be a few exceptions, where these symbols will denote something different than what is described below. For these exceptions, the definition will be given at the place of their occurrence.

Sets	
$\mathbb{R}, \mathbb{C}$	The set of all real and complex numbers, respectively.
$\mathbb{R}^{m\times n}$	The set of $m \times n$ real matrices.
$\mathbb{C}^{m\times n}$	The set of $m \times n$ matrices with complex-valued entries.
$\mathbb{Z}$	Set of integer numbers

Scalar Parameters	
j	subscript used for scatterer (multipath component).
i	subscript used for time sample.
l	subscript used for time sample.
k	subscript used for frequency sample. Also used for different LDP.
f_c	Carrier frequency.
c	Speed of light.
λ	Wavelength.
N_s	Number of scatterers.
Δt	Spacing between two consecutive time samples.
t_i	Time difference between time instant i and time instant 0.
N_t	Number of time samples.
Δf	Spacing between two consecutive frequency samples.
N_f	Number of frequency samples.
K	Number of different kinds of LDP.

$\mathcal{L}$	Log-likelihood function.
	Coordinates
$x_{mt} \equiv x_0$	x-coordinate of MT.[1]
$y_{mt} \equiv y_0$	y-coordinate of MT.[1]
x_i	x-coordinate of MT at time instant i.[2]
y_i	y-coordinate of MT at time instant i.[2]
x_{bs_j}	x-coordinate of BS j.
y_{bs_j}	y-coordinate of BS j.
x_{s_j}	x-coordinate of scatterer j.
y_{s_j}	y-coordinate of scatterer j.
	Motion parameters
υ	magnitude of speed.[3]
ω_υ	direction of speed.[3]
$\upsilon_x \equiv \upsilon_{x_0}$	speed component along the x-axis.[3]
$\upsilon_y \equiv \upsilon_{y_0}$	speed component along the y-axis.[3]
υ_i	magnitude of speed. [4]
ω_i	direction of speed.[4]
υ_{x_i}	speed component along the x-axis at time instant i.[4]
υ_{y_i}	speed component along the y-axis at time instant i.[4]
α_x	acceleration component along the x-axis at time instant i.[4]
α_y	acceleration component along the y-axis at time instant i.[4]
	LDP
$d_{mts,ij}$	distance between the MT and scatterer j at time instant i.
$d_{bs,j}$	distance between BS j(or BS 1) and scatterer j.
d_{ij}	length of path j at time instant i.
ϕ_{ij}	AoA of MPC j at time instant i.
ψ_{ij}	AoD of MPC j at time instant i.
$f_{d,ij}$	DS of MPC j at time instant i.

(Column) Vectors

$\mathbf{1}$	vector of all ones.
$\mathbf{x}_{kl} = vec\{\mathbf{X}_{kl}\}$	Transmitted signal matrix.[5]
$\mathbf{h}_{kl} = vec\{\mathbf{H}_{kl}\}$	Channel vector.[5]

[1] Static environment.
[2] Dynamic environment.
[3] Constant speed mobility model.
[4] Constant acceleration mobility model.
[5] At time k and frequency l.

$\mathbf{y}_{kl} = vec\{\mathbf{Y}_{kl}\}$	Received signal vector.[5]
$\mathbf{n}_{kl} = vec\{\mathbf{N}_{kl}\}$	Noise vector.[5]
$\mathbf{a}_R$	Rx array response.
$\mathbf{a}_T$	Tx array response.
$\mathbf{p}_{int}$	parameters of interest.
$\mathbf{p}_{nui}$	nuisance parameters.
Coordinate vectors	
$\mathbf{x}_{mt} = [x_0, \ldots, x_{N_t-1}]^t$	x-coordinates of the MT at all time instances.[6]
$\mathbf{y}_{mt} = [y_0, \ldots, y_{N_t-1}]^t$	y-coordinates of the MT at all time instances.[6]
$\mathbf{x}_{bs} = [x_{bs_1}, \ldots, x_{bs_{N_s}}]^t$	x-coordinates of all the BS.[7]
$\mathbf{y}_{bs} = [y_{bs_1}, \ldots, x_{bs_{N_s}}]^t$	y-coordinates of all the BS.[7]
$\mathbf{x}_s = [x_{s_1}, \ldots, x_{s_{N_s}}]^t$	x-coordinates of all the scatterers.
$\mathbf{y}_s = [y_{s_1}, \ldots, y_{s_{N_s}}]^t$	y-coordinates of all the scatterers.
LDP vectors	
$\boldsymbol{\phi} \equiv \boldsymbol{\phi}_i = [\phi_1, \ldots, \phi_{N_s}]^t$	AoA of all MPC for 1 time instant.[8]
$\boldsymbol{\phi} = [\boldsymbol{\phi}_0^t, \ldots, \boldsymbol{\phi}_{N_t-1}^t]^t$	AoA of all MPC for all time instances.
$\boldsymbol{\psi} \equiv \boldsymbol{\psi}_i = [\psi_1, \ldots, \psi_{N_s}]^t$	AoD of all MPC for 1 time instant.[8]
$\boldsymbol{\psi} = [\boldsymbol{\psi}_0^t, \ldots, \boldsymbol{\psi}_{N_t-1}^t]^t$	AoD of all MPC for all time instances.
$\mathbf{d}_i = [d_1, \ldots, d_{N_s}]^t$	Lengths of all MPC for 1 time instant.[8]
$\mathbf{d} = [\mathbf{d}_0^t, \ldots, \mathbf{d}_{N_t-1}^t]^t$	Lengths of all MPC for all time instances.
$\mathbf{d}_{mts,i} = [d_{mts,1}, \ldots, d_{mts,N_s}]^t$	MT-scatterers distances for 1 time instant.[8]
$\mathbf{d}_{mts} = [\mathbf{d}_{mts,0}^t, \ldots, \mathbf{d}_{mts,N_t-1}^t]^t$	MT-scatterers distances for all time instances.
$\mathbf{d}_{bs} = [d_{bs,1}, \ldots, d_{bs,N_s}]^t$	BS-scatterers distances .
$\mathbf{f}_{d,i} = [f_{d,1}, \ldots, f_{d,N_s}]^t$	DS of all MPC for 1 time instant.
$\mathbf{f}_d = [\mathbf{f}_{d,0}^t, \ldots, \mathbf{f}_{d,N_t-1}^t]^t$	DS of all MPC for all time instances.
$\boldsymbol{\theta}_k \in \{\boldsymbol{\phi}, \boldsymbol{\psi}, \mathbf{d}, \mathbf{f}_d\}$	LDP.
$\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K]^t$	All available LDP.
Vectors of Trigonometric functions of $\mathbf{z} = T(\boldsymbol{\phi}) + T(\boldsymbol{\psi}) + T(\boldsymbol{\omega})$	
$\mathbf{c}_\mathbf{z} = [cos(z_1), \ldots, cos(z_N)]^t$	Cosines of the angles.
$\mathbf{s}_\mathbf{z} = [cos(z_1), \ldots, cos(z_N)]^t$	Sines of the angles.

[6] If the MT is not moving it reduces to scalar.
[7] If the MT is communicating only with 1 BS it reduces to scalar.
[8] Static environment.

Matrices	
$\mathbf{J}$	FIM (also used for selection matrices in chapter 2).
$\mathbf{I}$	Identity matrix.
$\mathbf{H}$	CIR matrix.
$\mathbf{G}$	Transformation matrix (matrix of partial derivatives).
$\mathbf{X}_{kl}$	Transmitted signal matrix.[9]
$\mathbf{H}_{kl}$	Channel matrix.[9]
$\mathbf{Y}_{kl}$	Received signal matrix.[9]
$\mathbf{N}_{kl}$	Noise Matrix.[9]
$\mathbf{R}_{\mathbf{b}}$	Correlation Matrix of $\mathbf{b}$.
$\mathbf{C}_{\mathbf{b}}$	Covariance Matrix of $\mathbf{b}$.
$\mathbf{A}_R$	Rx array response to all incident waves.
$\mathbf{A}_T$	Tx array response to all incident waves.
Coordinate matrices	
$\mathbf{X}_{mt}$	x-coordinates of the MT.
$\mathbf{Y}_{mt}$	y-coordinates of the MT.
$\mathbf{X}_{bs}$	x-coordinates of the BS.
$\mathbf{Y}_{bs}$	y-coordinates of the BS.
$\mathbf{X}_s$	x-coordinates of the scatterers.
$\mathbf{Y}_s$	y-coordinates of the scatterers.
LDP matrices	
$\mathbf{\Phi}$	AoA.
$\mathbf{\Psi}$	AoD.
$\mathbf{D}$	lengths of paths.
$\mathbf{D}_{mts}$	MT-scatterers distances.
$\mathbf{D}_{bs}$	BS-scatterers distances.
$\mathbf{F}_d$	DS of MPC.
	Matrices of Trigonometric functions of $\mathbf{z} = T(\boldsymbol{\phi}) + T(\boldsymbol{\psi}) + T(\boldsymbol{\omega})$
$\mathbf{C}_{\mathbf{z}}$	Cosines of the angles.[10]
$\mathbf{S}_{\mathbf{z}}$	Sines of the angles.

[9] At time k and frequency l.

[10] The symbol $\mathbf{C}$ is also used for Covariance matrices. Its definition will always be clear from the context.

Chapter 1

Introduction

Localization, wireless or cellular geolocation and often positioning are some of the different names one can come across when browsing the existing rich literature on methods and algorithms that are utilized to find the exact geographical location of a Mobile Transceiver/Terminal (MT). Even more impressing than the number of different names assigned to this problem, is the even larger number of fundamentally different approaches to solve it along with the evolution of these approaches through the years. Although research on this specific topic of the area known as "parameter estimation" has been conducted since the late seventies [1], localization attracted much more interest after the U.S. Federal Communications Commission (FCC) announced that it is mandatory for all wireless service providers to be able to provide location information to public safety services in case of an emergency [2, 3]. However, that was just the initial motivation, since, during the attempt to meet with the FCC requirements in the predetermined time-interval, researchers envisioned new commercial services that could become feasible if the exact location of the MT is known to the provider or just to the user. Specifically, location-sensitive billing, increased data rate due to optimum resource allocation, cellular phone fraud detection, cargo tracking, navigational and yellow-pages services could be introduced by wireless service providers to attract new costumers and satisfy the demanding ones.

Inspired by the difficulty of the localization problem under realistic propagation conditions and motivated by the aforementioned new commercial

services, researches exploited the nature of the wireless channel in their attempt to estimate meaningful parameters that could in turn be used for geolocation. Amongst the numerous methods that were developed, the most commonly used and accepted are the geometrical ones. As the name indicates, geometrical methods are those that are based on the estimation of some location-dependent parameters (LDP)[1] and the exploitation of geometrical relations to express these LDP as a function of the MT position in order to estimate it. Commonly used LDP are the angle of arrival (AoA), the time of arrival (ToA), the time difference of arrival (TDoA), the Doppler shift (DS) and the received signal strength (RSS) [4,5]. Often, a combination of two or more of the above can be employed. Another wide class of localization methods is the so-called fingerprinting (FP) methods. The concept and the implementation of fingerprinting is totally different than the geometrical methods. This approach is based on comparing a set of estimated LDP that compose a "fingerprint" with the entries of an existing database that contains "fingerprints" for various discrete location points. FP was created as an alternative to traditional geometrical methods in difficult propagation environments, like eg. indoor and dense urban environments. In such environments, traditional geometrical methods fail to achieve high accuracy. By "traditional" we refer to the very first localization methods, that use trilateration and/or triangulation and were developed under the assumption of existence of a line-of-sight (LoS) component.

In parallel with FP, "modern" geometrical techniques were developed. These techniques take into account the difficulties faced when localizing in real propagation environments, as opposed to ideal environments and thus are able to overcome them. Albeit sometimes complex, they can estimate the MT position accurately in any environment that fulfills some realistic assumptions. This is exactly the reason why we were inspired by these approaches and decided to develop state-of-the-art geometrical methods. The details of these methods, the key ideas behind them and the assumptions they are based on will be explained in detail throughout this document.

1.1 Fundamentals of Geometrical Localization Methods

In this section we will present the fundamentals of traditional geometrical methods. As already mentioned, in these methods, localization is performed

[1] Location-dependent parameters can also be found under the name channel-dependent parameters in the literature.

in 2 steps. In the 1^{st} step, one or more (kinds of) LDP, are estimated in an adequate number of base stations (BS). The necessary for identification and at the same time sufficient for high accuracy, number of BS, depends on the number of different kinds of LDP that each BS can estimate and on the propagation environment. Once LDP estimates are available, standard geometric and trigonometric laws can be applied to form a system of equations that needs be solved to estimate the coordinates of the MT with respect to the coordinates of the BS, in a two or three dimensional plane. We briefly explain below the basic principles of various geometrical methods that are based on different kinds of LDP estimates. The analysis and the remarks of this section - with the exception of the RSS subsection- are based on the assumption of an ideal, error-free, static LoS environment. Therefore it can serve as a reference for any reader that does not master the fundamentals of localization. The sources of errors, their impact on the techniques and the attempts to mitigate them will be stated in the following sections.

1.1.1 Angle of Arrival Methods

In order to estimate the AoA of incident signals, a directional antenna, such as an adaptive phased array of two or more antenna elements, is required. The most straightforward method is to measure the phase difference between the signals, when impinging on different antenna elements and convert this to an AoA estimate. This method is known as interferometry. Another conceptually simple method is beamforming. It gives an estimation of AoA by electronically steering the main lobe of the antenna array in the direction of the arrival signal and measuring the input power. The main drawback of beamforming is that the angular resolution is limited by the beamwidth of the array[2].

Once the AoA (or AoD if we consider downlink transmission) have been estimated in a minimum of 2 base stations (BS), the MT can be localized. The position of the mobile device is precisely given by the intersection of the 2 lines of bearing, a technique well-known as triangulation. With respect to the figure 1.1, the lines of bearing for $\forall j$ are given by

$$y_{mt} = \frac{1}{\tan(\theta_j)} x_{mt} + y_{bs_j} - \frac{1}{\tan(\theta_j)} x_{bs_j} \tag{1.1}$$

[2]The half-power beamwidth of a linear array of antennas is approximately equal to 1/L, where L is the array aperture.

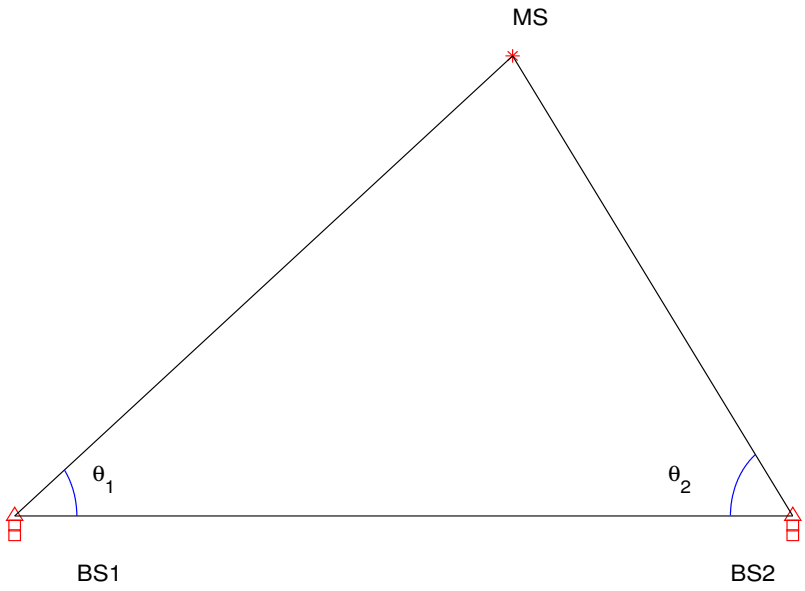

Figure 1.1: AoA localization with 2 BS (triangulation)

so that the position vector $\mathbf{p}_{mt} = [x_{mt}, y_{mt}]^t$ is given by

$$\mathbf{p}_{mt} = \begin{bmatrix} \frac{\tan(\theta_2)x_{bs_1} - \tan(\theta_1)x_{bs_2} - \tan(\theta_1)\tan(\theta_2)(y_{bs_1} - y_{bs_2})}{\tan(\theta_2) - \tan(\theta_1)} \\ \frac{x_{bs_1} - x_{bs_2} - (\tan(\theta_1)y_{bs_1} - \tan(\theta_2)y_{bs_2})}{\tan(\theta_2) - \tan(\theta_1)} \end{bmatrix} \tag{1.2}$$

To combat inaccuracies, more than two BS and highly directional antennas might be employed. The system of lines of bearing can then be solved by means of eg. Least Squares (LS). The main disadvantages of AoA techniques are that they require relatively large and complex hardware and periodic array calibration. The fact that the position estimate degrades as the MS moves further from the BS also renders the AoA techniques less attractive than the ToA and the TDoA techniques in many realistic scenarios.

1.1.2 Time of Arrival Methods

Based on the fact that electromagnetic waves propagate through vacuum -and approximately through any free from objects space medium- at the constant speed of light ($c \approx 3 \times 10^8 m/s$) the distance between the MT and BS j can be calculated as $d_j = c * t_j$ when the duration t_j of the travel of the signal originating from the BS and arriving at the MT, or vice versa,

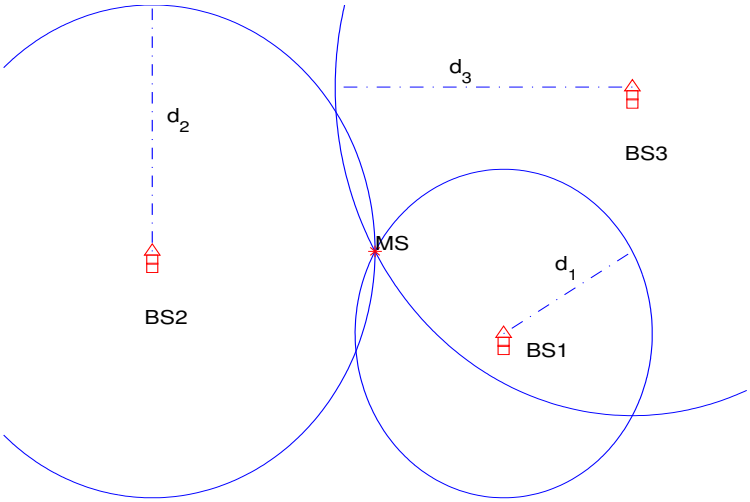

Figure 1.2: ToA localization with 3 BS (trilateration)

is available. The duration t_j is usually estimated during the acquisition or the tracking process. Once the distance d_j between BS j and the MT is known, a sphere (or circle if we assume a 2-D planar geometry) with radius d_j is recognized as the the geographic region on which the MT must lie. If the same procedure is performed at a minimum of 3 BS, the intersection of the spheres (circles) obtained gives the exact location of the MT, as shown in 1.2. This technique is known as trilateration. It should be noted that ToA technique requires a very accurate time synchronization among all transmitters and receivers involved, and a time-stamp to be included in the signal. Using the formula of the euclidean distance between 2 points in a 2-D plane

$$d_j = \sqrt{(x_{mt} - x_{bs_j})^2 + (y_{mt} - y_{bs_j})^2} \tag{1.3}$$

and assuming without loss of generality that BS 1 is located at the {0,0} point (origin), the unknown MT coordinates are given by

$$\begin{bmatrix} x_{mt} \\ y_{mt} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_{bs_2} & y_{bs_2} \\ x_{bs_3} & y_{bs_3} \end{bmatrix}^{-1} \begin{bmatrix} x_{bs_2}^2 + y_{bs_2}^2 + d_1^2 - d_2^2 \\ x_{bs_3}^2 + y_{bs_3}^2 + d_1^2 - d_3^2 \end{bmatrix} \tag{1.4}$$

1.1.3 Time-Difference of Arrival Methods

An efficient way of overcoming the two prerequisites for ToA localization, namely the synchronization and the labeling of the signal with time information, is to examine the difference, denoted herein τ, between the time of arrival of the signal at two different BS, rather than the absolute arrival time. A straightforward method of estimating that difference in time is to cross-correlate the signals arriving at a pair of BS. The cross-correlation of the signals received at BS 1 and BS 2 is given by

$$R_{1,2}(\tau) = \frac{1}{T}\int_0^T s_1(t)s_2(t+\tau)dt \tag{1.5}$$

and will have a pick for τ equal to the exact TDoA, in the absence of errors. Each TDoA estimate determines that in the 2-D case, the transmitter must lie on a hyperboloid[3], rather than a circle. As can be seen from figure 1.3, the intersection of two or more hyperboloids, whose generation requires three or more BS, gives the exact location of the MT. In the 3-D plane, a minimum of four BS is required to estimate the location. It should be mentioned that overcoming ToA method's restrictions by using TDoA, does not come without a cost. For the same number of BS, the performance of TDoA method is always worse than that of ToA, which can be intuitively explained by the fact that by formulating differences of estimates, the number of available data decreases by 1 (eg. 4 ToA estimates give rise to 3 TDoA estimates). Another drawback of the TDoA method is solving the non-linear equations, which has been proved to be cumbersome. Many different approaches can be found in the literature. In [6] an exact solution was presented, but only for the case when the number of TDoA measurements and the number of unknown coordinates are equal. In [7] Abel and Smith presented the "divide and conquer" method, which required sufficiently large Fisher Information while in [8] Foy explored the Taylor-series estimation, an iterative technique with good accuracy under the reasonable initial guess assumption. Last in [9] Chan and Ho presented an non-iterative efficient technique that gives an explicit solution. Below we give their solution for the simple case of three BS which is also equivalent to the solution given in [10].

Let each difference in distance be

$$d_{ij} = d_i - d_j = \sqrt{(x_{mt} - x_{bs_i})^2 + (y_{mt} - y_{bs_i})^2} - \sqrt{(x_{mt} - x_{bs_j})^2 + (y_{mt} - y_{bs_j})^2} \tag{1.6}$$

[3]The hyperboloid may be defined as the locus of points where the difference in the distance to two fixed points (called the foci) is constant.

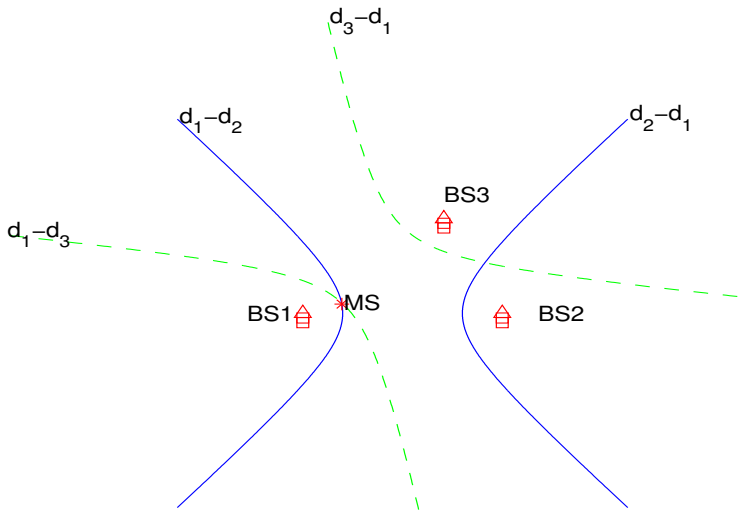

Figure 1.3: TDoA localization with 3 BS

and satisfy the following equation: $d_{ij} = d_{ik} - d_{jk}$, $\forall i, j, k$. Using BS 1 as a reference and introducing the terms $K_i = x_{bs_i}^2 + y_{bs_i}^2$, $x_{i1} = x_{bs_i} - x_{bs_1}$ and $y_{i1} = y_{bs_i} - y_{bs_1}$ we get the following solution for $\mathbf{p}_{mt}$ that depends on d_1.

$$\begin{bmatrix} x_{mt} \\ y_{mt} \end{bmatrix} = -\begin{bmatrix} x_{21} & y_{21} \\ x_{31} & y_{31} \end{bmatrix}^{-1} \left(\begin{bmatrix} d_{21} \\ d_{31} \end{bmatrix} d_1 + \frac{1}{2} \begin{bmatrix} d_{21}^2 - K_2 + K_1 \\ d_{31}^2 - K_3 + K_1 \end{bmatrix} \right) \quad (1.7)$$

Inserting this result into $d_1^2 = (x_{mt} - x_{bs_1})^2 + (y_{mt} - y_{bs_1})^2 = K_1 - 2x_{mt}x_{bs_1} - 2y_{mt}y_{bs_1} + x_{mt}^2 + y_{mt}^2$ we get a quadratic in d_1. Substituting the positive root back into (1.7) produces the final solution for $\mathbf{p}_{mt}$.

1.1.4 Hybrid Methods

The main disadvantage of single-LDP geometrical techniques is the so-called "hearability" problem. Since cellular wireless systems are designed to minimize the number of BS receiving a high-strength signal, the lack of signals with adequate Signal-to-Noise Ratio (SNR) can cause a significant variation in the accuracy of these techniques. In situations where no more than one or two BS can be used in the localization process, more than one kind of LDP should be estimated in order to get the unique location of the MT.

The need for estimating different kinds of LDP and fusing them to estimate the MT location gave birth to the so-called hybrid methods.

In [11], a ToA/AoA hybrid scheme that employs only 1 BS was introduced. In that method, the 1^{st} step is implemented by premultiplying the channel impulse response matrix H by an N_l-by-N_θ matrix of beamformer coefficients W to obtain the spatial impulse response of the antenna array H_θ:

$$H_\theta = W^H * H$$

where N_θ represents the number of "look" directions. The squared magnitudes of the entries in H_θ represent the received energy as a function of ToA and AoA. Specifically the dominant multipath components (MPC)[4] are estimated as the N_p highest peaks of the function obtained by taking the squared magnitudes of the entries of H_θ. Then AoA and ToA estimates are given by the row and column indexes respectively. Once the ToA t and the AoA θ of the LoS path are estimated, the position vector is simply given by

$$\begin{bmatrix} x_{mt} \\ y_{mt} \end{bmatrix} = \begin{bmatrix} x_{bs} + ctcos(\theta) \\ y_{bs} + ctsin(\theta) \end{bmatrix} \tag{1.8}$$

This method, besides having the advantage of using only one BS, achieved relatively good performance. In [12] the authors considered the TDoA/AoA method as a hybrid scheme to achieve high accuracy in a CDMA system. Their method, although requiring a minimum of 2 base stations, has tremendous advantages over other hybrid methods, mainly because it assumes that AoA is estimated only at the serving BS and secondly because there is no need for synchronization as in the ToA/AoA method. To solve the complex system of equations (1.9) they translated all positions to a convenient coordination system by placing BS 1 at the origin and the MT on the x-axis as shown in the following equations:

$$\begin{cases} d_{21} = d_2 - d_1 = \\ \sqrt{(x_{mt} - x'_{bs_2})^2 + (y_{mt} - y'_{bs_2})^2} - \sqrt{(x_{mt} - x'_{bs_1})^2 + (y_{mt} - y'_{bs_1})^2} \\ \theta = tan^{-1}\left(\frac{y_{mt} - y'_{bs_1}}{x_{mt} - x'_{bs_1}}\right) \end{cases}$$

$$\Rightarrow \quad \begin{cases} d_{21} = d_2 - d_1 = \\ \sqrt{(x_{mt} - x_{bs_2})^2 + (y_{mt} - y_{bs_2})^2} - \sqrt{x_{mt}^2 + y_{mt}^2} \\ y_{mt} = 0 \end{cases} \tag{1.9}$$

[4] More on localization algorithms for multipath environments will be presented in the next section.

where the prime is used to denote the coordinates before the translation. The solution then is

$$x_{mt} = \frac{x_{bs_2}^2 + y_{bs_2}^2 - d_{21}^2}{2(x_{bs_2} + d_{21})}, y_{mt} = 0 \tag{1.10}$$

where x_{bs_2} and y_{bs_2} are given by

$$x_{bs_2} = \sqrt{(x'_{bs_2} - x'_{bs_1})^2 + (y'_{bs_2} - y'_{bs_1})^2}\cos(\arctan\frac{y'_{bs_2} - y'_{bs_1}}{x'_{bs_2} - x'_{bs_1}} - \theta) \tag{1.11}$$

$$y_{bs_2} = \sqrt{(x'_{bs_2} - x'_{bs_1})^2 + (y'_{bs_2} - y'_{bs_1})^2}\sin(\arctan\frac{y'_{bs_2} - y'_{bs_1}}{x'_{bs_2} - x'_{bs_1}} - \theta) \tag{1.12}$$

1.1.5 Received Signal Strength Methods

RSS localization methods were first introduced in 1969 [1]. Their major advantage compared to other methods is the availability of RSS measurements in practically all systems [13]. Their major drawback is their low accuracy. The combination of the 2 renders RSS methods, the easier to implement, but at the same time, the least appealing. This is exactly why, with the exception of this section, we make no other reference to RSS throughout this document. At this point, it is unavoidable to mention sources of errors and inaccuracies, which will be discussed extensively in the next section. In an ideal environment, the distance between the MT and the BS can just be computed by using Friis equation for free space loss transmission

$$P_r = P_t + G_t + G_r + 20\log(\lambda/(4\pi d)), \tag{1.13}$$

therefore, there is no practical meaning in discussing it further. Instead lets briefly mention some of the first approaches to RSS localization. In [14] minimum LS estimation of the position in a locally linear model, led to Kalman filtering as an interesting choice to improve performance. The method was tested in a simulation environment at first but later also with the use of real measurement data [15]. However, in an evaluation of RSS methods [13], the authors question the improvement of the location estimate due to Kalman filtering in real cellular networks. To determine the location, RSS measurement must be modeled as a function of MT coordinates. In general, this task can be performed by using a RSS model[5] which is based either on a propagation model or on measurements. In [17], Weiss suggested a signal model described by (1.14) that includes two error terms: A common error

[5]RSS methods can be classified also as Statistical Modeling Localization Techniques [16].

term c for signals transmitted from all BS and a random zero-mean error X_σ which is composed of several factors that cause the variation of the measured signal power around the predicted mean. Such factors are multipath fading, shadowing, antenna pattern, thermal noise and interference.

$$p_i = c + p_{ave,i}(x, y) + X_\sigma,\ i = 1, ..., N. \tag{1.14}$$

N is the number of signals received above a certain threshold out of M base stations and $p_{ave,i}$ is the average power (in dB) given by the propagation model or measurements or a combination of both. For a propagation model one can choose any of the several alternatives that exist in the literature, like, for example, the log-distance models that have the following generic form

$$p_{ave}(r) = \beta - 10\alpha \log_{10}(r) \tag{1.15}$$

where α is the path-loss exponent and β is the intercept. A more complicated approach is to use a terrain-based model. These models require a digital map of the area and may be computationally complex. It is further possible to use only measurements in creating a model, if there is high resolution in the measurement data. The distance between a BS and the MT could be estimated by minimizing a cost function like the one in (1.16), which can be derived using the weighted least squares (WLS) algorithm. Once the distances between the N BS and the MT have been estimated, MT location is given by the intersection of circles of radius d_i centered on BS i respectively, exactly as in the ToA method.

$$Q(r) = [\mathbf{p} - \mathbf{p}_{ave}(r)]^T * W * [\mathbf{p} - \mathbf{p}_{ave}(r)] \tag{1.16}$$

where the weighting matrix $W = \mathbf{C}_{X_\sigma}^{-1}$ is equal to the inverse of the covariance matrix of the errors.

1.2 Fundamentals of Fingerprinting Methods

Primarily motivated by the fact that traditional network-based and satellite-assisted geometrical techniques perform poorly or even totally fail in dense urban and indoor environments, researchers realized the need to attack the localization problem from a totally different perspective. As a result the FP techniques were invented. Following the general guidelines given in [18], the concept of FP is based on the following tasks:

- A priori and periodical “off-line” tasks

  - Choose the appropriate location-dependent parameters that will contribute to the signal signature, the so-called "fingerprint"
  - Run measurements campaigns, (ray-tracing) simulations or both, in order to get the signal signatures, each of which corresponds to a specific unique location.
  - Store the signatures in a large database to use as a reference every time a localization task is performed.
  - Periodically calibrate the database by updating the signatures when significant changes have been made in the environment.

- Real-time "on-line" tasks
  - Evaluate the signature of the signal transmitted by the MS at the serving BS or possibly the serving and more BS.
  - Compare the signature with the stored signatures in order to find good matches and assign probabilities to them. The comparison is usually based on minimizing an appropriate cost function.
  - Choose the location for which the corresponding stored signature best matches the measured signature, i.e. has the highest probability to be the correct one.
  - If necessary, process further to resolve any ambiguity in location estimation.

The signal signature can be derived from any combination of amplitude, phase, delay, direction and polarization information not only of the direct (LoS) component but also of the MPC. Obviously the more parameters used, the higher the accuracy of the method will be, until maybe a saturation point is reached. However, adding parameters also leads to an increase in systems complexity- by requiring more storage space and higher computational power- and in the time needed for the location to be estimated.

One interesting choice for signal signature was described in a realization of the FP in [18]. Influenced by the fact that the received signal vector - assuming antenna array at the Rx- changes rapidly due to noise, movement of the mobile and other effects, the authors considered creating an approximate signal subspace in which all the array vectors must lie. That can be done by collecting a sufficient number of array vectors and using well-known methods. The subspace which is spanned by a set of q linear independent array vectors, where its dimension q depends on the number of distinct paths, can serve as a signal signature associated with each location. In addition to

the signal subspace the authors considered adding to the signal signature a set of differential time delays for the MPC, in order to improve the method's accuracy.

Storing and then processing the signal subspace can lead to high-complexity systems that might not be feasible in practical cases. A lot of undergoing work focuses on defining a fingerprint with as few parameters as possible, while at the same time keeping the accuracy above a certain threshold. A very simplistic approach, presented in [19], is to use just the RSS. Surprisingly simulations showed that the method has a relatively good performance. In [20] the authors considered the power delay profile (PDP) and also introduced the power spatial delay profile (PSDP) as two alternative fingerprints. While PDP takes into account the amplitude and delays of the MPC and completely ignores information contained in their phases, the PSDP matrix exploits additional spatial information contained in the phase-shifts.

As explained in the step-by-step description of FP, after the signal signature has been defined, the researcher will face the dilemma of running simulations or a measurement campaign to obtain the signatures for each and every location. Again there is a trade-off between cost and effort on one hand and accuracy on the other. For instance a database with very accurate signatures can be created and periodically calibrated with the use of a vehicle, that moves through the entire service area, a mobile station and a GPS receiver. However that would require much time, effort and is of course very costly. Simulating a slow time-varying wireless channel propagation model could be an alternative option, but the achieved accuracy will certainly not be as good as in the previous case. Another consideration is the cost function, sometimes referred to as matching score function. It should be carefully chosen in order to fully take advantage of the information contained in the LDP. Finally, one last important consideration when developing a FP technique is the resolving of ambiguities that may appear if the cost function has more than one minima. It can be mitigated in many ways, such as adding more parameters to the signature, measuring the MS signature in more than one BS, using tracking motion systems or even comparing the candidate locations with real locations on the map to find the most possible one.

1.3 Realistic approach to Localization

The previous section, albeit an excellent reference for understanding the concept of localization, is a bit misleading, since it assumes ideal conditions

and presents simple equations that can be solved to give the exact location of a MT. In what follows, we present a more realistic approach to localization. Possible sources of errors and inaccuracies will be identified and existing approaches to alleviate them will be briefly explained. The main source of errors and inaccuracies in geometrical localization methods, is the deviation of the real wireless propagation environment from the ideal environment assumed. Specifically, while most traditional methods were designed for pure LoS environments (one direct and no multipath components), in reality the signal propagates through a multipath or a strictly NLoS environment [21–23]. Although the NLoS environment is just a special case of the multipath environment, the approaches that have been adopted to eliminate or reduce their impact on the accuracy of localization methods are quite different. We will therefore study the impact and some existing localizations techniques designed for each one of the 2 different environments, separately.

1.3.1 Multipath Environment

If a LoS path between the MT and the BS does exist, multipath propagation is essentially the main cause of inaccuracies [21]. Since a LoS component allows the implementation of really simple and accurate localization methods like the ones presented in the previous section, the main objective of any algorithm, designed for estimating the MT location in such an environment, becomes the identification of the LoS component and the accurate estimation of the LDP corresponding to it. It therefore becomes apparent that the impact of multipath environment should be treated during the 1^{st} step of localization. If the receiver is unable to determine which is the first-arriving path in order to identify it as the LoS path and treats a MPC as the LoS one, the position error will be very high. This is explained below were we explain the impact of multipath environment on the LDP estimation.

Multipath propagation affects AoA estimates when the receiver cannot resolve closely-spaced multipath delays. In flat-fading channels the delay spread is very small relatively to the channel bandwidth BW, so the multipath components will all be correlated. However, since delay spread and angular spread are interrelated, the receiver must be able to handle small angular spread. In an attempt to find direction finding methods with good accuracy in multipath environments, researchers employed classical sophisticated methods, like the Maximum Likelihood (ML) and the Maximum Entropy (ME) methods or developed other high resolution ones like the MUSIC [24] and the ESPRIT [25] algorithms. The latter two enjoyed wide acceptance and we will present them briefly in the introduction of chapter 2,

since in that chapter we also develop an ESPRIT-based algorithm for LDP estimation.

In TDoA and ToA techniques, the presence of multipath components limits the time resolution of the TDoA (ToA) estimate to approximately $1/B_W$, when the conventional cross-correlation is used. To convert this resolution to accuracy consider the IS-95 standard for CDMA. The bandwidth of the signal is 1.25 MHz so the signal can propagate $c * (1/B_W) = 240$ meters in $1/B_W$ sec. The TDoA (ToA) estimation error can be much bigger for signals with narrower bandwidth, like in the GSM standard. Therefore, again high resolution estimation techniques are required for better accuracy. In [26] the Root-MUSIC algorithm is used to resolve multipath components that arrive within one chip interval of one another in a Spread Spectrum System. In [27] the total least square ESPRIT (TLS-ESPRIT) algorithm is applied to produce unbiased accurate time delay estimations. In this method, the author proposes to pass the channel output through a matched filter -an operation that is the same as autocorrelating the signal- and then perform a frequency domain deconvolution to convert the TDOA estimation problem to an estimation of frequencies of complex sinusoids in a white non-stationary noise. The TLS-ESPRIT can then be applied, to estimate the unknown frequencies. The above two algorithms clearly indicate the universality of the MUSIC and ESPRIT algorithms in LDP estimation problems. In [28] a concatenation of matched filtering, set-theoretic deconvolution and autoregressive modeling is employed to estimate the parameters of the multipath channel, while in [29, 30] the background theory of cyclostationary signals is presented and applied to the same problem.

Recently in [31], Qi et al considered a total different approach for positioning in multipath environments and derived an algorithm that utilizes TOA. Instead of trying to localize the MT, based solely on the ToA of the first arriving signal component -which propagates through the LoS path- they investigated the enhancement in performance when the ToA of the MPC are also being processed. They showed that the signal strength of those components and the prior statistics of their ToA play an important role in the enhancement. This approach will be better explained in the following analysis of existing techniques for NLoS localization, since in a strictly NLoS environment i.e. in the absence of a LoS component, only the LDP of the MPC can be estimated and utilized in the 2^{nd} step of the MT localization.

1.3.2 Non Line-of-Sight Environment

The problem of localizing in NLoS environments has been addressed in the most recent years and is still an open area of research. There exist 2 fundamentally different approaches to this problem: In the 1^{st} approach, there is an attempt to reduce the impact of the MPC, while in the 2^{nd} approach there is an attempt to benefit from it, by making various assumptions and essentially creating more complex algorithms. Various methods that adopt the 1^{st} approach assume the existence of a LoS path between the MT and at least few of the involved BS. These methods are based on identifying and removing the NLoS measurements [32, 33], weighting these measurements appropriately in order to minimize their impact [34] or utilizing these measurements assuming knowledge of the NLoS errors statistics [35]. Other methods that adopt the 1^{st} approach but are designed for strictly NLoS environments, attempt to solve a constrained optimization problem where the error bias leads to inequalities instead of equalities [36, 37]. On the other hand, methods that adopt the 2^{nd} approach utilize appropriate NLoS channel models to describe the environment and create a mapping between the LDP of the MPC and the the MT coordinates [38,39]. Usually these methods result in more complex algorithms, since, in order to make the channel models as accurate as possible, nuisance parameters need to be introduced and either be integrated out (marginalization) or be jointly estimated. All the methods presented in this document fall into this last category. We will therefore explain briefly herein only methods that adopt the 1^{st} approach.

For AoA positioning, the direction of the received signal at BS j can be expressed as

$$\phi_j = \theta_j + \tilde{\theta}_j, \, \forall j \tag{1.17}$$

where $\tilde{\theta}_j$ is the NLoS induced error. In figure 1.4, it can be clearly seen that if the signal arrives at one BS by a reflection, the direction of arrival is misestimated. This results in a distance error e_d, which is proportional not only to the angle error $\tilde{\theta}_j$ but also to the distance d between that BS and the MT. Specifically using the law of sines and having figure 1.4 as a reference, we get for the error of the estimation of BS 1:

$$e_{d_1} = d_1 * \sin(\psi_1) / \sin(180 - \phi_1 - \theta_2)$$

In contrast to different single-LDP methods, like TDoA and ToA, little work that adopts the 1^{st} approach has been done in improving the accuracy when AoA is used in a NLoS environment. This probably stems from the fact that in AoA methods, a small error in the estimated angle can cause an arbitrarily large error in the estimated MT location. On top of that, to mitigate

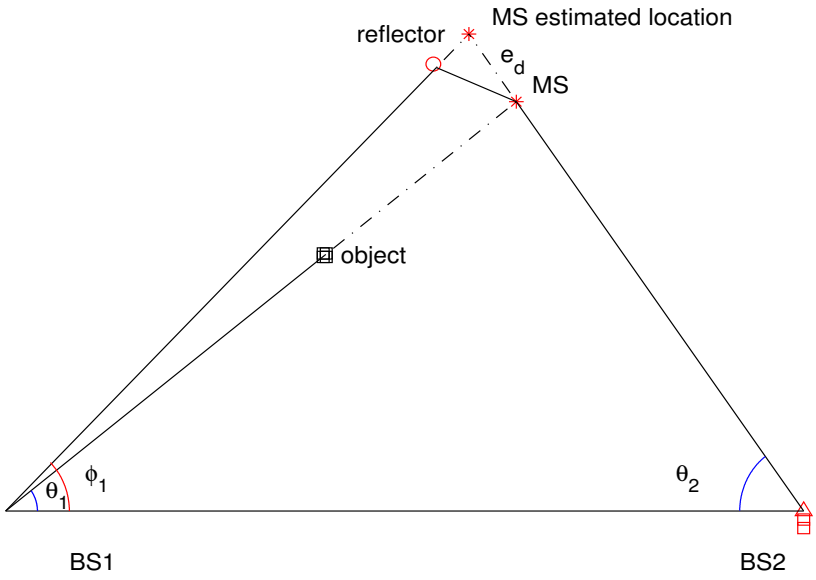

Figure 1.4: AoA location estimation error induced by NLoS environment

effectively the NLoS error, a lot more than 2 BS need to be employed, therefore the great advantage of using the minimum number of just 2 BS is lost. For the case of employing more than 2 BS in AoA positioning, a somehow efficient in reducing the estimation error algorithm was proposed in [32]. The author derived analytical expressions for the coordinates (x_{mt}, y_{mt}) by maximizing the joint distribution of the angle errors $\tilde{\theta}_i = \phi_i - \theta_i$ assuming that they are jointly Gaussian and that the induced error in distance is small compared to the distance between the MT and the BS. An iterative algorithm was then proposed, to identify and exclude NLoS signals in the following 4 steps:

- Estimate the coordinates (x_{mt}, y_{mt}), using all the available measurements.
- Calculate the error of angular measurement at BS j using the formula:

$$\psi_j = \frac{|(x_{mt} - x_{bs_j})sin\theta_i - (y_{mt} - y_{bs_j})cos\theta_j|}{\sqrt{(x_{mt} - x_{bs_j})^2 + (y_{mt} - y_{bs_j})^2}} \ , \ \forall j \tag{1.18}$$

  Then sort the $\tilde{\theta}_j$'s according to their absolute value and calculate the root mean square RMS$(\tilde{\theta}_j)$.

- Reject the BS whose $\tilde{\theta}_j$ is much greater than RMS($\tilde{\theta}_j$).
- Use the remaining BS to estimate (x_{mt}, y_{mt}) again.

In ToA and TDoA techniques NLoS propagation will bias the time measurements with a non-negative error term ϵ_{τ_j} so that the distance becomes:

$$d_j = \sqrt{(x_{mt} - x_{bs_j})^2 + (y_{mt} - y_{bs_j})^2} + \epsilon_{d_j} \tag{1.19}$$

where $\epsilon_{d_j} = c * \epsilon_{\tau_j}$. Most of the contributions cited at the beginning of this section concern methods for ToA NLoS localization. A somehow different method than the ones cited there, which still adopts the 1^{st} approach is based on exploiting the property of the NLoS errors (biases), which are always non-negative as shown in (1.19) and then searches for the true position by adding some constraints. This method was proposed in [40], where the authors extended the solution proposed in [9] to improve accuracy in the NLoS propagation. Their simulation revealed the algorithm's good performance on one hand but indicated the need for more than 5 BS to be employed on the other.

1.4 Geometrical Channel Models

We present in this section the 2-D channel models that enable us to express the LDP as a function of the MT coordinates, among other parameters. For LoS environments, it suffices to use simple trigonometric functions and euclidean distances. We present these relations in the following subsection for reference. On the other hand, for NLoS static environments, some assumptions for the propagation environment need to be made. To that end, we utilize the so-called single-bounce model (SBM) that will be presented in subsection 1.4.2. On top of that, to account for NLoS dynamic environments, we introduce in the same subsection, the dynamic single-bounce model (DSBM), which is the result of the integration of the SBM with a mobility model.

1.4.1 LoS Static Environment

In a LoS environment, the MT communicates with the BS through a direct component, as shown in fig. 1.5. The LDP commonly used in localization algorithms, like eg. the AoA, the AoD and the lengths of the direct paths between the MT and BS j, $1 \leq j < N_s$, are given by

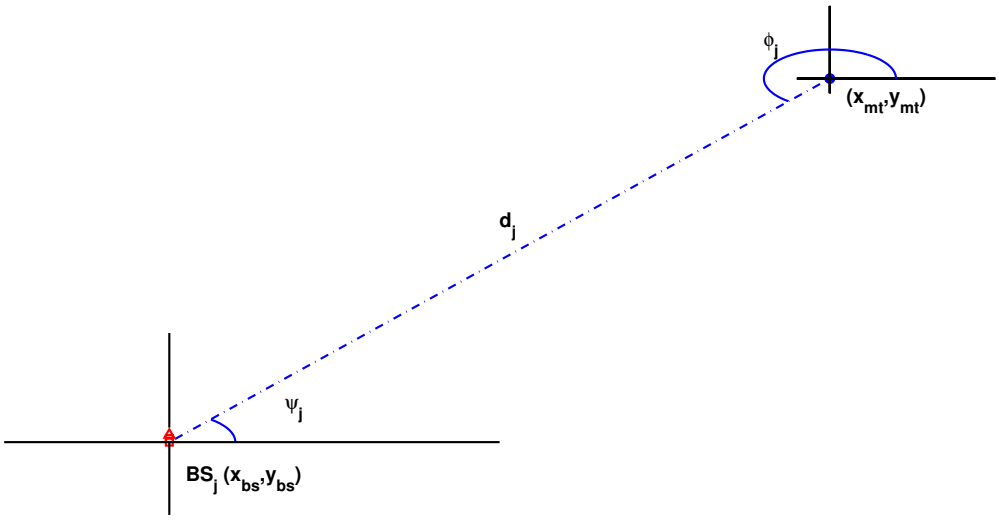

Figure 1.5: LDP in a LoS environment

$$\phi_j = \tfrac{\pi}{2}(1 - sgn\{x_{mt} - x_{bs_j}\}) + \tan^{-1}\tfrac{y_{mt}-y_{bs_j}}{x_{mt}-x_{bs_j}} \tag{1.20}$$

$$\psi_j = \phi_j \pm \pi \tag{1.21}$$

$$d_j = \sqrt{(y_{bs_j} - y_{mt})^2 + (x_{bs_j} - x_{mt})^2}. \tag{1.22}$$

The sign term that appears on the r.h.s of (1.20) stems from the fact that the AoA ϕ_j can take any value in $[0, 2\pi]$. This is explained in more detail in appendix B.

1.4.2 Single-Bounce NLoS Environment

Throughout this thesis , we utilize the single-bounce model (SBM), slightly different versions of which have been introduced and employed in localization methods by Miao et al [38, 41] and Jazzar and Caffery [39]. The SBM describes accurately numerous NLoS and multipath scenarios, despite the fact that it is very simple. Its wide applicability stems from the fact that in a physical propagation environment, the more bounces, the larger the attenuation will be, not only because the scatterer absorbs some of the signal's energy but also because more bounces usually implies a longer path length. Thus, if a limited number of NLoS signal components with non-negligible energy arrive at the receiver, it is reasonable to assume that they

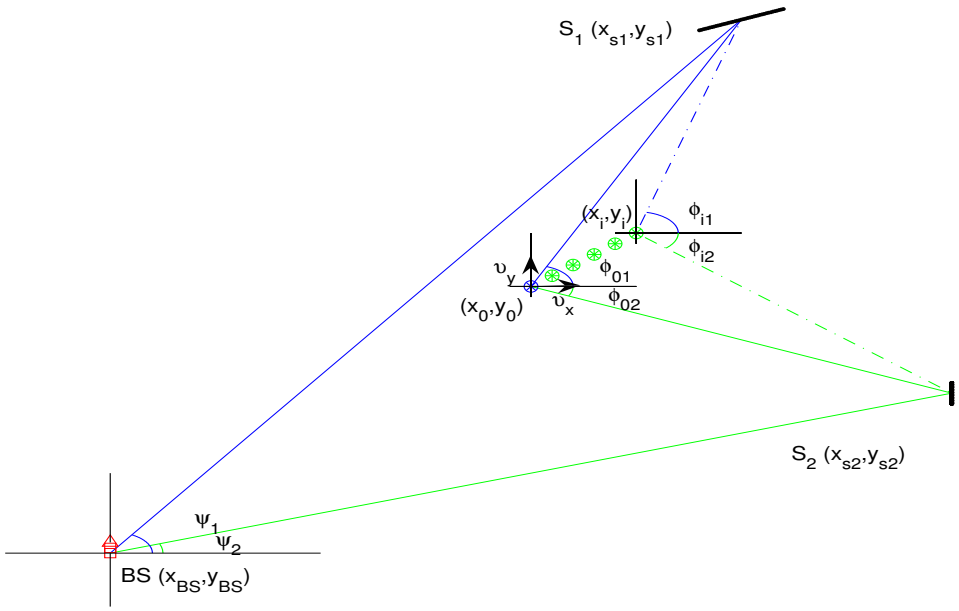

Figure 1.6: LDP in a NLoS environment: Dynamic single bounce model

have bounced only once. Define the following position vectors:

$$\mathbf{p}_{mt} = [x_{mt}, y_{mt}]^t \tag{1.23}$$
$$\mathbf{p}_{bs} = [\mathbf{x}_{bs}^t, \mathbf{y}_{bs}^t]^t \tag{1.24}$$
$$\mathbf{p}_s = [\mathbf{x}_s^t, \mathbf{y}_s^t]^t \tag{1.25}$$

Using the SBM we are able to express any LDP explicitly as a function of the entries of these vectors, i.e. as a function of the BS, MT and scatterers' coordinates, as follows

$$\phi_j = \tfrac{\pi}{2}(1 - sgn\{x_{s_j} - x_{mt}\}) + \tan^{-1}\tfrac{y_{s_j} - y_{mt}}{x_{s_j} - x_{mt}} \tag{1.26}$$
$$\psi_j = \tfrac{\pi}{2}(1 - sgn\{x_{s_j} - x_{bs_j}\}) + \tan^{-1}\tfrac{y_{s_j} - y_{bs_j}}{x_{s_j} - x_{bs_j}} \tag{1.27}$$

$$d_j = d_{mts,j} + d_{bs,j} \tag{1.28}$$
$$d_{mts,j} = \sqrt{(y_{s_j} - y_{mt})^2 + (x_{s_j} - x_{mt})^2} \tag{1.29}$$
$$d_{bs,j} = \sqrt{(y_{s_j} - y_{bs_j})^2 + (x_{s_j} - x_{bs_j})^2} \quad . \tag{1.30}$$

The aforementioned approaches consider a static propagation environment, i.e. they assume that the MT is not moving. We, on the other hand, are also interested in dynamically changing environments, where we assume that the MT is moving with speed that has magnitude $\upsilon_i = \sqrt{\upsilon_{x_i}^2 + \upsilon_{y_i}^2}$ and direction $\omega_i = \frac{\pi}{2}(1 - sgn\{\upsilon_{x_i}\}) + \tan^{-1}(\upsilon_{y_i}/\upsilon_{x_i})$. In such environments, the LDP[6] are time-varying. To account for this and at the same time benefit from it, we introduce the dynamic single bounce model (DSBM), which is a result of the integration of the SBM with an appropriate mobility model. Two mobility models are considered: The constant-speed model, for which the MT position at time i is given by

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \upsilon_x \\ \upsilon_y \end{bmatrix} t_i \tag{1.31}$$

and the constant-acceleration model, for which the MT position and speed at time i are given by the equations below

$$\begin{bmatrix} \upsilon_{x,i} \\ \upsilon_{y,i} \end{bmatrix} = \begin{bmatrix} \upsilon_{x,0} \\ \upsilon_{y,0} \end{bmatrix} + \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} t_i \tag{1.32}$$

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \upsilon_{x,0} \\ \upsilon_{y,0} \end{bmatrix} t_i + \frac{1}{2} \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} t_i^2 \tag{1.33}$$

where t_i is the time difference between time instants i and 0. Similarly to the static scenario, we can use the DSBM to express any time-varying LDP (including the Doppler Shift) explicitly as a function of the BS, MT and scatterers' coordinates, as follows

$$\phi_{ij} = \frac{\pi}{2}(1 - sgn\{x_{s_j} - x_i\}) + \tan^{-1}\frac{y_{s_j} - y_i}{x_{s_j} - x_i} \tag{1.34}$$

$$d_{ij} = d_{mts,ij} + d_{bs,j} \tag{1.35}$$

$$d_{mts,ij} = \sqrt{(y_{s_j} - y_i)^2 + (x_{s_j} - x_i)^2} \tag{1.36}$$

$$f_{d,ij} = \frac{f_c}{c}\upsilon_i \cos(\phi_{ij} - \omega_i) = \frac{f_c}{c}\frac{\upsilon_{x_i}(x_{s_j} - x_i) + \upsilon_{y_i}(y_{s_j} - y_i)}{\sqrt{(y_{s_j} - y_i)^2 + (x_{s_j} - x_i)^2}} \tag{1.37}$$

The AoD and the distances between the BS and the corresponding scatterers are still given by eq. (1.27) and (1.30) respectively. Using eq. (1.31) we

[6] In dynamic environments, LDP is used as an acronym for LMDP, where "M" stands for motion.

obtain the DSBM equations that express the time-varying LDP as a function of the initial position of the MT and its constant speed components

$$\phi_{ij} = \frac{\pi}{2}(1 - sgn\{x_{s_j} - x_0 - \upsilon_x t_i\}) + \tan^{-1}\frac{y_{s_j} - y_0 - \upsilon_y t_i}{x_{s_j} - x_0 - \upsilon_x t_i} \tag{1.38}$$

$$d_{mts,ij} = \sqrt{(y_{s_j} - y_0 - \upsilon_y t_i)^2 + (x_{s_j} - x_0 - \upsilon_x t_i)^2} \tag{1.39}$$

$$f_{d,ij} = \frac{f_c}{c}\frac{\upsilon_{x_i}(x_{s_j} - x_0 - \upsilon_x t_i) + \upsilon_{y_i}(y_{s_j} - y_0 - \upsilon_y t_i)}{\sqrt{(y_{s_j} - y_0 - \upsilon_y t_i)^2 + (x_{s_j} - x_0 - \upsilon_x t_i)^2}} \tag{1.40}$$

while using eq. (1.32)-(1.33), we obtain the DSBM equations that express the time-varying LDP as a function of the initial position of the MT, its initial speed and its constant acceleration components

$$\phi_{ij} = \frac{\pi}{2}(1 - sgn\{x_{s_j} - x_0 - \upsilon_{x_0} t_i - \frac{1}{2}\alpha_x t_i^2\}) + \tan^{-1}\frac{y_{s_j} - y_0 - \upsilon_{y_0} t_i - \frac{1}{2}\alpha_y t_i^2}{x_{s_j} - x_0 - \upsilon_{x_0} t_i - \frac{1}{2}\alpha_x t_i^2} \tag{1.41}$$

$$d_{mts,ij} = \sqrt{(y_{s_j} - y_0 - \upsilon_{y_0} t_i - \frac{1}{2}\alpha_y t_i^2)^2 + (x_{s_j} - x_0 - \upsilon_{x_0} t_i - \frac{1}{2}\alpha_x t_i^2)^2} \tag{1.42}$$

$$f_{d,ij} = \frac{f_c}{c}\frac{\upsilon_{x_i}(x_{s_j} - x_0 - \upsilon_{x_0} t_i - \frac{1}{2}\alpha_x t_i^2)) + \upsilon_{y_i}(y_{s_j} - y_0 - \upsilon_{y_0} t_i - \frac{1}{2}\alpha_y t_i^2)}{\sqrt{(y_{s_j} - y_0 - \upsilon_y t_i - \frac{1}{2}\alpha_y t_i^2)^2 + (x_{s_j} - x_0 - \upsilon_{x_0} t_i - \frac{1}{2}\alpha_x t_i^2)^2}} \tag{1.43}$$

1.5 Statistical Channel Models

We present herein channel models that can be utilized to express the channel impulse response (CIR)[7] matrix as a function of the LDP and possibly other random parameters. Statistical models based on the geometrical models described in the previous section, are also presented.

1.5.1 Double Directional Model for MIMO Channels

The double directional model (DDM) describes a time-variant, frequency-selective channel, taking into account the AoA, the AoD, the delays, the Doppler shifts (DS) and the powers of the paths at the steering directions

[7] CIR will also be used for channel transfer function that is obtained by taking the CIR FFT with respect to delay.

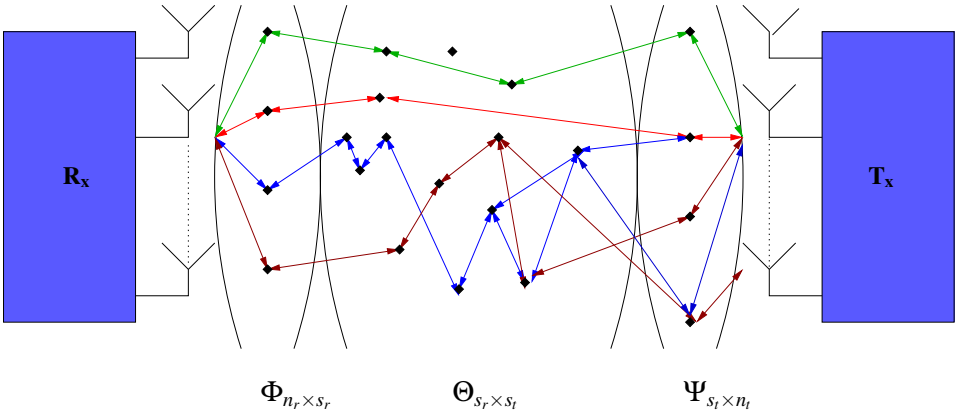

Figure 1.7: Double directional model for a MIMO channel

at the transmitter and receiver size. Its based on the assumption that the scattering environment does not change during the transmission, thus the variation in time is a result of the movement of the MT and/or the BS. Debbah et al. introduced this model in [42]. In [43], the same authors validated the model using the Maximum Entropy (ME) principle. Moreover they proved that the DDM encompasses the "Kronecker", the "Müller", the "Virtual Representation" and the "keyhole" models as special cases. According to the DDM, the $n_r \times n_t$ MIMO matrix $\mathbf{H}$ in the time-frequency domain, is given by

$$\mathbf{H}_{n_r \times n_t}(f,t) = \frac{1}{\sqrt{s_r s_t}} \mathbf{\Phi}_{n_r \times s_r}(t) \mathbf{P_r} (\mathbf{\Theta}_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f)) \mathbf{P_t} \mathbf{\Psi}^t_{s_t \times n_t}(t) \quad (1.44)$$

where n_r,n_t,s_r,s_t are the number of receiving and transmitting antennas and the number of scatterers in the area of the receiver and the transmitter respectively (See figure 1.7). The entries of $\mathbf{\Theta}$ are i.i.d. complex Gaussian with zero mean and unit variance. $\mathbf{P}_r$ and $\mathbf{P}_t$ are diagonal matrices containing the powers. The rest of the matrices on the r.h.s. of (1.44) are defined as follows[8]

$$\mathbf{\Phi} = \begin{bmatrix} e^{j(\beta_{r,11} + 2\pi \frac{f_c}{c} \upsilon_r \cos(\phi_{11} - \omega_r)t)} & \dots & e^{j(\beta_{r,1s_r} + 2\pi \frac{f_c}{c} \upsilon_r \cos(\phi_{1s_r} - \omega_r)t)} \\ \vdots & \ddots & \vdots \\ e^{j(\beta^r_{r,n_r1} + 2\pi \frac{f_c}{c} \upsilon_r \cos(\phi_{n_r1} - \omega_r)t)} & \dots & e^{j(\beta^r_{r,n_rs_r} + 2\pi \frac{f_c}{c} \upsilon_r \cos(\phi_{n_rs_r} - \omega_r)t)} \end{bmatrix} \quad (1.45)$$

[8]Symbol definitions in this subsection are quite different from the general definition given in the Symbol Index.

$$\mathbf{\Psi} = \begin{bmatrix} e^{j(\beta_{t,11}+2\pi\frac{f_c}{c}\upsilon_t\cos(\phi_{11}-\omega_t)t)} & \dots & e^{j(\beta_{t,1s_t}+2\pi\frac{f_c}{c}\upsilon_t\cos(\phi_{1s_t}-\omega_t)t)} \\ \vdots & \ddots & \vdots \\ e^{j(\beta^r_{t,n_t1}+2\pi\frac{f_c}{c}\upsilon_t\cos(\phi_{n_t1}-\omega_t)t)} & \dots & e^{j(\beta^r_{t,n_ts_t}+2\pi\frac{f_c}{c}\upsilon_r\cos(\phi_{n_ts_t}-\omega_t)t)} \end{bmatrix} \tag{1.46}$$

$$\mathbf{D} = \begin{bmatrix} e^{-j2\pi f\tau_{11}} & \dots & e^{-j2\pi f\tau_{1s_t}} \\ \vdots & \ddots & \vdots \\ e^{-j2\pi f\tau_{s_r1}} & \dots & e^{-j2\pi f\tau_{s_rs_t}} \end{bmatrix} \tag{1.47}$$

To explain the meaning of the entries of the above matrices, lets first define the subscripts:

$$q,\ 1 \le q \le n_r \tag{1.48}$$
$$i,\ 1 \le i \le s_r \tag{1.49}$$
$$j,\ 1 \le j \le s_t \tag{1.50}$$
$$p,\ 1 \le p \le n_t \tag{1.51}$$

The phase $\beta_{r,qi}$ represents the initial phase of the signal from scatterer i to receiving antenna q and ϕ_{qi} is the angle between a line perpendicular to the antenna array and the wavefront's direction (AoA). υ_r and ω_r are the receiver's speed magnitude and direction. The parameters $\beta_{t,pj}$ and ψ_{pj} along with υ_t and ω_t are defined in a similar way for the transmitter. τ_{ij} is the unknown delay required for the signal to propagate from scatterer j at the transmitter's side to scatterer i at the receiver's side.

The phases $2\pi\frac{f_c}{c}\upsilon_r\cos(\phi_{qi}-\omega_r)t$ and $2\pi\frac{f_c}{c}\upsilon_t\cos(\psi_{pj}-\omega_t)t$ represent the shift in frequency due to the movement of the receiver and the transmitter (Doppler effect). For the common scenario of Uniform Linear Arrays (ULA) at both ends the DDM becomes simpler since the number of parameters is reduced. Specifically the entries of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ become[9]:

$$[\mathbf{\Phi}]_{ij} = e^{j2\pi(\frac{d(i-1)\sin(\phi_{qi})}{\lambda}+\frac{f_c}{c}\upsilon_r cos(\phi_{qi}-\omega_r)t)} \tag{1.52}$$

$$[\mathbf{\Psi}]_{ij} = e^{j2\pi(\frac{d(j-1)\sin(\psi_{pj})}{\lambda}+\frac{f_c}{c}\upsilon_t cos(\psi_{pj}-\omega_t)t)} \tag{1.53}$$

where we have used the fact that under the far-field approximation, the initial phase $\beta_{r,qi}$ ($\beta_{t,pj}$) varies linearly with the sine of ϕ_{qi} (ψ_{pj}).

[9] Any one or two-dimensional antenna array can be considered instead of ULA with a simple modification of the exponent of the entries of these matrices.

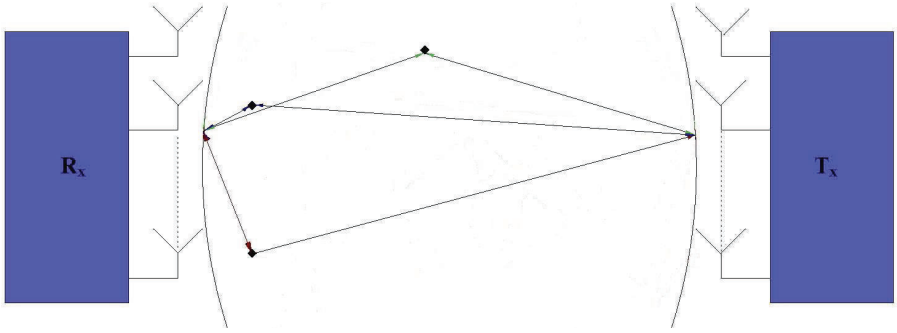

Figure 1.8: Double directional SBM for a MIMO Channel

1.5.2 Double Directional Model for MIMO Channels assuming Single-Bounce Environments

In this section we introduce a special case of the DDM that is more appropriate for NLoS localization problems, since it is far more simple than the general DDM. It can be used to describe environments, where each signal component bounces only once exactly as shown in fig. 1.8. Therefore, this model not only is based on but also can be integrated with the SBM. It is a slightly modified version of the random matrix model introduced in [44]. According to this model, the $n_r \times n_t$ MIMO matrix $\mathbf{H}$ in the time-frequency domain, is given by

$$\begin{aligned}\mathbf{H}_{kl} &= \tfrac{1}{\sqrt{P_{tot}}} \textstyle\sum_{j=1}^{N_s} \sqrt{P_j}\gamma_j e^{j2\pi l \Delta t f_{d,lj}} \mathbf{a}_R(\phi_{lj}) \mathbf{a}_T^t(\psi_{lj}) H_{TR,k} e^{-j2\pi k \Delta f \tau_{lj}} \\ &= \mathbf{A}_{R,l}(\mathbf{\Gamma} \odot (\mathbf{D}_k \mathbf{F}_{d,l})) \mathbf{A}_{T,l}^t = \mathbf{A}_{R,l} \mathbf{\Gamma} \mathbf{D}_k \mathbf{F}_{d,l} \mathbf{A}_{T,l}^t. \end{aligned} \tag{1.54}$$

The subscripts k, $1 \leq k \leq N_f$ and l, $0 \leq l < N_t - 1$ denote frequency and time sample respectively, i.e. $\mathbf{H}_{kl} = \mathbf{H}(f_k, t_l)$. The subscript j, $1 \leq j \leq N_s$ denotes scatterer or multipath signal component. The definition of all the parameters in the first representation (sum of rank 1 terms) of the channel matrix are given in table 1.1. The newly introduced matrices in the matrix representation are defined as follows

$$\mathbf{A}_{R,l} \triangleq [\mathbf{a}_R(\phi_{l1}), \ldots, \mathbf{a}_R(\phi_{lN_s})] \tag{1.55}$$

$$\mathbf{A}_{T,l} \triangleq [\mathbf{a}_T(\psi_{l1}), \ldots, \mathbf{a}_T(\psi_{lN_s})] \tag{1.56}$$

$$\mathbf{\Gamma} \triangleq H_{TR,k} diag(\boldsymbol{\gamma}) \tag{1.57}$$

$$\mathbf{D}_k \triangleq diag(\mathbf{d}_k) \tag{1.58}$$

$$\mathbf{F}_{d,l} \triangleq diag(\mathbf{f}_{d,l}) \tag{1.59}$$

Table 1.1: Parameters in terms composing $\mathbf{H}_{kl}$

Δt, Δf	sampling space in time and frequency
τ_{lj}	delay
$\gamma_j \sim \mathcal{CN}(0,1)$	complex amplitude
$P_j(\propto \tau_{lj}^{-a})$	power[10]
P_{tot}	power normalization constant
$H_{TR,k}$	filter's[11] transfer function
ϕ_{lj}, ψ_{lj}	AoA and AoD
$\mathbf{a}_R(\phi_{lj})$ and $\mathbf{a}_T(\psi_{lj})$	Rx and Tx array response

where

$$\boldsymbol{\gamma} \triangleq \frac{1}{\sqrt{P_{tot}}} diag([\sqrt{P_1}\gamma_1, \ldots, \sqrt{P_{N_s}}\gamma_{N_s}]) \tag{1.60}$$

$$\mathbf{d}_k \triangleq [e^{-j2\pi k\Delta f \tau_{l1}}, \ldots, e^{-j2\pi k\Delta f \tau_{lN_s}}] \tag{1.61}$$

$$\mathbf{f}_{d,l} \triangleq [e^{j2\pi l\Delta t f_{d,l1}}, \ldots, e^{j2\pi l\Delta t f_{d,lN_s}}]. \tag{1.62}$$

The last equality stems from the fact that in single-bounce scenario $\boldsymbol{\Gamma}$, $\mathbf{D}_k$ and $\mathbf{F}_{d,l}$ are diagonal matrices. It is worth pointing out that this model can in fact describe any NLoS environment with delays that correspond to more than 1 pair of angles (AoA-AoD) as long as each AoA is linked with one AoD. The model is not appropriate only when an angle at one side is linked to multiple angles on the other side.

1.5.3 Double Directional Model for MIMO Channels assuming Multipath Environments

The MIMO channel matrix given by eq. (1.44) and(1.54) is based on the assumption of a strictly NLoS environment. If a LoS component exists, then, in order to represent the multipath environment, the MIMO channel matrix can be written as a sum of the following 2 components

$$\mathbf{H}_{kl} = \mathbf{H}_{NL,kl} + \mathbf{H}_{L,kl}. \tag{1.63}$$

The NLoS component $\mathbf{H}_{NL,kl}$ is given by eq. (1.44) or (1.54) and the LoS component $\mathbf{H}_{L,kl}$ is given by

$$\begin{aligned}\mathbf{H}_{L,kl} &= \frac{\sqrt{P_0}}{\sqrt{P_{tot}}} e^{j\theta_n} e^{j2\pi f_{d,l0} t_l} \mathbf{a}_R(\phi_{l0}) \mathbf{a}_T^t(\psi_0) H_{TR,k} e^{-j2\pi f_k \tau_{l0}} \\ &= e^{j\theta_n} d_{k0} f_{d,l0} \mathbf{a}_{R,0} \mathbf{a}_{T,0}^t. \end{aligned} \tag{1.64}$$

In the equation above we have introduced the index $j = 0$ which will be used through this document for LDP that correspond to the LoS component and the unknown phase shift (due to phase noise) of the LoS path $\theta_n \sim U[0, 2\pi]$.

1.6 Input-Output Relationship for MIMO-OFDM Systems

The discrete input-output relationship of a $n_r \times n_t$ MIMO-OFDM system can be written as

$$\mathbf{Y}(f_k, t_l) = \mathbf{H}(f_k, t_l)\mathbf{X}(f_k, t_l) + \mathbf{N}(f_k, t_l) \tag{1.65}$$

or alternatively, using the subscript $(\cdot)_{kl}$ to denote frequency sample $f_k = k\Delta f$ and time sample $t_l = l\Delta t$,

$$\mathbf{Y}_{kl} = \mathbf{H}_{kl}\mathbf{X}_{kl} + \mathbf{N}_{kl}. \tag{1.66}$$

$\mathbf{H}_{kl}$ is the $n_r \times n_t$ channel matrix is given by eq. (1.63) for multipath environments and reduces to the one given by eq. (1.44) or (1.54) for strictly NLoS ones. However, Δf is now also equal to the sub-carrier spacing and N_f is also equal to the number of sub-carriers. $\mathbf{X}_{kl}$ is the $n_t \times N$ transmitted signal matrix, which, in our work, we consider to be composed of a training sequence of symbols. N is the number of OFDM symbols per sub-carrier transmitted during the channel's coherence time. N can be greater than 1 since, in most practical systems of interest, the product of symbol's period T with the number of sub-carriers N_f utilized by one user is much smaller than the channel's coherence time, i.e. $N_f T < 1/B_d$, where B_d is the Doppler Spread. Indeed, while the Doppler Spread usually ranges between 0 and some hundreds of Hz, so that the coherence time is at least $1msec$, the symbol's period is of the order of μsec and the number of sub-carriers utilized is at most a few tens. This essentially means that the channel is static during the transmission of consecutive (blocks of) symbols but changes between 2 consecutive time instances t_l, t_{l+1}, provided that Δt is comparable with

$1/B_d$. $\mathbf{Y}_{kl}$ is the $n_r \times N$ received signal matrix and $\mathbf{N}_{kl}$ is the $n_r \times N$ noise matrix, all at frequency k, $1 \leq k \leq N_f$ and time l, $1 \leq l \leq N_t$. The entries of $\mathbf{N}_{kl}$ are i.i.d complex Gaussian with 0 mean and variance σ_n^2.

The input-output relationship can be equivalently written in vectorized form, as follows:

$$\mathbf{y}_{kl} = (\mathbf{X}_{kl}^t \otimes \mathbf{I}_{\mathbf{n_r}})\mathbf{h}_{kl} + \mathbf{n}_{kl} \tag{1.67}$$

where $\mathbf{h}_{kl} = vec(\mathbf{H}_{kl})$ and $\mathbf{n}_{kl} = vec(\mathbf{N}_{kl})$. Using (1.63), the channel vector can also be expressed as a sum of 2 terms for multipath environments:

$$\mathbf{h}_{kl} = \mathbf{h}_{NL,kl} + \mathbf{h}_{L,kl} \tag{1.68}$$

where

$$\mathbf{h}_{NL,kl} = vec(\mathbf{H}_{NL,kl}) = (\mathbf{A}_{T,l}^t \boxtimes \mathbf{A}_{R,l})\mathbf{D}_{kl}\gamma \tag{1.69}$$

$$\mathbf{h}_{L,kl} = vec(\mathbf{H}_{L,kl}) = e^{j\theta} d_{0,kl}(\mathbf{a}_{T,l0}^t \otimes \mathbf{a}_{R,l0}). \tag{1.70}$$

1.7 Maximum Likelihood and Bayesian Estimation

Whenever researchers attack a parameter estimation problem, they face the dilemma of choosing between maximum likelihood (ML) and Bayesian estimation (BE). Since localization is essentially a parameter estimation problem, we faced this dilemma numerous times. In the ML approach the unknown parameters are treated as deterministic quantities. The estimation is then based on maximizing the density of the data conditioned on the parameter vector. This function is called the "likelihood" function and thus the method is called ML. In the Bayesian approach, the unknown parameters are treated as random variables. If their prior distributions are not known, a method known as Bayesian Inference [45] can be used to produce meaningful priors. The estimation can then be based on maximizing the density of the parameter vector conditioned on the data (Maximum a posteriori estimation). In general MAP estimators have superior performance than ML estimators, assuming that the priors used are the correct ones of course. However, it should be noted that when the priors are non-informative, as is the case of uniformly distributed parameters, the MAP and the ML methods are equivalent.

Another dilemma faced when trying to estimate unknown parameters in the presence of nuisance parameters, is that of choosing between marginalization and joint estimation. In the first approach, the nuisance parameters

are being integrated out of the likelihood or the posterior density, to obtain a density that depends only on the parameters of interest. In the second approach, the nuisance parameters are jointly estimated. Intuitively, one would expect that marginalization has superior performance than joint estimation, since in the latter case more parameters need to be estimated with the same amount of data. However, to the best of our knowledge, there is no theorem proving this. Some interesting results on parameter estimation in the presence of nuisance parameters can be found in [46–48].

In our work, we have mostly chosen ML estimators. The reason for this is that we tried to avoid introducing unnecessary assumptions that would lead to informative priors and thus, using the principle of maximum entropy (ME) we would always come up with uniform priors. Moreover, in our work we have mostly used joint estimation. The reason for this is that integrating out parameters like eg. the scatterers' coordinates has been proved cumbersome. The exception to this, is the case of Bayesian linear models, where an unknown mean (usually due to unknown complex amplitudes) is Gaussian distributed and the received signal is also Gaussian distributed, due to white noise. In such cases the vector containing these nuisance parameters can be easily integrated out, leading to a new Gaussian likelihood.

1.8 ML Location Estimation for SBM-based Methods

In this section we present a general ML estimator for the 2^{nd} step of any 2-step SBM-based localization algorithm. It is based on available LDP estimates. Let $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^t, \ldots, \boldsymbol{\theta}_K^t]^t$ denote the $N_{\boldsymbol{\theta}} = KN_tN_s$ vector containing the true values of the $K = \{3, 4\}$ different kind of LDP and $\hat{\boldsymbol{\theta}}$ denote the vector containing their available estimates. For the DSBM case (dynamic channel), $\boldsymbol{\theta}$ could contain a subset of 3 or all of the LDP given below

$$\boldsymbol{\theta}_k = \begin{cases} \mathbf{d},\ k = 1 \\ \boldsymbol{\phi},\ k = 2 \\ \boldsymbol{\psi},\ k = 3 \\ \mathbf{f}_d,\ k = 4 \end{cases} \tag{1.71}$$

For the SBM case, $N_t = 1$, $K = 3$ and $\boldsymbol{\theta}$ is composed of just the first 3 of the above vectors. Assuming that the estimates of the LDP are not perfect, but contain an error $\tilde{\boldsymbol{\theta}}$ we can write

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \tilde{\boldsymbol{\theta}}. \tag{1.72}$$

Table 1.2: Parameters of interest $\mathbf{p}_{int}$

$\mathbf{p}_{int} = \mathbf{p}_{mt} = [x_{mt}, y_{mt}]$	SBM
$\mathbf{p}_{int} = [x_0, y_0, \upsilon_x, \upsilon_y]$	DSBM with constant speed
$\mathbf{p}_{int} = [x_0, y_0, \upsilon_{x,0}, \upsilon_{y,0}, \alpha_x, \alpha_y]$	DSBM with constant acceleration

We further assume that $\hat{\boldsymbol{\theta}}$ is an unbiased estimator that possesses the asymptotic normality property and that a sufficiently large number of samples of the received signal was used in estimating $\boldsymbol{\theta}$. Due to these assumptions, $\tilde{\boldsymbol{\theta}} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\tilde{\boldsymbol{\theta}}})$ and thus $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \mathbf{C}_{\tilde{\boldsymbol{\theta}}})$.

The main objective of any localization method is to estimate the MT position. In dynamic scenarios it might also be desirable to estimate its speed. However, in section 1.4.2, we showed that for a NLoS environment that can be described by the SBM, the LDP depend not only on the MT coordinates and speed but also on the scatterers' coordinates. Since there is usually little interest in knowing the scatterers' position, we will treat their coordinates as nuisance parameters and denote the vector that contains them as $\mathbf{p}_{nui} = \mathbf{p}_s = [\mathbf{x}_s^t, \mathbf{y}_s^t]^t$. On the other hand we will denote the parameters of interest as $\mathbf{p}_{int}$. $\mathbf{p}_{int}$ can vary, depending on the scenario and the mobility model, as shown in table 1.2. Introducing the $N_{\mathbf{p}} \times 1$ vector of all the unknown parameters, $\mathbf{p} = [\mathbf{p}_{int}^t, \mathbf{p}_{nui}^t]^t$, we can rewrite the set of equations (1.26)-(1.28), (1.38)-(1.40) and (1.41)-(1.43) in a more compact way $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathbf{p})$ to show the dependence of the mean of the LDP estimates on the unknown parameters. Apart from the mean, it is likely that the covariance matrix $\mathbf{C}_{\tilde{\boldsymbol{\theta}}}$ also depends on the unknown parameters, i.e. it is likely that the accuracy of the method used to estimate the LDP from the received signal samples depends on the geometry of the environment. However, since we consider a broad class of estimators and not a specific one (thus we don't introduce any expression for the covariance matrix), we will not consider any such dependency. The pdf of $\hat{\boldsymbol{\theta}}$ conditioned on $\mathbf{p}$ is given by:

$$f(\hat{\boldsymbol{\theta}}|\mathbf{p}) = \frac{1}{(2\pi)^{\frac{1}{2}N_{\boldsymbol{\theta}}}(\det\mathbf{C}_{\tilde{\boldsymbol{\theta}}})^{1/2}} e^{-\frac{1}{2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^t \mathbf{C}_{\tilde{\boldsymbol{\theta}}}^{-1}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})} \tag{1.73}$$

To obtain a ML estimate of our parameters of interest, we need to maximize $f(\hat{\boldsymbol{\theta}}|\mathbf{p})$ -or equivalently a corresponding likelihood- with respect to $\mathbf{p}$. Define a log-likelihood obtained by taking the natural logarithm of $f(\hat{\boldsymbol{\theta}}|\mathbf{p})$ and

ignoring the constant terms as:

$$\mathcal{L} = \mathcal{L}(\boldsymbol{\theta}(\mathbf{p})) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t \mathbf{C}_{\tilde{\boldsymbol{\theta}}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \tag{1.74}$$

Maximizing $f(\hat{\boldsymbol{\theta}}|\mathbf{p})$ is equivalent to minimizing $\mathcal{L}$, therefore the ML estimate of $\mathbf{p}$ is given by

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}}\{\mathcal{L}\} \tag{1.75}$$

1.9 Cramer-Rao Bound

According to the Cramer-Rao Bound (CRB) for an unbiased estimator $\widehat{\mathbf{p}}$ of $\mathbf{p}$, the correlation matrix of the parameter estimation errors $\widetilde{\mathbf{p}}$ is bounded below by the inverse of the Fisher Information Matrix (FIM) $\mathbf{J}$ as shown below

$$R_{\widetilde{\mathbf{p}}\widetilde{\mathbf{p}}} = E\{(\widehat{\mathbf{p}} - \mathbf{p})(\widehat{\mathbf{p}} - \mathbf{p})^t\} \geq \mathbf{J}^{-1} \tag{1.76}$$

where the FIM is given by:

$$\boldsymbol{J} = E\Big\{ \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right) \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right)^t \Big\} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}} \mathbf{C}_{\tilde{\boldsymbol{\theta}}}^{-1} \frac{\partial \boldsymbol{\theta}}{\partial \mathbf{p}^t} = \mathbf{G} \mathbf{C}_{\tilde{\boldsymbol{\theta}}}^{-1} \mathbf{G}^t \tag{1.77}$$

$\mathcal{L}$ is the log-likelihood given by (1.74) and we have introduced the transformation matrix $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$. It is well known and can be easily shown, that the information contained in independent data (in our case LDP) can be summed up, so that the FIM for a hybrid localization method is given by:

$$\boldsymbol{J} = \sum_k \boldsymbol{J}_{\boldsymbol{\theta}_k} = \sum_k \frac{\partial \boldsymbol{\theta}_k^t}{\partial \mathbf{p}} \mathbf{C}_{\tilde{\boldsymbol{\theta}}_k}^{-1} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}^t} = \sum_k \frac{1}{\sigma_k^2} \frac{\partial \boldsymbol{\theta}_k^t}{\partial \mathbf{p}} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}^t} \tag{1.78}$$

where the 3rd equality holds only if the entries of $\boldsymbol{\theta}_k$, $\forall k$, are i.i.d. with variance σ_k^2. If $\mathbf{p} = [\mathbf{p}_{int}^t, \mathbf{p}_{nui}^t]^t$, i.e. if we are estimating the parameters of interest in the presence of nuisance parameters, which is the case for SBM-based methods, we can substitute $\frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}^t} = [\frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}_{int}^t} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}_{nui}^t}]$ in (1.78) to get

$$\mathbf{J} = \begin{bmatrix} \sum_k \frac{1}{\sigma_{\boldsymbol{\theta}_k}^2} \frac{\partial \boldsymbol{\theta}_k^t}{\partial \mathbf{p}_{int}} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}_{int}^t} & \sum_k \frac{1}{\sigma_{\boldsymbol{\theta}_k}^2} \frac{\partial \boldsymbol{\theta}_k^t}{\partial \mathbf{p}_{int}} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}_{nui}^t} \\ \sum_k \frac{1}{\sigma_{\boldsymbol{\theta}_k}^2} \frac{\partial \boldsymbol{\theta}_k^t}{\partial \mathbf{p}_{nui}} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}_{int}^t} & \sum_k \frac{1}{\sigma_{\boldsymbol{\theta}_k}^2} \frac{\partial \boldsymbol{\theta}_k^t}{\partial \mathbf{p}_{nui}} \frac{\partial \boldsymbol{\theta}_k}{\partial \mathbf{p}_{nui}^t} \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{bmatrix} \tag{1.79}$$

Based on the expressions for the FIM presented above, we can examine the identifiability, i.e. the ability to estimate all unknown parameters and the asymptotic performance of efficient estimators, as was demonstrated in [49].

1.10 Identifiability Concerns for Location Estimation

The entries of the FIM $\mathbf{J}$ are continuous functions of $\mathbf{p}$ everywhere in $\mathcal{R}^{N_\mathbf{p}}$. A point $\mathbf{p}^0$ is said to be a *regular point* of the matrix $\mathbf{J}$ if there exists an open neighborhood of $\mathbf{p}^0$ in which $\mathbf{J}$ has constant rank. Using this definition and making some weak assumptions, Rothenberg proved the following theorem in [50]:

Theorem 1. *Let $\mathbf{p}^0$ be a regular point of the FIM $\mathbf{J} = \mathbf{J}(\mathbf{p})$. Then $\mathbf{p}^0$ is locally identifiable if and only if $\mathbf{J}(\mathbf{p}^0)$ is non-singular.*

Assuming that a vector $\mathbf{p}$ containing the true values of the unknown parameters is a regular point, which in general is true, the above theorem tells us that the unknown parameters become identifiable when FIM evaluated at the true values is nonsingular. The following corollary is an immediate consequence of the above theorem and eq. (1.77):

Corollary 1. *In SBM-based localization methods, local identifiability of the parameter vector $\mathbf{p}$ can be achieved when the transformation matrix $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$ is square or wide ($N_{\boldsymbol{\theta}} \geq N_\mathbf{p}$) and has full rank $N_\mathbf{p}$.*

1.11 Performance Concerns for Location Estimation

In this section we study the impact on the performance of any localization method for cases in which exploiting new LDP comes at the cost of jointly estimating a new set of nuisance parameters. This happens when considering a dynamic rather than a static environment and thus the speed of the MT needs to be jointly estimated as mentioned above. This can also happen when the set of unexploited LDP depends deterministically on the entries of $\mathbf{p}$ but also on an unknown error term. For example there might be an unknown synchronization offset that needs to be taken into account for the delays (and thus for the path lengths) or an orientation/calibration offset that needs to be taken into account for the AoA and/or the AoD. The following theorem applies to all of the above cases and specifies the cases when the location ML estimation will be more accurate.

Theorem 2. *Introducing and exploiting new LDP $\boldsymbol{\theta}_2$ (data) that depend on the entries of the $N_{\mathbf{p}_1} \times 1$ parameter vector $\mathbf{p}_1$ that needs to be estimated due to the problem formulation (which might consist of parameters of interest*

and possibly some nuisance parameters as well) but also on the entries of new vector of nuisance parameters $\mathbf{p}_2$*, will lead to an enhancement of the (asymptotic) performance of the ML estimation only if the transformation matrix* $\mathbf{G}_{22} = \frac{\partial \boldsymbol{\theta}_2^t}{\partial \mathbf{p}_2}$ *is wide (*$N_{\boldsymbol{\theta}_2} > N_{\mathbf{p}_2}$*) and has full rank* $N_{\mathbf{p}_2}$*.*

Proof. Let $\boldsymbol{\theta}_1$ be the $N_{\boldsymbol{\theta}_1} \times 1$ vector containing the data that are already used in the estimation process and define $\boldsymbol{\theta} \triangleq [\boldsymbol{\theta}_1^t \boldsymbol{\theta}_2^t]^t$, $\mathbf{p} \triangleq [\mathbf{p}_1^t \mathbf{p}_2^t]^t$, $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$ and $\mathbf{G}_{ij} = \frac{\partial \boldsymbol{\theta}_i^t}{\partial \mathbf{p}_j}$. The FIM for the new problem is given by :

$$\begin{aligned}
\boldsymbol{J}_{new} &= E\Big\{ \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right) \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right)^t \Big\} = \mathbf{G}\mathbf{C}^{-1}\mathbf{G}^t \\
&= \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{21} \\ \mathbf{0} & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{C}_{\boldsymbol{\theta}_1|\mathbf{p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{G}_{11}^t & \mathbf{0} \\ \mathbf{G}_{21}^t & \mathbf{G}_{22}^t \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{G}_{11}\mathbf{C}_{\boldsymbol{\theta}_1|\mathbf{p}}^{-1}\mathbf{G}_{11}^t + \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t & \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t \\ \mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t & \mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t \end{bmatrix} \\
&\triangleq \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}
\end{aligned} \tag{1.80}$$

The first term of the sum composing $\boldsymbol{A}$ can be recognized as the FIM $\boldsymbol{J}_1$ of the original estimation problem while $\boldsymbol{D}$ is the FIM $\boldsymbol{J}_2$ for the estimation problem of $\mathbf{p}_2$.

If both of the conditions for $\boldsymbol{G}_{22}$ are met, we can use the inversion formula for 2×2 block matrices to obtain the $N_{\mathbf{p}_1} \times N_{\mathbf{p}_1}$ upper left submatrix of the inverse of $\boldsymbol{J}_{new}$ as follows:

$$\begin{aligned}
&[\boldsymbol{J}_{new}^{-1}]_{(1:L,1:L)} = (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} \\
&= (\ \boldsymbol{J}_1 + \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t - \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t\boldsymbol{J}_2^{-1}\mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t\ )^{-1}
\end{aligned} \tag{1.81}$$

To show that the performance is improved with the addition of new data, it suffices to show that $[\boldsymbol{J}_{new}]_{(1:L,1:L)} > \boldsymbol{J}_1$. This is true because the sum of the other two matrices on the r.h.s. of (1.81) results in a positive semidefinite matrix. This can be proved by defining $\mathbf{a} \triangleq \mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1/2}\mathbf{G}_{22}^t\mathbf{u}$ and $\mathbf{b} \triangleq \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t(\mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t)^{-1/2}\mathbf{u}$, where $\mathbf{u}$ is any non-zero vector and applying Cauchy-Schwartz inequality $||\mathbf{a}||^2||\mathbf{b}||^2 \geq (\mathbf{a}^t\mathbf{b})^2$, to get:

$$\mathbf{u}^t\ (\ \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t - \mathbf{G}_{21}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{22}^t\boldsymbol{J}_2^{-1}\mathbf{G}_{22}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}^{-1}\mathbf{G}_{21}^t\ )\ \mathbf{u} \geq 0 \tag{1.82}$$

If $N_{\boldsymbol{\theta}_2} = N_{\mathbf{p}_2}$ and $\boldsymbol{G}_2$ has full rank, (1.81) reduces to :

$$[\boldsymbol{J}_{new}^{-1}]_{(1:L,1:L)} = \boldsymbol{J}_1^{-1} \tag{1.83}$$

thus the performance of the ML estimation is exactly the same, while the complexity of the method increases.

If $N_{\boldsymbol{\theta}_2} < N_{\mathbf{p}_2}$ and $\boldsymbol{G}_2$ is full rank, $\boldsymbol{D}$ is singular and thus non-invertible. The parameter vector $\mathbf{p}_2$ is not identifiable, however $\mathbf{p}_1$ is still identifiable and $[\boldsymbol{J}_{new}^{-1}]_{(1:L,1:L)}$ can be derived by replacing the inverse of $\boldsymbol{D}^{-1}$ with its pseudo-inverse $\boldsymbol{D}^{+} = \mathbf{G}_{22}(\mathbf{G}_{22}^t\mathbf{G}_{22})^{-1}\mathbf{C}_{\boldsymbol{\theta}_2|\mathbf{p}}(\mathbf{G}_{22}^t\mathbf{G}_{22})^{-1}\mathbf{G}_{22}^t$ in (1.81). This results again in (1.83) and thus in no improvement in performance.

Finally, if $\boldsymbol{G}_{22}$ has rank $k < \min\{N_{\boldsymbol{\theta}_2}, N_{\mathbf{p}_2}\}$, $N_{\boldsymbol{\theta}_2} - k$ of its columns contain no additional information for $\mathbf{p}_2$. If they contain additional information only for $\mathbf{p}_1$ then, the information they contain for $\mathbf{p}_2$ (if any) could be removed by elementary column operations so that the corresponding entries of $\boldsymbol{\theta}_2$ can be included in $\boldsymbol{\theta}_1$, leading to a different partitioning of the FIM. That will improve performance, since new LDP that depend only on $\mathbf{p}_1$ and no new nuisance parameters, are exploited. If, however, they contain no additional information on any of the entries of $\mathbf{p}$ then they should be discarded. Either way, $\boldsymbol{G}_2$ becomes a $N_{\mathbf{p}_2} \times k$ full rank matrix and the last of the above cases applies. □

Theorem 2 can be applied in any ML localization approach. For example, proposition 1 in [35] can be derived using this theorem. The fact that the performance of a TOA method does not improve by including the TOA of the NLoS components, if these are modeled as the TOA corresponding to the LoS component plus an unknown error term, is a direct application of theorem 2 with $N_{\boldsymbol{\theta}_2} = N_{\mathbf{p}_2}$.

Chapter 2

Estimation of Location Dependent Parameters

2.1 Introduction

As already explained in the previous chapter, traditional geometrical localization techniques consist of two steps: First a set of location dependent parameters (LDP) are estimated in one or more BS. The location of the Mobile Terminal (MT) is then estimated by finding the values for the coordinates x_{mt} and y_{mt} that best fit the data (LDP estimates). This chapter contains a subspace-based method that can be utilized to simultaneously estimate different kinds of LDP for all the MPC. We consider a MT that communicates with a BS through a NLoS propagation environment. The MT moves and therefore, the channel impulse response (CIR) is affected by Doppler frequency shifts (DFS). We limit our study to a MIMO system and an OFDM signal. We parameterize the CIR matrix in such a way, that a 4-dimensional (4D) Unitary ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) algorithm can be utilized to jointly estimate 4 subsets of LDP, namely the angles of arrival (AoA), the angles of departure (AoD), the path lengths and the DFS.

ESPRIT algorithm was introduced in [25] as a computationally attractive estimation algorithm that exploits the rotational invariance of the signal subspace. Compared to MUSIC (Multiple Signal Classification), which is

another high resolution algorithm that was introduced in [24] and enjoyed wide acceptance, ESPRIT achieves a significant reduction in computational complexity by imposing a constraint on the structure of the antenna array. Due to this constraint, ESPRIT is able to output estimates directly in terms of generalized eigenvalues, without the need for a search over the array manifold. This will become more obvious below, where the steps of the 2 algorithms are given, exactly as they originally appeared in [24, 25] for the problem of AoA estimation.

MUSIC ALGORITHM

Suppose there are D incident waveforms received by an antenna array of M elements. The received $M \times 1$ vector is

$$x = Af + w \tag{2.1}$$

where f represents the $D \times 1$complex incident signals vector and w is the noise vector. The elements of A are known functions of the AoA and the array element locations. In fact the jth column of A is a "mode" vector of responses $\alpha(\theta_j)$ to the direction of arrival θ_j. The $M \times M$ covariance matrix of x is given by:

$$S = E\{xx^\dagger\} = AE\{ff^\dagger\}A^\dagger + E\{ww^\dagger\} = APA^\dagger + \lambda S_0 \tag{2.2}$$

under the assumption that incident signals and noise are uncorrelated. Skipping the analysis we directly give the algorithm in steps:

- Collect data and form the covariance matrix S
- Calculate eigenstructure of S in metric of S_0
- Estimate $D = M - N$ where N is the multiplicity of the minimum eigenvalue of S in the metric of S_0
- Evaluate
$$P_{MU}(\theta) = \frac{1}{\alpha^\dagger(\theta) E_N E_N^\dagger \alpha(\theta)} \tag{2.3}$$
as a function of θ , where E_N is the the $M \times N$ matrix whose columns are the N noise eigenvectors and $\alpha(\theta)$ is the continuum of the mode vectors.
- Pick D peaks of $P_{MU}(\theta)$ that give you the D AoA.

ESPRIT ALGORITHM

Conveniently describing the antenna array as being comprised of two identical subarrays Z_X and Z_Y physically displaced from each other by Δ, the output signal vector z can be written as

$$z = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} A \\ A\Phi \end{bmatrix} s + \begin{bmatrix} n_x \\ n_y \end{bmatrix} = \bar{A}s + n_z \tag{2.4}$$

where Φ is a diagonal $d \times d$ matrix of the phase delays between the doublet elements:

$$\Phi = diag\{[e^{j\gamma_1}, \ldots, e^{j\gamma_d}]\} \tag{2.5}$$

In the same fashion as before, we give the algorithm in steps:

- Collect data and compute the sample covariance matrix $R_z = \bar{A}R_s\bar{A}^\dagger + \sigma^2\Sigma_n$
- Compute the generalized eigendecomposition of R_z and Σ_n

$$R_z\bar{E} = \Sigma_n\bar{E}\Lambda \tag{2.6}$$

- Estimate the number of sources d using existing techniques
- Obtain the signal subspace $\mathcal{S}_z = \mathcal{R}\{E_S\} = \mathcal{R}\{\bar{A}\}$. Decompose E_S to get E_x and E_y
- Compute the eigendecomposition

$$E_{xy}^\dagger E_{xy} = \begin{bmatrix} E_x^\dagger \\ E_y^\dagger \end{bmatrix} \begin{bmatrix} E_x & E_y \end{bmatrix} = E\Lambda E^\dagger \tag{2.7}$$

  and partition E into $d \times d$ submatrices :

$$E = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} \tag{2.8}$$

- Calculate the eigenvalues of $\Psi = -E_{12}E_{22}^{-1}$ (TLS solution)

$$\phi_i = \lambda_i(-E_{12}E_{22}^{-1}),\ \forall\, k = 1, \ldots, d \tag{2.9}$$

- Estimate $\theta_i = f^{-1}(\phi_i)$ which for the case of AoA is

$$\theta_i = \sin^{-1}(c\arg(\phi_i)/(\omega_0\Delta)) \tag{2.10}$$

The ESPRIT algorithm has evolved significantly since it was first published. From the numerous publications that followed, we will restrict ourselves to the ones that inspired our method presented in the following sections. Specifically, in [51], Unitary ESPRIT was introduced and it was extended to the multidimensional case in [52]. Unitary ESPRIT improves accuracy by taking advantage of the unit magnitude property of the terms that represent the phase delays between the two identical subarrays. In [53] a thorough study and implementation of the ESPRIT algorithm in the joint estimation of 2 kinds of LDP, namely the AoA and the delays of MPC, was presented. A 2-D Unitary ESPRIT for MIMO systems was introduced in [54] for the joint estimation of the AoA and the AoD. The proposed 4D Unitary ESPRIT is an extension of the aforementioned work applicable to MIMO-OFDM systems that operate in dynamic environments.

2.2 Channel Model

The discrete input-output relationship of a $n_r \times n_t$ MIMO-OFDM system in the time-frequency domain is given in section 1.6 by eq. (1.66). For reference, we rewrite it below:

$$\mathbf{Y}_{kl} = \mathbf{H}_{kl}\mathbf{X}_{kl} + \mathbf{N}_{kl}. \tag{2.11}$$

The $n_r \times n_t$ channel matrix $\mathbf{H}_{kl}$ is given by

$$\begin{aligned}\mathbf{H}_{kl} &= \tfrac{1}{\sqrt{P_{tot}}} \textstyle\sum_{j=1}^{N_s} \sqrt{P_j}\gamma_j e^{j2\pi l\Delta t f_{d,j}} \mathbf{a}_R(\phi_j)\mathbf{a}_T^t(\psi_j) H_{TR,k} e^{-j2\pi k\Delta f \tau_j} \\ &= \mathbf{A}_R(\mathbf{\Gamma} \odot (\mathbf{D}_k\mathbf{F}_{d,l}))\mathbf{A}_T^t = \mathbf{A}_R\mathbf{\Gamma}\mathbf{D}_k\mathbf{F}_{d,l}\mathbf{A}_T^t \end{aligned} \tag{2.12}$$

and the rest of the parameters have been defined in sections 1.5.2 and 1.6. In what follows we will assume that $\mathbf{X}_{kl} = \mathbf{X}$, i.e. the training sequence transmitted over different subcarriers and at different time instants does not change. If this condition is not met, estimates of the channel matrix $\mathbf{H}_{kl}$ are required to serve as a starting point for the ESPRIT algorithm. The difference between eq. (1.54) and (2.12) is that in the latter the LDP are treated as constants, since their variation for a small observation time (of the order of $msec$) is negligible. We will further assume that both the Tx and the Rx are equipped with uniform linear arrays (ULA), so that their responses to the signal component at angle j are given respectively by

$$\mathbf{a}_R(\phi_j) = [1, e^{j2\pi\frac{f_c}{c}d_r\sin(\phi_j)}, \ldots, e^{j2\pi\frac{f_c}{c}d_r(n_r-1)\sin(\phi_j)}]^t \tag{2.13}$$

$$\mathbf{a}_T(\psi_j) = [1, e^{j2\pi\frac{f_c}{c}d_t\sin(\psi_j)}, \ldots, e^{j2\pi\frac{f_c}{c}d_t(n_t-1)\sin(\psi_j)}]^t \quad . \tag{2.14}$$

Any array that can be decomposed into 2 subarrays with identical elements separated by an arbitrary distance d can be considered instead.

2.3 Data Preprocessing

In order to implement the 4D Unitary ESPRIT algorithm and estimate the AoA, AoD, path lengths and DS directly from the received signal samples, we need to rewrite the input-output relationship in a form such that the channel matrix inherits a shift invariance property in all four dimensions. This is achieved by separating the set of data matrices and subsequently employing vectorization and concatenation operations. The separation can be done in an arbitrary way, but for ease of notation we will assume that consecutive samples are grouped together. The 4 necessary steps are described in detail below:

1. Split the set of the $N_f N_t$ received matrices (samples) into $L = L_f L_t$ disjoint subsets $\mathbb{S}_{l_f,l_t}$ of $|\mathbb{S}_{l_f,l_t}| \stackrel{\Delta}{=} M = M_f M_t$ samples each. If $N_f/L_f \notin \mathbb{Z}$ and/or $N_t/L_t \notin \mathbb{Z}$, the subsets $\mathbb{S}_{l_f,L_t}$, $1 \leq l_f \leq L_f$ and/or $\mathbb{S}_{L_f,l_t}$, $1 \leq l_t \leq L_t$ are padded with zero matrix entries. It follows that $M_f = \left\lceil \frac{N_f}{L_f} \right\rceil$ and $M_t = \left\lceil \frac{N_t}{L_t} \right\rceil$. This separation of the data is equivalent to a smoothing [52] that ensures that more than one measurement vector will be used in the final formulation.

2. Vectorize the received signal matrices to obtain received signal vectors $\mathbf{y}_{kl}$, with $N = n_r n_t$ elements each, as follows:

$$\mathbf{y}_{kl} = \mathbf{h}_{kl} + \mathbf{n}_{kl} \tag{2.15}$$

$$\mathbf{h}_{kl} \stackrel{\Delta}{=} (\mathbf{X}^t \mathbf{A}_T \boxtimes \mathbf{A}_R) \mathbf{\Gamma} \mathbf{F}_{d,1}^{(l_t-1)} \mathbf{D}_k \mathbf{f}_{d,m_t} \tag{2.16}$$

where $l_t = \left\lfloor \frac{l}{M_t} \right\rfloor$ and $m_t = l - \left\lfloor \frac{l}{M_t} \right\rfloor M_t$. For each of the subsets, i.e. for $1 \leq l_f \leq L_f$ and $1 \leq l_t \leq L_t$, do the following concatenations and vectorizations:

3a. Stack the column vectors $\mathbf{y}_{kl}$, $(l_t - 1)M_t + 1 \leq l \leq l_t M_t$ of the corresponding subset $\mathbb{S}_{L_f,l_t}$ to form the $n_r n_t \times M_t$ matrix $\bar{\mathbf{Y}}_{kl_t}$ and vectorize

again to get:

$$\bar{\mathbf{y}}_{kl_t} = \bar{\mathbf{h}}_{kl_t} + \bar{\mathbf{n}}_{kl_t} \tag{2.17}$$

$$\bar{\mathbf{h}}_{kl_t} \triangleq (\mathbf{F}^t_{d,1:M_t} \boxtimes \mathbf{X}^t\mathbf{A}_T \boxtimes \mathbf{A}_R)\bar{\mathbf{\Gamma}}_{l_f l_t}\mathbf{d}_{m_f} \tag{2.18}$$

$$\mathbf{F}_{d,1:M_t} \triangleq [\mathbf{f}_{d,1}, \ldots, \mathbf{f}_{d,M_t}] \tag{2.19}$$

$$\bar{\mathbf{\Gamma}}_{l_f l_t} \triangleq \mathbf{\Gamma}\mathbf{F}^{(l_t-1)}_{d,1}\mathbf{D}^{(l_f-1)}_1 \tag{2.20}$$

where $l_f = \left\lfloor \frac{k}{M_f} \right\rfloor$ and $m_f = l - \left\lfloor \frac{k}{M_f} \right\rfloor M_f$.

3b. In a similar way, stack the column vectors $\mathbf{y}_{kl_t}$, $(l_f - 1)M_f + 1 \leq k \leq l_f M_f$ to form $L = L_f L_t$ matrices $\bar{\bar{\mathbf{Y}}}_{l_f l_t}$ and vectorize again to get:

$$\bar{\bar{\mathbf{y}}}_{l_f l_t} = \bar{\bar{\mathbf{H}}}\bar{\gamma}_{l_f l_t} + \bar{\bar{\mathbf{n}}}_{l_f l_t} \tag{2.21}$$

$$\bar{\bar{\mathbf{H}}} \triangleq (\mathbf{D}^t_{1:M_f} \boxtimes \mathbf{F}^t_{d,1:M_t} \boxtimes \mathbf{X}^t\mathbf{A}_T \boxtimes \mathbf{A}_R) \tag{2.22}$$

$$\mathbf{D}_{1:M_f} \triangleq [\mathbf{d}_{(i-1)M_f+1}, \ldots, \mathbf{d}_{iM_f}] \tag{2.23}$$

$$\bar{\gamma}_{l_f l_t} \triangleq (\mathbf{D}^{(l_f-1)}_1 \odot \mathbf{F}^{(l_t-1)}_{d,1})\gamma \tag{2.24}$$

4. Stack all generated vectors in a big matrix of size $MN \times L$

$$\bar{\bar{\mathbf{Y}}} = [\bar{\bar{\mathbf{y}}}_{\mathbf{11}}, \ldots, \bar{\bar{\mathbf{y}}}_{\mathbf{L_f L_t}}] = \bar{\bar{\mathbf{H}}}\bar{\mathbf{\Gamma}} + \bar{\bar{\mathbf{N}}} \tag{2.25}$$

where

$$\bar{\mathbf{\Gamma}} = [\bar{\gamma}_{11}, \ldots, \bar{\gamma}_{L_f L_t}] \tag{2.26}$$

$$\bar{\bar{\mathbf{N}}} = [\bar{\bar{\mathbf{n}}}_{11}, \ldots, \bar{\bar{\mathbf{n}}}_{L_f L_t}]. \tag{2.27}$$

It should be noted that this type of formulation is possible due to the fact that $\mathbf{\Gamma}$, $\mathbf{D}_k$ and $\mathbf{F}_{d,l}$ are diagonal, which in turn is a consequence of the fact that each path is distinct. Also it becomes obvious that the matrix $\bar{\mathbf{\Gamma}}$ and not the original transmitted matrix $\mathbf{X}$ plays the role of the unknown signal in our model. An alternative approach to the one presented above would be to stack all $N_f N_t$ received matrices and create a $N_f N_t nr \times n_t$ matrix $\mathcal{H}$ and subsequently follow the approach in [53] and construct a Hankel matrix by stacking shifted versions of $\mathcal{H}$.

2.4 4D ESPRIT

Due to its structure (Khatri-Rao Product of 4 matrices), the newly constructed $\bar{\bar{\mathbf{H}}}$ satisfies the following invariance properties in the $r = 4$ dimensions

$$\mathbf{J}_{N_r}\bar{\bar{\mathbf{H}}}\mathbf{Z}_r = \mathbf{J}'_{N_r}\bar{\bar{\mathbf{H}}},\ 1 \leq r \leq 4 \tag{2.28}$$

where $N_r \in \{n_r, n_t, M_t, M_f\}$ denotes the number of samples in dimension r and

$$\mathbf{Z}_r = \begin{cases} diag\{e^{j2\pi \frac{f_c}{c} d_r \sin(\phi_i)}\}_{i=1}^{N_s}, \ r = 1 \\ diag\{e^{j2\pi \frac{f_c}{c} d_t \sin(\psi_i)}\}_{i=1}^{N_s}, \ r = 2 \\ diag\{e^{-j2\pi \Delta f \tau_i}\}_{i=1}^{N_s}, \ r = 3 \\ diag\{e^{j2\pi \Delta t f_{d,i}}\}_{i=1}^{N_s}, \ r = 4 \end{cases} \tag{2.29}$$

In every dimension r, the pairs of selection matrices satisfy the equation

$$\mathbf{J}_{N_r} = \mathbf{\Pi}_{\frac{M(N_r-1)}{N_r}} \mathbf{J}'_{N_r} \mathbf{\Pi}_{MN} \tag{2.30}$$

where $\mathbf{\Pi}$ is a symmetric permutation matrix (Givens reflection). It therefore suffices to define one of the two in each pair. For the 4 dimensions $\mathbf{J}_{N_r}$ is given by

$$\mathbf{J}_{n_r} = \mathbf{I}_{M_f} \otimes \mathbf{I}_{M_t} \otimes \mathbf{I}_{n_t} \otimes [\mathbf{0}_{(n_r-1)\times 1} \mathbf{I}_{n_r-1}] \tag{2.31}$$

$$\mathbf{J}_{n_t} = \mathbf{I}_{M_f} \otimes \mathbf{I}_{M_t} \otimes [\mathbf{0}_{(n_t-1)\times 1} \mathbf{I}_{n_t-1}] (\mathbf{X}^t)^+ \otimes \mathbf{I}_{n_r} \tag{2.32}$$

$$\mathbf{J}_{M_t} = \mathbf{I}_{M_f} \otimes [\mathbf{0}_{(M_t-1)\times 1} \mathbf{I}_{M_t-1}] \otimes \mathbf{I}_{n_t} \otimes \mathbf{I}_{n_r} \tag{2.33}$$

$$\mathbf{J}_{M_f} = [\mathbf{0}_{(M_f-1)\times 1} \mathbf{I}_{M_f-1}] \otimes \mathbf{I}_{M_t} \otimes \mathbf{I}_{n_t} \otimes \mathbf{I}_{n_r}. \tag{2.34}$$

These pairs of selection matrices correspond to choosing 2 subarrays (out of all possible choices for ULA) with maximum overlap. Extending that to the time and frequency sets of samples, these pairs result in 2 subsets of samples with maximum overlap. Other selections that would exploit the multiple invariances of the ULA and of the sets of samples could be considered instead. Multiple Invariance ESPRIT was studied in [55] but is beyond the scope of this work.

Following the guidelines for Unitary ESPRIT, lets define

$$\bar{\bar{\mathbf{H}}}' = \mathbf{Q}_{MN}^{\dagger} \bar{\bar{\mathbf{H}}} \tag{2.35}$$

where $\mathbf{Q}$ denotes a left $\mathbf{\Pi}$-real matrix (see App. C). Using the fact that $\mathbf{Q}$ is unitary, eq. (2.28) becomes

$$\begin{aligned} \mathbf{J}_{N_r} \mathbf{Q}_{MN} \mathbf{Q}_{MN}^{\dagger} \bar{\bar{\mathbf{H}}} \mathbf{Z}_r = \mathbf{J}'_{N_r} \mathbf{Q}_{MN} \mathbf{Q}_{MN}^{\dagger} \bar{\bar{\mathbf{H}}} \quad &\Leftrightarrow \\ \mathbf{J}_{N_r} \mathbf{Q}_{MN} \bar{\bar{\mathbf{H}}}' \mathbf{Z}_r = \mathbf{J}'_{N_r} \mathbf{Q}_{MN} \bar{\bar{\mathbf{H}}}' \quad &\Leftrightarrow \\ \mathbf{Q}_{MN}^{\dagger} \mathbf{J}_{N_r} \mathbf{Q}_{MN} \bar{\bar{\mathbf{H}}}' \mathbf{Z}_r = \mathbf{Q}_{MN}^{\dagger} \mathbf{J}'_{N_r} \mathbf{Q}_{MN} \bar{\bar{\mathbf{H}}}'. \end{aligned} \tag{2.36}$$

Subsequently using the relations $\mathbf{\Pi}^2 = \mathbf{I}$, $\mathbf{\Pi Q} = \mathbf{Q}^*$ and eq. (2.30) we get

$$\begin{aligned}\mathbf{Q}_{MN}^{\dagger}\mathbf{J}'_{N_r}\mathbf{Q}_{MN} &= \mathbf{Q}_{MN}^{\dagger}\mathbf{\Pi}^2_{\frac{M(N_r-1)}{N_r}}\mathbf{J}'_{N_r}\mathbf{\Pi}^2_{MN}\mathbf{Q}_{MN} \\ &= \mathbf{Q}_{MN}^{\dagger}\mathbf{\Pi}_{\frac{M(N_r-1)}{N_r}}\mathbf{J}_{N_r}\mathbf{\Pi}_{MN}\mathbf{Q}_{MN} \\ &= \mathbf{Q}_{MN}^{t}\mathbf{J}_{N_r}\mathbf{Q}_{MN}^{*} = (\mathbf{Q}_{MN}^{\dagger}\mathbf{J}_{N_r}\mathbf{Q}_{MN})^{*}. \quad (2.37)\end{aligned}$$

Therefore if we introduce

$$\mathbf{K}_{N_r,1} = Re\{\mathbf{Q}_{N_r}^{\dagger}\mathbf{J}'_{N_r}\mathbf{Q}_L\} \quad (2.38)$$

$$\mathbf{K}_{N_r,2} = Im\{\mathbf{Q}_{N_r}^{\dagger}\mathbf{J}'_{N_r}\mathbf{Q}_L\}. \quad (2.39)$$

we can write

$$\mathbf{Q}_{MN}^{\dagger}\mathbf{J}_{N_r}\mathbf{Q}_{MN} = \mathbf{K}_{N_r,1} - j\mathbf{K}_{N_r,2} \quad (2.40)$$

$$\mathbf{Q}_{MN}^{\dagger}\mathbf{J}'_{N_r}\mathbf{Q}_{MN} = \mathbf{K}_{N_r,1} + j\mathbf{K}_{N_r,2} \quad (2.41)$$

Substituting these in (2.36) we get the invariance equations for $\bar{\bar{\mathbf{H}}}'$

$$\begin{aligned}(\mathbf{K}_{N_r,1} - j\mathbf{K}_{N_r,2})\bar{\bar{\mathbf{H}}}'\mathbf{Z}_r = (\mathbf{K}_{N_r,1} + j\mathbf{K}_{N_r,2})\bar{\bar{\mathbf{H}}}' &\Leftrightarrow \\ \mathbf{K}_{N_r,1}\bar{\bar{\mathbf{H}}}'(\mathbf{Z}_r - \mathbf{I}) = j\mathbf{K}_{N_r,2}\bar{\bar{\mathbf{H}}}'(\mathbf{Z}_r + \mathbf{I}) &\Leftrightarrow \\ \mathbf{K}_{N_r,1}\bar{\bar{\mathbf{H}}}'\mathbf{\Omega}_r = \mathbf{K}_{N_r,2}\bar{\bar{\mathbf{H}}}' & \quad (2.42)\end{aligned}$$

where

$$\mathbf{\Omega}_r = j(\mathbf{I} - \mathbf{Z}_r)(\mathbf{Z}_r + \mathbf{I})^{-1} = \begin{cases} diag\{\tan(\pi\frac{f_c}{c}d_r\sin(\phi_i))\}_{i=1}^{N_s},\ r=1 \\ diag\{\tan(\pi\frac{f_c}{c}d_t\sin(\psi_i))\}_{i=1}^{N_s},\ r=2 \\ diag\{\tan(-\pi\Delta f\tau_i)\}_{i=1}^{N_s},\ r=3 \\ diag\{\tan(\pi\Delta t f_{d,i})\}_{i=1}^{N_s},\ r=4 \end{cases} \quad (2.43)$$

Observe that in eq. (2.42) all matrices except $\bar{\bar{\mathbf{H}}}'$ are real so we can proceed with the steps for Unitary ESPRIT. A simple, yet efficiency way to increase the number of the data used in the estimation process, while simultaneously decreasing the computational cost, is to transform $\bar{\bar{\mathbf{Y}}}$ into a centro-Hermitian matrix and subsequently into a real matrix, to get:

$$\bar{\bar{\mathbf{Y}}}_{\mathbf{r}} = \mathbf{Q}_{\mathbf{MN}}^{\dagger}[\bar{\bar{\mathbf{Y}}}\ \mathbf{\Pi}_{\mathbf{MN}}\bar{\bar{\mathbf{Y}}}^{*}\mathbf{\Pi}_{\mathbf{L}}]\mathbf{Q}_{\mathbf{2L}}. \quad (2.44)$$

For a detailed description of this operation, which is essentially a forward-backward averaging and a mapping to $\mathbb{R}^{MN\times 2L}$, please see section Appendix

C. Due to noise, $\bar{\bar{\mathbf{Y}}}_{\mathbf{r}}$ is full rank, instead of rank N_s. This affects the solution of any ESPRIT algorithm, by resulting in a number of estimates that is greater than the number of parameters that need to be estimated. To mitigate this, a rank reduction can be performed, using for example the SVD decomposition of $\bar{\bar{\mathbf{Y}}}_{\mathbf{r}}$. From the SVD, the N_s dominant left singular vectors, composing $\mathbf{E}_{N_s}$, can be derived. These vectors span the signal subspace, i.e. in the noiseless case $\mathbf{E}_{N_s}$ and $\bar{\bar{\mathbf{H}}}'$ span the same N_s-dimensional space. We can and thus write $\bar{\bar{\mathbf{H}}}' = \mathbf{E}_{N_s}\mathbf{U}$ for a non-singular square matrix $\mathbf{U}$. Substituting in (2.42) we get

$$\mathbf{K}_{N_r,1}\mathbf{E}_{N_s}\mathbf{\Theta}_r = \mathbf{K}_{N_r,2}\mathbf{E}_{N_s} \tag{2.45}$$

where

$$\mathbf{\Theta}_r = \mathbf{U}\mathbf{\Omega}_r\mathbf{U}^{-1} \tag{2.46}$$

The invariance equations (2.45) can be solved by means of Least-Squares (LS) or any of its more advanced variants, to obtain the matrices $\mathbf{\Theta}_r$. The LS solution yields

$$\mathbf{\Theta}_r = ((\mathbf{K}_{N_r,1}\mathbf{E}_{N_s})^t\mathbf{K}_{N_r,1}\mathbf{E}_{N_s})^{-1}(\mathbf{K}_{N_r,1}\mathbf{E}_{N_s})^t\mathbf{K}_{N_r,2}\mathbf{E}_{N_s} \tag{2.47}$$

while the Total Least-Squares (TLS) solution yields

$$\mathbf{\Theta}_r = -\mathbf{V}_{(1:N_s,N_s:2N_s)}\mathbf{V}^{-1}_{(N_s:2N_s,N_s:2N_s)} \tag{2.48}$$

where $\mathbf{V}$ is obtained from the SVD of $[\mathbf{K}_{N_r,1}\mathbf{E}_{N_s}\ \ \mathbf{K}_{N_r,2}\mathbf{E}_{N_s}]$, i.e.

$$[\mathbf{K}_{N_r,1}\mathbf{E}_{N_s}\ \ \mathbf{K}_{N_r,2}\mathbf{E}_{N_s}] = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\dagger}. \tag{2.49}$$

The final step of the algorithm is the eigendecomposition of each of the $\mathbf{\Theta}_r$ to obtain their N_s eigenvalues which are equal to the diagonal entries of the corresponding $\mathbf{\Omega}_r$. To avoid the need for pairing the LDP estimates, joint diagonalization or triangularization is highly recommended, since it can achieve automatic pairing. This is a direct consequence of the fact that the four $\mathbf{\Theta}_r$ share the same set of eigenvectors in the absence of noise. In the presence of noise, these sets are approximately the same. Following the work in [52], Simultaneous Schur Decomposition (SSD) will be utilized to compute the eigenvalues of the four $\mathbf{\Theta}_r$.

The SSD is an iterative procedure that tries to derive approximate upper triangular matrices simultaneously. Each iteration has $\frac{1}{2}N_s(N_s-1)$ steps. It starts with $\mathbf{R}_{r,0} = \mathbf{\Theta}_r$, $r = 1,\ldots,4$ and in each step j the matrices $\mathbf{R}_{r,j}$, are updated as follows

$$\mathbf{R}_{r,j} = \mathbf{B}^t_{i_1 i_2}\mathbf{R}_{r,j-1}\mathbf{B}_{i_1 i_2} \tag{2.50}$$

where the elementary Jacobi rotation matrix $\mathbf{B}_{i_1 i_2}(\alpha)$ is chosen to minimize the following cost function

$$e(\mathbf{B}_{i_1 i_2}) = \sum_{r=1}^{4} ||\mathcal{L}(\mathbf{R}_{r,j})||^2 = tr\{\sum_{r=1}^{4} \mathcal{L}(\mathbf{R}_{r,j})\mathcal{L}(\mathbf{R}_{r,j})^t\}. \tag{2.51}$$

$\mathcal{L}$ denotes the strictly lower triangular part of a matrix. These Jacobi matrices are constructed from identity matrices, by replacing four of their entries as follows

$$\beta_{i_1 i_2} = -\beta_{i_1 i_2} = \sin(\alpha) \tag{2.52}$$
$$\beta_{i_1 i_1} = \beta_{i_2 i_2} = \cos(\alpha). \tag{2.53}$$

It is obvious that finding the matrix $\mathbf{B}_{i_1 i_2}(\alpha)$ that minimizes $e(\mathbf{B}_{i_1 i_2})$ in each step is equivalent to finding α. The solution to this problem was given in [52]. After the $\frac{1}{2}N_s(N_s-1)$ steps have been completed, the final matrices $\mathbf{R}_{r,\frac{1}{2}N_s(N_s-1)}$ serve as starting points $\mathbf{R}_{r,0}$ and the operation can be repeated to yield matrices that are even closer to upper triangular. After just a few iterations, the algorithm outputs $\mathbf{I}$ as the Jacobi matrix in each step and thus the cost function can not be minimized further.

As already mentioned above, the final approximate upper triangular matrices produced, contain N_s eigenvalues of the corresponding $\mathbf{\Theta}_r$ on their main diagonals and thus the LDP estimates can be easily computed using eq. (2.43). Fig. 2.1 depicts a block diagram of the 4D ESPRIT algorithm described in this section along with the data preprocessing described in the previous section.

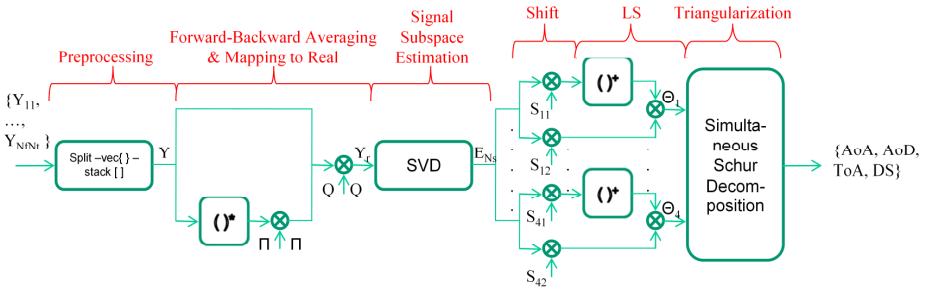

Figure 2.1: 4D ESPRIT block diagram

Table 2.1: MT and scatterers' coordinates

(x_{mt}, y_{mt})	(x_{s_1}, y_{s_1})	(x_{s_2}, y_{s_2})	(x_{s_3}, y_{s_3})	(x_{s_4}, y_{s_4})
$(30, 20)$m	$(20, 30)$m	$(35, 20)$m	$(40, 15)$m	$(15, 25)$m

2.5 Numerical Example

In this section we evaluate the performance of the proposed method in terms of the RMSE of the LDP estimates for a 2×4 MIMO system equipped with ULA on both sides. The transmitted signal propagates through $N_s = 4$ distinct NLoS paths. The coordinates of the corresponding four scatterers along with the coordinates of the MT are given in table 2.1. The BS is assumed to be placed at the origin. The magnitude of the speed of the MT is $|\upsilon| = 1.5$m/sec (average walking speed) and the direction is $-\frac{\pi}{3}$. $N_t = 40$ time samples with $\Delta t = 1$msec and $N_f = 8$ frequency samples (subcarriers) with $\Delta f = 10$MHz are considered. The impact of the choice of the data smoothing numbers, L_t and L_f, on the estimates of the different LDP was studied. The results indicated that there is a trade-off between performance for different subsets of LDP, rather than some optimal smoothing numbers that minimize RMSE for all of them. The only restriction on the smoothing numbers is that their product must satisfy $L \geq N_s$, so that the LDP become identifiable. The results shown in the figures correspond to $L_f = 2$ and $L_t = 8$. On a similar basis, for identifiability purposes, the largest of the four dimensions N_r must satisfy

$$\prod_{r=1}^{4} N_r \frac{N_{r,max} - 1}{N_{r,max}} \geq N_s. \tag{2.54}$$

The carrier frequency is $f_c = 1.9GHz$ and the transmitted pilot signal is the training matrix $\mathbf{X} = \mathbf{I}$. We run $N = 50$ independent trials and averaged the results, thus the RMSE is

$$RMSE(\mu_{ir}) = \sqrt{\frac{1}{N} \sum_{n=1}^{N} |\hat{\mu}_{ir} - \mu_{ir}|^2} \tag{2.55}$$

where the terms μ_{ir} depend on the LDP according to

$$\mu_{ir} = \frac{2}{\pi} \arctan(\omega_{ir}),\ 1 \leq i \leq N_s,\ 1 \leq r \leq 4. \tag{2.56}$$

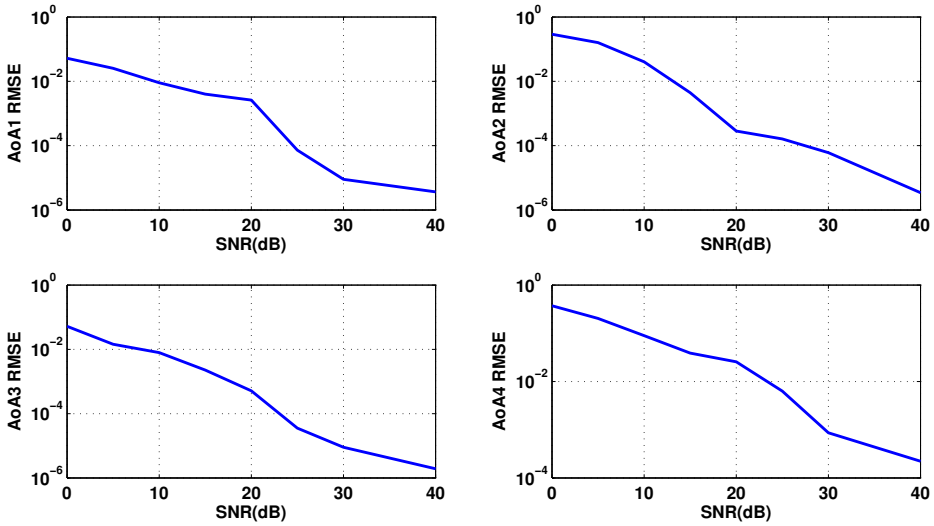

Figure 2.2: RMSE of sine of AoA $\sin(\phi)$

In the figures we plot the RMSE versus the Signal-to-Noise Ratio (SNR) at the receiver, which is given by

$$SNR = 10\log_{10}\left(\frac{E\{tr((\bar{\bar{\mathbf{H}}}\bar{\mathbf{\Gamma}})(\bar{\bar{\mathbf{H}}}\bar{\mathbf{\Gamma}})^{\dagger})\}}{E\{tr(\bar{\bar{\mathbf{N}}}\bar{\bar{\mathbf{N}}}^{\dagger})\}}\right) \tag{2.57}$$

Results show that the RMSE of the estimates are very small even for small to medium SNR ($5-10$dB). Its excellent performance along with its low computational cost make this algorithm an attractive solution to LDP estimation problems and motivates us to implement the 2^{nd} step of localization, which will be presented in the following 2 chapters, assuming the existence of accurate LDP estimates.

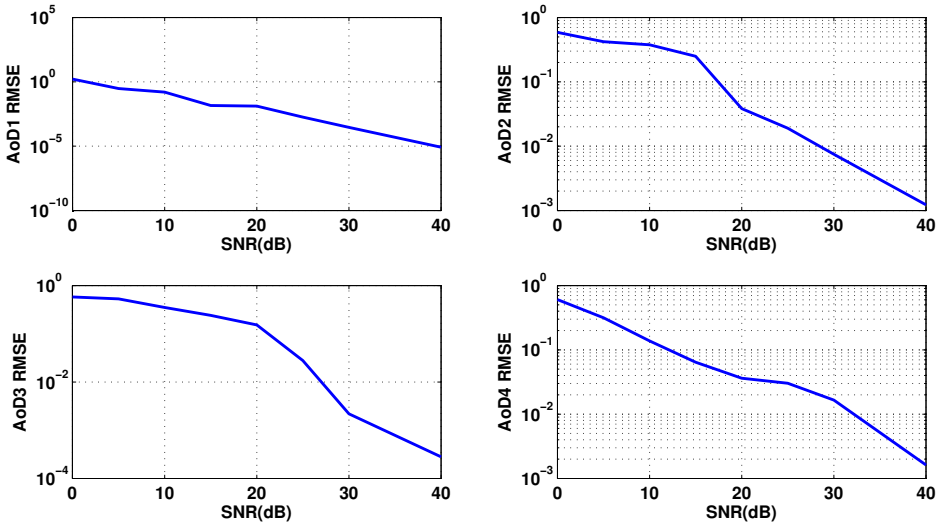

Figure 2.3: RMSE of sine of AoD $\sin(\psi)$

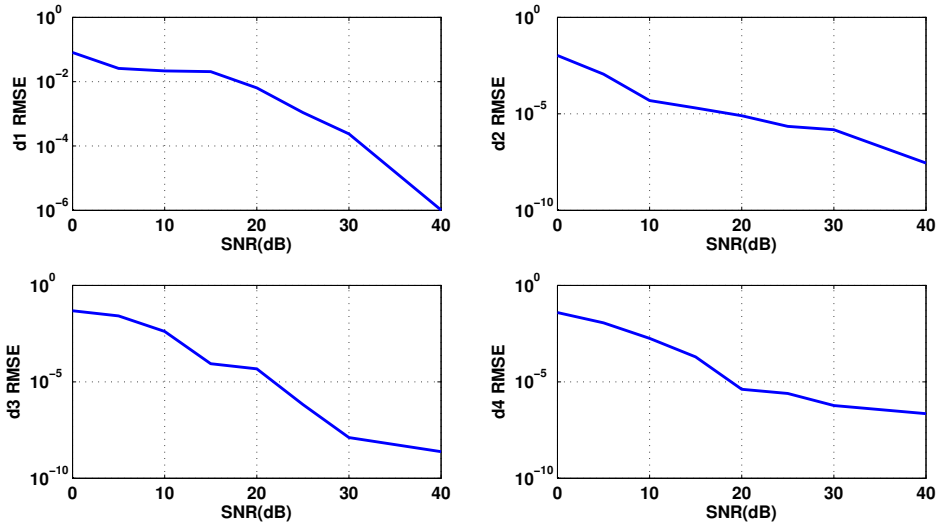

Figure 2.4: RMSE of delay times sample frequency spacing $\Delta f \tau$

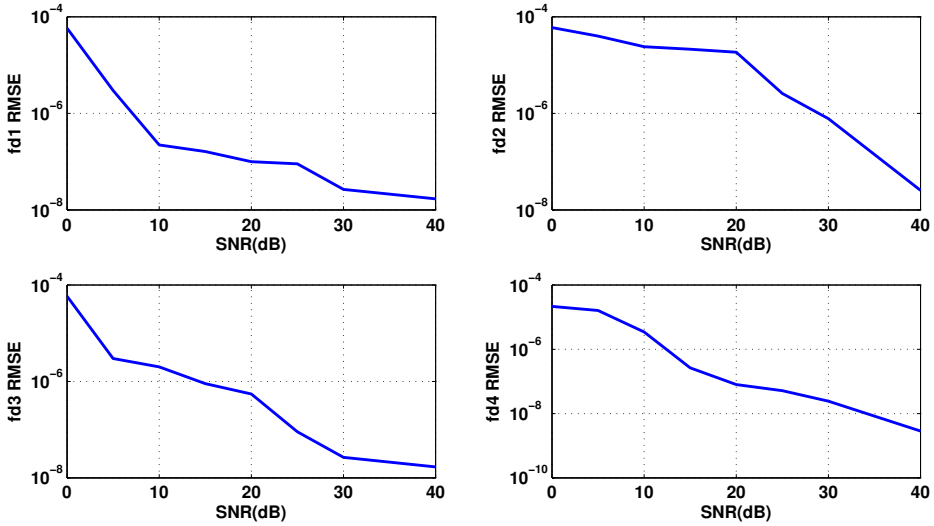

Figure 2.5: RMSE of DS times sample time spacing $\Delta t \mathbf{f}_d$

Chapter 3

Hybrid Localization for NLoS Static Environments

3.1 Introduction

In this chapter the fundamentals of Least Squares (LS) and Maximum Likelihood (ML) SBM-based localization are presented. This localization method was originally proposed [41] and developed further in [38]. The key ideas of those contributions are stated below. In our contribution, which consists of the rest of the sections, we extend the study on SBM-based localization method to include a WLS solution, a brief discussion on identifiability and a thorough study on the impact of network geometry on performance.

In [38], the authors proved that given the AoA, AoD and the length of a path j the MT coordinates should satisfy the following N_s straight line equations:

$$y_{mt} = k_j x_{mt} + b_j \tag{3.1}$$

where the 2 coefficients are given by:

$$k_j = \frac{\cos(\psi_j) + \cos(\phi_j)}{\sin(\psi_j) + \sin(\phi_j)} \tag{3.2}$$

$$b_j = -k_j(x_{s_j} - d_j \sin(\psi_j)) + y_{s_j} - d_j \cos(\psi_j). \tag{3.3}$$

This enabled them to derive a closed-form solution for the MT location, which for $N_s = 2$ is

$$\hat{\mathbf{p}}_{mt} = \left[\frac{b_2 - b_1}{k_1 - k_2}, \frac{k_1 b_2 - k_2 b_1}{k_1 - k_2}\right]^t \tag{3.4}$$

while for $N_s > 2$ the estimates that are chosen minimize the following expression

$$(\hat{x}_{mt}, \hat{y}_{mt}) = \underset{(x,y)}{\operatorname{argmin}} \sum_{i=j}^{N_s} (k_j x_{mt} + b_j - y_{mt})^2. \tag{3.5}$$

Using the LS criterion, the solution is given by

$$\hat{x}_{mt} = \frac{\sum_{j=1}^{N_s} b_j \sum_{j=1}^{N_s} k_j - N_s \sum_{j=1}^{N_s} b_j k_j}{N_s \sum_{j=1}^{N_s} k_j^2 - (\sum_{j=1}^{N_s} k_j)^2} \tag{3.6}$$

$$\hat{y}_{mt} = \frac{\sum_{j=1}^{N_s} b_j \sum_{j=1}^{N_s} k_j^2 - \sum_{j=1}^{N_s} k_j \sum_{j=1}^{N_s} b_j k_j}{N_s \sum_{j=1}^{N_s} k_j^2 - (\sum_{j=1}^{N_s} k_j)^2}. \tag{3.7}$$

Subsequently, the authors presented a formulation similar to the one in section 1.8, to obtain ML estimates. In order to compute the ML estimates, a standard built-in matlab function called “fmincon” was utilized. “fmincon” can be used to solve constrained nonlinear optimization problems. It is an iterative algorithm and therefore it requires initial points close to the true values to converge. Since for the LS estimates given above it was showed through simulations that the estimation error can be small, these estimates were considered as the initial points in the algorithm. The authors further showed through simulations that the ML solution not only outperforms the LS solution as expected, but can also achieve the CRB for reasonable values of the variances of the LDP errors.

3.2 LS Estimation for ToA/AoA/AoD Localization

The aforementioned existing method truly inspired us, mainly due to its simplicity and its high accuracy in environments when the single-bounce assumption is valid. One obvious extension to this method is to formulate a generic LS and for improved performance, a generic WLS algorithm for the joint estimation of parameters of interest and the nuisance parameters. To formulate a system of linear equations that can be solved to obtain a LS

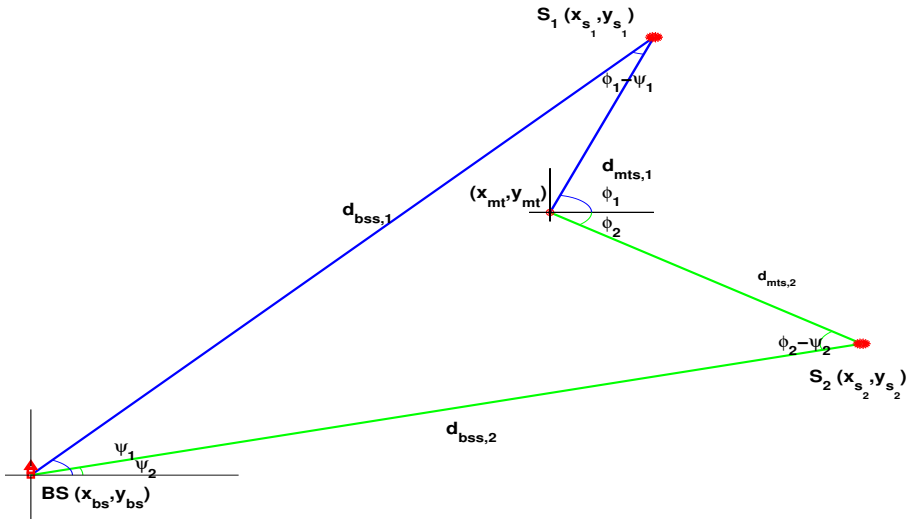

Figure 3.1: LDP in a NLoS environment: Static single bounce model

estimate of $\mathbf{p}$, we start with the sine and the cosine equations of the angles shown in figure 3.1 as follows

$$sin(\phi_j) = \frac{y_{s_j} - y_{mt}}{d_{mts_j}} \tag{3.8}$$

$$cos(\phi_j) = \frac{x_{s_j} - x_{mt}}{d_{mts_j}} \tag{3.9}$$

$$sin(\psi_j) = \frac{y_{s_j} - y_{bs_j}}{d_{bss_j}} \tag{3.10}$$

$$cos(\psi_j) = \frac{x_{s_j} - x_{bs_j}}{d_{bss_j}} \tag{3.11}$$

Solving for the distances and replacing the solutions in $d_j = d_{mts,j} + d_{bs,j}$, we get the following set of $4N_s$ linear equations

$$x_{s_j}(cos(\phi_j)+cos(\psi_j)) - x_{mt}cos(\psi_j) = cos(\phi_j)cos(\psi_j)d_j + x_{bs_j}cos(\phi_j) \tag{3.12}$$

$$y_{s_j}(sin(\phi_j)+sin(\psi_j)) - y_{mt}sin(\psi_j) = sin(\phi_j)sin(\psi_j)d_j + y_{bs_j}sin(\phi_j) \tag{3.13}$$

$$x_{s_j}sin(\psi_j) + y_{s_j}cos(\phi_j) - x_{mt}sin(\psi_j) = cos(\phi_j)sin(\psi_j)d_j + y_{bs_j}cos(\phi_j) \tag{3.14}$$

$$x_{s_j}sin(\phi_j) + y_{s_j}cos(\psi_j) - y_{mt}cos(\psi_j) = sin(\phi_j)cos(\psi_j)d_j + x_{bs_j}sin(\phi_j). \tag{3.15}$$

Obviously only $3N_s$ of these equations are linearly independent and thus there is no benefit in considering all of them but we ignore this fact in the formulation below. Defining the following diagonal matrices

$$\mathbf{C}_\phi = diag\{cos(\boldsymbol{\phi})\} \tag{3.16}$$
$$\mathbf{S}_\phi = diag\{sin(\boldsymbol{\phi})\} \tag{3.17}$$
$$\mathbf{C}_\psi = diag\{cos(\boldsymbol{\psi})\} \tag{3.18}$$
$$\mathbf{S}_\psi = diag\{sin(\boldsymbol{\psi})\} \tag{3.19}$$
$$\mathbf{D} = diag\{\mathbf{d}\} \tag{3.20}$$

the set of $4N_s$ equations can be put in vector form as shown below

$$\underbrace{\begin{bmatrix} -\mathbf{C}_\psi\mathbf{1} & \mathbf{0} & (\mathbf{C}_\phi + \mathbf{C}_\psi) & \mathbf{O} \\ \mathbf{0} & -\mathbf{S}_\psi\mathbf{1} & \mathbf{O} & (\mathbf{S}_\phi + \mathbf{S}_\psi) \\ -\mathbf{S}_\psi\mathbf{1} & \mathbf{0} & \mathbf{S}_\psi & \mathbf{C}_\phi \\ \mathbf{0} & -\mathbf{C}_\psi\mathbf{1} & \mathbf{S}_\phi & \mathbf{C}_\psi \end{bmatrix}}_{=\mathbf{A}} \underbrace{\begin{bmatrix} x_{mt} \\ y_{mt} \\ \mathbf{x}_s \\ \mathbf{y}_s \end{bmatrix}}_{=\mathbf{p}} = \underbrace{\begin{bmatrix} (\mathbf{D}\mathbf{C}_\psi + \mathbf{X}_{bs})\mathbf{C}_\phi\mathbf{1} \\ (\mathbf{D}\mathbf{S}_\psi + \mathbf{Y}_{bs})\mathbf{S}_\phi\mathbf{1} \\ (\mathbf{D}\mathbf{S}_\psi + \mathbf{Y}_{bs})\mathbf{C}_\phi\mathbf{1} \\ (\mathbf{D}\mathbf{C}_\psi + \mathbf{X}_{bs})\mathbf{S}_\phi\mathbf{1} \end{bmatrix}}_{=\mathbf{b}} . \tag{3.21}$$

We can then obtain a LS estimate for $\mathbf{p}$ given by:

$$\hat{\mathbf{p}}_{LS} = (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\mathbf{b}. \tag{3.22}$$

Should we desire higher accuracy at the cost of slightly higher computational complexity, we can always use Weighted Least Squares (WLS), to get

$$\hat{\mathbf{p}}_{WLS} = (\mathbf{A}^t\mathbf{C}_\mathbf{b}^{-1}\mathbf{A})^{-1}\mathbf{A}^t\mathbf{C}_\mathbf{b}^{-1}\mathbf{b}. \tag{3.23}$$

The weighting matrix $\mathbf{C}_\mathbf{b}^{-1}$ is the inverse of the covariance matrix of $\mathbf{b}$ and is given in appendix D. Since the entries of $\mathbf{A}$ also contain errors, it could be argued that Total Least Squares (TLS) or one of its more advanced versions (STLS etc) are more appropriate to solve eq. (3.21). However, assuming that the errors in the LDP estimates are small compared to their true values, the errors in $\mathbf{A}$ will be small compared to the errors in $\mathbf{b}$, since the latter contain error terms, which are the product of LDP errors times LDP true values. Therefore, despite the fact that WLS neglects the errors in $\mathbf{A}$, it is still a very attractive solution since it treats the errors in $\mathbf{b}$ in an optimal way.

3.3 Identifiability Concerns

In a static environment for which the AoA, the AoD and the path lengths are available, $N_{\boldsymbol{\theta}} = 3N_s$ LDP estimates are used in the estimation of $N_{\mathbf{p}} = 2N_s + 2$ unknown parameters. According to Corollary 1, identifiability is feasible if

$$3N_s \geq 2N_s + 2 \Leftrightarrow N_s \geq 2 \tag{3.24}$$

and the rank of the transformation matrix is $N_{\mathbf{p}}$, a condition that is usually met. This means that at least 2 distinct paths corresponding to different scatterers are required for location estimation. It also becomes apparent that if less than 3 LDP per scatterer are available, then the MT location is not identifiable independently of the richness of the channel (number of multipath components).

3.4 The impact of Network Geometry on Performance

In this section, we evaluate the performance of the SBM-based localization method in LoS and NLoS static environments. To do so, we compute and plot the Cramer-Rao bound (CRB). In all geometrical localization methods, the CRB (and the actual performance) depends on two factors: the accuracy of the available LDP estimates and the network geometry. The impact of the accuracy of the available LDP estimates on the CRB has been studied extensively for both LoS and NLoS environments. For the SBM-based methods, it was studied in [38]. However, in that contribution, the authors considered only a particular fixed setting (scatterers and BS locations) and thus the impact of network geometry on the accuracy was completely ignored. As a matter of fact, to the best of our knowledge, there exist no publications that address this topic. On the other hand, for LoS environments, there are many publications that deal with the impact of network geometry, especially for non-hybrid methods [8, 23, 56, 57]. For hybrid methods, the topic was addressed in [58, 59].

In contrast to all of the aforementioned existing studies, in [60] we derived expressions for the CRB as a function of distances and angles, which allowed for easy interpretation of the impact of the network geometry. As already mentioned, in an SBM-based localization method, the coordinates of the scatterers need to be jointly estimated with the coordinates of the MT. This increases the size of the transformation matrix $\mathbf{G}$, defined in section 1.9 from $2 \times N_{\boldsymbol{\theta}}$ to $(2N_s + 2) \times N_{\boldsymbol{\theta}}$ and leads to a complicated expression for

the FIM (see eq. 1.77), from which, it is impossible to demonstrate the impact of network geometry. However, after the straightforward derivation given in the following subsections, this becomes feasible. Contour maps in the numerical examples' subsection validate the conclusions drawn from the CRB expressions and serve as indicators on how the localization performance can be improved.

3.4.1 CRB for LoS Environments

In our work we have intentionally ignored localization under LoS conditions, since this topic has been studied extensively. However, in this section, we will derive the CRB for location estimation in LoS conditions to point out the similarities with the expression of the CRB for the NLoS case. Due to (1.21), it can be easily shown that

$$\sigma_{\boldsymbol{\phi}}^2 \boldsymbol{J}_{\boldsymbol{\phi}} = \sigma_{\boldsymbol{\psi}}^2 \boldsymbol{J}_{\boldsymbol{\psi}} \tag{3.25}$$

Therefore the FIM for this case becomes

$$\boldsymbol{J} = \frac{1}{\sigma_{\mathbf{d}}^2} \boldsymbol{J}_{\mathbf{d}} + \frac{1}{(\sigma_{\boldsymbol{\phi}}^2 + \sigma_{\boldsymbol{\psi}}^2)} \boldsymbol{J}_{\boldsymbol{\phi}} \tag{3.26}$$

Using eq. (1.20)-(1.22), we obtain for the 4 entries of the FIM

$$\mathbf{j}_{11} = \tfrac{N_s}{\sigma_d^2} - \textstyle\sum_i \alpha_i \sin^2(\phi_i) \tag{3.27}$$

$$\mathbf{j}_{12} = \mathbf{j}_{21} = \textstyle\sum_i \alpha_i \sin(\phi_i)\cos(\phi_i) \tag{3.28}$$

$$\mathbf{j}_{22} = \tfrac{N_s}{\sigma_d^2} - \textstyle\sum_i \alpha_i \cos^2(\phi_i) \tag{3.29}$$

where

$$\alpha_i = \frac{1}{\sigma_d^2} - \frac{1}{(\sigma_\phi^2 + \sigma_\psi^2) d_i^2} = \frac{1}{\sigma_d^2} - \frac{1}{\sigma_{\phi\psi}^2 d_i^2}. \tag{3.30}$$

The FIM for the LoS scenario is a 2×2 matrix and thus it can easily be inverted to get the $CRB_{pos} = tr\{\mathbf{J}^{-1}\}$. The derivation is simple and the result is given below

$$CRB_{pos} = \frac{2\sum_i \frac{1}{\sigma_d^2} - \frac{\alpha_i}{2}}{(\sum_i \frac{1}{\sigma_d^2} - \frac{\alpha_i}{2})^2 - (\sum_i \frac{\alpha_i \cos 2\phi_i}{2})^2 - (\sum_i \frac{\alpha_i \sin 2\phi_i}{2})^2} \tag{3.31}$$

Introducing the diagonal matrix of cosines of AoA, $\mathbf{C}_{2\phi}$ (see also definition in the index of symbols) and the matrix

$$\mathbf{A} = diag\{[a_1, \ldots, a_{N_s}]\} = \frac{1}{\sigma_d^2}\mathbf{I} - \frac{1}{\sigma_{\phi\psi}^2}\mathbf{D}^{-2} \tag{3.32}$$

we can rewrite the CRB expression in a way that allows for easy comparison with the one for the NLoS case

$$CRB_{pos} = \frac{2\mathbf{1}^t(\frac{1}{\sigma_d^2}\mathbf{I} - \frac{1}{2}\mathbf{A})\mathbf{1}}{\frac{1}{\sigma_d^2}N_s\mathbf{1}^t(\frac{1}{\sigma_d^2}\mathbf{I} - \mathbf{A})\mathbf{1} + \mathbf{1}^t\mathbf{A}(\mathbf{1}\mathbf{1}^t - \breve{\mathbf{C}}_{\delta 2\phi})\mathbf{A}\mathbf{1}} \tag{3.33}$$

where $\breve{\mathbf{C}}_{\delta 2\phi}$ is a symmetric matrix whose $\{i, j\}$ entry is equal to $cos(2\phi_i - 2\phi_j)$.

3.4.2 CRB for NLoS Environments

To compute the CRB in a NLoS static environment, the $(2N_s+2)\times(2N_s+2)$ FIM needs to be inverted. This is feasible even for large values of N_s, if we write the FIM as a 2×2 block matrix, exactly as we did in eq. (1.79) and then use blockwise inversion. Besides we only need to focus on the upper left 2×2 submatrix of its inverse, the trace of which gives the best possible accuracy, i.e. the CRB for the MT position.

$$CRB_{pos} = tr\{[\boldsymbol{J}^{-1}]_{1:2,1:2}\} \tag{3.34}$$

Using blockwise inversion we can obtain the upper left submatrix of the inverse of $\mathbf{J}$, given by the Schur complement of $\mathbf{J}_{22}$

$$[\mathbf{J}^{-1}]_{1:2,1:2} = (\mathbf{J}_{11} - \mathbf{J}_{12}(\mathbf{J}_{22})^{-1}\mathbf{J}_{21})^{-1} \stackrel{\Delta}{=} \mathbf{J}_p^{-1}. \tag{3.35}$$

The solution is derived in appendix E and the 4 entries of $\mathbf{J}_p$ are given by eq. (E.37)-(E.39). The CRB for the position estimate is then simply given by

$$\begin{aligned} CRB_{pos} &= \frac{2\mathbf{1}^t\bar{\mathbf{J}}_{det}^{-1}\mathbf{1}}{\mathbf{1}^t\bar{\mathbf{J}}_{det}^{-1}(\mathbf{Q}'_{\boldsymbol{\phi}+\boldsymbol{\psi}}\mathbf{1}\mathbf{1}^t\mathbf{Q}_{\boldsymbol{\phi}+\boldsymbol{\psi}} - \mathbf{S}_{\boldsymbol{\phi}+\boldsymbol{\psi}}\mathbf{1}\mathbf{1}^t\mathbf{S}_{\boldsymbol{\phi}+\boldsymbol{\psi}})\bar{\mathbf{J}}_{det}^{-1}\mathbf{1}} \\ &= \frac{2\mathbf{1}^t\bar{\mathbf{J}}_{det}^{-1}\mathbf{1}}{\mathbf{1}^t\bar{\mathbf{J}}_{det}^{-1}(\mathbf{1}\mathbf{1}^t - \breve{\mathbf{C}}_{\delta\phi+\delta\psi})\bar{\mathbf{J}}_{det}^{-1}\mathbf{1}} \end{aligned} \tag{3.36}$$

where $\breve{\mathbf{C}}_{\delta\phi+\delta\psi}$ is a symmetric matrix whose $\{i, j\}$ entry is equal to $cos(\phi_i - \phi_j + \psi_i - \psi_j)$ and the rest of the matrices are given in eq. (E.40)-(E.44).

3.4.3 Demonstration and Comparison

First one can observe from eq. (3.33) and (3.36) that for both LoS and NLoS environments, the CRB depends on distances through the matrices $\mathbf{A}$ and $\bar{\mathbf{J}}_{det}$. This is no big surprise, since this hybrid method utilizes angles. In

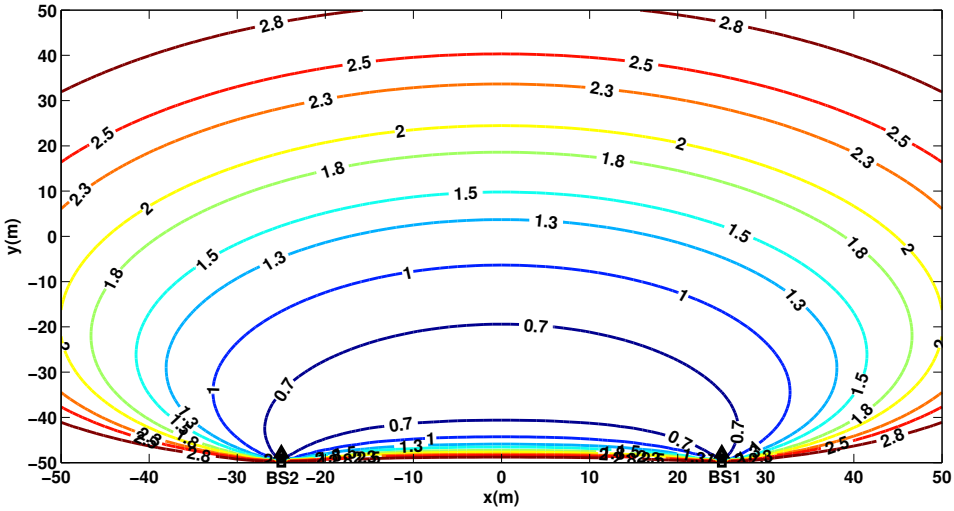

Figure 3.2: CRB vs MT position for 2 distant BS LoS environment

contrast to ToA-based methods where distances do not impact performance (only their errors do of course), in AoA methods, the greater the distances the signal components cover, the worse the performance. Similarly here, by taking the partial derivative of the LoS (NLoS) CRB with respect to any d_i ($d_{mts,i}$), it can be proved that performance worsens when the MT moves away from the BS (the scatterers). To demonstrate this, consider the following example, where the MT communicates with 2 BS. The following contour maps show the CRB for a region of $10^4 m^2$. The contour lines are based on the c.d.f. of the CRB. Numbering the lines in increasing order of the corresponding CRB, contour line j encloses $(j/10)100\%, j = 1 : 9$ of the total area, i.e., for a contour line j we have $p(CRB < CRB(j)) = j/10$. As can be clearly seen, in a NLoS environment, performance degrades as the distance between the MT and the scatterers increases. Accepting commonly used channel models like the circular one for macrocells and the elliptical one for pico and microcells [61–65], we conclude that the performance will not vary much in different environments. This is because in the circular model, the scatterers are located inside a circle with the center placed at the MT position, and in the elliptical one, the scatterers are located inside an ellipse with the foci placed at the MT and BS position. Thus, in both models, the distance between the scatterers and the MT is comparable.

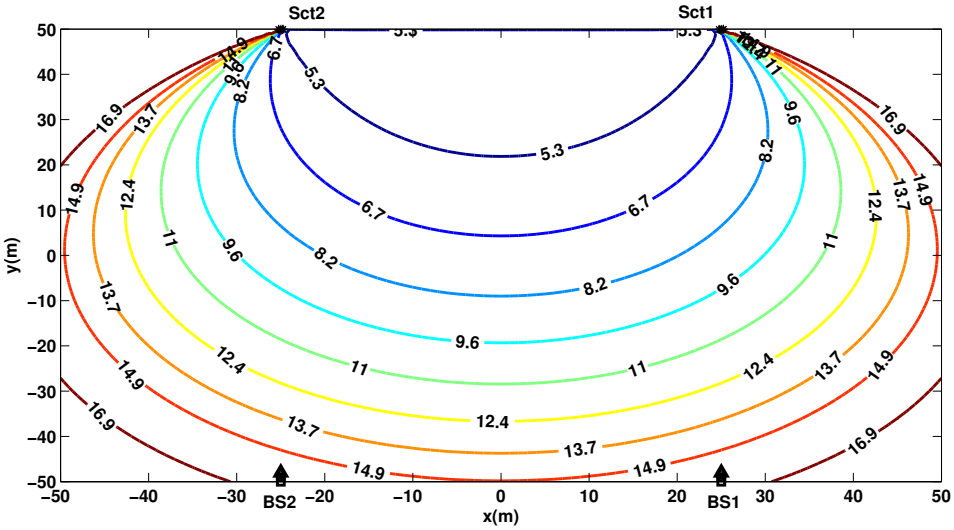

Figure 3.3: CRB vs MT position for 2 BS - 2 distant scatterers NLoS environment

More interesting than the impact of distances is the impact of the angles of the MPC. The NLoS CRB_{pos} depends solely on the sums and the differences of AoA with the corresponding AoD. This is no big surprise either, since due to symmetry in a 1 BS scenario, we would expect to obtain the same performance if we exchange the position of the BS with that of the MT. Furthermore, one can observe the similarity between the 2 CRB expressions, by replacing ψ in the NLoS CRB_{pos}, using the LoS condition (1.21). The differences of AoA with AoD do not depend on the angles anymore and the sums are equal to two times the AoA plus a constant $c \in \{-2\pi, 0, 2\pi\}$ so that $\breve{\mathbf{C}}_{\delta\phi+\delta\psi} = \breve{\mathbf{C}}_{\delta 2\phi}$. The only difference is the extra term in the denominator of the LoS CRB expression that does not exist in the NLoS one due to the fact that in this expression estimation of $2N_s$ more parameters is assumed. One last comment about these expressions concerns the terms involving the matrices denoted by $\breve{\mathbf{C}}$. Due to these terms the denominators decrease and thus the CRB increases. Both CRB can be maximized with respect to angles if $\breve{\mathbf{C}}_{\delta 2\phi} = \breve{\mathbf{C}}_{\delta\phi+\delta\psi} = \mathbf{1}\mathbf{1}^t$. This corresponds to collocated BS for the LoS case and collocated scatterers for the 1 BS NLoS case. While for the LoS the CRB remains finite due to the extra term[1], for the NLoS

[1]This means that localization with this hybrid method is possible even with 1 BS, as

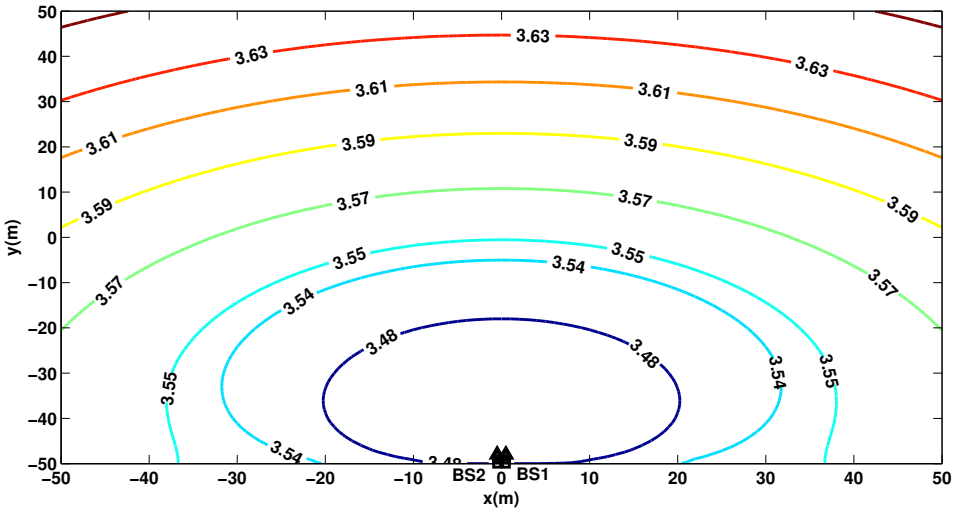

Figure 3.4: CRB vs MT position for 2 collocated BS LoS environment

case the CRB goes to infinity and thus it is impossible to estimate the MT location. The significance of these matrices is demonstrated with the following contour maps for 1 and 2 BS scenarios. Comparing fig. (3.2) with fig. (3.4), we observe that indeed, in a LoS environment, performance decreases for closely located BS, but not significantly (same order of magnitude) due to the first term in the denominator that depends only on distances. On the other hand, in NLoS environments, the impact of the relative position of the scatterers will have different impact on the performance depending on the number of BS employed in the localization method. In fig. (3.6), we can observe the huge impact of collocated scatterers in a 1 BS environment. In such an environment localization is practically impossible. However, if in such an environment 2 BS are used, each one of which communicates with the MT via one of the scatterers, the result is totally different. Comparing fig. (3.3) with fig. (3.7) we observe that for this case the CRB actually improves for collocated scatterers. The cdf of the CRB for the 4 different NLoS environments considered (1 or 2 BS with collocated or distant scatterers) are plotted in fig.(3.8). From that last figure it becomes apparent that while for collocated scatterers it is preferable to have communications with more than 1 BS, in environments with distant scatterers communication with just

expected.

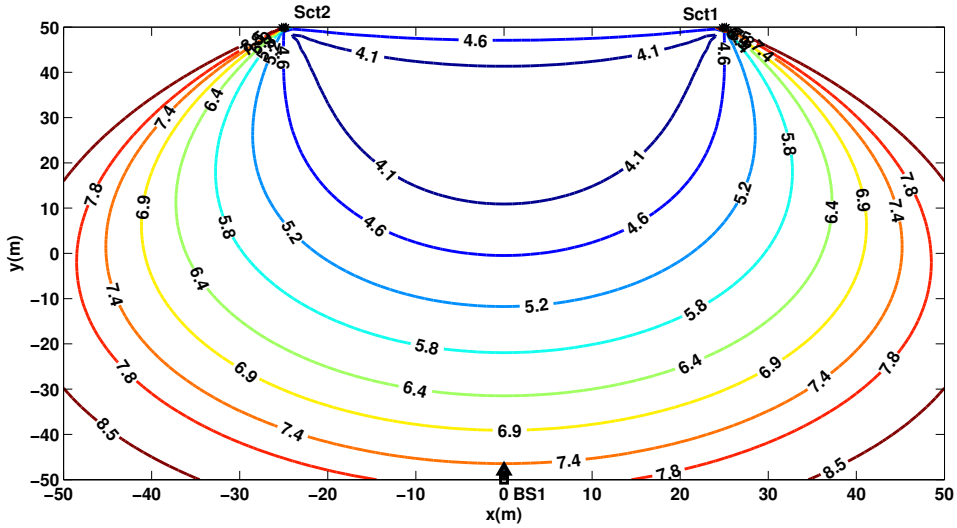

Figure 3.5: CRB vs MT position for 1 BS - 2 distant scatterers NLoS environment

1 BS via multiple paths can lead to better performance.

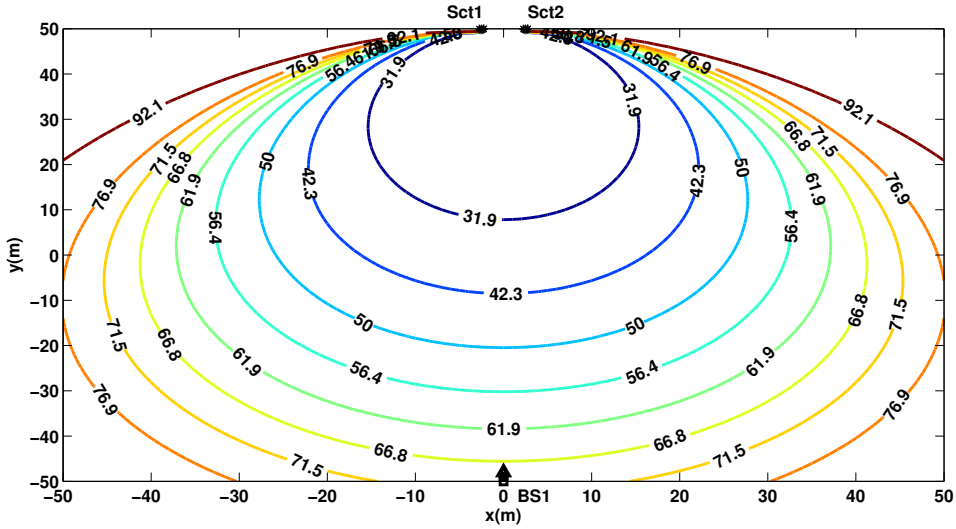

Figure 3.6: CRB vs MT position for 1 BS - 2 collocated scatterers NLoS environment

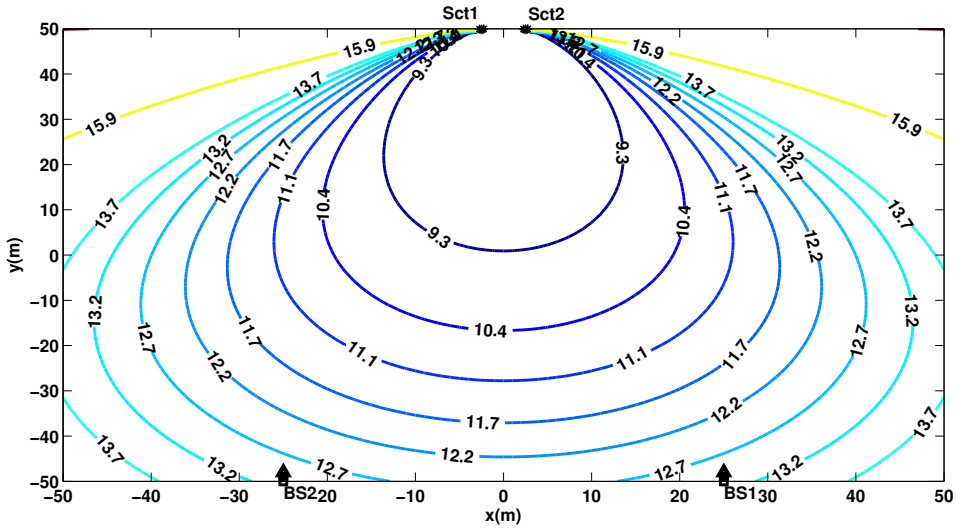

Figure 3.7: CRB vs MT position for 2 BS - 2 collocated scatterers NLoS environment

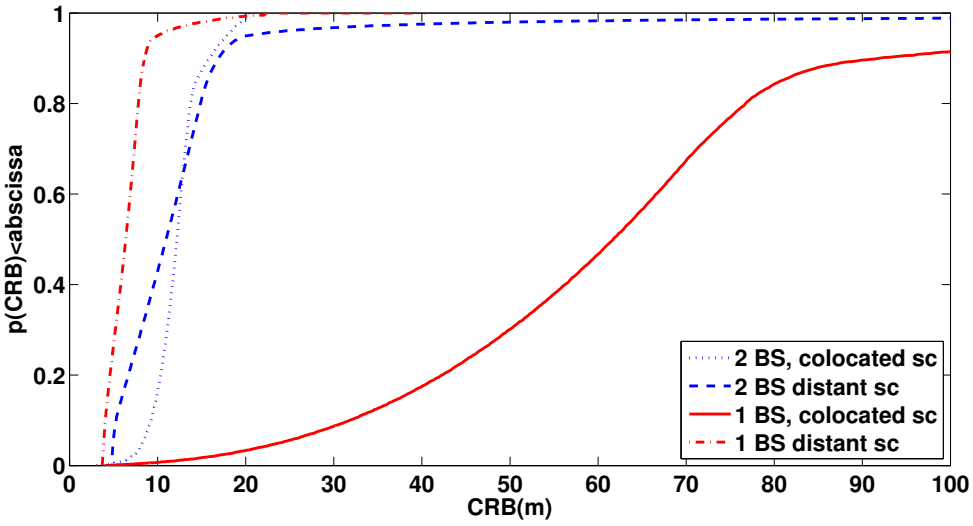

Figure 3.8: CRB cdf for 4 different NLoS scenarios

Chapter 4

Hybrid Localization for NLoS Dynamic Environments

4.1 Introduction

In this chapter, two hybrid localization methods that utilize the DSBM are presented and their advantage over the SBM-based method is demonstrated. Both methods are applicable to NLOS environments that change dynamically due to the movement of the MT and both result in Least Squares (LS) as well as Maximum Likelihood (ML) estimates. The LS estimates are given in the following two sections while to obtain ML estimates one could use the formulation presented in section 1.8. The two methods assume knowledge of different kinds of location-dependent parameters (LDP), like the path length d which is proportional to the delay, the AoD ψ at the BS and the normalized (with respect to the carrier frequency and the inverse of the speed of light) Doppler Shift f_d. The first method further assumes knowledge of the AoA ϕ of different paths at the MT. Despite the fact that this appears to be a small difference, the two methods actually differ a lot. Due to the lack of availability of AoA in the second method, the parameters that need to be estimated become identifiable solely due to the variation in time of the LDP. A detailed discussion on identifiability and performance of the 2 methods is

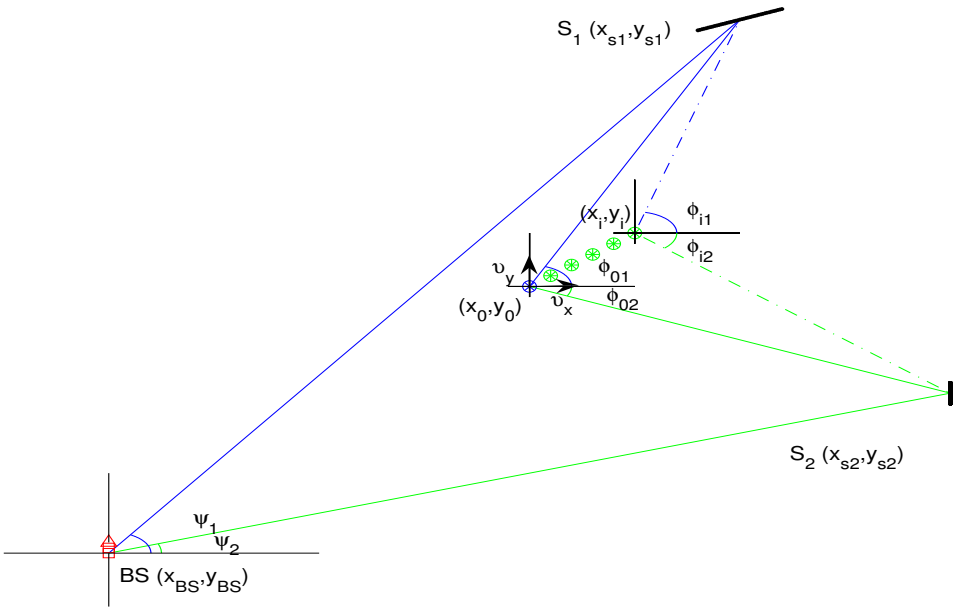

Figure 4.1: LDP in a NLoS environment: Dynamic single bounce model

presented in sections 4.4 and 4.5, while in the last section the 2 methods are compared. This chapter is a collection and an extension of the methods and the results presented in [49,66,67] and some preliminary results that will be presented in [60].

4.2 LS Estimation for ToA/AoA/AoD/DS Localization

In this method $K = 4$ and $\boldsymbol{\theta} = [\mathbf{d}^t, \boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{f}_d^t]^t$. Similarly to the static case scenario, we can formulate a system of linear equations that can be solved to obtain a (W)LS estimate of the speed and position of the MT. The LS formulation is the same as in the static case, so similarly to eq. (3.8-3.15) and introducing the subscript i to denote the time dependency, the sines and the cosines of the angles shown in fig. 4.1 are given by

$$sin(\phi_{ij}) = \frac{y_{s_j} - y_i}{d_{mts_{ij}}} \tag{4.1}$$

$$cos(\phi_{ij}) = \frac{x_{s_j} - x_i}{d_{mts_{ij}}} \tag{4.2}$$

$$sin(\psi_{ij}) = \frac{y_{s_j} - y_{bs_j}}{d_{bss_{ij}}} \tag{4.3}$$

$$cos(\psi_{ij}) = \frac{x_{s_j} - x_{bs_j}}{d_{bss_{ij}}} \tag{4.4}$$

where the distances that appear in the denominators satisfy the following equation

$$d_{ij} = d_{bss_{ij}} + d_{mts_{ij}}. \tag{4.5}$$

Furthermore we can write any DS as

$$f_{d,ij} = \upsilon_x cos(\phi_{ij}) + \upsilon_y sin(\phi_{ij}). \tag{4.6}$$

Solving for the distances and replacing the solutions in d_{ij} we get the following set of $4N_sN_t$ linear equations (from this set of equations only $3N_sN_t$ are linearly independent):

$$x_{s_j}(cos(\phi_{ij}) + cos(\psi_{ij})) - x_i cos(\psi_{ij}) = cos(\phi_{ij})cos(\psi_{ij})d_{ij} + x_{bs_j}cos(\phi_{ij}) \tag{4.7}$$

$$y_{s_j}(sin(\phi_{ij}) + sin(\psi_{ij})) - y_i sin(\psi_{ij}) = sin(\phi_{ij})sin(\psi_{ij})d_{ij} + y_{bs_j}sin(\phi_{ij}) \tag{4.8}$$

$$x_{s_j}sin(\psi_{ij}) + y_{s_j}cos(\phi_{ij}) - x_i sin(\psi_{ij}) = cos(\phi_{ij})sin(\psi_{ij})d_{ij} + y_{bs_j}cos(\phi_{ij}) \tag{4.9}$$

$$x_{s_j}sin(\phi_{ij}) + y_{s_j}cos(\psi_{ij}) - y_i cos(\psi_{ij}) = sin(\phi_{ij})cos(\psi_{ij})d_{ij} + x_{bs_j}sin(\phi_{ij}). \tag{4.10}$$

The solution for dynamic environments comes as an extension of the solution for the static one. If the constant speed mobility model is considered, the matrix $\mathbf{A}$ in eq. (3.21) needs to be augmented with the following matrix

$$\mathbf{A}_\upsilon = \begin{bmatrix} -\mathbf{C}_\psi(\mathbf{t} \otimes \mathbf{1}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{S}_\psi(\mathbf{t} \otimes \mathbf{1}) \\ -\mathbf{S}_\psi(\mathbf{t} \otimes \mathbf{1}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{C}_\psi(\mathbf{t} \otimes \mathbf{1}) \\ \mathbf{C}_\phi\mathbf{1} & \mathbf{S}_\phi\mathbf{1} \end{bmatrix} \tag{4.11}$$

to get

$$\underbrace{\left[\begin{array}{c|c} \mathbf{A} & \mathbf{A}_\upsilon \\ \mathbf{O} & \end{array}\right]}_{=\mathbf{A}'} \underbrace{\begin{bmatrix} \mathbf{p} \\ \upsilon_x \\ \upsilon_y \end{bmatrix}}_{=\mathbf{p}'} = \underbrace{\begin{bmatrix} \mathbf{b} \\ \mathbf{F}_d\mathbf{1} \end{bmatrix}}_{=\mathbf{b}'}. \tag{4.12}$$

The weighted least square solution follows immediately

$$\hat{\mathbf{p}}'_{WLS} = (\mathbf{A}'^t \mathbf{C}_{\mathbf{b}'}^{-1} \mathbf{A}')^{-1} \mathbf{A}'^t \mathbf{C}_{\mathbf{b}'}^{-1} \mathbf{b}'. \tag{4.13}$$

Similarly if the constant acceleration mobility model is used, we can further augment $\mathbf{A}'$ with the matrix

$$\mathbf{A}_\alpha = \begin{bmatrix} -\mathbf{C}_\psi(\frac{1}{2}(\mathbf{t} \odot \mathbf{t}) \otimes \mathbf{1}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{S}_\psi(\frac{1}{2}(\mathbf{t} \odot \mathbf{t}) \otimes \mathbf{1}) \\ -\mathbf{S}_\psi(\frac{1}{2}(\mathbf{t} \odot \mathbf{t}) \otimes \mathbf{1}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{C}_\psi(\frac{1}{2}(\mathbf{t} \odot \mathbf{t}) \otimes \mathbf{1}) \\ \mathbf{C}_\phi(\mathbf{t} \otimes \mathbf{1}) & \mathbf{S}_\phi(\mathbf{t} \otimes \mathbf{1}) \end{bmatrix} \tag{4.14}$$

to get

$$\underbrace{\left[\begin{array}{c|c} \mathbf{A}' & \mathbf{A}_\alpha \end{array}\right]}_{=\mathbf{A}''} \underbrace{\begin{bmatrix} \mathbf{p}' \\ \alpha_x \\ \alpha_y \end{bmatrix}}_{=\mathbf{p}''} = \mathbf{b}'. \tag{4.15}$$

the WLS solution of which is

$$\hat{\mathbf{p}}''_{WLS} = (\mathbf{A}''^t \mathbf{C}_{\mathbf{b}'}^{-1} \mathbf{A}'')^{-1} \mathbf{A}''^t \mathbf{C}_{\mathbf{b}'}^{-1} \mathbf{b}'. \tag{4.16}$$

The matrix $\mathbf{C}_{\mathbf{b}'}^{-1}$ in eq. (4.13) and (4.16) is the inverse of the covariance matrix of $\mathbf{b}'$ and is given in appendix F. Similarly to the static environment, the motive behind choosing the WLS solution and not TLS or one of its more advanced versions is that the errors in $\mathbf{A}'$ ($\mathbf{A}''$) are usually small compared to the errors in $\mathbf{b}'$. Therefore, despite the fact that WLS neglects the errors in $\mathbf{A}'$ ($\mathbf{A}''$), it is still a very attractive solution since it treats the errors in $\mathbf{b}'$ in an optimal way.

4.3 LS Estimation for ToA/AoD/DS Localization

In contrast to the method presented in the previous section, where we have derived a LS solution based on knowledge of all 4 subsets of LDP, herein we do not consider knowledge of the AoA (angle of the impinging wave at the MT, considering downlink transmission). We do so in order to account for numerous realistic scenarios in which AoA estimates might not be available or might be totally unreliable due to e.g. lack of calibration of the antenna array at the receiver or completely modified antenna pattern because of the way the user is holding the device. Although the Doppler shift of each MPC

explicitly depends on the AoA through the cosine of its difference with the direction of movement ω, we show how to effectively extract useful - for localization purposes - information from this LDP and combine it with the rest of the available information to form a 2-step LS solution. This was achieved in [67] through a geometric interpretation of the DS that enabled the use of trigonometric laws and led to a linear system of equations. In this section however, we take an equivalent approach that will be presented in [60] and utilizes time derivatives.

In this method $K = 3$ and $\boldsymbol{\theta} = [\mathbf{d}^t, \boldsymbol{\psi}^t, \mathbf{f}_d^t]^t$. In the 1^{st} step of the method the nuisance parameters in polar coordinates will be estimated, i.e. instead of considering $\mathbf{p}_{nui} = [\mathbf{x}_s^t, \mathbf{y}_s^t]^t$ we consider $\mathbf{p}'_{nui} = [\mathbf{d}_{bs}^t, \mathbf{p}_{\upsilon,\alpha}^t]^t$, where $\mathbf{p}_{\upsilon,\alpha}^t$ is a vector of unknown parameters that depend on the speed (or the speed and the acceleration) of the MT. In the 2^{nd} step these estimates are utilized to formulate a ToA localization problem and estimate the parameters of interest $\mathbf{p}_{int} = [\mathbf{p}_0, \boldsymbol{\upsilon}_0, \boldsymbol{\alpha}]^t$ using an extension of the lines-of-position (LoP) method that was presented in [68] for static channels. The first step is based on the following observations: First we observed that $f_{d,ij}$ can be written as a function of $d_{mts_{ij}}$

$$f_{d,ij} = \frac{s_{ij}}{d_{mts,ij}} \tag{4.17}$$

where $s_{ij} = \upsilon_x(x_{s_j} - x_i) + \upsilon_y(y_{s_j} - y_i)$. It follows that

$$f_{d,ij} d_{ij} = f_{d,ij} d_{bs,j} + s_{ij}. \tag{4.18}$$

Next we observed that the 1^{st} derivative of s_{ij} w.r.t. time for the constant speed model

$$\frac{ds_{ij}}{dt} = -\upsilon^2 \tag{4.19}$$

and the 2^{nd} derivative for the constant acceleration model

$$\frac{ds_{ij}}{dt} = -\upsilon_i^2 + \alpha_x(x_{sj} - x_i) + \alpha_y(y_{sj} - y_i) \tag{4.20}$$

$$\frac{d^2 s_{ij}}{dt^2} = -3(\upsilon_{x_0}\alpha_x + \upsilon_{y_0}\alpha_y) - 3\alpha^2 t_i \tag{4.21}$$

do not depend on the unknown coordinates. This can be extended to higher derivatives if the MT position as a function of time is a higher degree polynomial. We can therefore differentiate eq. (4.18) n times[1] to get

$$\frac{d^{(n)}(f_{d,ij} d_{ij})}{dt^{(n)}} = \frac{d^{(n)} f_{d,ij}}{dt^{(n)}} d_{bs,j} + \frac{d^{(n)} s_{ij}}{dt^{(n)}}. \tag{4.22}$$

[1] n depends on the mobility model utilized.

Table 4.1: Parameters that depend on the mobility model

mobility model	$\mathbf{V}$	$\mathbf{p}_{\upsilon,\alpha}$
const. speed ($n=1$)	$-\mathbf{1}$	υ^2
const. acceleration ($n=2$)	$-3[\mathbf{1}\|\mathbf{t}]$	$[(\upsilon_{x_0}\alpha_x+\upsilon_{y_0}\alpha_y),\alpha^2]^t$

Subsequently we can put the above set of $N_s(N_t-n)$ equations in vector form, using $\mathbf{d}=(\mathbf{1}\otimes\mathbf{d}_{bs})+\mathbf{d}_{ms}$ and $\mathbf{f}_d=\mathbf{D}_{ms}^{-1}\mathbf{s}$ as follows

$$\frac{d^{(n)}(\mathbf{f}_d\odot\mathbf{d})}{dt^{(n)}}=\frac{d^{(n)}\mathbf{f}_d}{dt^{(n)}}(\mathbf{1}\otimes\mathbf{d}_{bs})+\frac{d^{(n)}\mathbf{s}}{dt^{(n)}}=\frac{d^{(n)}}{dt^{(n)}}\mathbf{F}_d(\mathbf{1}\otimes\mathbf{I})\mathbf{d}_{bs}+\mathbf{V}\mathbf{p}_{\upsilon,\alpha} \quad (4.23)$$

where, for the 2 mobility models considered, $\mathbf{V}$ and $\mathbf{p}_{\upsilon,\alpha}$ are given in table 4.1 and $\mathbf{t}=[t_0,\ldots,t_{N_t-1}]^t$ is the vector containing all time instances. To make computations feasible, derivatives w.r.t time need to be replaced by differences of time samples. To that end we introduce the following " time difference" matrices

$$\mathbf{R}_t=\mathbf{R}_n\otimes\mathbf{I}. \quad (4.24)$$

where the $N_t-n\times N_t$ matrix $\mathbf{R}_n$ is given by

$$\mathbf{R}_n=\begin{cases}[\mathbf{0}|\mathbf{I}]-[\mathbf{I}|\mathbf{0}], & n=1\\ [\mathbf{O}|\mathbf{I}]-2[\mathbf{0}|\mathbf{I}|\mathbf{0}]+[\mathbf{I}|\mathbf{O}], & n=2\end{cases} \quad (4.25)$$

When $\mathbf{R}_t$ multiplies on the left the vectors (matrices) that contain LDP in (4.23), the resulting vectors (matrices) contain the differences (or differences of differences) w.r.t to time of their entries. We can now rewrite (4.23) as follows :

$$\underbrace{[\mathbf{R}_t\mathbf{F}_d(\mathbf{1}\otimes\mathbf{I}),-(dt)^n\mathbf{V}]}_{\mathbf{Z}}\mathbf{p}'_{nui}=\underbrace{\mathbf{R}_t(\mathbf{f}_d\odot\mathbf{d})}_{\mathbf{w}} \quad (4.26)$$

If time sampling is not uniform, $(dt)^n$ should be replaced by the vector $\mathbf{R}_t\mathbf{t}$ and the product with $\mathbf{V}$ becomes a Hadamamard product of $\mathbf{R}_t\mathbf{t}$ with each of the columns of $\mathbf{V}$. Eq. (4.26) can be solved using eg. WLS and the solution for $\mathbf{p}'_{nui}=[\mathbf{d}_{bs}^t,\mathbf{p}_{\upsilon,\alpha}^t]^t$ is

$$\hat{\mathbf{p}}'_{nui}=(\mathbf{Z}^t(\mathbf{C}_\mathbf{w})^{-1}\mathbf{Z})^{-1}\mathbf{Z}^t(\mathbf{C}_\mathbf{w})^{-1}\mathbf{w} \quad . \quad (4.27)$$

If we further define $\mathbf{Z}_f=\mathbf{R}_t\mathbf{F}_d(\mathbf{1}\otimes\mathbf{I})$ and $\mathbf{P}=\mathbf{I}-\mathbf{V}(\mathbf{V}^t\mathbf{C}_\mathbf{w}^{-1}\mathbf{V})^{-1}\mathbf{V}^t\mathbf{C}_\mathbf{w}^{-1}$, it can be easily shown that

$$\hat{\mathbf{d}}_{bss}=(\mathbf{Z}_f^t\mathbf{C}_\mathbf{w}^{-1}\mathbf{P}\mathbf{Z}_f)^{-1}\mathbf{Z}_f^t\mathbf{C}_\mathbf{w}^{-1}\mathbf{P}\mathbf{w}. \quad (4.28)$$

In the two equations above we have introduced the covariance matrix of $\mathbf{w}$, which, assuming i.i.d. LDP, is given by

$$\mathbf{C_w} = \mathbf{R}_t(\sigma_{f_d}^2 \mathbf{d}\mathbf{d}^t \odot \mathbf{I} + \sigma_d^2 \mathbf{f}_d \mathbf{f}_d^t \odot \mathbf{I} + \sigma_{f_d}^2 \sigma_d^2 \mathbf{I})\mathbf{R}_t^t. \tag{4.29}$$

We can now proceed with the 2^{nd} step of the method. Combining the estimated distances between the scatterers and the BS with the available AoD, we can obtain estimates for the scatterers' locations. Since also the distances between the scatterers and the MT can also be estimated using eq. (4.5), the problem of estimating the location of the MT at time i becomes equivalent to solving a ToA localization problem with N_s BS (which are the scatterers in this problem formulation), all of which are in a LoS environment. As mentioned above, one very appealing LS solution for this problem, is the so-called lines-of-position (LoP). We can modify this method to estimate $\mathbf{p}_{int}$. According to the LoP method, the coordinates of the MT at time i satisfy:

$$(c_{j'} - c_j)x_i + (s_{j'} - s_j)y_i = b_{ij'} - b_{ij}. \tag{4.30}$$

where

$$c_j = x_{bs_j} + \hat{d}_{bs_j} \cos(\psi_j) \tag{4.31}$$

$$s_j = y_{bs_j} + \hat{d}_{bs_j} \sin(\psi_{ij}) \tag{4.32}$$

$$\begin{aligned} b_{ij} &= -\tfrac{1}{2}(\hat{d}_{mts,ij}^2 - \hat{d}_{bs_j}^2 - 2\underbrace{(x_{bs_j}\cos(\psi_j) + y_{bs_j}\sin(\psi_j))}_{q_j}\hat{d}_{bs_j} - \underbrace{(x_{bs_j}^2 + y_{bs_j}^2)}_{p_{bs_j}}) \\ &= \tfrac{1}{2}(d_{ij}^2 - 2(d_{ij} + q_j)\hat{d}_{bs_j} - p_{bs_j}) \end{aligned} \tag{4.33}$$

In contrast to the 1^{st} step that required differences with respect to time, eq. (4.30) indicates that differences with respect to different paths are required. Therefore, in order to put the set of equations in vector form, we introduce a "space difference" matrix

$$\mathbf{R}_s = \mathbf{I} \otimes \mathbf{R}_1 \tag{4.34}$$

which, by multiplying on the left a vector (matrix) that contains LDP, gives the differences of the entries (that correspond to different paths) of the vector (matrix). The matrix $\mathbf{R}_1$ is a $N_s - 1 \times N_s$ matrix that has the same form as the one defined in eq. (4.25) for $n = 1$. After a few algebraic computations shown in app. G we derive the following

$$\mathbf{A}\mathbf{p}_{int} = \mathbf{b} \tag{4.35}$$

where the matrix $\mathbf{A}$ is defined in eq. (G.20) and the vector $\mathbf{b}$ is defined in eq. (G.9). Therefore the WLS estimate for the parameters of interest is given by

$$\hat{\mathbf{p}}_{int} = (\mathbf{A}^t \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{A})^{-1} \mathbf{A}^t \mathbf{C}_{\mathbf{b}}^{-1} \mathbf{b} \tag{4.36}$$

where we have introduced the covariance matrix of $\mathbf{b}$. Its expression (a sum of many terms) along with its lengthy derivation can be found in app. H.

4.4 Identifiability Concerns

Since in a DSBM model we consider that the environment changes only due to the movement of the MT, the AoD are time-invariant LDP. Therefore, even though $N_s N_t$ AoD estimates might be available, only N_s of them can potentially lead to independent equations. Thus, to study identifiability, we consider $N_{\boldsymbol{\theta}} = (K-1)N_s N_t + N_s$ and not $N_{\boldsymbol{\theta}} = K N_s N_t$, where K denotes the total number of different kinds of LDP. Consider first a MT that is moving with constant speed. The number of parameters that needs to be estimated is $N_{\mathbf{p}} = 2N_s + 2$. According to Corollary 1 in section 1.10, the first condition that must be met for identifiability is

$$(K-1)N_s N_t + N_s \geq 2N_s + 4 \tag{4.37}$$

The second condition is that the transformation matrix must have rank $N_{\mathbf{p}}$. Instead of computing the rank, we will take a different approach and require at least $N_{\mathbf{p}}$ data equations to be linearly independent under some assumptions for N_t. The only restriction on the number of time samples N_t of each LDP is that its product with Δt, i.e. the total observation time $N_t \Delta t$, should remain sufficiently small so that the constant speed assumption is realistic. So if the sampling rate is quite high, N_t can be quite large too. However, considering a really small observation time during which the LDP vary very little, would result in fewer than N_t linearly independent equations per LDP per path. Specifically, for the constant speed model, the small observation time assumption leads to exactly 2 linearly independent equations per LDP per path. To prove this, consider the Taylor series expansion of any LDP $\theta_{k,i} = \theta_k(\mathbf{p}_i)$ around the parameters' vector at time instant t_0, $\mathbf{p}_0$. Since the observation time is small, we can ignore higher order terms and consider the LDP variation to be linear, i.e.

$$\theta_{k,i} = \theta_{k,0} + (\mathbf{p}_i - \mathbf{p}_0)^t \left. \frac{d\theta_k}{d\mathbf{p}} \right|_{\mathbf{p}=\mathbf{p}_0} \tag{4.38}$$

Table 4.2: $N_{\boldsymbol{\theta}}$ - $N_{\mathbf{p}}$ pairs

$\{N_\theta, N_{\mathbf{p}}\}$		K		
		2	3	4
N_s	1	$\{3,6\}$	$\{5,6\}$	$\{7,6\}$
	2	$\{6,8\}$	$\{10,8\}$	$\{14,8\}$
	3	$\{9,10\}$	$\{15,10\}$	$\{21,10\}$
	4	$\{12,12\}$	$\{20,12\}$	$\{28,12\}$

Furthermore, since 2 entries of $\mathbf{p}$, the MT coordinates, vary linearly with time and the rest remain constant, each $\theta_{k,i}$ is also a linear function of time

$$\theta_{k,i} = \underbrace{\theta_{k,0}}_{c_0} + \underbrace{\left(\upsilon_x \frac{\partial \theta_k}{\partial x} + \upsilon_y \frac{\partial \theta_k}{\partial y} \right)\Big|_{\mathbf{p}=\mathbf{p}_0}}_{c_1} t_i. \tag{4.39}$$

The above set of equations can be written in vector form as

$$\underbrace{\begin{bmatrix} 1 & t_0 \\ \vdots & \vdots \\ 1 & t_{N_t-1} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} \theta_{k,0} \\ \vdots \\ \theta_{k,N_t-1} \end{bmatrix} \tag{4.40}$$

The number of linearly independent equations is equal to the rank of $\mathbf{A}$, which is obviously 2. The values of $N_{\boldsymbol{\theta}}$ and $N_{\mathbf{p}}$ for $N_t = 2$ are given in table 4.2. From that table we can conclude that for the worst case -in terms of identifiability- scenario[2], the unknown parameters can be estimated if all 4 different kinds of LDP are available for just 1 scatterer, or if 3 kinds of LDP are available for 2 distinct scatterers or if just 2 of them are available for 4 distinct scatterers.

In [49], a different scenario in which, N_t time samples lead to N_t linearly independent equations per LDP per path, was considered. Identifiability was investigated in terms of the rank of the FIM (or equivalently of the transformation matrix $\mathbf{G}$). The rank was computed analytically (using Gaussian elimination) and verified by simulations (for reasonable values of $N_t \Delta t$ that

[2]By worst-case scenario we mean the one in which many time samples lead to only 2 linearly independent equations per LDP per path

Table 4.3: Rank of the transformation matrix

$N_s = 1 \Leftrightarrow N_{\mathbf{p}} = 6, N_t \geq 4$			
$\boldsymbol{\theta}$	$[\boldsymbol{\phi}^t, \mathbf{d}^t]^t$	$[\mathbf{d}^t, \mathbf{f}_d^t]^t$	$[\boldsymbol{\phi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$
$rank(\mathbf{G})$	6	6	6
$\boldsymbol{\theta}$	$[\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{d}^t]^t$	$[\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{f}_d^t]^t$	$[\boldsymbol{\psi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$
$rank(\mathbf{G})$	6	5	6

do not lead to singularities). The results are given in table 4.3 for environments with just 1 resolvable path and different subsets of the LDP being available. Comparing these results with the ones for the worst case scenario, we observe that according to Corollary 1, for the general case, identifiability is feasible for all cases of $K = 3$ except 1 and it is even feasible for 2 cases of $K = 2$. However, as will be shown in the next section, in many of these cases the performance can be poor.

If we allow some linear time-variation of the MT speed by introducing a constant acceleration term, the restriction of small observation time can be loosened. However, it is still of great interest to study the worst-case scenario, thus we again consider that the observation time is small enough so that the LDP time-variation can be considered linear and (4.38) holds. We study again identifiability using the number of possibly linearly independent equations. In this scenario, 2 of the entries of of $\mathbf{p}$, the MT coordinates, are a quadratic function of time and 2 more, the speed components, are linear function of time. Therefore, each $\theta_{k,i}$ is also a quadratic function of time

$$\theta_{k,i} = \underbrace{\theta_{k,0}}_{c_0} + \underbrace{\left(v_x \frac{\partial \theta_k}{\partial x} + v_y \frac{\partial \theta_k}{\partial y} + \alpha_x \frac{\partial \theta_k}{\partial v_x} + \alpha_y \frac{\partial \theta_k}{\partial v_y} \right)\Big|_{\mathbf{p}=\mathbf{p}_0}}_{c_1} t_i + \underbrace{\frac{1}{2}\left(\alpha_x \frac{\partial \theta_k}{\partial x} + \alpha_y \frac{\partial \theta_k}{\partial y} \right)\Big|_{\mathbf{p}=\mathbf{p}_0}}_{c_2} t_i^2 \quad . \tag{4.41}$$

The above set of equations can be written in vector form as

$$\underbrace{\begin{bmatrix} 1 & t_0 & t_0^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{N_t-1} & t_{N_t-1}^2 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \theta_{k,0} \\ \vdots \\ \theta_{k,N_t-1} \end{bmatrix} \tag{4.42}$$

Table 4.4: $N_{\boldsymbol{\theta}}$ - $N_{\mathbf{p}}$ pairs

$\{N_\theta, N_{\mathbf{p}}\}$		K		
		2	3	4
N_s	1	$\{4,8\}$	$\{7,8\}$	$\{10,8\}$
	2	$\{8,10\}$	$\{14,10\}$	$\{20,10\}$
	3	$\{12,12\}$	$\{21,12\}$	$\{30,12\}$
	4	$\{16,14\}$	$\{28,14\}$	$\{40,14\}$

The number of linearly independent equations is equal to the rank of $\mathbf{A}$, which in this case is 3. Therefore, introducing acceleration increases not only the number of parameters that needs to be estimated but also the minimum number of possible independent equations. The values of $N_{\boldsymbol{\theta}}$ and $N_{\mathbf{p}} = 2N_s + 4$ for $N_t = 3$ are given in table 4.4. Comparing the $N_{\boldsymbol{\theta}}$-$N_{\mathbf{p}}$ pairs for the 2 mobility model cases, we notice that the only difference in terms of worst-case scenario identifiability is that using the constant acceleration mobility model makes estimation feasible even when $K = 2$ different kinds of LDP are available for $N_s = 3$ distinct scatterers. Of course, for more general scenarios where the number of linearly independent eq. scales with N_t for both methods, there is no advantage in using different models.

4.5 Performance Concerns

In some of the cases studied in the previous section, although the FIM is invertible, it is ill-conditioned. This has a huge impact on the performance of the ML estimators, since the CRB becomes very large (noise is amplified). To avoid scenarios like these, in which the parameters are "ill-identified" since the estimation error will be so large that their estimated value is meaningless, we will compare the condition number of the FIM for different cases and study its impact on the CRB through a numerical example. In this example we consider a picocell so that the distance between the MT and the BS is a few tens of meters ($[x_{BS}, y_{BS}] = [0,0]$, $[x_0, y_0] = [30,20]$) while the MT is moving with average walking speed ($[\upsilon_x, \upsilon_y] = [2, -1.5]$). $N_s = 4$ MPC are available and the scatterers' coordinates are drawn from a uniform distribution with support region $[x_{BS}, x_{BS} + 2x_0] \times [y_{BS}, y_{BS} + 2y_0]$. The total observation time is $1sec$. The results are averaged for 10^3 different samples of scatterers' coordinates. The condition number of the FIM $\boldsymbol{J}$ is

defined as:

$$c_{\boldsymbol{\theta}} = \frac{\lambda_{max}}{\lambda_{min}} \tag{4.43}$$

where λ_{max} (λ_{min}) is the maximum (minimum) eigenvalue of $\boldsymbol{J}$. The CRB for the position and speed are given respectively by

$$CRB_{p,\boldsymbol{\theta}} = \sqrt{tr([\mathbf{J}^{-1}]_{(1:2,1:2)})} \tag{4.44}$$

$$CRB_{sp,\boldsymbol{\theta}} = \sqrt{tr([\mathbf{J}^{-1}]_{(3:4,3:4)})} \quad . \tag{4.45}$$

The FIM $\mathbf{J}$ is defined in eq. (1.77). Since we are not interested in the impact of $\mathbf{C}_{\tilde{\boldsymbol{\theta}}}$ on the performance, we will assume that $\mathbf{C}_{\tilde{\boldsymbol{\theta}}} = \boldsymbol{I}$. Any scaling by a positive scalar does not change the condition number, while if the diagonal entries change slightly, the condition number changes but generally its order of magnitude remains the same. In table 4.5 we give the condition numbers and the CRB for cases of interest, normalized with respect to the same quantities for the case when all LDP are available (method 1), $c_{\boldsymbol{\theta}_{all}}$, $CRB_{p,\boldsymbol{\theta}_{all}}$ and $CRB_{sp,\boldsymbol{\theta}_{all}}$ respectively. All quantities are in dB, i.e. the entries are defined as follows

$$c_n = 10\log_{10}\left(\frac{c_{\boldsymbol{\theta}}}{c_{\boldsymbol{\theta}_{all}}}\right) \tag{4.46}$$

$$CRB_{p,n} = 10\log_{10}\left(\frac{CRB_{p,\boldsymbol{\theta}}}{CRB_{p,\boldsymbol{\theta}_{all}}}\right) \tag{4.47}$$

$$CRB_{sp,n} = 10\log_{10}\left(\frac{CRB_{sp,\boldsymbol{\theta}}}{CRB_{sp,\boldsymbol{\theta}_{all}}}\right) \tag{4.48}$$

The cases shown in table 4.5 are the ones with the best performance over all the cases for which local identifiability is feasible. For the cases not shown in this table, the condition number is increased by at least an order of magnitude and that leads to a similar degradation in performance, when fewer LDP are available. The only exceptions are actually the 3 cases shown in this table. However for 2 of these cases (method 2 and method not utilizing DS) the CRB for both the position and the speed is more than doubled and only for the case when the AoD is not available the degradation is very small.

4.6 Comparison of the 2 proposed methods

In this section we will further evaluate and compare the performance of the 2 methods ($K = 4$ and $K = 3$) for a different scenario than the one considered in the previous section. Specifically we consider a vehicle moving with low

Table 4.5: Condition Number Ratio and CRB Ratio

$N_s = 4, N_t = 10$			
$\boldsymbol{\theta}$	$[\boldsymbol{\phi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$	$[\boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{d}^t]^t$	$[\boldsymbol{\psi}^t, \mathbf{d}^t, \mathbf{f}_d^t]^t$
$c_n(dB)$	1	14.3	4.3
$RMSE_{p,n}(dB)$	0.7	18.7	7.1
$RMSE_{sp,n}(dB)$	0.3	9	3

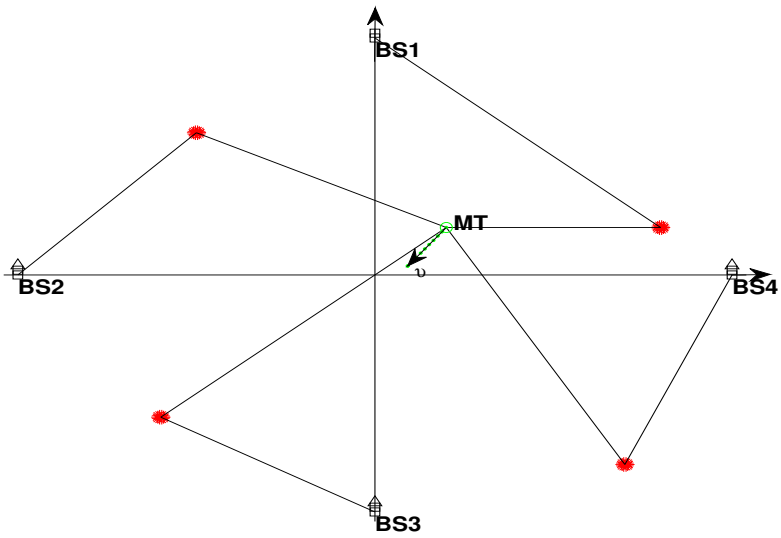

Figure 4.2: NLoS Environment with 4 BS

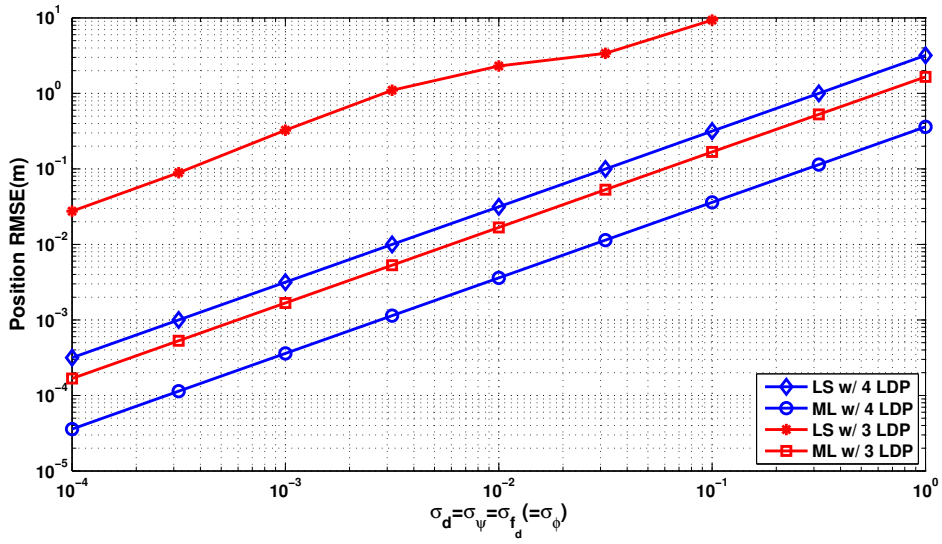

Figure 4.3: Position RMSE vs SNR for various methods

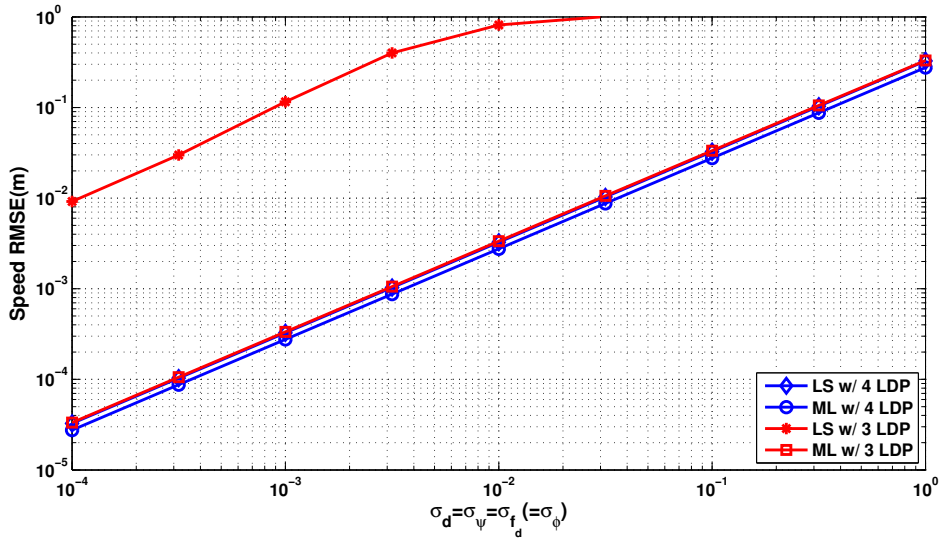

Figure 4.4: Speed RMSE vs SNR for various methods

Table 4.6: BS, MT and scatterers' coordinates

(x_{bs_1}, y_{bs_1})	(x_{bs_2}, y_{bs_2})	(x_{bs_3}, y_{bs_3})	(x_{bs_4}, y_{bs_4})
$(0, 50)$m	$(-50, 0)$m	$(0, -50)$m	$(50, 0)$m
(x_{s_1}, y_{s_1})	(x_{s_2}, y_{s_2})	(x_{s_3}, y_{s_3})	(x_{s_4}, y_{s_4})
$(40, 10)$m	$(-25, 30)$m	$(-30, -30)$m	$(35, -40)$m

constant speed (less than $40km/h$) inside a micro-cell (city environment) and receiving a signal from 4 BS. All of the received signal components propagate through a NLoS environment. The coordinates of the BS and the corresponding scatterers are given in table 4.6 and the environment is shown in figure 4.2. The MT is located at $\{x_0, y_0\} = \{25, 25\}$ and moving with speed $\{u_x, u_y\} = \{6, 9\}m/sec$. The total observation time is $1sec$, during which $N_t = 10$ LDP of each kind and for each path are estimated and used in the localization process. We run N independent experiments and averaged the results. In figures 4.3 and 4.4 we plot the RMSE of the MT position and speed respectively given by

$$RMSE_p = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{x}_{0,k} - x_0)^2 + (\hat{y}_{0,k} - x_0)^2} \tag{4.49}$$

$$RMSE_{sp} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{v}_{x,k} - v_x)^2 + (\hat{v}_{y,k} - v_y)^2} \tag{4.50}$$

versus the standard deviation σ of all LDP estimates (we consider σ to be equal for all LDP just for demonstration purposes). It can be observed that when only 3 LDP are available, the LS estimation for the speed and the position is accurate only for high SNR (defined as $\frac{1}{\sigma^2}$). This is because the method depends on the variation in time of the LDP and is very sensitive to noise. This does not seem to cause the same problem to the ML estimates, although the performance is slightly worse than the $1st$ method, as was also indicated in the previous section using the CRB. As a matter of fact, although the scenario is different, the increase in RMSE due to lack of AoA knowledge that was computed in this section is approximately the same with the increase in the CRB which was computed in the previous section.

Chapter 5

Direct Location Estimation for MIMO-OFDM systems

5.1 Introduction

Direct Location Estimation (DLE) was originally introduced in [69] under the name Direct Position Determination (DPD). As the name indicates, in contrast to the traditional 2-step localization methods, DLE is a method that processes the received signal samples and outputs estimates of the MT location (and possibly other parameters) directly, without explicitly estimating the location dependent parameters (LDP). The performance of 2-step approaches, has been proved to converge to the CRB for high SNR and sufficient number of data samples [70]. However, in existing wireless communication systems, high SNR is not always guaranteed. Furthermore, if the channel varies rapidly, the number of data samples that can be used in the estimation process is very limited. The motivation behind DLE was exactly to find a method that can localize efficiently even under these circumstances. DLE is able to achieve that by taking advantage of the fact that all LDP estimates correspond to the same MT location.

In the aforementioned contributions, DLE was applied to LoS static environments, where the narrow-band signal transmitted by the MT is intercepted by L BS. The only disadvantage of the method compared to 2-step methods was the increase in feedback information. If DLE is used, sampled

versions of the received signals from the various BS need to be transmitted to a central processing unit, while if traditional methods are used, only the LDP estimates need to be transmitted. The DLE method was further extended in [71] to include calibration errors and errors due to multipath propagation. We on the other hand, have proposed and implemented DLE for MIMO-OFDM systems that operate in strictly NLoS or multipath environments, but instead of considering errors due to the multipath components, we exploit the information contained in them. Furthermore, we assume that the MT is communicating only with 1 BS, thus not only there is no need for feeding information to a central unit, but also we overcome the so-called hearibility problem, which has been a major concern for localization algorithms [3].

DLE becomes feasible if a unique invertible mapping exists between the parameters of interest (MT position, speed etc) and the LDP. Such a mapping is simple to derive when a LoS signal component exists in the received signal, thus, this assumption was made in the existing literature. In urban environments a LoS signal component rarely exists. When it does exist, the received signal is also composed of various MPC. Most often however, the signal propagates in a rich scattering environment under strict NLoS conditions. In order to express all the LDP that the MIMO channel matrix depends on, i.e. the AoA, the AoD, the delays and the DS, as a function of the MT's coordinates, an appropriate geometrical representation of the propagation environment is required. To that end, we based our approach on the Single-Bounce Model (SBM), which we also employed in the 2-step localization methods presented in the previous chapters. It is due to this widely acceptable channel model, that we were able to create the mapping needed to implement the DLE method.

5.2 DLE for NLoS environments

In this section we assume that the signal propagates in a strictly NLoS environment that changes due to the movement of the MT and that each signal component has bounced exactly once. Therefore the model described in section 1.5.2 can be applied here. We are interested in estimating jointly the MT's coordinates at time t_0, x_0 and y_0 and its constant speed components υ_x and υ_y directly from the received signal matrices $\mathbf{Y}_{kl}$. The mobility model with constant acceleration is not considered here. For ease of notation we express the mapping of the vector $\mathbf{p}$ to the LDP vector $\boldsymbol{\theta}$, along with its

unique inverse, in a more compact way:

$$\boldsymbol{\theta} = \boldsymbol{\theta}(\mathbf{p}) \ , \ \mathbf{p} = \boldsymbol{\theta}^{-1}(\boldsymbol{\theta}(\mathbf{p})). \tag{5.1}$$

As already mentioned in the introduction, this mapping allows us to do a direct position and speed estimation based on the received signal. Let $\mathbb{S}_{\mathbf{Y}} = \{\mathbf{Y}_{11}, \ldots, \mathbf{Y}_{N_f N_t}\}$ and $\mathbb{S}_{\mathbf{H}} = \{\mathbf{H}_{11}, \ldots, \mathbf{H}_{N_f N_t}\}$ be the set of all received signal matrices and the set of the corresponding channel matrices respectively. To implement ML estimation, we further define the following log-likelihood:

$$\mathcal{L} \stackrel{\Delta}{=} \mathcal{L}(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}) = \ln(f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p})). \tag{5.2}$$

Using (5.1), $f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p})$ becomes equivalent to $f(\mathbb{S}_{\mathbf{Y}}|\boldsymbol{\theta})$ and thus can be derived as follows:

$$\begin{aligned}
f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}) &\equiv f(\mathbb{S}_{\mathbf{Y}}|\boldsymbol{\theta}) = \int_{\mathbb{C}^{N_s}} f(\mathbb{S}_{\mathbf{Y}}|\boldsymbol{\theta}, \boldsymbol{\gamma}) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} \\
&= \int_{\mathbb{C}^{N_s}} f(\mathbb{S}_{\mathbf{Y}}|\mathbb{S}_{\mathbf{H}}) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} = \int_{\mathbb{C}^{N_s}} \prod_{k=1}^{N_f} \prod_{l=1}^{N_t} f(\mathbf{Y}_{kl}|\mathbf{H}_{kl}) f(\boldsymbol{\gamma}) d\boldsymbol{\gamma} \\
&= \int_{\mathbb{C}^{N_s}} \prod_{k=1}^{N_f} \prod_{l=1}^{N_t} \left(\frac{1}{(\pi\sigma^2)^{n_r n_t}} e^{-\frac{1}{\sigma^2}|\mathbf{Y}_{kl} - \mathbf{H}_{kl}\mathbf{X}_{kl}|^2} \right) \frac{1}{(\pi\sigma_\gamma^2)^{N_s}} e^{-\frac{1}{\sigma_\gamma^2}\boldsymbol{\gamma}^\dagger \boldsymbol{\gamma}} d\boldsymbol{\gamma} \\
&= \int_{\mathbb{C}^{N_s}} \left(\frac{1}{(\pi\sigma^2)^{n_r n_t}} e^{-\frac{1}{\sigma^2} \sum_{k=1}^{N_f} \sum_{l=1}^{N_t} |\mathbf{Y}_{kl} - \mathbf{H}_{kl}\mathbf{X}_{kl}|^2} \right) \frac{1}{(\pi\sigma_\gamma^2)^{N_s}} e^{-\frac{1}{\sigma_\gamma^2}\boldsymbol{\gamma}^\dagger \boldsymbol{\gamma}} d\boldsymbol{\gamma}
\end{aligned} \tag{5.3}$$

where we have used the io relationship given by eq. (1.66). The solution of the above integral for $\sigma_\gamma^2 = 1$ and $\mathbf{X}_{k,l} = \mathbf{I}_{n_t}$, $\forall k, l$ was derived in [72] and is given in the next chapter. The extension to the more general case is trivial and the result is given below (given $\mathbf{p}$, each $\mathbf{Y}_{k,l}$ is linear in $\boldsymbol{\gamma}$ and $\mathbf{N}_{k,l}$, which are both Gaussian, hence $f(\mathbf{Y}_{kl}|\mathbf{p})$ is Gaussian distributed and $\mathbb{S}_{\mathbf{Y}}$ will also be Gaussian distributed):

$$f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}) \propto \det((\sigma_\gamma^2 \mathbf{V}\mathbf{V}^\dagger + \sigma^2 \mathbf{I})^{-1}) e^{-\mathbf{y}^\dagger (\sigma_\gamma^2 \mathbf{V}\mathbf{V}^\dagger + \sigma^2 \mathbf{I})^{-1} \mathbf{y}} \tag{5.4}$$

where $\mathbf{y}$ is a $N n_r N_f N_t \times 1$ vector containing all the received signal samples, i.e.

$$\mathbf{y} = [\mathbf{y}_{11}^t, \ldots, \mathbf{y}_{N_f N_t}^t]^t \tag{5.5}$$

and $\mathbf{V}$ is a $N n_r N_f N_t \times N_s$ block matrix given by

$$\mathbf{V} = [\mathbf{V}_{11}^\dagger, \ldots, \mathbf{V}_{N_f N_t}^\dagger]^\dagger \tag{5.6}$$

Each $Nn_r \times N_s$ submatrix $\mathbf{V}_{kl}$ is given by

$$\mathbf{V}_{kl} = (\mathbf{X}_{kl}^t \otimes \mathbf{I}_{\mathbf{n_r}}) \underbrace{(\mathbf{A}_{T,l} \boxtimes \mathbf{A}_{R,l})\mathbf{D}_{kl}}_{\mathbf{Q}_{kl}} \tag{5.7}$$

To simplify notation lets introduce the block diagonal matrix $\boldsymbol{\mathcal{X}}$ that depends only on the training sequence of symbols

$$\boldsymbol{\mathcal{X}} = \begin{bmatrix} (\mathbf{X}_{11}^t \otimes \mathbf{I}_{\mathbf{n_r}}) & \mathbf{0} & \dots \\ \mathbf{0} & \ddots & \mathbf{0} \\ \dots & \mathbf{0} & (\mathbf{X}_{N_f N_t}^t \otimes \mathbf{I}_{\mathbf{n_r}}) \end{bmatrix} \tag{5.8}$$

and the matrix $\mathbf{Q}$

$$\mathbf{Q} = [\mathbf{Q}_{11}^t, \dots, \mathbf{Q}_{N_f N_t}^t]^t. \tag{5.9}$$

We can then write

$$\mathbf{V} = \boldsymbol{\mathcal{X}}\mathbf{Q} \tag{5.10}$$

Lets further define the conditional covariance matrix of the data vector $\mathbf{y}$ as

$$\mathbf{C}_{\mathbf{y}|\mathbf{p}} \triangleq \sigma_\gamma^2 \mathbf{V}\mathbf{V}^\dagger + \sigma^2 \mathbf{I} \tag{5.11}$$

Since $\mathbf{A}_{T,l}$,$\mathbf{A}_{R,l}$ and $\mathbf{D}_{G,kl}$ depend on $\mathbf{p}$, so does $\mathbf{V}$. Therefore $\mathbf{C}_{\mathbf{y}|\mathbf{p}}$ also depends on the parameters we need to estimate, although this dependency is not explicitly shown in (5.11). Substituting (5.11) in (5.4) and the result in (5.2) we get

$$\mathcal{L} = -\ln(\det(\mathbf{C}_{\mathbf{y}|\mathbf{p}})) - \mathbf{y}^\dagger \mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1} \mathbf{y}. \tag{5.12}$$

The ML estimate of $\mathbf{p}$ is then simply given by

$$\widehat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmax}}\{\mathcal{L}\} \ . \tag{5.13}$$

5.3 DLE for Multipath Environments

In the presence of a LoS component, the MIMO channel matrix is given as a sum of 2 components. The NLoS component depends on the unknown complex amplitudes γ, which were assigned a complex Gaussian distribution according to the principle of Maximum Entropy and thus were easily integrated out in the previous section. In a similar fashion, to make the channel representation as realistic as possible, we introduced in eq. (1.64) the phase noise term $e^{j\theta}$. Therefore, in a multipath environment we are faced again

with the dilemma of integrating out this new unknown parameter or estimating it jointly with the rest of the unknown parameters. In this approach we choose joint estimation, since integration does not lead to a closed-form solution. Due to this choice, in this problem formulation we need to estimate $\mathbf{p}_{int}$ in the presence of nuisance parameters $\mathbf{p}_{nuis} = [\mathbf{x}_s^t, \mathbf{y}_s^t, \theta]^t$, i.e. our goal becomes to estimate the $(2N_s + 5) \times 1$ vector:

$$\mathbf{p} = [\mathbf{p}_{int}^t, \mathbf{p}_{nuis}^t]^t. \tag{5.14}$$

Under the Bayesian framework [45] and having the principle of Maximum Entropy as a guiding rule, we can assign uniform probability density to the phase θ. Thus, the use of these non-informative priors will lead to no improvement in the accuracy of the estimation method, and a Maximum a-Posteriori (MAP) estimator becomes equivalent to a ML estimator, which is also implemented for this environment. To formulate the ML estimation problem more precisely, let us introduce the following vectors:

$$\mathbf{h}_L = [\mathbf{h}_{L,11}^t, \ldots, \mathbf{h}_{L,N_f N_t}^t]^t \tag{5.15}$$

$$\mathbf{m}_\mathbf{y} = \boldsymbol{\mathcal{X}} \mathbf{h}_L \tag{5.16}$$

$$\mathbf{y}' = \mathbf{y} - \mathbf{m}_\mathbf{y} \tag{5.17}$$

We can directly apply the results of the previous section on the centered random vector $\mathbf{y}'$ to get the new log-likelihood

$$\mathcal{L} = -\ln(\det(\mathbf{C}_{\mathbf{y}|\mathbf{p}})) - (\mathbf{y} - \mathbf{m}_\mathbf{y})^\dagger \mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1} (\mathbf{y} - \mathbf{m}_\mathbf{y}). \tag{5.18}$$

The ML estimate of $\mathbf{p}$ is once given by maximizing the above function, i.e.

$$\widehat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmax}}\{\mathcal{L}\} \ . \tag{5.19}$$

5.4 Cramer-Rao Bound

Since in a multipath environment both the conditional mean and the conditional covariance on the r.h.s of (5.18) depend on the unknown parameters, using the definition of the FIM given by the first equality of eq. (1.77), will lead to a sum of 2 terms for each entry. The solution for the i', j' entry is given by [73, eq.(8.34)]:

$$J_{i'j'} = tr\Big\{ \mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1} \frac{\partial \mathbf{C}_{\mathbf{y}|\mathbf{p}}}{\partial p_{i'}} \mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1} \frac{\partial \mathbf{C}_{\mathbf{y}|\mathbf{p}}}{\partial p_{j'}} \Big\} + 2Re\Big\{ \frac{\partial \mathbf{m}_\mathbf{y}^\dagger}{\partial p_{i'}} \mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1} \frac{\partial \mathbf{m}_\mathbf{y}}{\partial p_{j'}} \Big\}. \tag{5.20}$$

If a strictly NLoS environment is considered, the second term on the r.h.s of (5.20) is equal to zero. The partial derivative of the conditional covariance matrix is

$$\frac{\partial \mathbf{C}_{\mathbf{y}|\mathbf{p}}}{\partial p_{i'}} = \boldsymbol{\mathcal{X}} \left(\frac{\partial \mathbf{Q}}{\partial p_{i'}} \mathbf{Q}^{\dagger} + \mathbf{Q} \frac{\partial \mathbf{Q}^{\dagger}}{\partial p_{i'}} \right) \boldsymbol{\mathcal{X}}^{\dagger} \tag{5.21}$$

and $\frac{\partial \mathbf{Q}}{\partial p_{i'}}$ is constructed by concatenating the following submatrices:

$$\frac{\partial \mathbf{Q}_{kl}}{\partial p_{i'}} = (\frac{\partial \mathbf{A}_{T,l}}{\partial p_{i'}} \boxtimes \mathbf{A}_{R,l} + \mathbf{A}_{T,l} \boxtimes \frac{\partial \mathbf{A}_{R,l}}{\partial p_{i'}}) \mathbf{D}_{kl} + (\mathbf{A}_{T,l} \boxtimes \mathbf{A}_{R,l}) \frac{\partial \mathbf{D}_{kl}}{\partial p_{i'}}. \tag{5.22}$$

The partial derivatives of $\mathbf{A}_{T,l}$ and $\mathbf{A}_{R,l}$ are given by

$$\frac{\partial \mathbf{A}_{T,l}}{\partial p_{i'}} = \mathbf{1}_{\{2j+3,2j+4\}}(i') [\mathbf{0}_{n_t \times (j-1)}, \frac{\partial \mathbf{a}_T(\psi_j)}{\partial \psi_j}, \mathbf{0}_{n_t \times (N_s - j)}] \frac{\partial \psi_j}{\partial p_{i'}} \tag{5.23}$$

$$\frac{\partial \mathbf{A}_{R,l}}{\partial p_{i'}} = \mathbf{1}_{\{1,\ldots,4,2j+3,2j+4\}}(i') [\mathbf{0}_{n_r \times (j-1)}, \frac{\partial \mathbf{a}_R(\phi_{lj})}{\partial \phi_{lj}}, \mathbf{0}_{n_r \times (N_s - j)}] \frac{\partial \phi_{lj}}{\partial p_{i'}} \tag{5.24}$$

while the partial derivative of $\mathbf{D}_{kl}$ is given by

$$\frac{\partial \mathbf{D}_{kl}}{\partial p_{i'}} = \mathbf{1}_{\{1,\ldots,4,2j+3,2j+4\}}(i') \left[-\frac{1}{\tau_{lj}} \frac{\partial \tau_{lj}}{\partial p_{i'}} + j2\pi \left(t_l \frac{\partial f_{d,lj}}{\partial p_{i'}} - f_k \frac{\partial \tau_{lj}}{\partial p_{i'}} \right) \right] \mathbf{D}_{kl} \tag{5.25}$$

Similarly, the partial derivative of the conditional mean is

$$\frac{\partial \mathbf{m}_{\mathbf{y}}}{\partial p_{i'}} = \boldsymbol{\mathcal{X}} \frac{\partial \mathbf{h}_L}{\partial p_{i'}} \tag{5.26}$$

and $\frac{\partial \mathbf{h}_L}{\partial p_{i'}}$ is constructed by concatenating the following subvectors:

$$\begin{aligned} \frac{\partial \mathbf{h}_{L,kl}}{\partial p_{i'}} =& \mathbf{1}_{\{2Ns+5\}}(i') j\theta \mathbf{h}_{L,kl} + \mathbf{1}_{\{1,\ldots,4\}}(i') \\ & \times \left[\left(-\frac{1}{\tau_{l0}} \frac{\partial \tau_{l0}}{\partial p_{i'}} + j2\pi \left(t_l \frac{\partial f_{d,l0}}{\partial p_{i'}} - f_k \frac{\partial \tau_{l0}}{\partial p_{i'}} \right) \right) \mathbf{h}_{L,kl} \right. \\ & \left. + \left(\frac{\partial \mathbf{a}_T(\psi_{l0})^t}{\partial \psi_{l0}} \otimes \mathbf{I}_{n_r} \right) \mathbf{a}_R(\phi_{l0}) \frac{\partial \psi_{l0}}{\partial p_{i'}} + (\mathbf{a}_{T,l0}^t \otimes \mathbf{I}_{n_r}) \frac{\partial \mathbf{a}_R(\phi_{l0})}{\partial \phi_{l0}} \frac{\partial \phi_{l0}}{\partial p_{i'}} \right]. \end{aligned} \tag{5.27}$$

In the above equations we introduced the indicator function, defined as:

$$\mathbf{1}_{\mathbb{S}}(i') \triangleq \begin{cases} 1, & i' \in \mathbb{S} \\ 0, & i' \notin \mathbb{S} \end{cases} \tag{5.28}$$

Table 5.1: BS, MT and scatterers' coordinates

(x_{bs}, y_{bs})	(x_0, y_0)	(x_{s_1}, y_{s_1})	(x_{s_2}, y_{s_2})	(x_{s_3}, y_{s_3})
$(0,0)$m	$(30,20)$m	$(20,30)$m	$(5,5)$m	$(40,15)$m

The need for this indicator stems from the fact that all LDP of a MPC j depend only on the corresponding scatterer j. The partial derivatives of the LDP with respect to the entries of $\mathbf{p}$ can be found in app. I. Deriving $\frac{\partial \mathbf{a}_T(\psi_j)}{\partial \psi_j}$ and $\frac{\partial \mathbf{a}_R(\phi_{lj})}{\partial \phi_{lj}}$ is trivial once the geometry of the antenna arrays is known. For the case of ULA, the solution will be given in the results section. Constructing the conditional mean and covariance according to eq. (5.26) and (5.21) and substituting the result in (5.20) we get an expression for the FIM which is valid for any geometry of the arrays at the transmitter and the receiver.

5.5 Simulation Results

In this section we compute and we plot the CRB for three different environment: A LoS environment (information from NLoS components is either not available or not used in the estimation process), a multipath environment with 2 NLoS and a LoS path and a strictly NLoS environment with 3 paths. The power normalization constant ensures that the channel's energy remains the same independently of the environment or the number of available paths. The coordinates of the BS, the MT and the scatterers considered, correspond to a picocell that fits the elliptical model [61, 64, 65] and are given in table 5.1. The magnitude of the speed of the MT is $|\upsilon| = 1.5$m/sec (average walking speed) and we average the simulation results for 20 different directions of the speed, drawn independently from a uniform distribution with support region $[0, 2\pi]$. The $N_t = 40$ time samples are uniformly spaced and $t_{tot} = t_{N_t-1} - t_0$ is 100msec. Also only $N_f = 2$ subcarriers are considered with spacing of $\Delta_f = 10MHz$ and the carrier frequency is $f_c = 1.9GHz$. The transmitted signal is the training matrix $\mathbf{X}_{kl} = \mathbf{I}_{n_t},\ \forall k, l$. The array response of the receiver's ULA to signal component j arriving at time l, is

$$\mathbf{a}_R(\phi_{lj}) = [1, e^{j2\pi \frac{f_c}{c} d_r \sin(\phi_{lj})}, \ldots, e^{j2\pi \frac{f_c}{c} d_r (n_r-1) \sin(\phi_{lj})}]^t \tag{5.29}$$

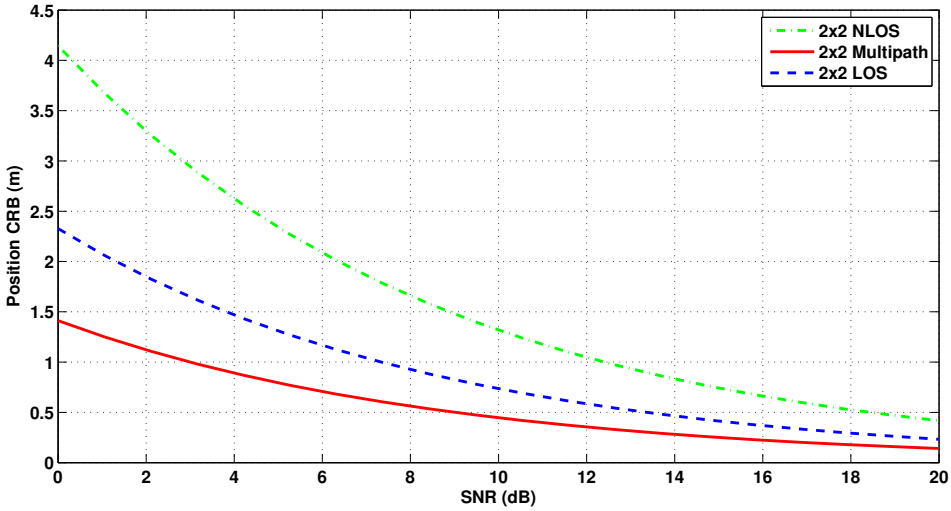

Figure 5.1: Position CRB vs SNR, various environments

and its partial derivative with respect to ϕ_{lj} is

$$\frac{\partial \mathbf{a}_R}{\partial \phi_{lj}} = j2\pi \frac{f_c}{c} d_r \cos(\phi_{lj})[0, 1, \ldots, (n_r - 1)]^t \odot \mathbf{a}_R(\phi_{lj}) \tag{5.30}$$

where d_r is the distance between two adjacent antenna elements. Replacing d_r with d_t, ϕ with ψ and n_r with n_t, we get the transmitter's array response (and the corresponding derivative). In our simulations we considered $d_r = d_t = \lambda/2$. In figures 5.1 and 5.2 we plot the position and speed root mean square error (RMSE) respectively for an efficient estimator (i.e. the CRB), versus the received SNR for a 2×2 MIMO system. The SNR is defined as:

$$SNR = 10 \log_{10} \left(\frac{E\{tr(\mathbf{H}\mathbf{X}\mathbf{X}^\dagger \mathbf{H}^\dagger)\}}{E\{tr(\mathbf{N}\mathbf{N}^\dagger)\}} \right) = 10 \log_{10} \left(\frac{\sigma_\gamma^2}{\sigma^2} \right) \tag{5.31}$$

where we introduced the matrices

$$\mathbf{H} = [\mathbf{H}_{11}, \ldots, \mathbf{H}_{N_f N_t}] \tag{5.32}$$

$$\mathbf{X} = [\mathbf{X}_{11}^t, \ldots, \mathbf{X}_{N_f N_t}^t]^t \tag{5.33}$$

$$\mathbf{N} = [\mathbf{N}_{11}, \ldots, \mathbf{N}_{N_f N_t}]. \tag{5.34}$$

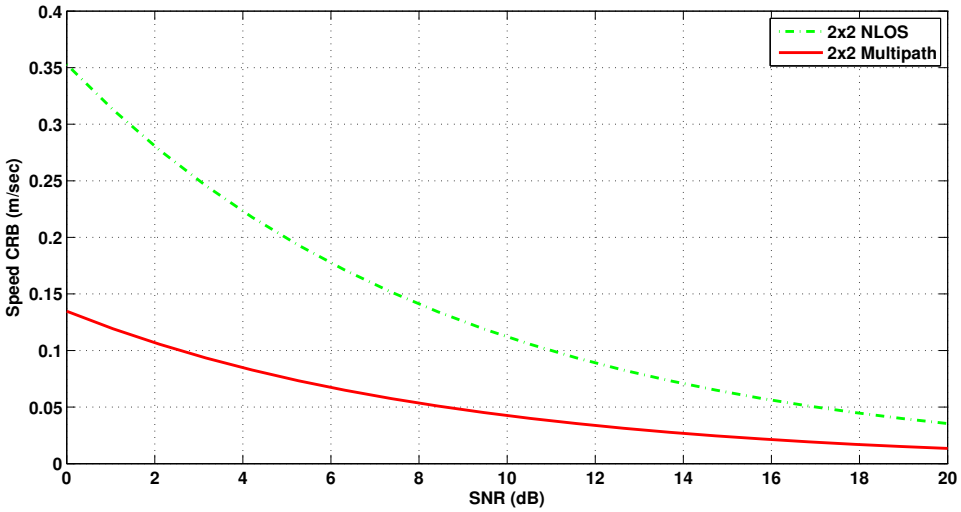

Figure 5.2: Speed CRB vs SNR, various environments

The position and speed RMSE are defined as:

$$RMSE_p = \sqrt{\sigma^2_{\widetilde{x_0}} + \sigma^2_{\widetilde{y_0}}} = \sqrt{tr([\mathbf{J}^{-1}]_{(1:2,1:2)})} \tag{5.35}$$

$$RMSE_{sp} = \sqrt{\sigma^2_{\widetilde{v_x}} + \sigma^2_{\widetilde{v_y}}} = \sqrt{tr([\mathbf{J}^{-1}]_{(3:4,3:4)})} \tag{5.36}$$

It can be noticed that the estimation error is very small even for a strictly NLoS environment. Moreover, if the NLoS signal components are considered along with the LoS component, the position RMSE is significantly reduced (e.g. 40% at $10dB$) and speed estimation becomes feasible. In figure 5.3 the effect of increasing the number of antennas on position accuracy is depicted, for the Multipath environment only. For a MISO system, $RMSE_p < 1m$ for $SNR > 11dB$, while a 2×2 system can achieve the same accuracy with an SNR of $3dB$. The effect is similar for the other two environments, however position is not identifiable for a MISO system in a NLoS environment.

Finally in figure 5.4, the position's RMSE as a function of the MT speed is plotted. It can be seen that the movement of the MT has a huge impact on localization accuracy, especially for the NLoS environment, where the error is reduced by more than 50% when the speed is increased to $2m/sec$, largely independently of the direction of movement. In the figures we can notice that for the 2×4 system, $RMSE_p$ is less than $1m$ and $RMSE_{sp}$ is

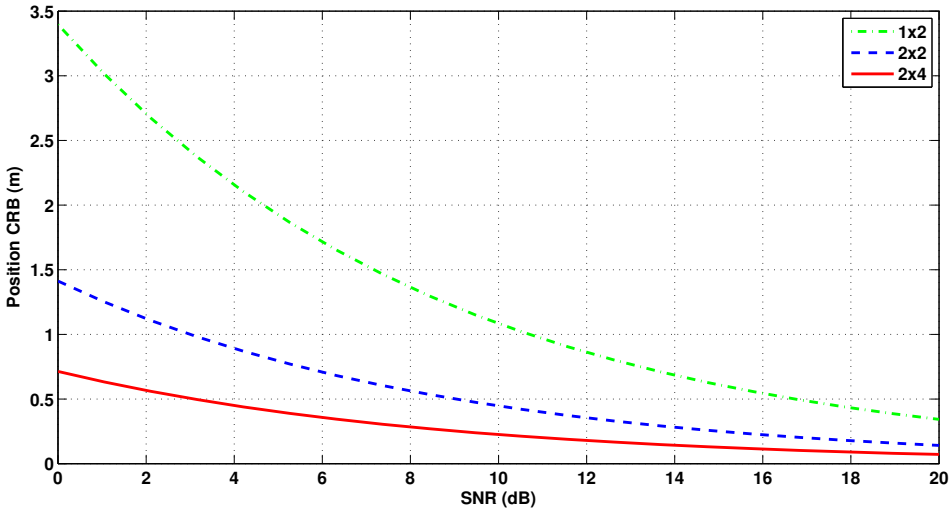

Figure 5.3: Position CRB vs SNR, various systems

less than $0.1m/sec$ for $SNR \geq 5.5dB$. The great enhancement in performance due to the increase in transmitting antennas (while keeping the same transmitting power) is obvious. Approximately the same enhancement can be alternatively achieved by doubling the time samples N_t, instead of n_t.

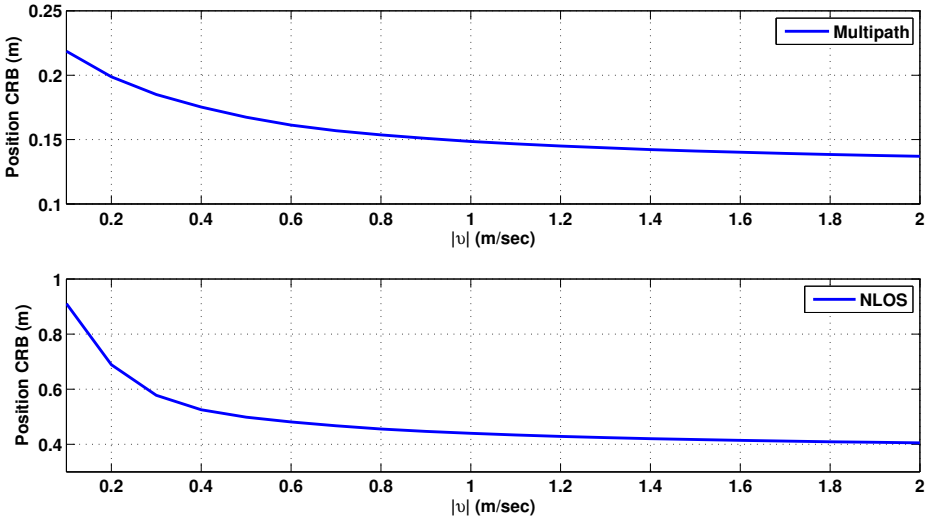

Figure 5.4: Position CRB vs Speed Magnitude

Chapter 6

Mobile Terminal Tracking for MIMO-OFDM systems

6.1 Introduction

In contrast to the localization methods presented in the previous chapters, the algorithm presented herein estimates solely the magnitude and direction of speed[1], thus the characterization "tracking". Moreover, this algorithm is quite different from the aforementioned methods, since it utilizes double directional model (DDM) in its more generic version, as described in section 1.5.1. The major advantage of using the DDM is that no assumption on the scattering environment (number of bounces) is made. The major disadvantage is that many new nuisance parameters appear in this channel representation (see eq. (6.1) below) and on top of that, it is not possible to express all of these nuisance parameters explicitly as a function of the MT location. This is exactly the reason why the proposed algorithm is used to estimate only the speed vector.

$$\mathbf{H}_{n_r \times n_t}(f,t) = \frac{1}{\sqrt{s_r s_t}} \mathbf{\Phi}_{n_r \times s_r}(t) \mathbf{P_r} (\mathbf{\Theta}_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f)) \mathbf{P_t} \mathbf{\Psi}_{s_t \times n_t}(t) \quad (6.1)$$

Although this algorithm is presented last, it was our first attempt to solve the NLoS localization problem. It was originally published in [72]. As will be

[1] Speed components are assumed to be constant.

shown in the simulation results section, this algorithm requires multiple antennas at both ends (preferably more than 2), in order to exploit thoroughly the space dimension and overcome the difficulty of estimating the speed vector in the presence of numerous nuisance parameters. MIMO systems have attracted much interest from an information-theoretic perspective, as it was proved that they increase the capacity linearly with the number of antennas [74]. This work illustrates the huge advantage of MIMO when used for MT localization. Specifically, antenna arrays have been extensively used in the past for estimating the AoA and/or the AoD. This utility of MIMO systems will play a key role in the proposed tracking method.

6.2 Bayesian Estimation of the Speed Vector

We consider herein the estimation of the magnitude υ_r and the direction ω_r of the MT speed via the transmission of a training sequence. The training sequence transmitted by the BS and received by the MT, consists of a set of N orthogonal sub-vectors of size n_t each and is assumed to be transmitted within the channel's coherence time T[2]. Each sub-vector is given by $\mathbf{x}_l = [\mathbf{0}^t_{(l-1)}, x, \mathbf{0}^t_{(n_t-l)}]^t$, $l = 1, \ldots, N$, where x is the training symbol which is known at the receiver. This training sequence was derived in [75] and was proved to maximize the capacity, when used for channel estimation in MIMO-OFDM systems with cyclic prefix.

The discrete-time model describing the io relationship of a time-variant frequency-selective MIMO channel is given by eq. (1.65), which we rewrite here for reference:

$$\mathbf{Y}(f_k, t_l) = \mathbf{H}(f_k, t_l)\mathbf{X}(f_k, t_l) + \mathbf{N}(f_k, t_l). \tag{6.2}$$

Without loss of generality we will choose $N = n_t$ and the training symbol $x = 1$ so that the input training matrix is $\mathbf{X} = \mathbf{I}_{n_t}$. This reduces (6.2) to the simpler form:

$$\mathbf{Y}(f_k, t_k) = \mathbf{H}(f_k, t_l) + \mathbf{N}(f_k, t_l). \tag{6.3}$$

To formulate an estimator, we define first the joint conditional density of all

[2]We assume that the random matrices composing $\mathbf{H}$, do not change within T and choose to send a set of orthogonal training sub-vectors $\mathbf{x}$ that span the column space of the MIMO matrix during this interval.

received matrices

$$
\begin{aligned}
&f(\{\mathbf{Y}_{11}\dots\mathbf{Y}_{N_fN_t}\}|\{\mathbf{H}_{11}\dots\mathbf{H}_{N_fN_t}\}) \stackrel{\Delta}{=} f(\mathbb{S}_{\mathbf{Y}}|\mathbb{S}_{\mathbf{H}}) \\
&= \prod_{k=1}^{N_f}\prod_{l=1}^{N_t}\frac{1}{(\pi\sigma^2)^{n_rn_t}}e^{-\frac{|\mathbf{Y}(f_k,t_l)-\mathbf{H}(f_k,t_l)|^2}{\sigma^2}} \\
&= \prod_{k=1}^{N_f}\prod_{l=1}^{N_t}\prod_{p=1}^{n_t}\prod_{q=1}^{n_r}\frac{1}{(\pi\sigma^2)}e^{-\frac{|y_{p,q}(f_k,t_l)-h_{p,q}(f_k,t_l)|^2}{\sigma^2}}
\end{aligned}
\tag{6.4}
$$

Using this density, an ML estimator for the speed vector $\mathbf{p}_\upsilon = [\upsilon_r, \omega_r]^t$, will be formulated. The ML estimator is actually equivalent to the Maximum a-Posteriori (MAP) estimator, if uniform a-priori distributions are assigned to the speed and the direction. Let $\mathbf{p}_{int}$ denote the parameters of interest and let $\mathbf{p}_{nuis}$ be the vector of all the rest unknown parameters in $\mathbf{H}$ (nuisance parameters). These two vectors will be explicitly defined later. For the time being it suffices to state that $\mathbf{p}_{nuis} = [\boldsymbol{\theta}, \tilde{\mathbf{p}}_{nuis}]$, where $\boldsymbol{\theta} = vec(\boldsymbol{\Theta})$ and $\boldsymbol{\Theta}$, along with all the other matrices composing each $\mathbf{H}(f_k, t_l)$, are defined in section 1.5.1. Under the Bayesian framework [45], we can infer on the a-priori distributions of all the random variables composing the entries of $\mathbf{H}$, therefore also composing $\mathbf{p}_{nuis}$. Specifically the pmf of the number of scatterers $s_{rt} \in \{s_r, s_t\}$ at both sides will be $P[s_{rt} = S_{rt}] = \frac{1}{N_fN_t}$. This choice is justified by the fact that using N_t time samples and N_f frequency samples, at most N_fN_t scatterers can be identified. The priors of the continuous random variables $\forall q = 1,\dots,n_r, i = 1,\dots,s_r, j = 1,\dots,s_t, p = 1,\dots,n_t$, are given below:

$$\omega_r, \omega_t, \phi_{qi}, \beta^r_{qi}, \psi_{jp}, \beta^t_{jp} \sim \mathcal{U}[0, 2\pi] \tag{6.5}$$

$$\upsilon_r \sim \mathcal{U}[0, \upsilon^r_{max}]\ ,\ \upsilon_t \sim \mathcal{U}[0, \upsilon^t_{max}] \tag{6.6}$$

$$\tau_{ij} \sim \mathcal{U}[0, \tau_{max}]\ ,\ \theta_{ij} \sim \mathcal{CN}(0, 1) \tag{6.7}$$

$$p^r_{ii} \sim \mathcal{U}[0, p^r_{max}]\ ,\ p^t_{jj} \sim \mathcal{U}[0, p^t_{max}] \tag{6.8}$$

The fact that most parameters are uniformly distributed stems from the fact that no assumptions on the environment have been made. The uniform distribution is the least-informative one and thus expresses this lack of knowledge of the propagation setting. The complex Gaussian distribution for the θ, was derived in [42] using the principle of maximum entropy (ME) and assuming knowledge of the 2^{nd} order statistics of the channel. According to ML estimation,

$$\hat{\mathbf{p}}_{int} = \underset{\mathbf{p}_{int}}{\operatorname{argmax}}\, f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) \tag{6.9}$$

where $f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int})$ is the joint density of all the received data matrices $\mathbf{Y}(f_k, t_l)$ conditioned only on the parameters of interest. We can obtain this density by marginalizing over all the nuisance parameters, according to:

$$f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) = \int_{\mathcal{A}} f(\mathbb{S}_{\mathbf{Y}}|\mathbb{S}_{\mathbf{H}}) \cdot f(\mathbf{p}_{nuis}) d\mathbf{p}_{nuis} \tag{6.10}$$

since $f(\mathbb{S}_{\mathbf{Y}}|\mathbb{S}_{\mathbf{H}}) = f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}, \mathbf{p}_{nuis})$. Let $\mathcal{A}$ denote the vector space where the vector $\mathbf{p}_{nuis}$ lies. $f(\mathbf{p}_{nuis})$ is the product of the a-priori densities of all nuisance parameters, some of which are conditioned on s_r, s_t or both. Since all the parameters, except $\boldsymbol{\theta}$, are uniformly distributed, we can write:

$$f(\mathbf{p}_{nuis}) = O(s_r, s_t) e^{-\boldsymbol{\theta}^\dagger \boldsymbol{\theta}} \tag{6.11}$$

We can proceed by marginalizing over the Gaussian vector $\boldsymbol{\theta}$ as follows:

$$\begin{aligned}
f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) &= \int_{\tilde{\mathcal{A}}} \int_{\mathcal{C}^{s_r s_t}} f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}, \mathbf{p}_{nuis}) f(\mathbf{p}_{nuis}) d\boldsymbol{\theta} d\tilde{\mathbf{p}}_{nuis} \\
&= \int_{\tilde{\mathcal{A}}} \int_{\mathcal{C}^{s_r s_t}} \prod_{k,l,p,q} \left(\frac{1}{(\pi\sigma^2)} e^{-\frac{|y_{p,q}(f_i,t_j) - h_{p,q}(f_i,t_j)|^2}{\sigma^2}} \right) O(s_r, s_t) e^{-\boldsymbol{\theta}^\dagger \boldsymbol{\theta}} d\boldsymbol{\theta} d\tilde{\mathbf{p}}_{nuis} \\
&= \int_{\tilde{\mathcal{A}}} \int_{\mathcal{C}^{s_r s_t}} \prod_{m=1}^{M} \left(\frac{1}{(\pi\sigma^2)} e^{-\frac{|y_m - \sum_{s=1}^{S} c_{m,s}\theta_s|^2}{\sigma^2}} \right) e^{-\sum_{s=1}^{S} |\theta_s|^2} d\boldsymbol{\theta}\, O(s_r, s_t) d\tilde{\mathbf{p}}_{nuis} \\
&= \int_{\tilde{\mathcal{A}}} \mathcal{J}_{\boldsymbol{\theta}} O(s_r, s_t) d\tilde{\mathbf{p}}_{nuis}
\end{aligned} \tag{6.12}$$

where $\tilde{\mathbf{p}}_{nuis}$ lies on a space $\tilde{\mathcal{A}}$ that depends on the priors and the vector $\boldsymbol{\theta}$ lies on $\mathcal{C}^{s_r s_t}$. In the second equality above we used the fact that according to the DDM channel matrix representation, each one of its entries $h_{p,q}(f_k, t_l)$ can be expressed as:

$$\begin{aligned}
h_{p,q}(f_k, t_l) &= \frac{1}{\sqrt{s_r s_t}} \sum_{i=1}^{s_r} \sum_{j=1}^{s_t} p_{ii}^r \phi_{pi}(t_l) \psi_{jq}(t_l) p_{jj}^t d_{ij}(f_k) \theta_{ij} \\
&= \sum_{i=1}^{s_r} \sum_{j=1}^{s_t} c_{pqij}(f_k, t_l) \theta_{ij} = \sum_{s=1}^{S} c_{m,s} \theta_s
\end{aligned} \tag{6.13}$$

where, for ease of notation, we introduced the subscript $m = 1, \ldots, M$ to replace the dependency on p, q, k, l and the subscript $s = 1, \ldots, S$ to replace ij. Obviously, $M = N_f N_t n_t n_r$ and $S = s_r s_t$. Introducing $\mathbf{c}_m \stackrel{\Delta}{=} [c_{m,1}, \ldots, c_{m,S}]^t$, $\mathbf{c'}_m \stackrel{\Delta}{=} y_m \mathbf{c}_m^*$ and $\mathbf{C}_m \stackrel{\Delta}{=} \mathbf{c}_m^* \mathbf{c}_m^t$ we can show that:

$$e^{-\frac{1}{\sigma^2}|y_m - \sum_{s=1}^{S} c_{m,s}\theta_s|^2} = e^{-\frac{|y_m|^2}{\sigma^2}} e^{-\frac{1}{\sigma^2}\left[-\mathbf{c'}_m^\dagger \boldsymbol{\theta} - \boldsymbol{\theta}^\dagger \mathbf{c'}_m + \boldsymbol{\theta}^\dagger \mathbf{C}_m \boldsymbol{\theta}\right]} \tag{6.14}$$

To further simplify our notation, we will use the subscript "*all*" to denote a vector or matrix that is equal to the sum of all M vectors or matrices, respectively, which are denoted by the same symbol. After substituting the r.h.s. of (6.14) in $\mathcal{J}_{\boldsymbol{\theta}}$ we can proceed as follows:

$$\begin{aligned}\mathcal{J}_{\boldsymbol{\theta}} &= \int_{\mathcal{C}^{s_r s_t}} \frac{1}{(\pi\sigma^2)^M} e^{-\frac{1}{\sigma^2}\sum_{m=1}^{M}|y_m|^2} e^{-\frac{1}{\sigma^2}\sum_{m=1}^{M}\left[-\mathbf{c'}_m^{\dagger}\boldsymbol{\theta}-\boldsymbol{\theta}^{\dagger}\mathbf{c'}_m+\boldsymbol{\theta}^{\dagger}\mathbf{C}_m\boldsymbol{\theta}\right]} e^{-\boldsymbol{\theta}^{\dagger}\boldsymbol{\theta}} d\boldsymbol{\theta} \\ &= \frac{1}{(\pi\sigma^2)^M} e^{-\frac{1}{\sigma^2}\mathbf{y}^{\dagger}\mathbf{y}} \int_{\mathcal{C}^{s_r s_t}} e^{-\frac{1}{\sigma^2}\left[-\mathbf{c'}_{all}^{\dagger}\boldsymbol{\theta}-\boldsymbol{\theta}^{\dagger}\mathbf{c'}_{all}+\boldsymbol{\theta}^{\dagger}\mathbf{C}_{all}\boldsymbol{\theta}+\sigma^2\boldsymbol{\theta}^{\dagger}\boldsymbol{\theta}\right]} d\boldsymbol{\theta} \end{aligned} \tag{6.15}$$

where the data vector $\mathbf{y} = [y_1, \ldots, y_M]^t$ contains all the received signal values, over different time and frequency samples in different receiving antennas and for different transmitted training vectors. Also $\mathbf{C}_{all} = \sum_{m=1}^{M}\mathbf{C}_m = \sum_{m=1}^{M}\mathbf{c}_m^*\mathbf{c}_m^t$ is an $S \times S$ Hermitian matrix of rank $r = \min\{M, S\}$. Let $\mathbf{C'}_{all} = \mathbf{C}_{all} + \sigma^2\mathbf{I}$. $\mathbf{C'}_{all}$ is also a Hermitian matrix of rank S (full rank). Therefore it's inverse exists and is also Hermitian and positive definite. Using this fact we can integrate over $\boldsymbol{\theta}$ and get an explicit expression for $\mathcal{J}_{\boldsymbol{\theta}}$:

$$\begin{aligned}\mathcal{J}_{\boldsymbol{\theta}} &= \frac{\det(\mathbf{C'}_{all}^{-1})}{(\pi\sigma^2)^{(M-S)}} e^{-\frac{1}{\sigma^2}\mathbf{y}^{\dagger}\mathbf{y}} e^{\frac{1}{\sigma^2}(\mathbf{c'}_{all}^{\dagger}\mathbf{C'}_{all}^{-1}\mathbf{c'}_{all})} \\ &= \frac{1}{(\pi\sigma^2)^{(M-S)}} \det((\mathbf{C}_G^*\mathbf{C}_G^t + \sigma^2\mathbf{I})^{-1}) e^{-(\mathbf{y}^{\dagger}(\mathbf{C}_G^t\mathbf{C}_G^*+\sigma^2\mathbf{I})^{-1}\mathbf{y})} \end{aligned} \tag{6.16}$$

where we have introduced the $S \times M$ matrix $\mathbf{C}_G$:

$$\mathbf{C}_G = \begin{bmatrix} \uparrow & \ldots & \uparrow \\ \mathbf{c}_1 & \ldots & \mathbf{c}_M \\ \downarrow & \ldots & \downarrow \end{bmatrix} \tag{6.17}$$

and we have applied the matrix inversion lemma. After careful inspection we can write $\mathbf{C}_G$ as [3]

$$\mathbf{C}_G = \begin{bmatrix} \mathbf{C}_{L,(1,1)} & \ldots & \mathbf{C}_{L,(N_f,N_t)} \end{bmatrix} \tag{6.18}$$

with each submatrix given by :

$$\mathbf{C}_{L,(k,l)} = diag(\mathbf{d}_k)(\bar{\boldsymbol{\Psi}}_l \otimes \bar{\boldsymbol{\Phi}}_l^t) \tag{6.19}$$

and $\mathbf{d}_k = vec(\mathbf{D}(f_k))$, $\bar{\boldsymbol{\Psi}}_l = \mathbf{P_t}\boldsymbol{\Psi}(t_l)$ and $\bar{\boldsymbol{\Phi}}_l^t = \boldsymbol{\Phi}(t_l)\mathbf{P_r}$. From (6.18) and (6.19) the dependency of $\mathbf{C}_G$ on $\tilde{\mathbf{p}}_{nuis}$ becomes apparent. Substituting $\mathcal{J}_{\boldsymbol{\theta}}$

[3] This representation is not unique. By permuting the rows and/or the columns we get $S!M!$ equivalent representations but we should permute the elements of $\mathbf{y}$ as well.

in $f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int})$ we finally obtain:

$$f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) = \int_{\tilde{\mathcal{A}}} O'(s_r, s_t)\det((\mathbf{C}_G^*\mathbf{C}_G^t + \sigma^2\mathbf{I})^{-1})e^{-(\mathbf{y}^\dagger(\mathbf{C}_G^t\mathbf{C}_G^* + \sigma^2\mathbf{I})^{-1}\mathbf{y})}d\tilde{\mathbf{p}}_{nuis} \tag{6.20}$$

Furthermore if we consider that $N_f = 1$, the r.h.s. of eq. (6.20) does not depend on $\boldsymbol{\tau}$ either. This is not surprising, since if $N_f = 1$ we can replace $(\boldsymbol{\Theta}_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f))$ with a new matrix $\boldsymbol{\Theta}'_{s_r \times s_t}$ that has the same distribution.

Notice that so far no explicit categorization of the parameters has been made (except from the unknown amplitudes contained in $\boldsymbol{\theta}$ that are nuisance parameters). One can define $\mathbf{p}_{int}$ to contain just the parameters that need to be estimated for tracking the mobile, namely the speed magnitude and direction, or even choose to include some nuisance parameters and jointly estimate all. Whether the Bayesian estimation, based on the marginal pdf, or the joint estimation will yield better results is not trivial to show analytically. Marginalization would require integration over a subspace with many dimensions and is not guaranteed to always result in high accuracy. On the other hand, joint estimation would lead to an algorithm with very high computational complexity, since we would need to keep track of a multivariate density. We choose to sacrifice optimality for efficiency and do a step-by-step Bayesian estimation by recognizing that the AoA $\boldsymbol{\phi}$ and the AoD $\boldsymbol{\psi}$ must be estimated prior to or jointly with υ_r and ω_r for the tracking method to give accurate results. The proposed algorithm for a MIMO system is summarized below:

- Set $\mathbf{p}_{int} = [sr, \boldsymbol{\phi}]$, $\tilde{\mathbf{p}}_{nuis} = [s_t, \boldsymbol{\psi}, \mathbf{p}^r, \mathbf{p}^t]$ and use (6.20) to find the AoA, using only one observation, $N_f = N_t = 1$.
- Set $\mathbf{p}_{int} = [st, \boldsymbol{\psi}]$, $\tilde{\mathbf{p}}_{nuis} = [s_r, \boldsymbol{\phi}, \mathbf{p}^r, \mathbf{p}^t]$ and use (6.20) to find the AoD, using only one observation, $N_f = N_t = 1$.
- Set $\mathbf{p}_{int} = [\upsilon_r, \omega_r]$, $\tilde{\mathbf{p}}_{nuis} = [\upsilon_t, \alpha_t, \mathbf{p}^r, \mathbf{p}^t, \boldsymbol{\tau}]$ and use (6.20) to estimate our true parameters of interest, using all observations.

The reason for using only spatial (and no temporal) information to estimate AoA and AoD is that the terms due to the Doppler frequency shift cancel out of the expression (this is why $\upsilon_r, \omega_r, \upsilon_t, \omega_t$ are not contained in $\tilde{\mathbf{p}}_{nuis}$) leading to fewer nuisance parameters, therefore easier implementation and higher estimation accuracy. This is essentially an ML estimator for direction finding. Instead of implementing this ML algorithm in steps one and two of the above method, subspace-based algorithms like the ones in [24], [25]

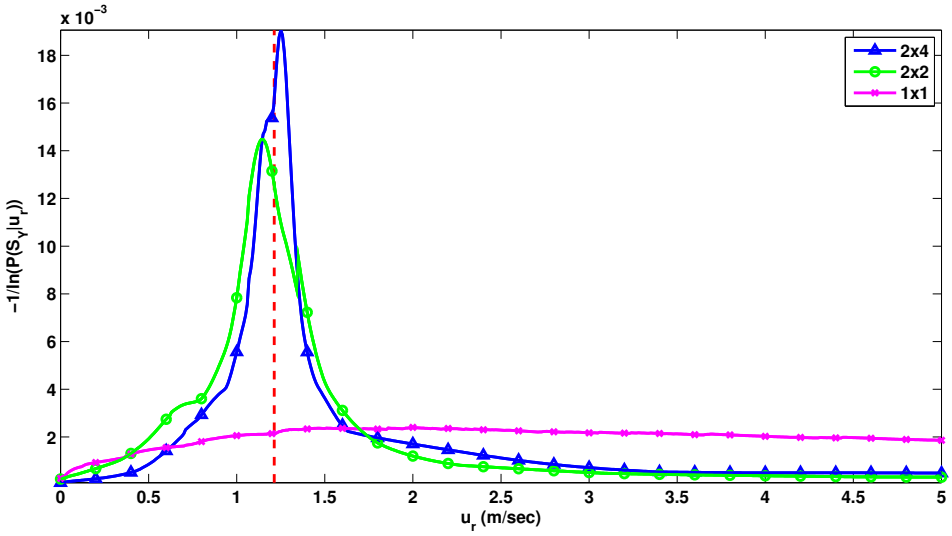

Figure 6.1: Log-likelihood for speed magnitude estimation at high SNR

could be employed. For a system with a single transmit and a single receive antenna (SISO), only the third step can be implemented, since the channel matrix reduces to a scalar that does not depend on angles.

6.3 Simulation Results

In most practical scenarios of interest, the BS is fixed. We will therefore assume that $\upsilon_t = \alpha_t = 0$[4]. We further assume that the main lobe of the transmitting and the receiving antenna array is steered to a direction perpendicular to the array and has a beamwidth of 180^o. The energy of the signal components transmitted to or received from other directions is negligible. This implies that the AoA and the AoD are either in $[-\pi/2, \pi/2]$ or in $[\pi/2, 3\pi/2]$. The power gains of the steering directions are also assumed to be known, i.e. calibration has preceded the estimation procedure. To compute the value of the multidimensional integral in (6.20), Monte Carlo simulations have been performed. Normally 100 iterations are enough for the algorithm to converge to the true density. To make our graphs more clear and emphasize our results, we have plotted the 1-dimensional normalized log-likelihoods $-\frac{1}{\ln f(\mathbb{S}_{\mathbf{Y}}|\upsilon_r)}$ and $-\frac{1}{\ln f(\mathbb{S}_{\mathbf{Y}}|\omega_r)}$ as a function of υ_r and ω_r,

[4]Or $\upsilon_r = \alpha_r = 0$ if we had considered transmission in the uplink.

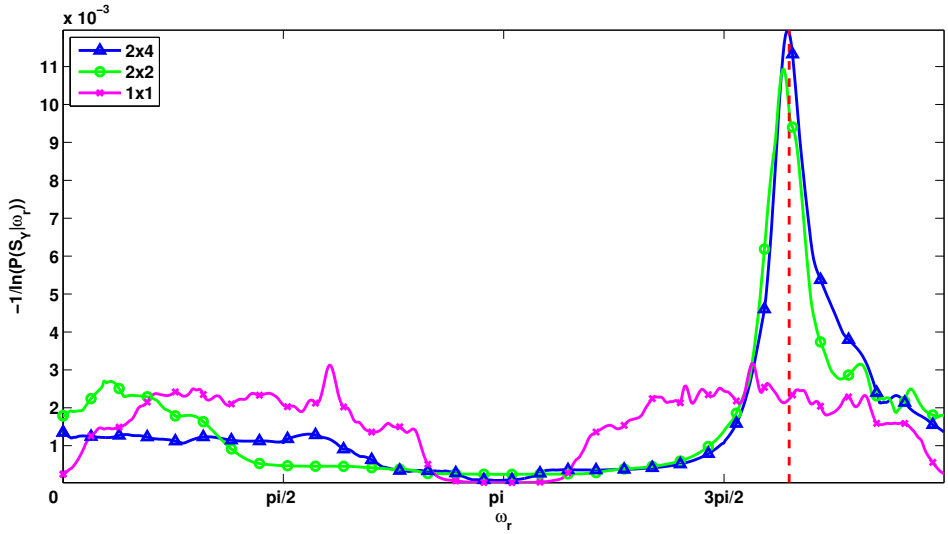

Figure 6.2: Log-likelihood for speed direction estimation at high SNR

respectively, for different $n_r \times n_t$ systems. In other words, once we derive the 2-dimensional log-likelihood $f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_\upsilon)$ we marginalize it further, just for demonstration purposes. The vertical dashed line in the figures depicts the true value of the estimated parameter (υ_r or ω_r).

In figures 6.1 and 6.2 the advantage of MIMO over SISO at high SNR ($20dB$) is clearly illustrated. With just one antenna at each side of the communication link, it is impossible to track the mobile. With as many as 2 antennas at each side, the estimation error becomes already very small. In figures 6.3 and 6.4 we show that even with noisy measurements (SNR= $10dB$), υ_r and ω_r can be estimated correctly in a 4×4 and a 2×8 system. On the other hand in a 2×2 or a 2×4 system our parameters of interest are slightly misestimated at low SNR as shown in figure 6.5. As expected, decreasing the SNR or the number of antennas leads to an increase of the estimation error. It further results in a second peak in the log-likelihood corresponding to ω_r. This stems from the fact that our expression (r.h.s. of 6.20) depends on ω_r only through its cosine and $cos(\phi_l - \omega_r) = cos(\omega_r - \phi_l)$, $\forall l$. Thus if there exist AoA that cannot be estimated in the first 2 steps of the algorithm (SISO case or MIMO with just a few antennas) or if their effect cannot be removed successfully by integration (very noisy estimates), this ambiguity cannot be resolved. These results illustrate clearly that when

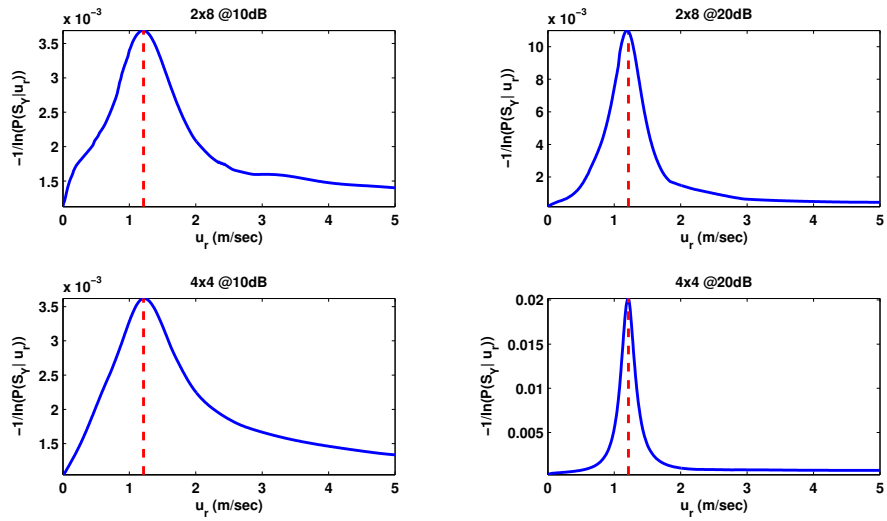

Figure 6.3: Log-likelihood for speed magnitude estimation at various SNR

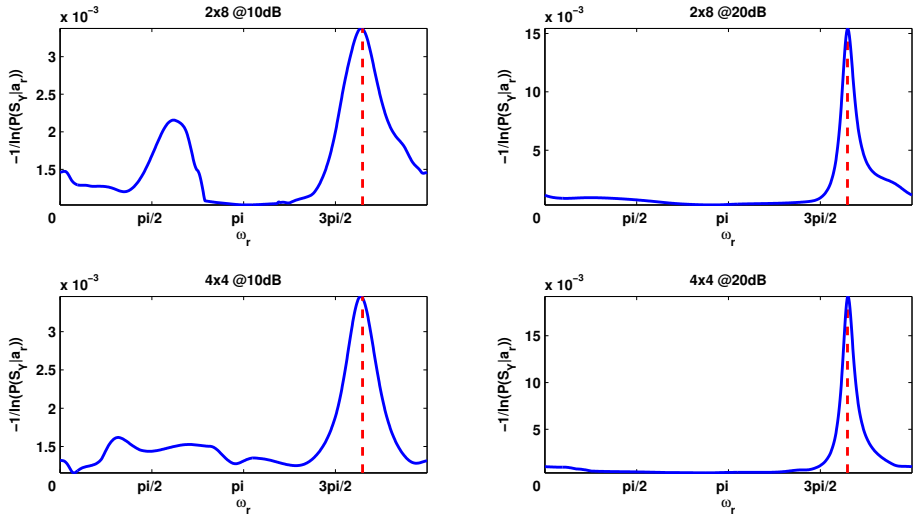

Figure 6.4: Log-likelihood for speed direction estimation at various SNR

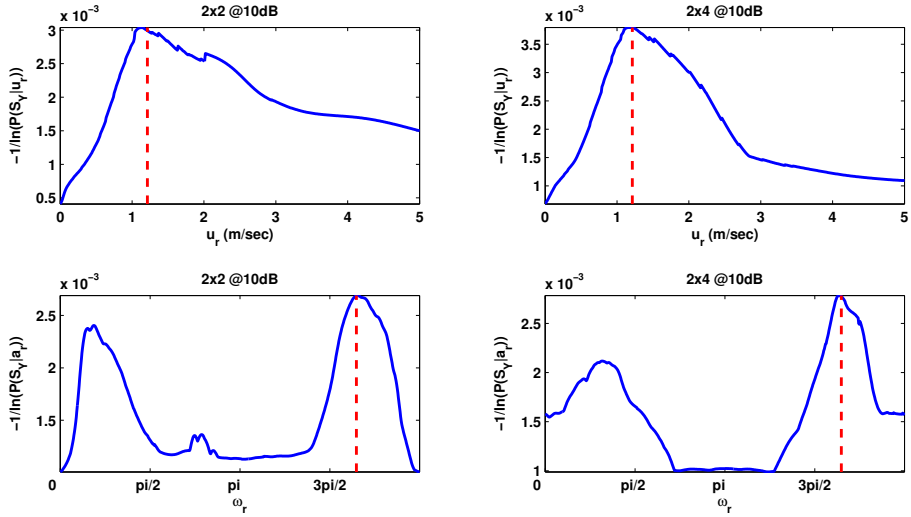

Figure 6.5: Log-likelihoods for speed magnitude and direction estimation at low SNR

a complex model like the DDM -and not its simplified version that assumes single-bounces and distinct paths- is used, one should seek solutions that utilize MIMO systems, which operate at high SNR and are equipped with large number of antennas at both ends. These requirements must be fulfilled in order to be able to localize (or more precisely to track) the MT.

Chapter 7

Conclusions and Future Work

Conducting research in the wider area of parameter estimation and in the more narrow area of localization, has undoubtedly been fruitful. Many results have been obtained, including but not limited to algorithms based on closed-form solutions, performance bounds, theorems on identifiability and performance for the problems of estimating parameters of interest in the presence of nuisance parameters. Furthermore many intuitive conclusions have been reached. We will try to summarize these conclusions below:

- "Traditional" geometrical localization methods that utilize the wireless network infrastructure, can not achieve sufficient accuracy when the signal propagates in NLoS environments. One needs to choose between fingerprinting and advanced geometrical methods.

- Choosing between fingerprinting and advanced geometrical methods is still under investigation. The two approaches are fundamentally different and thus it is difficult to compare them taking into account performance, complexity, cost for the user and cost for the provider. Under the assumptions made in this work, low-complexity geometrical localization algorithms (eg. LS) were derived, rendering this approach the most appealing.

- Within the geometrical methods framework, there also exist two fundamentally different approaches to attack the NLoS localization problem. The first approach introduces some errors that are induced from the NLoS propagation and tries to mitigate them. The second one introduces a more accurate but also more complex representation of the environment and thus leads to new unknown parameters. These nuisance parameters need to be eliminated or jointly estimated. Choosing between these two approaches, creates a performance-complexity trade-off. The second approach leads to improved accuracy at the cost of increased complexity. On the other hand, in the first approach at least some for the communications links must be LoS for the estimates to be accurate. Due to the non-satisfactory performance of the first approach in strictly NLoS environments, we focused our study on the second approach and searched for appropriate channel models.

- Choosing an appropriate -for NLoS localization- channel model, is still an open issue. Making too many assumptions on the environment leads to simplistic models, which in turn leads to powerful localization algorithms. However, their applicability is limited. On the other hand, using generic models leads to a very high-dimensional parameter estimation problem that is cumbersome to solve. Even when it's solved, either the performance is low or the complexity incredibly high. Combining this with the previous remark, it becomes obvious that there is actually a performance-complexity-applicability trade-off. Overall, we believe that the best way to attack the NLoS localization problem is to start with a simple model and try to generalize it. We therefore consider the SBM to be an excellent solution, that simplifies analysis without being unrealistic.

- On top of choosing an appropriate channel model, if the channel is time-varying, one should choose an appropriate mobility model. Similar conclusions as before can be reached for the choice of mobility model. We again strongly believe, that if an attempt to exploit the information in the time-variation of the LDP is made, a rather simple model should be used. It suffices to say that exploiting this new dimension, namely the variation in time, leads to a huge improvement in performance (and identifiability), at the cost of increased complexity that depends on the mobility model. Constant-speed and constant acceleration models have been proved to be good candidates.

- Depending on the model introduced, there exist several different kinds of LDP that the researcher must chose from and utilize in the localization algorithm. Of course choosing all available LDP results in higher accuracy at a marginal increment of complexity. However, not always all possible LDP are available and/or not always all the LDP estimates are trustworthy. We attacked the NLoS problem assuming lack of angle information at the MT (lack of AoA if DL is considered), which is a case of great practical interest. We showed that localization is feasible even for that scenario. One LDP that we have intentionally ignored in our work, is the RSS. Despite the fact that it is available in almost all existing systems, it can barely lead to any improvement if accurate ToA estimates are available.

- Identifiability and performance are related, when both are derived using the CRB. We studied both and this enabled us to classify estimators that utilize different kinds of LDP, based on whether they render the parameters identifiable or not and based on their performance with respect to the best achievable performance.

- In multipath environments, exploring the LDP of the MPC additionally to exploring the LDP of the LoS component, can enhance accuracy of the SBM-based localization methods.

- In the absence of knowledge on the priors of the unknown parameters, the least informative, which are the uniform ones, should be assumed. A MAP estimator then becomes equivalent to a ML estimator.

- The relative performance of choosing between elimination (marginalization) or joint estimation (maximization) of the nuisance parameters is still an open issue. We chose the former for cases when the nuisance can be easily integrated out of the likelihood and the latter for all other cases.

- Once the localization method has been defined, algorithms for obtaining the estimates can be derived and implemented. Independently of the choice of the algorithm, the best attainable performance for the defined method can be determined, using theoretical bounds like eg. the CRB. This is of great value, since it enables the researcher to get an overall idea on what accuracy to roughly expect. The actual performance of any algorithm used for a specific method, along with the bounds for the method's performance depend on 3 factors:

  - The standard deviation of the LDP estimates, or equivalently, the SNR (and the BW) of the received signal.
  - The number of data samples (LDP estimates), which depends on the observation duration and the sampling rate.
  - The geometric configuration of the environment.

  We have studied extensively the impact, of all of these factors, on the CRB of NLoS localization methods. This enables us to be able to recognize -almost immediately- the scenarios under which a NLoS localization method can output meaningful estimates.

- MIMO-OFDM was certainly not a random choice. Besides being the technology adopted for 4-G wireless communications, it has certain benefits for localization algorithms. To start with, MIMO systems, having multiple antennas at both ends, can exploit the space dimension for direction finding. OFDM signals on the other hand, due to the cyclic prefix, can transform a frequency-selective channel into a group of parallel frequency-flat channels. This essentially enables us to write the discrete time-frequency io relation as $\mathbf{Y}(f_k, t_l) = \mathbf{H}(f_k, t_l)\mathbf{X}(f_k, t_l) + \mathbf{N}(f_k, t_l)$. This was a key relation in deriving a 4D Unitary ESPRIT algorithm for the estimation of the LDP from the received signal. Furthermore, using the SBM model to parameterize $\mathbf{H}(f_k, t_l)$, this also proved to be a key relation in deriving the DLE algorithm.

- DLE is a method that combines the 2 steps of localization methods into 1. By doing so, it takes advantage of the fact that all LDP estimates correspond to the same MT location. While 2-step approaches can asymptotically - for high SNR and sufficient number of data samples- achieve the CRB, in (rapidly) time-varying environments this can not always assumed to be the case. We therefore find the DLE an attractive solution for such environments.

It's true that there are still many questions that need to be answered and many uninvestigated issues that need to be addressed, in order to formulate a complete solution for the NLoS localization problem. A few rather trivial extensions of the work presented in this document, include but are not limited to the following:

- Redefine the same methods for 3D environments. Introduce the new LDP, like the elevation angle and the new unknown parameters, like the position and speed components on the z-axis.

- Introduce the RSS as an available LDP for the cases when ToA estimates have limited resolution due to eg. small BW. Redefine the same methods in a way that the path lengths estimates come from RSS measurements. New nuisance parameters, like eg. the path-loss exponent should then be considered.

- Introduce TDoA as an alternative to ToA for the cases when there is no synchronization between the MT and the BS. Define new methods and derive closed-formed solutions.

- Compare the 2 different approaches for localizing in NLoS environments, the one based on channel modeling with the one that attempts to mitigate the errors in the NLoS components.

- Design and implement algorithms, like EM, to solve the ML optimization problem.

- Investigate further the elimination vs joint estimation choice for the nuisance parameters. Prove why one is preferable over the other for the location estimation problem.

- Study the degradation in the performance of the 1^{st} step of localization and in the performance of the DLE method due to non-calibrated arrays.

- Extend the study on the impact of network geometry to the DSBM case. Investigate how the location of the scatterers and the BS relatively to the direction of speed affects the performance of DSBM-based methods.

- Improve and complete the ToA/AoD/DS localization method of chapter 3. Improvement is necessary for the method to perform accurately with noise-impaired LDP estimates.

- Derive the CRB for the 4D Unitary ESPRIT and compare it with the actual performance.

- Simulate the actual RMSE for the SBM-based DLE method and compare it with the CRB.

- Compare the DSBM method with a method that utilizes SBM and then uses filtering (KF, EKF etc) to improve performance.

More important than extending the results obtained so far is to generalize the SBM (DSBM) method and broaden its applicability. After all, the main objective of this work, has always been to derive a method with excellent performance in real propagation environments. So at this point, although the argument supporting the single-bounce assumption is valid, we are forced to respond to the question:

"*Can SBM be utilized when the signal components have bounced more than once?*"

Surprisingly enough, the answer is: "*Yes*". After careful thought, it occurred to us that the SBM can be used to describe any environment by introducing the concept of "virtual scattering". Without getting into details, it suffices to say that introducing an unknown bias $\widetilde{\tau^0} \in \mathcal{R}$ only for the ToA measurements and possibly an orientation offset $\widetilde{\psi^0} \in \{0, \pi\}$ only for the AoD estimates, will render the SBM capable of describing any environment while increasing the complexity only slightly compared to using generic channel models. This is exactly why we advocate above that the best way to attack the NLoS localization problem is to start with a simple model and try to generalize it.

Appendices

Appendix A

Matrix Algebra and Statistics

Useful equalities and inequalities between random and/or deterministic vectors and matrices are provided below. Some of these are standard rules that can be found in textbooks, like eg. [76] while the others can be easily proved. Following the rules indicated in the Index of Symbols, if a matrix $\mathbf{A}$ is defined, a_{ij} will denote its $\{i,j\}$ entry and $\mathbf{a} = vec\{\mathbf{A}\}$. Similarly, $\mathbf{a}_t = vec\{\mathbf{A^t}\}$. If a vector $\mathbf{a}$ is defined, a_i will denote its i^{th} entry and $\mathbf{A} = diag\{\mathbf{a}\}$.

- $|\mathbf{A}|^2 = tr\{\mathbf{A}\mathbf{A}^\dagger\} = \mathbf{a}^\dagger\mathbf{a}$
- $\mathbf{A}\mathbf{b} = (\mathbf{I} \otimes \mathbf{b}^t)\mathbf{a}_t$
- $(\mathbf{a} \odot \mathbf{b})(\mathbf{c}^t \odot \mathbf{d}^t) = \mathbf{a}\mathbf{c}^t \odot \mathbf{b}\mathbf{d}^t$
- $\mathbf{1} \otimes \mathbf{A}\mathbf{B} = (\mathbf{1} \otimes \mathbf{A})\mathbf{B}$
- $\mathbf{A}\mathbf{b} \otimes \mathbf{C}\mathbf{d} = (\mathbf{A} \otimes \mathbf{C}\mathbf{d})\mathbf{b} = (\mathbf{A}\mathbf{b} \otimes \mathbf{C})\mathbf{d}$
- $\mathbf{A}\mathbf{b} \odot \mathbf{c} = \mathbf{C}\mathbf{A}\mathbf{b}$
- $\mathbf{B} \odot \mathbf{a}\mathbf{c^t} = \mathbf{A}\mathbf{B}\mathbf{C}$

Now let $\mathbf{a} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$. It follows that

- $\mathcal{E}\{\mathbf{A}\} = \mathbf{O}$
- $\mathcal{E}\{\mathbf{ABA}\} = \sigma^2\mathbf{B} \odot \mathbf{I}$
- $\mathcal{E}\{\mathbf{ABACA}\} = \mathbf{O}$
- $$\mathcal{E}\{\mathbf{ABACADA}\} = \sigma^4 \left[(\mathbf{B} \odot \mathbf{I})\mathbf{C}(\mathbf{D} \odot \mathbf{I}) \right.$$
$$\left. +\mathbf{B} \odot \mathbf{C}^t \odot \mathbf{D}(\mathbf{B} \odot (\mathbf{C} \odot \mathbf{I}) \odot \mathbf{D}) \odot \mathbf{I} + 3(\mathbf{B} \odot \mathbf{I})(\mathbf{C} \odot \mathbf{I})(\mathbf{D} \odot \mathbf{I})\right]$$

Similarly if $\mathbf{a} \sim \mathcal{N}(\mathbf{m}, \sigma^2\mathbf{I})$. It follows that

- $\mathcal{E}\{\mathbf{A}\} = \mathbf{M}$
- $\mathcal{E}\{\mathbf{ABA}\} = \sigma^2\mathbf{B} \odot (\sigma^2\mathbf{I} + \mathbf{m}\mathbf{m}^t)$

Appendix B

Mapping of a tangent $y = \tan(x)$ to its argument

It's a well known fact that the inverse trigonometric functions satisfy the injection (one-to-one) property only if their ranges are a subset of the domains of the original trigonometric functions. For the case of the inverse tangent, its range is usually restricted to a principal branch of $-\frac{\pi}{2} < x \leq \frac{\pi}{2}$. We can extend this principal branch to $-\pi < x \leq \pi$, if we utilize the sign of $\cos(x)$, since it is positive in the right half plane and negative in the left one. We can therefore write for the mapping of a tangent $y = \tan(x)$ to its argument

$$x = \frac{\pi}{2}(1 - sgn\{\cos(x)\}) + tan^{-1}(y) \tag{B.1}$$

Appendix C

Mapping of Complex to Real Matrices

In this section we briefly explain how to map a complex matrix into a real one. For more details within the ESPRIT context the interested reader can refer to [51] where the authors restate the results originally published in [77].

Let $\mathbf{\Pi}$ denote the permutation matrices obtained by reversing the order of the rows of $\mathbf{I}$, eg.

$$\mathbf{\Pi} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{bmatrix}. \tag{C.1}$$

A complex matrix $\bar{\boldsymbol{A}} \in \mathbb{C}^{m \times n}$ is called Centro-Hermitian if

$$\mathbf{\Pi}_m \boldsymbol{A}^* \mathbf{\Pi}_n = \boldsymbol{A}. \tag{C.2}$$

A complex matrix $\boldsymbol{Q} \in \mathbb{C}^{m \times n}$ is called left $\mathbf{\Pi}$-real if it satisfies

$$\mathbf{\Pi}_m \boldsymbol{Q}^* = \boldsymbol{Q}. \tag{C.3}$$

The following unitary matrices

$$\mathbf{Q} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{\Pi} & -j\boldsymbol{P}i \end{bmatrix} \tag{C.4}$$

$$\mathbf{Q} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I} & \mathbf{0} & j\mathbf{I} \\ \mathbf{0}^t & \sqrt{2} & \mathbf{0}^t \\ \mathbf{\Pi} & \mathbf{0} & -j\mathbf{\Pi} \end{bmatrix} \tag{C.5}$$

are typical examples of left $\mathbf{\Pi}$-real of even and odd order respectively. From any complex matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, we can construct a Centro-Hermitian matrix $\bar{\boldsymbol{A}} \in \mathbb{C}^{m\times 2n}$ as follows

$$\bar{\boldsymbol{A}} = [\boldsymbol{A} \ \mathbf{\Pi}_m \boldsymbol{A} \mathbf{\Pi}_n] \,. \tag{C.6}$$

Subsequently, we can map the Centro-hermitian matrix $\bar{\boldsymbol{A}}$ to a real one $\bar{\bar{\boldsymbol{A}}}$, by using two arbitrary non-singular left $\mathbf{\Pi}$-real matrices,

$$\bar{\bar{\boldsymbol{A}}} = \mathbf{Q}_m^{\dagger} \bar{\boldsymbol{A}} \mathbf{Q}_{2n}. \tag{C.7}$$

Appendix D

Optimal Weighting Matrix for SBM WLS

In the WLS formulation and solution presented in section 3.2 we have omitted the hats ($\hat{\cdot}$) of all symbols denoting estimates for clarity. Herein, since we are interested in deriving the covariance matrix of vector $\hat{\mathbf{b}}$, true values and estimates will be involved in the expressions. So we need to replace $\mathbf{b}$ in eq. (3.21) with $\hat{\mathbf{b}} = \mathbf{b} + \tilde{\mathbf{b}}$ and rewrite it so that any symbol representing an estimate has a ($\hat{\cdot}$) as follows:ting an estimate has a ($\hat{\cdot}$) as follows:

$$\hat{\mathbf{b}} = \begin{bmatrix} (\hat{\mathbf{D}}\hat{\mathbf{C}}_{\psi} + \mathbf{X}_{bs})\hat{\mathbf{C}}_{\phi}\mathbf{1} \\ (\hat{\mathbf{D}}\hat{\mathbf{S}}_{\psi} + \mathbf{Y}_{bs})\hat{\mathbf{S}}_{\phi}\mathbf{1} \\ (\hat{\mathbf{D}}\hat{\mathbf{S}}_{\psi} + \mathbf{Y}_{bs})\hat{\mathbf{C}}_{\phi}\mathbf{1} \\ (\hat{\mathbf{D}}\hat{\mathbf{C}}_{\psi} + \mathbf{X}_{bs})\hat{\mathbf{S}}_{\phi}\mathbf{1} \end{bmatrix} \triangleq \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \hat{\mathbf{b}}_3 \\ \hat{\mathbf{b}}_4 \end{bmatrix} \tag{D.1}$$

To compute $\mathbf{C}_{\hat{\mathbf{b}}} = \mathbf{C}_{\tilde{\mathbf{b}}}$, the prior distributions of the LDP estimates (or errors) are required. In the ML estimator presented in section 1.8 it was assumed the vector $\tilde{\theta}$ containing the errors of all available LDP estimates, is distributed according to $\tilde{\boldsymbol{\theta}} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\tilde{\boldsymbol{\theta}}})$ (assumption A). However, in chapter 2, an ESPRIT-based algorithm that estimates the delays, the DS and the sines of the AoA and the AoD, was derived. Therefore, instead of considering the errors of the angles to be Gaussian distributed, we could alternatively consider the errors in the estimates of their trigonometric functions to be Gaussian distributed (assumption B). Independently of the choosing A or

B, we will assume that LDP estimates of the same kind are i.i.d. and are also independent of the LDP estimates of other kinds.It follows that for assumption A

$$\tilde{\mathbf{d}} \sim \mathcal{N}(\mathbf{0}, \sigma_d^2\mathbf{I}) \Leftrightarrow \hat{\mathbf{d}} \sim \mathcal{N}(\mathbf{d}, \sigma_d^2\mathbf{I}) \tag{D.2}$$

$$\tilde{\boldsymbol{\phi}} \sim \mathcal{N}(\mathbf{0}, \sigma_\phi^2\mathbf{I}) \Leftrightarrow \hat{\boldsymbol{\phi}} \sim \mathcal{N}(\boldsymbol{\phi}, \sigma_\phi^2\mathbf{I}) \tag{D.3}$$

$$\tilde{\boldsymbol{\psi}} \sim \mathcal{N}(\mathbf{0}, \sigma_\psi^2\mathbf{I}) \Leftrightarrow \hat{\boldsymbol{\psi}} \sim \mathcal{N}(\boldsymbol{\psi}, \sigma_\psi^2\mathbf{I}) \tag{D.4}$$

and thus for the trigonometric function of the errors in the AoA estimates we have

$$\mathcal{E}\{\tilde{\mathbf{s}}_\phi\} = \mathbf{0} \tag{D.5}$$

$$\mathcal{E}\{\tilde{\mathbf{c}}_\phi\} \approx (1 - \tfrac{\sigma_\phi^2}{2})\mathbf{1} \triangleq m_{\mathbf{c}_\phi} \tag{D.6}$$

$$\mathbf{R}_{\tilde{\mathbf{s}}_\phi} = \mathbf{C}_{\tilde{\mathbf{s}}_\phi} = \tfrac{1}{2}\mathbf{I} \tag{D.7}$$

$$\mathbf{R}_{\tilde{\mathbf{c}}_\phi} = (\mathbf{m}_{\mathbf{c}_\phi} - \mathbf{m}_{\mathbf{c}_\phi}^2)\mathbf{I} + m_{\mathbf{c}_\phi}^2\mathbf{1}\mathbf{1}^t \tag{D.8}$$

$$\mathbf{C}_{\tilde{\mathbf{c}}_\phi} = (\mathbf{m}_{\mathbf{c}_\phi} - \mathbf{m}_{\mathbf{c}_\phi}^2)\mathbf{I} \tag{D.9}$$

The mean value and the correlation matrix of the trigonometric functions of the AoD are given by replacing σ_ϕ with σ_ψ and $\mathbf{m}_{\mathbf{c}_\phi}$ with $\mathbf{m}_{\mathbf{c}_\psi}$ in the equations above. If assumption B is made, i.e. if it is assumed that trigonometric functions are estimated directly, then

$$\mathcal{E}\{\tilde{\mathbf{s}}_\phi\} = \mathbf{0} \tag{D.10}$$

$$\mathcal{E}\{\tilde{\mathbf{c}}_\phi\} = \mathbf{0} \tag{D.11}$$

$$\mathbf{R}_{\tilde{\mathbf{s}}_\phi} = \mathbf{C}_{\tilde{\mathbf{s}}_\phi} = \sigma_{\mathbf{s}_\phi}^2\mathbf{I} \tag{D.12}$$

$$\mathbf{R}_{\tilde{\mathbf{c}}_\phi} = \mathbf{C}_{\tilde{\mathbf{c}}_\phi} = \sigma_{\mathbf{c}_\phi}^2\mathbf{I} \tag{D.13}$$

and similarly for the AoD by replacing the variances corresponding to the AoA with the ones corresponding to the AoD. We will first proceed with the computation assuming B is true. As can be seen from eq. (D.1), $\tilde{\mathbf{b}}$ is composed of 4 subvectors $\tilde{\mathbf{b}}_i$. Under assumption B, $\mathcal{E}\{\tilde{\mathbf{b}}_i\} = \mathbf{0}$, $\forall i$. To avoid repeating the same derivation for the cross-covariances of all pairs of subvectors, lets define the matrices $\mathbf{Q}_\phi, \mathbf{Q}'_\phi \in \{\mathbf{S}_\phi, \mathbf{C}_\phi\}$ and $\mathbf{Q}_\psi, \mathbf{Q}'_\psi \in \{\mathbf{S}_\psi, \mathbf{C}_\psi\}$, along with their estimates and the error of their estimates and also $\mathbf{Z}_{bs}, \mathbf{Z}'_{bs} \in \{\mathbf{X}_{bs}, \mathbf{Y}_{bs}\}$. We can now derive simultaneously the solution for all submatrices of $\mathbf{C}_{\tilde{\mathbf{b}}}$, since, with the aid of these matrices any $\hat{\mathbf{b}}_i$ can be written as

$$\hat{\mathbf{b}}_i = (\hat{\mathbf{D}}\hat{\mathbf{Q}}_\psi + \mathbf{Z}_{bs})\hat{\mathbf{Q}}_\phi\mathbf{1}. \tag{D.14}$$

Substituting each matrix of LDP estimates by the sum of a matrix of LDP true values and a matrix of errors, we get

$$\hat{\mathbf{b}}_i = (\mathbf{D}\mathbf{Q}_\psi + \mathbf{Z}_{bs})\mathbf{Q}_\phi\mathbf{1} + (\tilde{\mathbf{D}}\mathbf{Q}_\psi + \mathbf{D}\tilde{\mathbf{Q}}_\psi + \tilde{\mathbf{D}}\tilde{\mathbf{Q}}_\psi)\mathbf{Q}_\phi\mathbf{1} + (\hat{\mathbf{D}}\hat{\mathbf{Q}}_\psi + \mathbf{Z}_{bs})\tilde{\mathbf{Q}}_\phi\mathbf{1}. \tag{D.15}$$

From this equation it becomes obvious that

$$\begin{aligned}\tilde{\mathbf{b}}_i &= (\tilde{\mathbf{D}}\mathbf{Q}_\psi + \mathbf{D}\tilde{\mathbf{Q}}_\psi + \tilde{\mathbf{D}}\tilde{\mathbf{Q}}_\psi)\mathbf{Q}_\phi\mathbf{1} + (\hat{\mathbf{D}}\hat{\mathbf{Q}}_\psi + \mathbf{Z}_{bs})\tilde{\mathbf{Q}}_\phi\mathbf{1} \\ &\approx (\tilde{\mathbf{D}}\mathbf{Q}_\psi + \mathbf{D}\tilde{\mathbf{Q}}_\psi)\mathbf{Q}_\phi\mathbf{1} + (\hat{\mathbf{D}}\hat{\mathbf{Q}}_\psi + \mathbf{Z}_{bs})\tilde{\mathbf{Q}}_\phi\mathbf{1}. \end{aligned}\tag{D.16}$$

It is straightforward to show that for $i = i'$

$$\begin{aligned}\mathbf{C}_{\tilde{\mathbf{b}}_i\tilde{\mathbf{b}}_i} &= \mathbf{C}_{\tilde{\mathbf{d}}} \odot (\mathbf{Q}_\psi\mathbf{Q}_\phi\mathbf{1}\mathbf{1}^t\mathbf{Q}_\phi\mathbf{Q}_\psi) + \mathbf{C}_{\tilde{\mathbf{q}}_\psi} \odot (\mathbf{D}\mathbf{Q}_\phi\mathbf{1}\mathbf{1}^t\mathbf{Q}_\phi\mathbf{D}) \\ &+\mathbf{R}_{\hat{\mathbf{d}}} \odot \mathbf{R}_{\hat{\mathbf{q}}_\psi} \odot \mathbf{C}_{\hat{\mathbf{q}}_\phi} + \mathbf{D}\mathbf{Q}_\psi\mathbf{C}_{\hat{\mathbf{q}}_\phi}\mathbf{Z}_{bs} + \mathbf{Z}_{bs}\mathbf{C}_{\hat{\mathbf{q}}_\phi}\mathbf{Q}_\psi\mathbf{D} + \mathbf{Z}_{bs}\mathbf{C}_{\hat{\mathbf{q}}_\phi}\mathbf{Z}_{bs}\end{aligned}\tag{D.17}$$

for $\mathbf{Q}_\psi = \mathbf{Q}'_\psi$, i.e. when $(i, i') \in \{(1,4),(4,1),(2,3),(3,2)\}$

$$\mathbf{C}_{\tilde{\mathbf{b}}_i\tilde{\mathbf{b}}_{i'}} = \mathbf{C}_{\tilde{\mathbf{d}}} \odot (\mathbf{Q}_\psi\mathbf{Q}_\phi\mathbf{1}\mathbf{1}^t\mathbf{Q}'_\phi\mathbf{Q}_\psi) + \mathbf{C}_{\tilde{\mathbf{q}}_\psi} \odot (\mathbf{D}\mathbf{Q}_\phi\mathbf{1}\mathbf{1}^t\mathbf{Q}'_\phi\mathbf{D}), \tag{D.18}$$

for $\mathbf{Q}_\phi = \mathbf{Q}'_\phi$, i.e. when $(i, i') \in \{(1,3),(3,1),(2,4),(4,2)\}$

$$\begin{aligned}\mathbf{C}_{\tilde{\mathbf{b}}_i\tilde{\mathbf{b}}_{i'}} &= \mathbf{C}_{\tilde{\mathbf{d}}} \odot (\mathbf{Q}_\psi\mathbf{Q}_\phi\mathbf{1}\mathbf{1}^t\mathbf{Q}_\phi\mathbf{Q}'_\psi) + \mathbf{D}\mathbf{Q}_\psi\mathbf{C}_{\tilde{\mathbf{q}}_\phi}\mathbf{Z}_{bs} \\ &+\mathbf{Z}_{bs}\mathbf{C}_{\tilde{\mathbf{q}}_\phi}\mathbf{Q}'_\psi\mathbf{D} + \mathbf{Z}_{bs}\mathbf{C}_{\tilde{\mathbf{q}}_\phi}\mathbf{Z}_{bs}\end{aligned}\tag{D.19}$$

and last for $(i, i') \in \{(1,2),(2,1),(3,4),(4,3)\}$, $\mathbf{C}_{\tilde{\mathbf{b}}_i\tilde{\mathbf{b}}_{i'}} = \mathbf{O}$. Using these submatrices, $\mathbf{C}_{\tilde{\mathbf{b}}}$ can be constructed and thus optimal weighted least squares can be applied to solve eq. (3.21).

Appendix E

CRB for NLoS Static Environments

To compute the entries of $\mathbf{G}$ lets introduce some key variables

$$\mathbf{D}_{mts} \triangleq ((\mathbf{Y}_s - y_{mt}\mathbf{I})^2 + (\mathbf{X}_s - x_{mt}\mathbf{I})^2)^{\frac{1}{2}} \tag{E.1}$$

$$\mathbf{D}_{bs} \triangleq ((\mathbf{Y}_s - \mathbf{Y}_{bs})^2 + (\mathbf{X}_s - \mathbf{X}_{bs})^2)^{\frac{1}{2}} \tag{E.2}$$

$$\mathbf{c}_{\boldsymbol{z}} \triangleq [\cos(z_1), \ldots, \cos(z_{Ns})]^t \tag{E.3}$$

$$\mathbf{s}_{\boldsymbol{z}} \triangleq [\sin(z_1), \ldots, \sin(z_{Ns})]^t \tag{E.4}$$

where the last two vectors are defined for every vector $\mathbf{z}$ that contains angles. If the MT communicates only with 1 BS through a multipath environment, then $\mathbf{Y}_{bs} = y_{bs}\mathbf{I}$ and $\mathbf{X}_{bs} = x_{bs}\mathbf{I}$. Lets further define the vectors and matrices containing partial derivatives

$$\mathbf{D}_{\mathbf{x}_s} \triangleq \frac{\partial \mathbf{d}^t}{\partial \mathbf{x}_s} = \mathbf{C}_{\boldsymbol{\phi}} + \mathbf{C}_{\boldsymbol{\psi}} \tag{E.5}$$

$$\mathbf{D}_{\mathbf{y}_s} \triangleq \frac{\partial \mathbf{d}^t}{\partial \mathbf{y}_s} = \mathbf{S}_{\boldsymbol{\phi}} + \mathbf{S}_{\boldsymbol{\psi}} \tag{E.6}$$

$$\mathbf{d}_x^t \triangleq \frac{\partial \mathbf{d}^t}{\partial x_{mt}} = -\mathbf{c}_{\boldsymbol{\phi}}^t = -\mathbf{1}^t\mathbf{C}_{\boldsymbol{\phi}} \tag{E.7}$$

$$\mathbf{d}_y^t \triangleq \frac{\partial \mathbf{d}^t}{\partial y_{mt}} = -\mathbf{s}_{\boldsymbol{\phi}}^t = -\mathbf{1}^t\mathbf{S}_{\boldsymbol{\phi}} \tag{E.8}$$

$$\mathbf{\Phi}_{\mathbf{x}_s} \triangleq \frac{\partial \boldsymbol{\phi}^t}{\partial \mathbf{x}_s} = -\mathbf{S}_{\boldsymbol{\phi}}\mathbf{D}_{mts}^{-1} \tag{E.9}$$

$$\mathbf{\Phi}_{\mathbf{y}_s} \triangleq \frac{\partial \boldsymbol{\phi}^t}{\partial \mathbf{y}_s} = \mathbf{C}_{\boldsymbol{\phi}}\mathbf{D}_{mts}^{-1} \tag{E.10}$$

$$\boldsymbol{\phi}_x^t \triangleq \frac{\partial \boldsymbol{\phi}^t}{\partial x_{mt}} = \mathbf{s}_{\boldsymbol{\phi}}^t\mathbf{D}_{mts}^{-1} = \mathbf{1}^t\mathbf{S}_{\boldsymbol{\phi}}\mathbf{D}_{mts}^{-1} \tag{E.11}$$

$$\boldsymbol{\phi}_y^t \triangleq \frac{\partial \boldsymbol{\phi}^t}{\partial y_{mt}} = -\mathbf{c}_{\boldsymbol{\phi}}^t\mathbf{D}_{mts}^{-1} = -\mathbf{1}^t\mathbf{C}_{\boldsymbol{\phi}}\mathbf{D}_{mts}^{-1} \tag{E.12}$$

$$\mathbf{\Psi}_{\mathbf{x}_s} \triangleq \frac{\partial \boldsymbol{\psi}^t}{\partial \mathbf{x}_s} = -\mathbf{S}_{\boldsymbol{\psi}}\mathbf{D}_{bs}^{-1} \tag{E.13}$$

$$\mathbf{\Psi}_{\mathbf{y}_s} \triangleq \frac{\partial \boldsymbol{\psi}^t}{\partial \mathbf{y}_s} = \mathbf{C}_{\boldsymbol{\psi}}\mathbf{D}_{bs}^{-1} \tag{E.14}$$

$$\boldsymbol{\psi}_x^t \triangleq \frac{\partial \boldsymbol{\psi}^t}{\partial x_{mt}} = \mathbf{0}^t \tag{E.15}$$

$$\boldsymbol{\psi}_x^t \triangleq \frac{\partial \boldsymbol{\psi}^t}{\partial y_{mt}} = \mathbf{0}^t \quad . \tag{E.16}$$

Based on this we can compute each of the submatrices in (3.35). $\mathbf{J}_{22}$ is a 2×2 block matrix, each $N_s \times N_s$ submatrix of which is diagonal. Thus, we can again use block inversion and the solution is very simple since it resembles the solution of the 2×2 matrix inversion problem. Let

$$\mathbf{J}_{22} = \begin{bmatrix} \mathbf{J}_{22a} & \mathbf{J}_{22b} \\ \mathbf{J}_{22b} & \mathbf{J}_{22d} \end{bmatrix}. \tag{E.17}$$

Then

$$\mathbf{J}_{22}^{-1} = \begin{bmatrix} \mathbf{J}_{22d}\mathbf{J}_{det}^{-1} & -\mathbf{J}_{22b}\mathbf{J}_{det}^{-1} \\ -\mathbf{J}_{22b}\mathbf{J}_{det}^{-1} & \mathbf{J}_{22a}\mathbf{J}_{det}^{-1} \end{bmatrix} \tag{E.18}$$

where $\mathbf{J}_{det} = \mathbf{J}_{22a}\mathbf{J}_{22d} - \mathbf{J}_{22b}^2$ and the 3 different submatrices composing $\mathbf{J}_{22}$ are given by:

$$\mathbf{J}_{22a} = \sigma_d^{-2}\mathbf{D}_{\mathbf{x}_s}^2 + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{x}_s}^2 + \sigma_\psi^{-2}\mathbf{\Psi}_{\mathbf{x}_s}^2 \tag{E.19}$$

$$\mathbf{J}_{22b} = \sigma_d^{-2}\mathbf{D}_{\mathbf{x}_s}\mathbf{D}_{\mathbf{y}_s} + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{x}_s}\mathbf{\Phi}_{\mathbf{y}_s} + \sigma_\psi^{-2}\mathbf{\Psi}_{\mathbf{x}_s}\mathbf{\Psi}_{\mathbf{y}_s} \tag{E.20}$$

$$\mathbf{J}_{22d} = \sigma_d^{-2}\mathbf{D}_{\mathbf{y}_s}^2 + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{y}_s}^2 + \sigma_\psi^{-2}\mathbf{\Psi}_{\mathbf{y}_s}^2 \tag{E.21}$$

$\mathbf{J}_{21} = \mathbf{J}_{12}^t$ can also be expressed as a 2×2 block matrix,

$$\mathbf{J}_{21} = \begin{bmatrix} \mathbf{j}_{21a} & \mathbf{j}_{22b} \\ \mathbf{j}_{22c} & \mathbf{j}_{22d} \end{bmatrix}. \tag{E.22}$$

each element of which is a vector:

$$\mathbf{j}_{12a} = \sigma_d^{-2}\mathbf{D}_{\mathbf{x}_s}\mathbf{d}_x + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{x}_s}\boldsymbol{\phi}_x \tag{E.23}$$

$$\mathbf{j}_{12b} = \sigma_d^{-2}\mathbf{D}_{\mathbf{y}_s}\mathbf{d}_x + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{y}_s}\boldsymbol{\phi}_x \tag{E.24}$$

$$\mathbf{j}_{12c} = \sigma_d^{-2}\mathbf{D}_{\mathbf{x}_s}\mathbf{d}_y + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{x}_s}\boldsymbol{\phi}_y \tag{E.25}$$

$$\mathbf{j}_{12d} = \sigma_d^{-2}\mathbf{D}_{\mathbf{y}_s}\mathbf{d}_y + \sigma_\phi^{-2}\mathbf{\Phi}_{\mathbf{y}_s}\boldsymbol{\phi}_y \tag{E.26}$$

Last,

$$\mathbf{J}_{11} = \begin{bmatrix} \sigma_d^{-2}\mathbf{d}_x^t\mathbf{d}_x + \sigma_\phi^{-2}\boldsymbol{\phi}_x^t\boldsymbol{\phi}_x & \sigma_d^{-2}\mathbf{d}_x^t\mathbf{d}_y + \sigma_\phi^{-2}\boldsymbol{\phi}_x^t\boldsymbol{\phi}_y \\ \sigma_d^{-2}\mathbf{d}_x^t\mathbf{d}_y + \sigma_\phi^{-2}\boldsymbol{\phi}_x^t\boldsymbol{\phi}_y & \sigma_d^{-2}\mathbf{d}_y^t\mathbf{d}_y + \sigma_\phi^{-2}\boldsymbol{\phi}_y^t\boldsymbol{\phi}_y \end{bmatrix}. \tag{E.27}$$

Substituting the 4 submatrices given by(E.18),(E.22),(E.27) into (3.35) we obtain after some algebraic computations the following 4 elements for $\mathbf{J}_p$:

$$j_{p11} = \mathbf{d}_x^t\mathbf{F}_D\mathbf{d}_x + \boldsymbol{\phi}_x^t\mathbf{F}_\Phi\boldsymbol{\phi}_x + 2\mathbf{d}_x^t\mathbf{F}\boldsymbol{\phi}_x \tag{E.28}$$

$$j_{p22} = \mathbf{d}_y^t\mathbf{F}_D\mathbf{d}_y + \boldsymbol{\phi}_y^t\mathbf{F}_\Phi\boldsymbol{\phi}_y + 2\mathbf{d}_y^t\mathbf{F}\boldsymbol{\phi}_y \tag{E.29}$$

$$j_{p12} = j_{p21} = \mathbf{d}_x^t\mathbf{F}_D\mathbf{d}_y + \boldsymbol{\phi}_x^t\mathbf{F}_\Phi\boldsymbol{\phi}_y + \mathbf{d}_x^t\mathbf{F}\boldsymbol{\phi}_y + \mathbf{d}_y^t\mathbf{F}\boldsymbol{\phi}_x \tag{E.30}$$

where we have introduced the matrices

$$\mathbf{F}_D = \sigma_d^{-2}\mathbf{I} - \sigma_d^{-4}\left(\mathbf{D}_{\mathbf{x}_s}^2\mathbf{J}_{22a} + \mathbf{D}_{\mathbf{y}_s}^2\mathbf{J}_{22d} - 2\mathbf{D}_{\mathbf{x}_s}\mathbf{D}_{\mathbf{y}_s}\mathbf{J}_{22b}\right)\mathbf{J}_{det}^{-1} \tag{E.31}$$

$$\mathbf{F}_\Phi = \sigma_\phi^{-2}\mathbf{I} - \sigma_\phi^{-4}\left(\boldsymbol{\Phi}_{\mathbf{x}_s}^2\mathbf{J}_{22a} + \boldsymbol{\Phi}_{\mathbf{y}_s}^2\mathbf{J}_{22d} - 2\boldsymbol{\Phi}_{\mathbf{x}_s}\boldsymbol{\Phi}_{\mathbf{y}_s}\mathbf{J}_{22b}\right)\mathbf{J}_{det}^{-1} \tag{E.32}$$

$$\mathbf{F} = -\sigma_d^{-2}\sigma_\phi^{-2}\left(\mathbf{D}_{\mathbf{x}_s}\boldsymbol{\Phi}_{\mathbf{x}_s}\mathbf{J}_{22a} + \mathbf{D}_{\mathbf{y}_s}\boldsymbol{\Phi}_{\mathbf{y}_s}\mathbf{J}_{22d} - (\mathbf{D}_{\mathbf{x}_s}\boldsymbol{\Phi}_{\mathbf{y}_s} + \mathbf{D}_{\mathbf{y}_s}\boldsymbol{\Phi}_{\mathbf{x}_s})\mathbf{J}_{22b}\right)\mathbf{J}_{det}^{-1} \tag{E.33}$$

Replacing the vectors that contain partial derivatives into the entries of $\mathbf{J}_{int}$, we obtain

$$j_{p11} = \mathbf{1}^t(\mathbf{C}_{\boldsymbol{\phi}}^2\mathbf{F}_D + \mathbf{D}_{mts}^{-2}\mathbf{S}_{\boldsymbol{\phi}}^2\mathbf{F}_\Phi - 2\mathbf{D}_{mts}^{-1}\mathbf{C}_{\boldsymbol{\phi}}\mathbf{S}_{\boldsymbol{\phi}}\mathbf{F})\mathbf{1} \tag{E.34}$$

$$j_{p22} = \mathbf{1}^t(\mathbf{S}_{\boldsymbol{\phi}}^2\mathbf{F}_D + \mathbf{D}_{mts}^{-2}\mathbf{C}_{\boldsymbol{\phi}}^2\mathbf{F}_\Phi + 2\mathbf{D}_{mts}^{-1}\mathbf{C}_{\boldsymbol{\phi}}\mathbf{S}_{\boldsymbol{\phi}}\mathbf{F})\mathbf{1} \tag{E.35}$$

$$j_{p12} = j_{p21} = \mathbf{1}^t\left(\mathbf{C}_{\boldsymbol{\phi}}\mathbf{S}_{\boldsymbol{\phi}}\mathbf{F}_D - \mathbf{D}_{mts}^{-2}\mathbf{C}_{\boldsymbol{\phi}}\mathbf{S}_{\boldsymbol{\phi}}\mathbf{F}_\Phi + \mathbf{D}_{mts}^{-1}(\mathbf{C}_{\boldsymbol{\phi}}^2 - \mathbf{S}_{\boldsymbol{\phi}}^2)\mathbf{F}\right)\mathbf{1} \tag{E.36}$$

Finally, replacing the matrices that contain the partial derivatives from (E.5)-(E.13) into the submatrices of $\mathbf{J}_{22}$ and the matrices $\mathbf{F}$ and subsequently the results into the entries of $\mathbf{G}$, we obtain

$$j_{p11} = \mathbf{1}^t(\mathbf{Q}'_{\boldsymbol{\phi}+\psi}\bar{\mathbf{J}}_{det}^{-1})\mathbf{1} \tag{E.37}$$

$$j_{p22} = \mathbf{1}^t(\mathbf{Q}_{\boldsymbol{\phi}+\psi}\bar{\mathbf{J}}_{det}^{-1})\mathbf{1} \tag{E.38}$$

$$j_{p12} = j_{p21} = -\mathbf{1}^t(\mathbf{S}_{\boldsymbol{\phi}+\psi}\bar{\mathbf{J}}_{det}^{-1})\mathbf{1} \tag{E.39}$$

where

$$\begin{aligned}\bar{\mathbf{J}}_{det} &= (\sigma_d^2\sigma_\phi^2\sigma_\psi^2)\mathbf{D}_{bs}^2\mathbf{D}_{mts}^2\mathbf{C}_I^{-1}\mathbf{J}_{det} \\ &= ((\sigma_\psi^2\mathbf{D}_{bs}^2 + \sigma_\phi^2\mathbf{D}_{mts}^2)\mathbf{Q}_{\boldsymbol{\phi}-\boldsymbol{\psi}} + \sigma_d^2\mathbf{Q}'_{\boldsymbol{\phi}-\boldsymbol{\psi}})\end{aligned} \tag{E.40}$$

and we have introduced

$$\mathbf{Q}_{\boldsymbol{\phi}-\boldsymbol{\psi}} = \mathbf{I} + \mathbf{C}_{\boldsymbol{\phi}-\boldsymbol{\psi}} \tag{E.41}$$

$$\mathbf{Q}'_{\boldsymbol{\phi}-\boldsymbol{\psi}} = \mathbf{I} - \mathbf{C}_{\boldsymbol{\phi}-\boldsymbol{\psi}} \tag{E.42}$$

$$\mathbf{Q}_{\boldsymbol{\phi}+\boldsymbol{\psi}} = \mathbf{I} + \mathbf{C}_{\boldsymbol{\phi}+\boldsymbol{\psi}} \tag{E.43}$$

$$\mathbf{Q}'_{\boldsymbol{\phi}+\boldsymbol{\psi}} = \mathbf{I} - \mathbf{C}_{\boldsymbol{\phi}+\boldsymbol{\psi}} \quad . \tag{E.44}$$

Appendix F

Optimal Weighting Matrix for $K = 4$ DSBM WLS

The results for the optimal weighting matrix derived in appendix D for the SBM-based localization method, can be easily extended to the DSBM-based localization method with $K = 4$ kinds of LDP available. We again need to introduce hats $(\hat{\cdot})$ in our notation since both true values of LDP and their estimates will be involved in the expressions. This time we need to replace $\mathbf{b}'$ in eq. (4.12) with $\hat{\mathbf{b}}' = \mathbf{b}' + \tilde{\mathbf{b}}'$ and rewrite it so that any symbol representing an estimate has a $(\hat{\cdot})$ as follows:ting an estimate has a $(\hat{\cdot})$ as follows:

$$\hat{\mathbf{b}}' = \begin{bmatrix} \hat{\mathbf{b}} \\ \mathbf{F}'_d \mathbf{1} \end{bmatrix} \tag{F.1}$$

To compute $\mathbf{C}_{\hat{\mathbf{b}}'} = \mathbf{C}_{\tilde{\mathbf{b}}'}$, we will use the prior distributions of the LDP estimates and errors introduced in appendix D and further introduce the priors for the DS estimates and errors

$$\tilde{\mathbf{f}}_d \sim \mathcal{N}(\mathbf{0}, \sigma^2_{f_d}\mathbf{I}) \Leftrightarrow \hat{\mathbf{f}}_d \sim \mathcal{N}(\mathbf{f}_d, \sigma^2_{f_d}\mathbf{I}) \tag{F.2}$$

As before, we assume that LDP estimates (and errors) of different kind are independent. Thus, the vector $\tilde{\mathbf{f}}_d$ containing the errors in DS estimates is independent of $\tilde{\mathbf{b}}$ and since both have zero mean, it follows that $\mathcal{E}\{\tilde{\mathbf{f}}_d \tilde{\mathbf{b}}^t\} = \mathbf{O}$. Since the covariance matrix of $\mathbf{b}$ was derived in appendix D, it follows

immediately that

$$\mathbf{C}_{\tilde{\mathbf{b}}'} = \begin{bmatrix} \mathbf{C}_{\tilde{\mathbf{b}}} & \mathbf{O} \\ \mathbf{O} & \sigma_{f_d}^2 \mathbf{I} \end{bmatrix}. \tag{F.3}$$

Using this covariance matrix an optimal WLS solution to the DSBM-based localization method with $K = 4$ kinds of LDP available, can be derived.

Appendix G

Formulation of the 2^{nd} step of $K = 3$ DSBM WLS

In this appendix we will derive eq. (4.35), i.e. we will formulate the WLS algorithm for the 2^{nd} step of the $K = 3$ DSBM-based localization algorithm. To do so, we need to put the set of eq. (4.30) in a meaningful vector form. Lets start with the r.h.s. and define the following vectors

$$\mathbf{b}_i \stackrel{\Delta}{=} [b_{i1}, \ldots, b_{iN_s}]^t \tag{G.1}$$

$$\mathbf{b} \stackrel{\Delta}{=} [\mathbf{b}_1^t, \ldots, \mathbf{b}_{N_t}^t]^t \tag{G.2}$$

$$\mathbf{q} \stackrel{\Delta}{=} \mathbf{1} \otimes [q_1, \ldots, q_{N_s}]^t \tag{G.3}$$

$$\mathbf{p}_{bs} \stackrel{\Delta}{=} \mathbf{1} \otimes [p_{bs_1}, \ldots, p_{bs_{N_s}}]^t. \tag{G.4}$$

According to eq. (4.33)we can write

$$\begin{aligned} \mathbf{b}_{(N_s)} &= -\frac{1}{2}(\mathbf{d} \odot \mathbf{d} - 2(\mathbf{d} + \mathbf{q}) \odot (\mathbf{1} \otimes \hat{\mathbf{d}}_{bs}) - \mathbf{p}_{bs}) \\ &= -(\mathbf{d} \odot (\frac{1}{2}\mathbf{d} - (\mathbf{1} \otimes \hat{\mathbf{d}}_{bs})) - \mathbf{q} \odot (\mathbf{1} \otimes \hat{\mathbf{d}}_{bs}) - \frac{1}{2}\mathbf{p}_{bs}) \\ &= \mathbf{d} \odot ((\mathbf{1} \otimes \hat{\mathbf{d}}_{bs}) - \frac{1}{2}\mathbf{d}) + \mathbf{q} \odot (\mathbf{1} \otimes \hat{\mathbf{d}}_{bs}) + \frac{1}{2}\mathbf{p}_{bs}). \end{aligned} \tag{G.5}$$

It should be noted that the subscript (N_s) is used to denote that the MT can communicate with up to N_s BS through 1 signal component (path) for

each BS. If the MT is communicating with just 1 BS through several MPC, the expression for the above vector simplifies to:

$$\mathbf{b}_{(1)} = \mathbf{d} \odot ((\mathbf{1} \otimes \hat{\mathbf{d}}_{bs}) - \frac{1}{2}\mathbf{d}) \tag{G.6}$$

From the 1^{st} step of $K = 3$ DSBM WLS method, the estimates of the distances between scatterers and the corresponding BS are given by

$$\begin{aligned}
\hat{\mathbf{d}}_{bs} &= (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1} \mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{w} \\
&= (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1} \mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{R}_t (\mathbf{f}_d \odot \mathbf{d}) \\
&= (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1} (\mathbf{1}^t \otimes \mathbf{I}) \mathbf{F}_d^t \mathbf{R}_t^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{R}_t \mathbf{F}_d \mathbf{d} \\
&= (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1} (\mathbf{1}^t \otimes \mathbf{I}) (\mathbf{R}_t^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{R}_t \odot \mathbf{f}_d \mathbf{f}_d^t) \mathbf{d} \\
&= \underbrace{(\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1} (\mathbf{1}^t \otimes \mathbf{I}) \mathbf{B}_f}_{\mathbf{G}_f} \mathbf{d}
\end{aligned} \tag{G.7}$$

Using the above expression we can write

$$\begin{aligned}
\mathbf{1} \otimes \hat{\mathbf{d}}_{bs} &= \mathbf{1} \otimes (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1} (\mathbf{1}^t \otimes \mathbf{I}) \mathbf{B}_f \mathbf{d} \\
&= (\mathbf{1} \otimes (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1}) (\mathbf{1}^t \otimes \mathbf{I}) \mathbf{B}_f \mathbf{d} \\
&= \underbrace{(\mathbf{1}\mathbf{1}^t \otimes (\mathbf{Z}_f^t \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{P} \mathbf{Z}_f)^{-1})}_{\mathbf{H}_f} \mathbf{B}_f \mathbf{d}
\end{aligned} \tag{G.8}$$

Substituting this quantity in eq. (G.5) and using the "space difference" matrix given by eq. (4.34) we get the r.h.s of eq. (4.35)

$$\begin{aligned}
\mathbf{b} &\triangleq \mathbf{R}_s \mathbf{b}_{(N_s)} = \mathbf{R}_s (\mathbf{d} \odot (\mathbf{H}_f \mathbf{B}_f \mathbf{d} - \frac{1}{2}\mathbf{d}) + \mathbf{q} \odot (\mathbf{H}_f \mathbf{B}_f \mathbf{d}) + \frac{1}{2}\mathbf{p}_{bs}) \\
&= \mathbf{R}_s (\mathbf{d} \odot (\underbrace{(\mathbf{H}_f \mathbf{B}_f - \frac{1}{2}\mathbf{I})}_{\mathbf{\Gamma}_f} \mathbf{d}) + \underbrace{\mathbf{Q}\mathbf{H}_f \mathbf{B}_f}_{\mathbf{\Delta}_f} \mathbf{d} + \frac{1}{2}\mathbf{p}_{bs}) \\
&= \mathbf{R}_s (\mathbf{d} \odot \mathbf{\Gamma}_f \mathbf{d} + \mathbf{\Delta}_f \mathbf{d} + \frac{1}{2}\mathbf{p}_{bs}).
\end{aligned} \tag{G.9}$$

or equivalently, for the case of 1 BS,

$$\mathbf{b} = \mathbf{R}_s (\mathbf{d} \odot \mathbf{\Gamma}_f \mathbf{d}). \tag{G.10}$$

To procede with the derivation of the l.h.s. lets define the matrix of time instances $\mathbf{T}_n$, which depends on the order of the mobility model n. It is

given by

$$\mathbf{T}_n = \begin{cases} [\mathbf{1}|\mathbf{t}], & n=1 \\ [\mathbf{1}|\mathbf{t}|\frac{1}{2}(\mathbf{t}\odot\mathbf{t})], & n=2 \end{cases} \tag{G.11}$$

For the constant speed mobility model, the $N_t \times 1$ vectors of the MT coordinates are given by:

$$\mathbf{x}_{mt} = x_0\mathbf{1} + v_x\mathbf{t} = [\mathbf{1}|\mathbf{t}] \begin{bmatrix} x_0 \\ v_x \end{bmatrix} = \mathbf{T}_1\mathbf{p}_{int,x} \tag{G.12}$$

$$\mathbf{y}_{mt} = y_0\mathbf{1} + v_y\mathbf{t} = [\mathbf{1}|\mathbf{t}] \begin{bmatrix} y_0 \\ v_y \end{bmatrix} = \mathbf{T}_1\mathbf{p}_{int,y} \tag{G.13}$$

where we have introduced the vectors containing the parameters of interest along a specific axis (eg. x-axis). Similarly for the constant acceleration mobility model

$$\mathbf{x}_{mt} = \mathbf{T}_2\mathbf{p}_{int,x} \tag{G.14}$$

$$\mathbf{y}_{mt} = \mathbf{T}_2\mathbf{p}_{int,y} \tag{G.15}$$

Let

$$\mathbf{s} \stackrel{\Delta}{=} [s_1, \ldots, s_{N_s}]^t \tag{G.16}$$

$$\mathbf{c} \stackrel{\Delta}{=} [c_1, \ldots, c_{N_s}]^t. \tag{G.17}$$

From eq. (4.31-4.32) we can write

$$\mathbf{c} = \mathbf{x}_{bs} + \mathbf{C}_\psi\hat{\mathbf{d}}_{bs} \tag{G.18}$$

$$\mathbf{s} = \mathbf{y}_{bs} + \mathbf{S}_\psi\hat{\mathbf{d}}_{bs}. \tag{G.19}$$

According to eq. (4.30), each entry of $\mathbf{x}_{mt}$ should be multiplied with each entry of $\mathbf{c}$ and each entry of $\mathbf{y}_{mt}$ should be multiplied with each entry of $\mathbf{s}$, in order to form the $N_tN_s \times 1$ vector, the product of which with $\mathbf{R}_s$ will be

equal to the l.h.s. of eq. (4.35)

$$\begin{aligned}
\mathbf{R}_s(\mathbf{x} \otimes \mathbf{c} + \mathbf{y} \otimes \mathbf{c}) &= \mathbf{R}_s(\mathbf{T}_n \mathbf{p}_{int,x} \otimes \mathbf{c} + \mathbf{T}_n \mathbf{p}_{int,y} \otimes \mathbf{s}) \\
&= \mathbf{R}_s((\mathbf{T}_n \otimes \mathbf{c})\mathbf{p}_{int,x} + (\mathbf{T}_n \otimes \mathbf{s})\mathbf{p}_{int,y}) \\
&= \mathbf{R}_s \left([(\mathbf{T}_n \otimes \mathbf{c})|(\mathbf{T}_n \otimes \mathbf{s})] \begin{bmatrix} \mathbf{p}_{int,x} \\ \mathbf{p}_{int,y} \end{bmatrix} \right) \\
&= \mathbf{R}_s([(\mathbf{T}_n \otimes \mathbf{c})|(\mathbf{T}_n \otimes \mathbf{s})]\mathbf{p}_{int}) \\
&= \mathbf{R}_s((\mathbf{T}_n \otimes [\mathbf{x}_{bs} + \mathbf{C}_\psi \hat{\mathbf{d}}_{bs}|\mathbf{y}_{bs} + \mathbf{S}_\psi \hat{\mathbf{d}}_{bs}])\mathbf{p}_{int}) \\
&= \mathbf{R}_s \underbrace{(\mathbf{T}_n \otimes [\mathbf{x}_{bs} + \mathbf{C}_\psi \mathbf{G}_f \mathbf{d}|\mathbf{y}_{bs} + \mathbf{S}_\psi \mathbf{G}_f \mathbf{d}])}_{\mathbf{A}_{(N_s)}} \mathbf{p}_{int} \\
&= \underbrace{\mathbf{R}_s \mathbf{A}_{(N_s)}}_{\mathbf{A}} \mathbf{p}_{int}
\end{aligned} \tag{G.20}$$

Again the subscript (N_s) is used to denote that the MT can communicate with up to N_s BS through 1 signal component (path) for each BS. If the MT is communicating with just 1 BS through several MPC, the expression for the $\mathbf{A}_{(.)}$ simplifies to:

$$\mathbf{A}_{(1)} = \mathbf{T}_n \otimes [\mathbf{C}_\psi \mathbf{G}_f \mathbf{d}|\mathbf{S}_\psi \mathbf{G}_f \mathbf{d}] \tag{G.21}$$

Appendix H

Optimal Weighting Matrix for $K = 3$ DSBM WLS

We derive herein the covariance matrix of $\hat{\mathbf{b}}'$, given by

$$\hat{\mathbf{b}}' = (\hat{\mathbf{\Gamma}}_f \hat{\mathbf{d}} \odot \hat{\mathbf{d}}) \tag{H.1}$$

where

$$\mathbf{\Gamma}_f = \mathbf{H}(\mathbf{R}_t^t \mathbf{C}_{\hat{\mathbf{w}}}^{-1} \mathbf{P} \mathbf{R}_t \odot \hat{\mathbf{f}}_d \hat{\mathbf{f}}_d^t) - \frac{1}{2}\mathbf{I} \tag{H.2}$$

and the matrices composing $\mathbf{\Gamma}_f$ have already been defined in section 4.3. Throughout this derivation we make extensive use of the formulas given in app. A with referencing them. The vector $\hat{\mathbf{b}}'$ contains errors and so we can write

$$\hat{\mathbf{b}}' = \mathbf{b}' + \tilde{\mathbf{b}}' \Rightarrow \mathcal{E}\{\hat{\mathbf{b}}'\} = \mathbf{b}' + \mathcal{E}\{\tilde{\mathbf{b}}'\}$$

where the error vector and its mean are given by

$$\begin{aligned} \tilde{\mathbf{b}}' = \mathbf{\Gamma}_f \mathbf{d} \odot \tilde{\mathbf{d}} + \mathbf{\Gamma}_f \tilde{\mathbf{d}} \odot \mathbf{d} + \mathbf{\Gamma}_f \tilde{\mathbf{d}} \odot \tilde{\mathbf{d}} + \tilde{\mathbf{\Gamma}}_f \mathbf{d} \odot \mathbf{d} \\ + \tilde{\mathbf{\Gamma}}_f \mathbf{d} \odot \tilde{\mathbf{d}} + \tilde{\mathbf{\Gamma}}_f \tilde{\mathbf{d}} \odot \mathbf{d} + \tilde{\mathbf{\Gamma}}_f \tilde{\mathbf{d}} \odot \tilde{\mathbf{d}} \end{aligned} \tag{H.3}$$

and we have introduced

$$\tilde{\mathbf{\Gamma}}_f = \mathbf{H}(\mathbf{X}_r \odot \tilde{\mathbf{\Phi}}) \tag{H.4}$$

$$\tilde{\mathbf{\Phi}} = \mathbf{f}_d \tilde{\mathbf{f}}_d^t + \tilde{\mathbf{f}}_d \mathbf{f}_d^t + \tilde{\mathbf{f}}_d \tilde{\mathbf{f}}_d^t \tag{H.5}$$

with mean

$$\mathcal{E}\{\tilde{\mathbf{\Gamma}}_f\} = \sigma_{f_d}^2 \mathbf{H}(\mathbf{X}_r \odot \mathbf{I}). \tag{H.6}$$

It is straightforward to show that

$$\mathcal{E}\{\tilde{\mathbf{b}}'\} = \left[\sigma_{f_d}^2 \mathbf{D}\mathbf{H}(\mathbf{X}_r \odot \mathbf{I})\mathbf{D} + \sigma_d^2 \mathbf{\Gamma}_f \odot \mathbf{I} + \sigma_{f_d}^2 \sigma_d^2 (\mathbf{H}(\mathbf{X}_r \odot \mathbf{I})) \odot \mathbf{I}\right] \mathbf{1}. \tag{H.7}$$

It becomes obvious that the (W)LS estimator will be biased unless we subtract the mean of the error $\mathcal{E}\{\tilde{\mathbf{b}}\}$ from $\hat{\mathbf{b}}$. Since

$$\mathbf{C}_{\hat{\mathbf{b}}'} = \mathbf{C}_{\tilde{\mathbf{b}}'} = \mathbf{R}_{\tilde{\mathbf{b}}'} - \mathcal{E}\{\tilde{\mathbf{b}}'\}\mathcal{E}\{\tilde{\mathbf{b}}'^t\} \tag{H.8}$$

we will again restrict ourselves to deriving the correlation matrix $\mathbf{R}_{\tilde{\mathbf{b}}'}$ of the error vector. The product of $\tilde{\mathbf{b}}'$ obtained from (H.3), with its transpose, leads to a sum of 49 terms. Out of 49 terms, only 25 are non-zero (pairs of terms with 1 error vector $\tilde{\mathbf{d}}$ and pairs of terms with none or 2 error vectors $\tilde{\mathbf{d}}$). From these 25 terms we only need to compute the 7 covariance terms and 9 cross-covariance terms (the other 9 cross-covariance terms are the transposes of the ones computed). Furthermore, to reduce computations significantly, without introducing a noticeable approximation error, we will ignore terms that depend only on moments of order 6 or higher[1] and thus we only need to compute approximately the 6 covariance terms and the 8 cross-covariance terms. The 6 covariance terms are given by

$$\mathcal{E}\{\mathbf{\Gamma}_f \mathbf{d}\mathbf{d}^t \mathbf{\Gamma}_f^t \odot \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t\} = \sigma_d^2 \mathbf{\Gamma}_f \mathbf{d}\mathbf{d}^t \mathbf{\Gamma}_f^t \odot \mathbf{I} \tag{H.9}$$

$$\mathcal{E}\{\mathbf{\Gamma}_f \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t \mathbf{\Gamma}_f^t \odot \mathbf{d}\mathbf{d}^t\} = \sigma_d^2 \mathbf{\Gamma}_f \mathbf{\Gamma}_f^t \odot \mathbf{d}\mathbf{d}^t \tag{H.10}$$

$$\mathcal{E}\{\tilde{\mathbf{\Gamma}}_f \mathbf{d}\mathbf{d}^t \tilde{\mathbf{\Gamma}}_f^t \odot \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t\} = \sigma_d^2 \mathbf{H}\mathbf{C}_{\tilde{\mathbf{\Gamma}}_2}\mathbf{H}^t \odot \mathbf{I} \tag{H.11}$$

$$\mathcal{E}\{\tilde{\mathbf{\Gamma}}_f \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t \tilde{\mathbf{\Gamma}}_f^t \odot \mathbf{d}\mathbf{d}^t\} = \sigma_d^2 \mathbf{H}\mathbf{C}_{\tilde{\mathbf{d}}\tilde{\mathbf{d}}^t}\mathbf{H}^t \odot \mathbf{d}\mathbf{d}^t \tag{H.12}$$

$$\mathcal{E}\{\mathbf{\Gamma}_f \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t \mathbf{\Gamma}_f^t \odot \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t\} = \sigma_d^4 \mathbf{C}_{\mathbf{d}_4} \tag{H.13}$$

$$\mathcal{E}\{\tilde{\mathbf{\Gamma}}_f \mathbf{d}\mathbf{d}^t \tilde{\mathbf{\Gamma}}_f^t \odot \mathbf{d}\mathbf{d}^t\} = \mathbf{H}(\mathbf{C}_{\tilde{\mathbf{\Gamma}}_2} + \mathbf{C}_{\tilde{\mathbf{\Gamma}}_4})\mathbf{H}^t \odot \mathbf{d}\mathbf{d}^t \tag{H.14}$$

[1] Similarly we could also ignore all the terms that depend only on moments of order 4 or higher

and the 8 cross-covariance terms are given by

$$\mathcal{E}\{\boldsymbol{\Gamma}_f\mathbf{d}\tilde{\mathbf{d}}^t\boldsymbol{\Gamma}_f^t \odot \tilde{\mathbf{d}}\mathbf{d}^t\} = \sigma_d^2(\boldsymbol{\Gamma}_f\mathbf{d}\mathbf{1}^t \odot \mathbf{I})\boldsymbol{\Gamma}_f^t\mathbf{D} \tag{H.15}$$

$$\mathcal{E}\{\boldsymbol{\Gamma}_f\mathbf{d}\mathbf{d}^t\tilde{\boldsymbol{\Gamma}}_f^t \odot \tilde{\mathbf{d}}\tilde{\mathbf{d}}^t\} = \sigma_d^2\boldsymbol{\Gamma}_f\mathbf{d}\mathbf{d}^t\mathcal{E}\{\tilde{\boldsymbol{\Gamma}}_f^t\} \odot \mathbf{I} \tag{H.16}$$

$$\mathcal{E}\{\boldsymbol{\Gamma}_f\mathbf{d}\tilde{\mathbf{d}}^t\tilde{\boldsymbol{\Gamma}}_f^t \odot \tilde{\mathbf{d}}\mathbf{d}^t\} = \sigma_d^2(\boldsymbol{\Gamma}_f\mathbf{d}\mathbf{1}^t \odot \mathbf{I})\mathcal{E}\{\boldsymbol{\Gamma}_f^t\}\mathbf{D} \tag{H.17}$$

$$\mathcal{E}\{\boldsymbol{\Gamma}_f\tilde{\mathbf{d}}\mathbf{d}^t\tilde{\boldsymbol{\Gamma}}_f^t \odot \mathbf{d}\tilde{\mathbf{d}}^t\} = \sigma_d^2\mathbf{D}\boldsymbol{\Gamma}_f(\mathbf{1}\mathbf{d}^t\mathcal{E}\{\boldsymbol{\Gamma}_f^t\} \odot \mathbf{I}) \tag{H.18}$$

$$\mathcal{E}\{\boldsymbol{\Gamma}_f\tilde{\mathbf{d}}\tilde{\mathbf{d}}^t\tilde{\boldsymbol{\Gamma}}_f^t \odot \mathbf{d}\mathbf{d}^t\} = \sigma_d^2\boldsymbol{\Gamma}_f\mathcal{E}\{\tilde{\boldsymbol{\Gamma}}_f^t\} \odot \mathbf{d}\mathbf{d}^t \tag{H.19}$$

$$\mathcal{E}\{\tilde{\boldsymbol{\Gamma}}_f\mathbf{d}\tilde{\mathbf{d}}^t\tilde{\boldsymbol{\Gamma}}_f^t \odot \tilde{\mathbf{d}}\mathbf{d}^t\} = \sigma_d^2\mathbf{C}_{\tilde{\mathbf{d}}^t}\mathbf{H}^t\mathbf{D} \tag{H.20}$$

$$\mathcal{E}\{\tilde{\boldsymbol{\Gamma}}_f\mathbf{d}\tilde{\mathbf{d}}^t\boldsymbol{\Gamma}_f^t \odot \mathbf{d}\tilde{\mathbf{d}}^t\} = \sigma_d^2\mathbf{D}\mathcal{E}\{\boldsymbol{\Gamma}_f\}\mathbf{d}\mathbf{1}^t(\boldsymbol{\Gamma}_f^t \odot \mathbf{I}) \tag{H.21}$$

$$\mathcal{E}\{\tilde{\boldsymbol{\Gamma}}_f\tilde{\mathbf{d}}\mathbf{d}^t\tilde{\boldsymbol{\Gamma}}_f^t \odot \tilde{\mathbf{d}}\mathbf{d}^t\} = \sigma_d^2\mathbf{C}_{\tilde{\mathbf{d}}}\mathbf{H}^t\mathbf{D} \tag{H.22}$$

where we have introduced the matrices

$$\begin{aligned}\mathbf{C}_{\tilde{\boldsymbol{\Gamma}}_2} = \sigma_{f_d}^2(&\mathbf{F}_d\mathbf{X}_r(\mathbf{d}\mathbf{d}^t \odot \mathbf{I})\mathbf{X}_r\mathbf{F}_d + \mathbf{X}_r\mathbf{F}_d\mathbf{d}\mathbf{d}^t\mathbf{F}_d\mathbf{X}_r \odot \mathbf{I} \\ &+ (\mathbf{X}_r\mathbf{F}_d\mathbf{d}\mathbf{d}^t \odot \mathbf{I})\mathbf{X}_r\mathbf{F}_d + \mathbf{F}_d\mathbf{X}_r(\mathbf{d}\mathbf{d}^t\mathbf{F}_d\mathbf{X}_r \odot \mathbf{I}))\end{aligned} \tag{H.23}$$

$$\begin{aligned}\mathbf{C}_{\tilde{\boldsymbol{\Gamma}}_4} = \sigma_{f_d}^4(&(\mathbf{X}_r \odot \mathbf{I})\mathbf{d}\mathbf{d}^t(\mathbf{X}_r \odot \mathbf{I}) + \mathbf{X}_r \odot \mathbf{d}\mathbf{d}^t \odot \mathbf{X}_r \\ &+ \mathbf{X}_r(\mathbf{d}\mathbf{d}^t \odot \mathbf{I})\mathbf{X}_r + 3(\mathbf{X}_r \odot \mathbf{I})(\mathbf{d}\mathbf{d}^t \odot \mathbf{I})(\mathbf{X}_r \odot \mathbf{I}))\end{aligned} \tag{H.24}$$

$$\begin{aligned}\mathbf{C}_{\tilde{\mathbf{d}}\tilde{\mathbf{d}}^t} = \sigma_{f_d}^2(&\mathbf{F}_d\mathbf{X}_r\mathbf{X}_r\mathbf{F}_d + \mathbf{X}_r\mathbf{F}_d\mathbf{F}_d\mathbf{X}_r \odot \mathbf{I} \\ &+ (\mathbf{X}_r\mathbf{F}_d \odot \mathbf{I})\mathbf{X}_r\mathbf{F}_d + \mathbf{F}_d\mathbf{X}_r(\mathbf{F}_d\mathbf{X}_r) \odot \mathbf{I})\end{aligned} \tag{H.25}$$

$$\begin{aligned}\mathbf{C}_{\mathbf{d}_4} = &(\boldsymbol{\Gamma}_f \odot \mathbf{I})\mathbf{1}\mathbf{1}^t(\boldsymbol{\Gamma}_f^t \odot \mathbf{I}) + (\boldsymbol{\Gamma}_f\boldsymbol{\Gamma}_f^t) \odot \mathbf{I} \\ &+ \boldsymbol{\Gamma}_f \odot \mathbf{1}\mathbf{1}^t \odot \boldsymbol{\Gamma}_f^t + 3(\boldsymbol{\Gamma}_f \odot \mathbf{I})(\boldsymbol{\Gamma}_f^t \odot \mathbf{I})\end{aligned} \tag{H.26}$$

$$\begin{aligned}\mathbf{C}_{\tilde{\mathbf{d}}^t} = \sigma_{f_d}^2(&(\mathbf{H}\mathbf{F}_d\mathbf{X}_r \odot \mathbf{D})\mathbf{X}_r\mathbf{F}_d + \mathbf{H} \odot \mathbf{1}\mathbf{d}^t\mathbf{F}_d\mathbf{X}_r \odot \mathbf{F}_d\mathbf{X}_r \\ &+ (\mathbf{H} \odot \mathbf{1}\mathbf{d}^t\mathbf{F}_d\mathbf{X}_r \odot \mathbf{I})\mathbf{X}_r\mathbf{F}_d + \mathbf{H}\mathbf{F}_d\mathbf{X}_r \odot \mathbf{1}\mathbf{d}^t \odot \mathbf{F}_d\mathbf{X}_r)\end{aligned} \tag{H.27}$$

$$\begin{aligned}\mathbf{C}_{\tilde{\mathbf{d}}} = \sigma_{f_d}^2(&(\mathbf{H}\mathbf{F}_d\mathbf{X}_r \odot \mathbf{D})\mathbf{X}_r\mathbf{F}_d + \mathbf{H} \odot \mathbf{1}\mathbf{d}^t\mathbf{F}_d\mathbf{X}_r \odot \mathbf{F}_d\mathbf{X}_r \\ &+ (\mathbf{H} \odot \mathbf{1}\mathbf{d}^t \odot \mathbf{F}_d\mathbf{X}_r)\mathbf{X}_r\mathbf{F}_d + \mathbf{H}\mathbf{F}_d\mathbf{X}_r \odot \mathbf{1}\mathbf{d}^t\mathbf{F}_d\mathbf{X}_r \odot \mathbf{I})\end{aligned} \tag{H.28}$$

Summing up the terms give by eq. (H.9)-(H.22), we obtain $\mathbf{C}_{\hat{\mathbf{b}}'}$. The optimal weighting matrix for the WLS solution given by eq. (4.36) is simply given by $\mathbf{C}_{\hat{\mathbf{b}}} = \mathbf{R}_s\mathbf{C}_{\hat{\mathbf{b}}'}\mathbf{R}_s^t$.

Appendix I

Derivatives of LDP w.r.t. the Unknown Parameters

In what follows we consider a dynamic environment where the MT moves with constant speed. In such an environment, the DSBM can be utilized to describe the most significant signal components and therefore the LDP of the corresponding paths are given by eq. (1.27) and (1.38)-(1.40). The computation of the partial derivatives of the AoA ϕ, AoD ψ, path lengths d and DS f_d is straightforward and the results are given below [66]:

$$\frac{\partial \phi_{ij}}{\partial y_{s_j}} = -\frac{\partial \phi_{ij}}{\partial y_0} = -\frac{1}{t_i}\frac{\partial \phi_{ij}}{\partial \upsilon_y} = \frac{x_{s_j} - x_0 - \upsilon_x t_i}{d^2_{mts,ij}} \tag{I.1}$$

$$-\frac{\partial \phi_{ij}}{\partial x_{s_j}} = \frac{\partial \phi_{ij}}{\partial x_0} = \frac{1}{t_i}\frac{\partial \phi_{ij}}{\partial \upsilon_x} = \frac{y_{s_j} - y_0 - \upsilon_y t_i}{d^2_{mts,ij}} \tag{I.2}$$

$$\frac{\partial \psi_{ij}}{\partial y_{s_j}} = \frac{x_{s_j} - x_{bs_j}}{d^2_{bs,j}} \tag{I.3}$$

$$\frac{\partial \psi_{ij}}{\partial x_{s_j}} = -\frac{y_{s_j} - y_{bs_j}}{d^2_{bs,j}} \tag{I.4}$$

$$\frac{\partial \psi_{ij}}{\partial y_0} = \frac{\partial \psi_{ij}}{\partial \upsilon_y} = \frac{\partial \psi_{ij}}{\partial x_0} = \frac{\partial \psi_{ij}}{\partial \upsilon_x} = 0 \tag{I.5}$$

$$\frac{\partial d_{ij}}{\partial y_{sj}} = \frac{y_{s_j} - y_0 - \upsilon_y t_i}{d_{mts,ij}} + \frac{y_{s_j} - y_{bs_j}}{d_{bs,j}} \tag{I.6}$$

$$\frac{\partial d_{ij}}{\partial y_0} = \frac{1}{t_i}\frac{\partial d_{ij}}{\partial \upsilon_y} = -\frac{y_{s_j} - y_0 - \upsilon_y t_i}{d_{mts,ij}} \tag{I.7}$$

$$\frac{\partial d_{ij}}{\partial x_{sj}} = \frac{x_{s_j} - x_0 - \upsilon_x t_i}{d_{mts,ij}} + \frac{x_{s_j} - x_{bs_j}}{d_{bs,j}} \tag{I.8}$$

$$\frac{\partial d_{ij}}{\partial x_0} = \frac{1}{t_i}\frac{\partial d_{ij}}{\partial \upsilon_x} = -\frac{x_{s_j} - x_0 - \upsilon_x t_i}{d_{mts,ij}} \tag{I.9}$$

$$\frac{\partial f_{d,ij}}{\partial y_{sj}} = -\frac{\partial f_{d,ij}}{\partial y_0} = \frac{f_c}{c}\frac{1}{d_{mts,ij}^{3/2}}\left[\upsilon_y(x_{s_j} - x_0 - \upsilon_x t_i)^2 - \upsilon_x(x_{s_j} - x_0 - \upsilon_x t_i)\right.$$
$$\left.(y_{s_j} - y_0 - \upsilon_y t_i)\right] \tag{I.10}$$

$$\frac{\partial f_{d,ij}}{\partial x_{sj}} = -\frac{\partial f_{d,ij}}{\partial x_0} = \frac{f_c}{c}\frac{1}{d_{mts,ij}^{3/2}}\left[\upsilon_x(y_{s_j} - y_0 - \upsilon_y t_i)^2 - \upsilon_y(x_{s_j} - x_0 - \upsilon_x t_i)\right.$$
$$\left.(y_{s_j} - y_0 - \upsilon_y t_i)\right] \tag{I.11}$$

$$\frac{\partial f_{d,ij}}{\partial \upsilon_y} = \frac{f_c}{c}\frac{1}{d_{mts,ij}^{3/2}}\left[(y_{s_j} - y_0 - \upsilon_y t_i) + \upsilon_y t_i((y_{s_j} - y_0 - \upsilon_y t_i)^2 - 1)\right.$$
$$\left.+ \upsilon_x t_i(x_{s_j} - x_0 - \upsilon_x t_i)(y_{s_j} - y_0 - \upsilon_y t_i)\right] \tag{I.12}$$

$$\frac{\partial f_{d,ij}}{\partial \upsilon_x} = \frac{f_c}{c}\frac{1}{d_{mts,ij}^{3/2}}\left[(x_{s_j} - x_0 - \upsilon_x t_i) + \upsilon_x t_i((x_{s_j} - x_0 - \upsilon_x t_i)^2 - 1)\right.$$
$$\left.+ \upsilon_y t_i(x_{s_j} - x_0 - \upsilon_x t_i)(y_{s_j} - y_0 - \upsilon_y t_i)\right] \tag{I.13}$$

Appendix J

Résumé en français

J.1 Introduction

Le chapitre d'introduction a deux objectifs: D'une part, nous introduisons les principes fondamentaux des méthodes de localisation. Suivi par les sources principales d'erreurs rencontrées dans les méthodes de localisation et les tentatives actuelles pour les résoudre. D'autre part, nous présentons des notions qui nous utilisons dans ce document, comme par exemple, les modèles de canaux géométriques et statistiques, les expressions utiles pour un système MIMO-OFDM, l' estimation ML de la position et une discussion sur l'identifiabilité et la performance, sera présentée dans certains chapitres suivants. Pour des raisons de brièveté d'espace, nous allons présenter juste les notions qui nous utilisons.

J.1.1 Modèles du Canal

Afin de localiser le MT, les modèles 2-D du canal qui nous permettent d'exprimer les LDP en fonction de les coordonnées du MT et d'autres paramètres, sont nécessaires. Pour les environnements LoS, il suffit d'utiliser les fonctions trigonométriques simples et les distances euclidiennes. Toutefois, pour les environnements NLoS statiques, certaines hypothèses de l'environnement de propagation doivent être prises en compte. Á cette fin, nous utilisons le modèle SBM (seule réflexion) et nous introduisons le modèle DSBM (dynamique seule réflexion), qui est le résultat de la fusion de la SBM avec

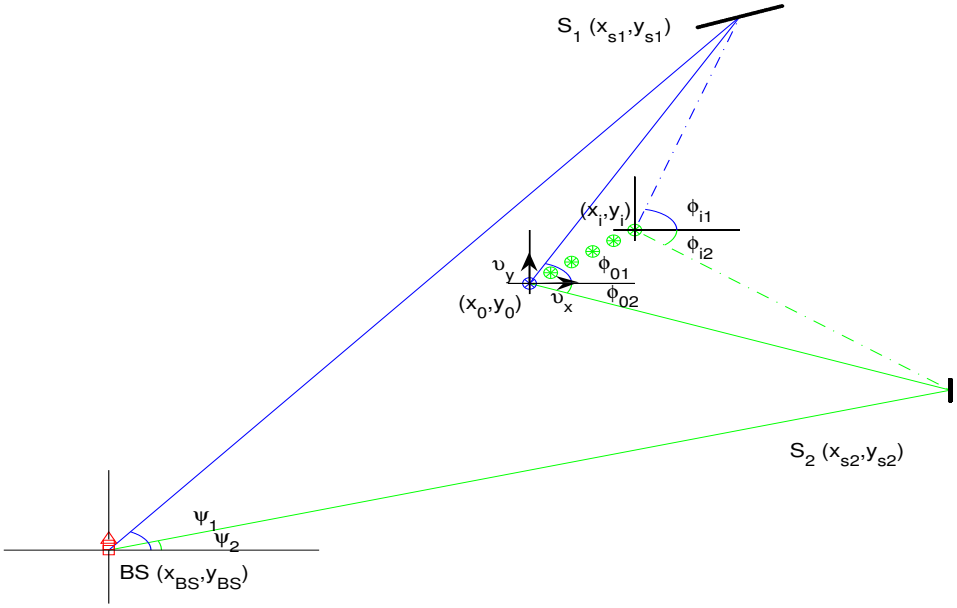

Figure J.1: LDP in a NLoS environment: Dynamic single bounce model

un modèle de mobilité. La utilisation de la SBM nous permet d'exprimer les LDP explicitement en fonction des coordonnées des BS, des MT et des diffuseurs, comme le suivant

$$\phi_j = \tfrac{\pi}{2}(1 - sgn\{x_{s_j} - x_{mt}\}) + \tan^{-1}\tfrac{y_{s_j} - y_{mt}}{x_{s_j} - x_{mt}} \tag{J.1}$$

$$\psi_j = \tfrac{\pi}{2}(1 - sgn\{x_{s_j} - x_{bs_j}\}) + \tan^{-1}\tfrac{y_{s_j} - y_{bs_j}}{x_{s_j} - x_{bs_j}} \tag{J.2}$$

$$d_j = d_{mts,j} + d_{bs,j} \tag{J.3}$$

$$d_{mts,j} = \sqrt{(y_{s_j} - y_{mt})^2 + (x_{s_j} - x_{mt})^2} \tag{J.4}$$

$$d_{bs,j} = \sqrt{(y_{s_j} - y_{bs_j})^2 + (x_{s_j} - x_{bs_j})^2} \quad . \tag{J.5}$$

Ces approches considèrent un environnement statique de propagation, c'est à dire qu'elles supposent que le MT ne bouge pas. En revanche, nous sommes également intéressés par l'évolution des environnements dynamiques, où l'on suppose que le MT se déplace à une vitesse qui a une magnitude $\upsilon_i = \sqrt{\upsilon_{x_i}^2 + \upsilon_{y_i}^2}$ et une direction $\omega_i = \frac{\pi}{2}(1 - sgn\{\upsilon_{x_i}\}) + \tan^{-1}(\upsilon_{y_i}/\upsilon_{x_i})$. Dans tels environnements, les LPD sont variables dans le temps. Pour en tenir

compte et en même temps en bénéficier, nous introduisons le DSBM qui est le résultat de la fusion de la SBM avec un modèle de mobilité approprié. Deux modèles de mobilité sont considérés: le modèle "vitesse constante " et le modèle " accélération constante". Si la vitesse est constante, la position de MT a l' instant i est donnée par

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \upsilon_x \\ \upsilon_y \end{bmatrix} t_i. \tag{J.6}$$

où t_i est la différence de temps entre les instants i et 0. L'extension du modèle au cas d'accélération constante est triviale. Egalemement au scénario statique, nous pouvons utiliser la DSBM pour exprimer une variable dans le temps LDP (y compris le décalage Doppler) explicitement en fonction des coordonnées des BS, MT et des diffuseurs, comme le suivant

$$\phi_{ij} = \frac{\pi}{2}(1 - sgn\{x_{s_j} - x_0 - \upsilon_x t_i\}) + \tan^{-1}\frac{y_{s_j} - y_0 - \upsilon_y t_i}{x_{s_j} - x_0 - \upsilon_x t_i} \tag{J.7}$$

$$d_{mts,ij} = \sqrt{(y_{s_j} - y_0 - \upsilon_y t_i)^2 + (x_{s_j} - x_0 - \upsilon_x t_i)^2} \tag{J.8}$$

$$f_{d,ij} = \frac{f_c}{c}\frac{\upsilon_{x_i}(x_{s_j} - x_0 - \upsilon_x t_i) + \upsilon_{y_i}(y_{s_j} - y_0 - \upsilon_y t_i)}{\sqrt{(y_{s_j} - y_0 - \upsilon_y t_i)^2 + (x_{s_j} - x_0 - \upsilon_x t_i)^2}}. \tag{J.9}$$

L'AoD et les distances entre les BS et les diffuseurs correspondants sont encore données par les équations (J.2) et (J.5) respectivement.

En plus des modèles géométriques, les modèles statistiques du canal sont nécessaires pour exprimer la matrice de la réponse impulsionnelle du canal (CIR) en fonction de la LDP et éventuellement d'autres paramètres. Un modèle très général est le modèle double directionnel (DDM), qui décrit un variant dans le temps, le canal sélectif en fréquence, en tenant compte des AoA, des AoD, des retards, des effets Doppler (DS) et des puissances des chemins aux instructions de direction à l'émetteur et le récepteur de taille. L'hypothèse que l'environnement de diffusion ne change pas au cours de la transmission est prise en consideration, donc la variation dans le temps est le résultat du mouvement de la MT et/ou de la BS. Selon le DDM, la $n_r \times n_t$ matrice MIMO $\mathbf{H}$ dans le domaine temps-fréquence, est donnée par

$$\mathbf{H}_{n_r \times n_t}(f,t) = \frac{1}{\sqrt{s_r s_t}}\mathbf{\Phi}_{n_r \times s_r}(t)\mathbf{P_r}(\mathbf{\Theta}_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f))\mathbf{P_t}\mathbf{\Psi}^t_{s_t \times n_t}(t) \tag{J.10}$$

où n_r,n_t,s_r et s_t sont le nombre d'antennes de réception et de transmission et le nombre de diffuseurs dans le domaine du récepteur et l'émetteur, respectivement (voir la figure 1.7). Les entrées de $\mathbf{\Theta}$ sont i.i.d. gaussiennes

complexes de moyenne nulle et de variance unité. $\mathbf{P}_r$ et $\mathbf{P}_t$ sont des matrices diagonales contenant les puissances. Le reste des matrices sur la r.h.s. de (J.10) sont définis dans les équations (1.45) - (1.47).

Il faut introduire un cas particulier du DDM qui est plus approprié pour les problèmes de localisation NLoS, car il est beaucoup plus simple que le DDM. Il peut être utilisé pour décrire les environnements, où chaque composante de signal rebondit qu'une seule fois. Selon ce modèle, le $n_r \times n_t$ matrice MIMO $\mathbf{H}$ dans le domaine temps-fréquence, est donnée par

$$\begin{aligned}\mathbf{H}_{kl} &= \tfrac{1}{\sqrt{P_{tot}}} \textstyle\sum_{j=1}^{N_s} \sqrt{P_j}\gamma_j e^{j2\pi l\Delta t f_{d,lj}} \mathbf{a}_R(\phi_{lj})\mathbf{a}_T^t(\psi_{lj}) H_{TR,k} e^{-j2\pi k\Delta f \tau_{lj}} \\ &= \mathbf{A}_{R,l}(\mathbf{\Gamma} \odot (\mathbf{D}_k\mathbf{F}_{d,l}))\mathbf{A}_{T,l}^t = \mathbf{A}_{R,l}\mathbf{\Gamma}\mathbf{D}_k\mathbf{F}_{d,l}\mathbf{A}_{T,l}^t. \end{aligned} \tag{J.11}$$

La dernière égalité découle du fait que, dans un scénario de rebond unique $\mathbf{\Gamma}$, $\mathbf{D}_k$ et $\mathbf{F}_{d,l}$ sont matrices diagonales. Les indices k, $1 \le k \le N_f$ et l, $0 \le l < N_t - 1$ désignent l'échantillon de fréquence et de temps respectivement, soit $\mathbf{H}_{kl} = \mathbf{H}(f_k, t_l)$. L'indice j, $1 \le j \le N_s$ désigne le diffuseur ou le composant du signal par trajets multiples. La définition de tous les paramètres de la première représentation (somme de rang 1 termes) de la chaîne de matrice est donnée par le tableau 1.1. Les matrices nouvellement introduites dans la deuxime représentation sont définies dans l'équation (1.55).

Les matrices du canal MIMO données par les équations (J.10) et (J.11) sont fondées sur l'hypothèse d'un environnement strictement NLoS. Si une composante LoS existe, alors, afin de représenter l'environnement multi-trajets, la matrice du canal peut s' écrire comme la somme des 2 composantes suivants

$$\mathbf{H}_{kl} = \mathbf{H}_{NL,kl} + \mathbf{H}_{L,kl}. \tag{J.12}$$

La composante NLoS $\mathbf{H}_{NL,kl}$ est donnée par l'équation (J.10) ou (J.11) et la composante LoS $\mathbf{H}_{L,kl}$ est donnée par

$$\mathbf{H}_{L,kl} = \frac{\sqrt{P_0}}{\sqrt{P_{tot}}} e^{j\theta_n} e^{j2\pi f_{d,l0} t_l} \mathbf{a}_R(\phi_{l0})\mathbf{a}_T^t(\psi_0) H_{TR,k} e^{-j2\pi f_k \tau_{l0}}. \tag{J.13}$$

Dans l'équation ci-dessus, nous avons introduit l'indice $j = 0$ qui sera utilisé dans ce document pour LDP qui correspond à la composante LoS et le décalage de phase inconnu (à cause de bruit de phase) de la trajectoire de LoS $\theta_n \sim U[0, 2\pi]$.

Si nous avons besoin d'exprimer le signal reçu directement en fonction du LPD, une relation discrète entrée-sortie (io) est nécessaire. La relation io d'un $n_r \times n_t$ système MIMO-OFDM peut s'écrire comme

$$\mathbf{Y}_{kl} = \mathbf{H}_{kl}\mathbf{X}_{kl} + \mathbf{N}_{kl}. \tag{J.14}$$

$\mathbf{H}_{kl}$ est la $n_r \times n_t$ matrice du canal, donnée par l'équation (J.12) pour les environnements à trajets multiples et se réduit à celle donnée par l'équation (J.10) ou (J.11) pour les environnements strictement NLoS. $\mathbf{X}_{kl}$ est la $n_t \times N$ matrice du signal transmis, que nous considérons composée d'une séquence formative des symboles. N est le nombre de symboles OFDM par sous-porteuse émis pendant le temps de cohérence du canal. $\mathbf{Y}_{kl}$ est la $n_r \times N$ matrice du signal reçu et $\mathbf{N}_{kl}$ est la $n_r \times N$ matrice du bruit. Toutes les matrices sont données à la fréquence k, $1 \leq k \leq N_f$ et temps l, $1 \leq l \leq N_t$. Les entrées de $\mathbf{N}_{kl}$ sont iid gaussiennes complexes de moyenne null et de variance σ_n^2.

J.1.2 Estimation ML

Chaque fois que des chercheurs s'attaquent à un problème d'estimation de paramètres, ils sont confrontés au dilemme de choisir entre la ML et l'estimation bayésienne (BE). Dans l'approche ML, les paramètres inconnus sont traités comme des quantités déterministes. L'estimation est alors basée sur la maximisation de la densité des données, conditionnée par le vecteur de paramètres. Dans l'approche bayésienne, les paramètres inconnus sont traités comme des variables aléatoires. Si leurs distributions à priori ne sont pas connues, une méthode connue sous le nom Inférence bayésienne peut être utilisée pour produire des a-priori significatifs. L'estimation peut alors être basée sur la maximisation de la densité du vecteur des paramètres, conditionnée par les données (Estimation MAP). En général, les estimateurs MAP ont des performances supérieures à celles des estimateurs ML, en supposant que les a-priori utilisés sont les bons. Cependant, nous notons que lorsque les a-priori sont non-informatifs, comme c'est le cas de paramètres uniformément répartis, l'estimation MAP et l'estimation ML sont équivalentes.

Dans cette section nous présentons un estimateur ML général pour la 2^e étape de tout algorithme de localisation SBM à base des 2 étapes. Il est basé sur les estimations de LDP disponibles. Soit $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^t, \ldots, \boldsymbol{\theta}_K^t]^t$ le $N_{\boldsymbol{\theta}} = KN_tN_s$ vecteur contenant les vraies valeurs de la $K = \{3, 4\}$ types différents de LPD et $\hat{\boldsymbol{\theta}}$ le vecteur contenant les estimations disponibles. Pour le cas DSBM, $\boldsymbol{\theta}$ peut contenir un sous-ensemble de 3 ou l'ensemble des LDP, donnée ci-dessous

$$\boldsymbol{\theta}_k = \begin{cases} \mathbf{d}, \ k = 1 \\ \boldsymbol{\phi}, \ k = 2 \\ \boldsymbol{\psi}, \ k = 3 \\ \mathbf{f}_d, \ k = 4 \end{cases} \tag{J.15}$$

Pour le cas SBM, $N_t = 1$, $K = 3$ et $\boldsymbol{\theta}$ est composé uniquement des 3 premièrs vecteurs ci-dessus. En supposant que les estimations de LDP ne sont pas parfaites, mais elles contiennent une erreur $\tilde{\boldsymbol{\theta}}$, nous pouvons écrire

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \tilde{\boldsymbol{\theta}}. \tag{J.16}$$

Nous supposons que $\hat{\boldsymbol{\theta}}$ est un estimateur sans biais qui possède la propriété de normalité asymptotique et qu'un nombre suffisant d'échantillons du signal reçu a été utilisée dans l'estimation de $\boldsymbol{\theta}$. En raison de ces hypothèses, $\tilde{\boldsymbol{\theta}} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\tilde{\boldsymbol{\theta}}})$ et donc $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \mathbf{C}_{\tilde{\boldsymbol{\theta}}})$.

L'objectif principal de toute méthode de localisation est d'estimer la position du MT. Dans les scénarios dynamiques, il serait également souhaitable d'estimer sa vitesse. Toutefois, dans la section 1.4.2, nous avons montré que pour un environnement NLoS qui peut être décrit par la SBM, les LPD ne dépendent pas seulement des coordonnées et de la vitesse du MT mais aussi des coordonnées des diffuseurs. Comme il y a en général peu d'intérêt à connaître la position des diffuseurs, nous allons traiter leurs coordonnées en tant que paramètres de nuisance et designer le vecteur qui les contient comme $\mathbf{p}_{nui} = \mathbf{p}_s = [\mathbf{x}_s^t, \mathbf{y}_s^t]^t$. D'autre part, nous designerons les paramètres d'intérêt comme $\mathbf{p}_{int}$. $\mathbf{p}_{int}$ peut varier, selon le scénario et le modèle de mobilité, comme donné au tableau 1.2. En introduisant le $N_{\mathbf{p}} \times 1$ vecteur de tous les paramètres inconnus, $\mathbf{p} = [\mathbf{p}_{int}^t, \mathbf{p}_{nui}^t]^t$, on peut réécrire l'ensemble des équations (J.1) - (J.5) et (J.7) - (J.9) d'une manière plus compacte $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathbf{p})$ pour montrer la dépendance de la moyenne des estimations LDP aux paramètres inconnus. En dehors de la moyenne, il est probable que la matrice de covariance $\mathbf{C}_{\tilde{\boldsymbol{\theta}}}$ dépend également des paramètres inconnus, c'est à dire qu'il est probable que la précision de la méthode utilisée pour estimer le LPD à partir des échantillons du signal reçu dépend de la géométrie de l'environnement. Cependant, puisque nous considérons une classe générale d'estimateurs et non pas un estimateur particulier (et donc nous ne présentons aucune expression de la matrice de covariance), nous ne considérerons pas une telle dépendance. Le pdf de $\hat{\boldsymbol{\theta}}$ conditionné sur $\mathbf{p}$ est donné par:

$$f(\hat{\boldsymbol{\theta}}|\mathbf{p}) = \frac{1}{(2\pi)^{\frac{1}{2}N_{\boldsymbol{\theta}}}(\det \mathbf{C}_{\tilde{\boldsymbol{\theta}}})^{1/2}} e^{-\frac{1}{2}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^t \mathbf{C}_{\tilde{\boldsymbol{\theta}}}^{-1}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})} \tag{J.17}$$

Pour obtenir une estimation ML de nos paramètres d'intérêt, nous devons maximiser $f(\hat{\boldsymbol{\theta}}|\mathbf{p})$ par rapport à $\mathbf{p}$. Définir une log-vraisemblance obtenue en prenant le logarithme naturel de $f(\hat{\boldsymbol{\theta}}|\mathbf{p})$ et en ignorant les termes constants en tant que:

$$\mathcal{L} = \mathcal{L}(\boldsymbol{\theta}(\mathbf{p})) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t \mathbf{C}_{\tilde{\boldsymbol{\theta}}}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \tag{J.18}$$

La maximisation de $f(\hat{\boldsymbol{\theta}}|\mathbf{p})$ est équivalente à la minimisation de $\mathcal{L}$, donc l'estimation ML de $\mathbf{p}$ est donnée par

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}}\{\mathcal{L}\} \tag{J.19}$$

Afin d'évaluer la performance d'un estimateur ML, le CRB peut être utilisé, car les estimateurs ML sont asymptotiquement efficaces. Selon le CRB pour un estimateur sans biais $\widehat{\mathbf{p}}$, la matrice de corrélation des erreurs d'estimation des paramètres $\widetilde{\mathbf{p}}$ est limitée par l'inverse de la FIM $\mathbf{J}$ comme indiqué ci-dessous

$$R_{\widetilde{\mathbf{p}}\widetilde{\mathbf{p}}} = E\{(\widehat{\mathbf{p}} - \mathbf{p})(\widehat{\mathbf{p}} - \mathbf{p})^t\} \geq \mathbf{J}^{-1} \tag{J.20}$$

où la FIM est donnée par:

$$\boldsymbol{J} = E\Big\{ \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right) \left(\frac{\partial \mathcal{L}}{\partial \mathbf{p}}\right)^t \Big\} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}} \mathbf{C}_{\widetilde{\boldsymbol{\theta}}}^{-1} \frac{\partial \boldsymbol{\theta}}{\partial \mathbf{p}^t} = \mathbf{G}\mathbf{C}_{\widetilde{\boldsymbol{\theta}}}^{-1}\mathbf{G}^t. \tag{J.21}$$

$\mathcal{L}$ est la log-vraisemblance donnée par (J.18) et nous avons introduit la matrice de transformation $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$.

Le CRB peut être utilisé aussi pour étudier l'identifiabilité comme le suivant: Les entrées de la FIM $\mathbf{J}$ sont des fonctions continues de $\mathbf{p}$ partout dans $\mathcal{R}^{N_\mathbf{p}}$. Un point de $\mathbf{p}^0$ est consideré un point régulier de la matrice $\mathbf{J}$ si il existe un voisinage ouvert de $\mathbf{p}^0$ ou $\mathbf{J}$ est de rang constant. En utilisant cette définition et en faisant certaines hypothèses faibles, le théorème suivant est prouvé

Theorem 3. *Soit $\mathbf{p}^0$ un point régulier de la FIM $\mathbf{J} = \mathbf{J}(\mathbf{p})$. Alors $\mathbf{p}^0$ est localement identifiable si et seulement si $\mathbf{J}(\mathbf{p}^0)$ est non-singulière.*

En supposant que le vecteur $\mathbf{p}$ contenant les vraies valeurs des paramètres inconnus est un point régulier, ce qui est vrai en général, le théorème ci-dessus nous dit que les paramètres inconnus deviennent identifiables lorsque FIM évaluée à la valeur réelle est non-singulière. Le corollaire suivant est une conséquence immédiate du théorème ci-dessus et de l'équation (J.21):

Corollary 2. *Dans les méthodes de localisation basées sur SBM, l' identifiabilité locale du vecteur paramètre $\mathbf{p}$ peut être réalisée lorsque la matrice de transformation $\mathbf{G} = \frac{\partial \boldsymbol{\theta}^t}{\partial \mathbf{p}}$ est de forme carrée ou large ($N_{\boldsymbol{\theta}} \geq N_\mathbf{p}$) et a plein rang $N_\mathbf{p}$.*

Enfin, on pourrait utiliser le CRB pour évaluer la performance des estimateurs ML dans les cas où l'exploitation de nouvelles LDP est au prix

d'estimer conjointement un nouvel ensemble de paramètres de nuisance. Cela se produit lorsque l'examen d'une dynamique plutôt que d'un environnement statique et donc la vitesse de la MT doit être conjointement estimée comme mentionné ci-dessus. Cela peut également se produire lorsque l'ensemble des LDP non-exploitées dépend non seulement des entrées de $\mathbf{p}$, mais aussi d'un terme d'erreur inconnue. Par exemple, il peut y avoir un offset de synchronisation inconnue qui doit être prise en compte pour les délais (et donc pour les longueurs de chemins) ou un offset de l'orientation et de la calibration qui doit être prise en compte pour les AoA et/ou les AoD. Le théorème suivant s'applique à toutes les cas ci-dessus et demontre les cas où l'estimation de la position ML sera plus précise.

Theorem 4. *L' exploitation de nouvelles LDP $\boldsymbol{\theta}_2$ (données) qui dépendent non seulement des entrées du $N_{\mathbf{p}_1} \times 1$ vecteur de paramètres $\mathbf{p}_1$, mais aussi des entrées du nouveau vecteur de paramètres de nuisance $\mathbf{p}_2$, entraîne une amélioration de la performance (asymptotique) de l'estimation ML seulement si la matrice de transformation $\mathbf{G}_{22} = \frac{\partial \boldsymbol{\theta}_2^t}{\partial \mathbf{p}_2}$ est large ($N_{\boldsymbol{\theta}_2} > N_{\mathbf{p}_2}$) et a rang plein $N_{\mathbf{p}_2}$.*

J.2 Estimation des LDP

Les techniques traditionnelles géométrique de localisation s'articulent en deux étapes: d'abord un ensemble de LDP sont estimés dans un ou plusieurs BS. La position du MT est alors estimée par trouver les valeurs des coordonnées x_{mt} et y_{mt} qui correspondent le mieux aux estimations LDP. Le chapitre 2 contient une méthode sous-espace-basée, qui peut être utilisée pour estimer simultanément différents types de LDP pour tous les MPC. Nous considérons un MT qui communique avec un BS dans un environnement de propagation NLoS. Le MT se déplace et, en conséquense, le CIR est affecté par DFS. Nous limiterons notre étude au système MIMO et au signal OFDM. Nous paramétrons la matrice CIR de telle manière, qu'un algorithme 4-D unitaire ESPRIT peut être utilisé pour estimer conjointement 4 sous-ensembles des LDP, les AoA, les AoD, les longueurs de chemins et les DFS.

J.2.1 Modèle du canal

La relation discrète d'entrée-sortie d'un $n_r \times n_t$ système MIMO-OFDM dans le domaine temps-fréquence est donnée par l'équation (J.14). Toutefois, contrairement à l'équation (J.11), les LDP sont traitées comme des constantes,

vu que leur variation pour une période d'observation brève (de l'ordre de ms) est négligeable. En conséquense, la $n_r \times n_t$ matrice du canal $\mathbf{H}_{kl}$ est donnée par

$$\begin{aligned}\mathbf{H}_{kl} &= \frac{1}{\sqrt{P_{tot}}} \sum_{j=1}^{N_s} \sqrt{P_j}\gamma_j e^{j2\pi l \Delta t f_{d,j}} \mathbf{a}_R(\phi_j)\mathbf{a}_T^t(\psi_j) H_{TR,k} e^{-j2\pi k \Delta f \tau_j} \\ &= \mathbf{A}_R(\mathbf{\Gamma} \odot (\mathbf{D}_k \mathbf{F}_{d,l}))\mathbf{A}_T^t = \mathbf{A}_R \mathbf{\Gamma} \mathbf{D}_k \mathbf{F}_{d,l} \mathbf{A}_T^t \end{aligned} \quad \text{(J.22)}$$

et le reste des paramètres sont définis dans les sections 1.5.2 et 1.6. Par la suite, nous supposons que $\mathbf{X}_{kl} = \mathbf{X}$, c'est à dire la séquence formative transmise par sous-porteuses différentes et à des instants différents, ne change pas. Nous supposons en outre que les Tx et Rx sont équipés de ULA. Chaque antenne qui peut être décomposé en 2 sous antennes avec des éléments identiques séparés par une distance arbitraire d peut être considérée à la place de ULA.

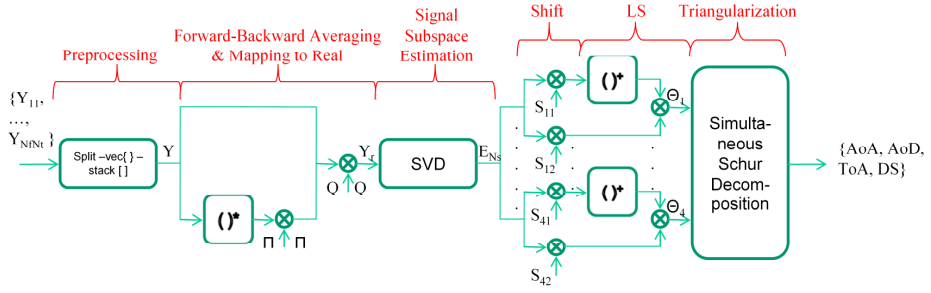

Figure J.2: 4D ESPRIT block diagram

Pour mettre en oeuvre l'algorithme ESPRIT 4D unitaire, nous avons besoin de réécrire la relation d'entrée-sortie sous une forme telle que la matrice du canal hérite une propriété d'invariance de changement dans les 4 dimensions. Le prétraitement des données et les étapes de 4D unitaire ESPRIT algorithme sont décrits en détail dans le chapitre 2. Ils peuvent être résumés sur le diagramme de la fig. J.2.

J.2.3 Exemple numérique

Dans cette section, nous évaluons la performance de la méthode proposée en termes de RMSE des prévisions LDP pour un système 2×4 MIMO

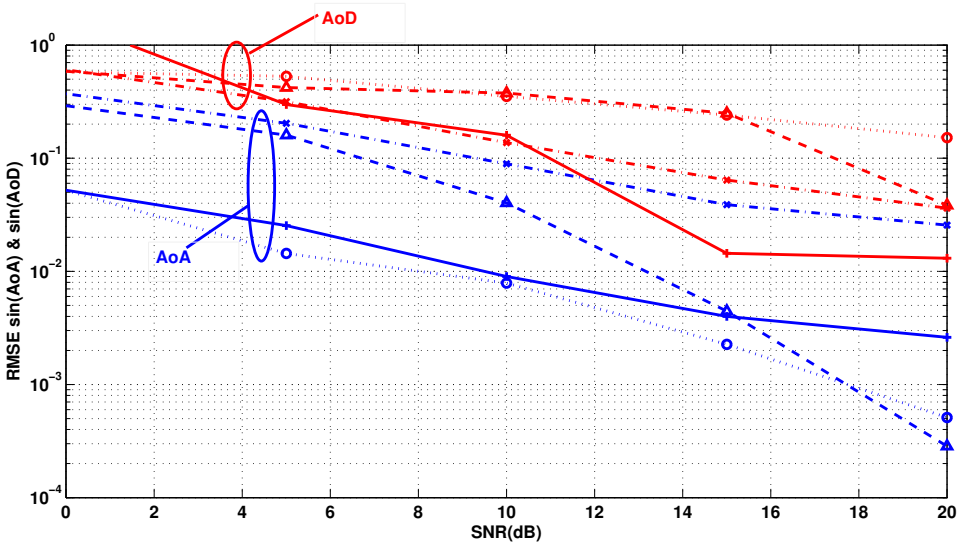

Figure J.3: RMSE of sine of AoA and sine of AoD

équipé d'ULA des deux côtés. Le signal transmis se propage par $N_s = 4$ chemins distincts NLoS. Les coordonnées des 4 diffuseurs correspondants ainsi que les coordonnées du MT sont donnés dans le tableau 2.1. Les valeurs des paramètres prises en considération se trouvent dans la section 2.5. Nous lançons $N = 50$ essais indépendants et moyennons les résultats. Nous calculons les RMSE, comee les suivant

$$RMSE(\mu_{ir}) = \sqrt{\frac{1}{N}\sum_{n=1}^{N} |\hat{\mu}_{ir} - \mu_{ir}|^2} \tag{J.23}$$

où les termes μ_{ir} dépendent des LDP en fonction de

$$\mu_{ir} = \frac{2}{\pi}\arctan(\omega_{ir}),\ 1 \leq i \leq N_s,\ 1 \leq r \leq 4. \tag{J.24}$$

Dans les figures, nous traçons les RMSE en fonction de SNR, défini comme

$$SNR = 10\log_{10}\left(\frac{E\{tr((\bar{\bar{\mathbf{H}}}\bar{\bar{\mathbf{\Gamma}}})(\bar{\bar{\mathbf{H}}}\bar{\bar{\mathbf{\Gamma}}})^{\dagger})\}}{E\{tr(\bar{\bar{\mathbf{N}}}\bar{\bar{\mathbf{N}}}^{\dagger})\}}\right) \tag{J.25}$$

Les résultats montrent que l'erreur quadratique moyenne des estimations est très petite, même pour les petites et moyennes SNR ($5 - 10$ dB). Grâce

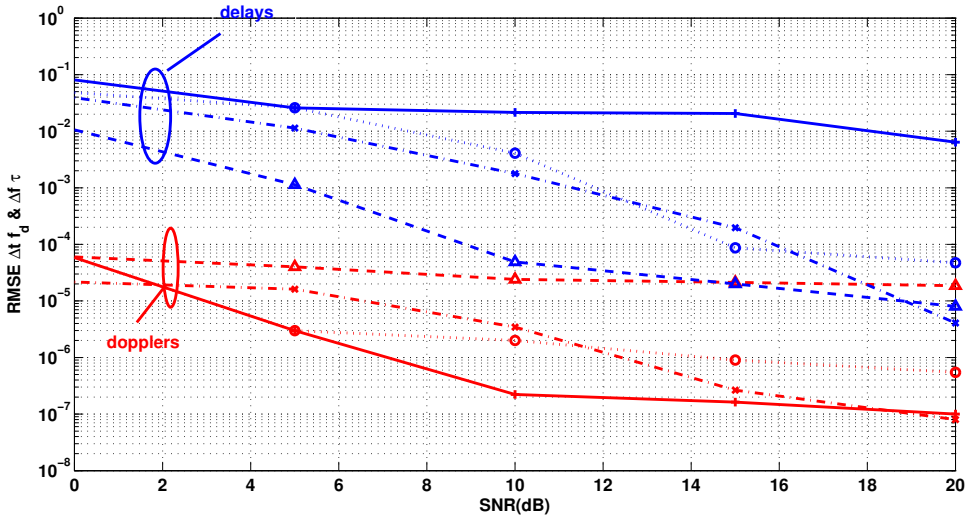

Figure J.4: RMSE of scaled delays and scaled DS

à son excellente performance et son coût de calcul réduit, cet algorithme estune solution interéssante pour des problèmes d'estimation des LDP et nous motive à concevoir des algorithmes pour la 2^e étape de la localisation, en supposant l'existence d'estimations exactes des LDP.

J.3 Localisation hybride pour les environnements NLoS statiques

Dans le chapitre 3, les fondements de la LS et la localisation ML basée sur le SBM sont présentés. Notre contribution consiste en une extension de la méthode de localisation basée sur SBM pour inclure une solution WLS, une brève discussion sur l'identifiabilité et une étude approfondie de l'impact de la géométrie du réseau sur les performances.

J.3.1 Estimation LS pour ToA/AoA/AoD localisation

En vue de formuler un système d'équations linéaires qui peut être résolu pour obtenir une estimation LS de $\mathbf{p}$, nous débutons avec les sinus et cosinus de l'équation des angles indiqués dans la figure 3.1 et nous les résoluons pour

les distances. Après quelques calculs nous obtenons

$$\underbrace{\begin{bmatrix} -\mathbf{C}_\psi\mathbf{1} & \mathbf{0} & (\mathbf{C}_\phi+\mathbf{C}_\psi) & \mathbf{O} \\ \mathbf{0} & -\mathbf{S}_\psi\mathbf{1} & \mathbf{O} & (\mathbf{S}_\phi+\mathbf{S}_\psi) \\ -\mathbf{S}_\psi\mathbf{1} & \mathbf{0} & \mathbf{S}_\psi & \mathbf{C}_\phi \\ \mathbf{0} & -\mathbf{C}_\psi\mathbf{1} & \mathbf{S}_\phi & \mathbf{C}_\psi \end{bmatrix}}_{=\mathbf{A}} \mathbf{p} = \underbrace{\begin{bmatrix} (\mathbf{D}\mathbf{C}_\psi+\mathbf{X}_{bs})\mathbf{C}_\phi\mathbf{1} \\ (\mathbf{D}\mathbf{S}_\psi+\mathbf{Y}_{bs})\mathbf{S}_\phi\mathbf{1} \\ (\mathbf{D}\mathbf{S}_\psi+\mathbf{Y}_{bs})\mathbf{C}_\phi\mathbf{1} \\ (\mathbf{D}\mathbf{C}_\psi+\mathbf{X}_{bs})\mathbf{S}_\phi\mathbf{1} \end{bmatrix}}_{=\mathbf{b}} \quad \text{(J.26)}$$

et donc on peut alors obtenir une estimation LS et une estimation WLS de $\mathbf{p}$, données respectivement par

$$\hat{\mathbf{p}}_{LS} = (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\mathbf{b} \quad \text{(J.27)}$$

$$\hat{\mathbf{p}}_{WLS} = (\mathbf{A}^t\mathbf{C}_\mathbf{b}^{-1}\mathbf{A})^{-1}\mathbf{A}^t\mathbf{C}_\mathbf{b}^{-1}\mathbf{b}. \quad \text{(J.28)}$$

J.3.2 Impact de la géométrie du réseau sur la performance

Dans cette section, nous évaluons la performance de la méthode de localisation basée sur le BSM, aux environnements LoS et NLoS statiques. Nous dérivons des expressions pour le CRB en fonction des distances et des angles, ce qui facilite l'interprétation de l'impact de la géométrie du réseau. La FIM pour le scénario LoS est une 2×2 matrice et donc elle peut être facilement inversée pour obtenir le $CRB_{pos} = tr\{\mathbf{J}^{-1}\}$. Le calcul est simple et le résultat est donné ci-dessous

$$CRB_{pos} = \frac{2\mathbf{1}^t(\frac{1}{\sigma_d^2}\mathbf{I} - \frac{1}{2}\mathbf{A})\mathbf{1}}{\frac{1}{\sigma_d^2}N_s\mathbf{1}^t(\frac{1}{\sigma_d^2}\mathbf{I} - \mathbf{A})\mathbf{1} + \mathbf{1}^t\mathbf{A}(\mathbf{1}\mathbf{1}^t - \breve{\mathbf{C}}_{\delta 2\phi})\mathbf{A}\mathbf{1}} \quad \text{(J.29)}$$

où $\breve{\mathbf{C}}_{\delta 2\phi}$ est une matrice symétrique dont l'entrée $\{i,j\}$ est égale à $cos(2\phi_i - 2\phi_j)$ et $\mathbf{A} = \frac{1}{\sigma_d^2}\mathbf{I} - \frac{1}{\sigma_{\phi\psi}^2}\mathbf{D}^{-2}$. Pour calculer le CRB dans un environnement statique NLoS, la $(2N_s+2)\times(2N_s+2)$ FIM doit être inversée, mais nous avons seulement besoin de se concentrer á la partie supérieure gauche de 2×2 sous-matrice de son inverse, la trace de ce qui donne la meilleure précision possible, c'est à dire, $CRB_{pos} = tr\{[\boldsymbol{J}^{-1}]_{1:2,1:2}\}$. Après une dérivation longue, nous obtenons

$$CRB_{pos} = \frac{2\mathbf{1}^t\bar{\mathbf{J}}_{det}^{-1}\mathbf{1}}{\mathbf{1}^t\bar{\mathbf{J}}_{det}^{-1}(\mathbf{1}\mathbf{1}^t - \breve{\mathbf{C}}_{\delta\phi+\delta\psi})\bar{\mathbf{J}}_{det}^{-1}\mathbf{1}} \quad \text{(J.30)}$$

où $\breve{\mathbf{C}}_{\delta\phi+\delta\psi}$ est une matrice symétrique dont l'entrée $\{i,j\}$ est égale à $cos(\phi_i - \phi_j + \psi_i - \psi_j)$ et les matrices restantes sont données par les équations (E.40)

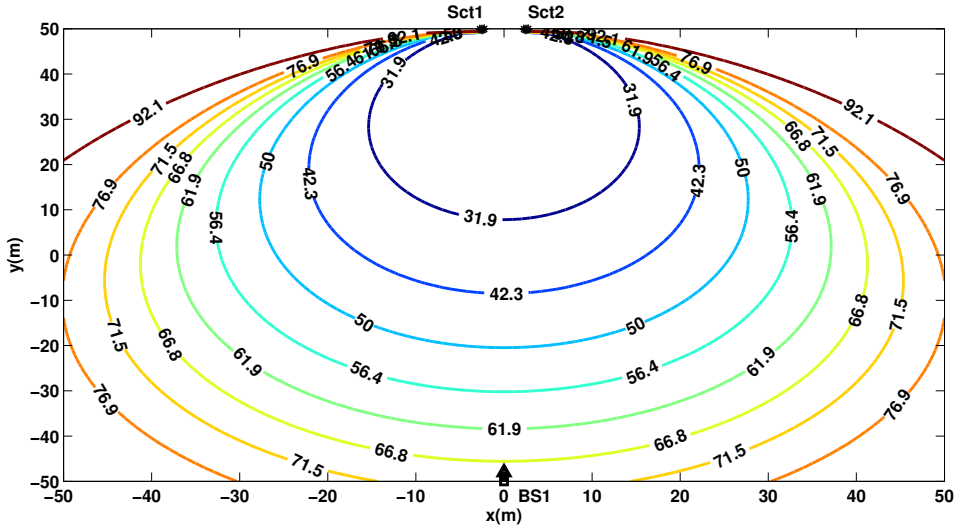

Figure J.5: CRB vs MT position for 1 BS - 2 collocated scatterers NLoS environment

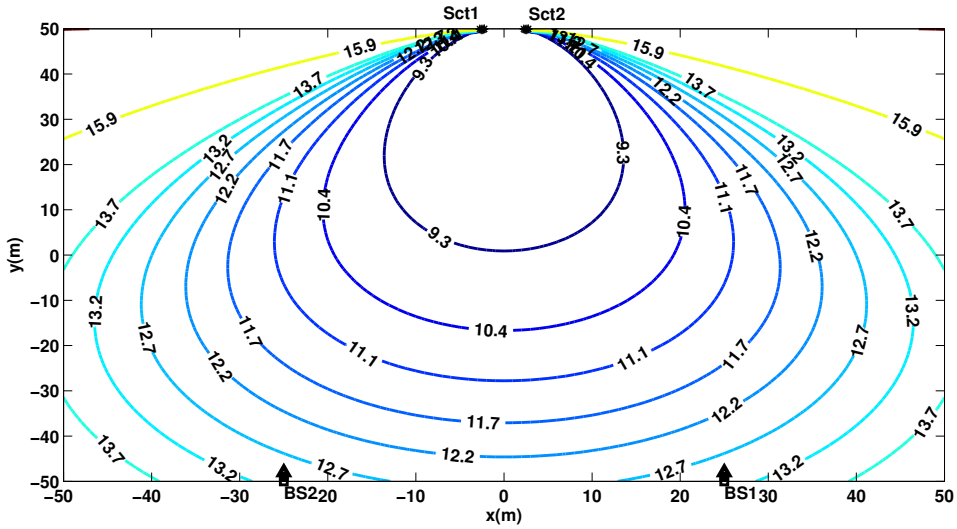

Figure J.6: CRB vs MT position for 2 BS - 2 collocated scatterers NLoS environment

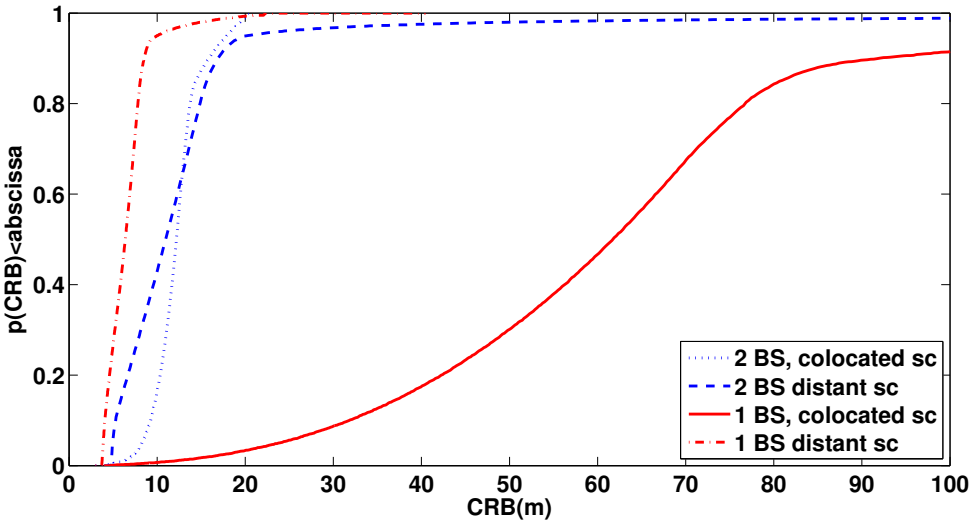

Figure J.7: CRB cdf for 4 different NLoS scenarios

- (E.44). La similitude entre les 2 expressions des CRB devient claire en remplaçant ψ dans les NLoS CRB_{pos}, en utilisant la condition LoS (1.21). Les différences des AoA avec AoD ne dépendent plus des angles et les sommes sont égales à deux fois les AoA, plus une constante $c \in \{-2\pi, 0, 2\pi\}$ de sorte que $\breve{\mathbf{C}}_{\delta\phi+\delta\psi} = \breve{\mathbf{C}}_{\delta 2\phi}$. La seule différence est le terme supplémentaire dans le dénominateur de l'expression de LoS CRB. Une autre remarque concerne les termes impliquant les matrices notées $\breve{\mathbf{C}}$. À cause de ces termes, les dénominateurs sont réduits, et donc les CRB sont augmentées. Les CRB peuvent être maximisées si $\breve{\mathbf{C}}_{\delta 2\phi} = \breve{\mathbf{C}}_{\delta\phi+\delta\psi} = \mathbf{1}\mathbf{1}^t$. Cela correspond aux BS colocalisées pour le cas LoS et aux diffuseurs colocalisées pour le cas 1 BS NLoS. Alors pour le cas LoS, le CRB reste fini, pour le cas NLoS, la CRB tend vers l'infini et il est donc impossible d'estimer la position du MT. La signification de ces matrices est démontrée dans les figures pour les scénarios de 1 et 2 BS. Les courbes de niveau sont basées sur la cdf de la CRB. Pour une ligne de contour $j = 1 : 9$ nous avons $p(CRB < CRB(j)) = j/10$. En comparant la figure J.5 et J.6, nous observons que la performance est vraiment mauvaise lorsque le MT communique avec 1 BS via 2 diffuseurs colocalisées mais pas quand 2 BS sont utilisées, de façon que chacune communique avec le MT via l'un des diffuseurs. Les cdf des CRB pour les 4 environnements NLoS différents considérés (1 ou 2 BS avec

diffuseurs à distance ou colocalisés) sont tracées sur la figure J.7. De cette derniere figure, il devient évident que, si pour diffuseurs colocalisées, il est préférable d'avoir des communications avec plus de 1 BS, mais dans les environnements avec les diffuseurs à distance, seulement 1 BS peut mener à de meilleures performances.

J.4 Localisation hybride pour les environnements NLoS dynamiques

Dans le chapitre 4, deux méthodes hybrides de localisation qui utilisent le DSBM sont présentées et leur avantage sur la méthode à base de SBM est démontré. Les deux méthodes sont applicables à des environnements NLoS qui changent dynamiquement à cause du mouvement de la MT. Les deux méthodes supposent la connaissance des différents types de LDP, comme la longueur du chemin d qui est proportionnelle au délai, l'AoD ψ à la BS et le décalage Doppler f_d. La première méthode suppose en outre la connaissance de l'AoA ϕ de chemins différents à la MT. Malgré le fait que cela semble être une petite différence, les deux méthodes effectivement diffèrent beaucoup. À cause du manque de disponibilité des AoA à la deuxième méthode, les paramètres qui doivent être estimés deviennent identifiables uniquement grâce à la variation dans le temps de la LDP.

J.4.1 Estimation LS pour ToA/AoA/AoD/DS localisation

Dans la 1^{ere} methode, $K = 4$ et $\boldsymbol{\theta} = [\mathbf{d}^t, \boldsymbol{\phi}^t, \boldsymbol{\psi}^t, \mathbf{f}_d^t]^t$. Également au scénario de cas statique, nous pouvons formuler un système d'équations linéaires qui peut être résolu pour obtenir une estimation (W)LS de la vitesse et de la position du MT. La solution pour les environnements dynamiques représente une extension de la solution pour les environnements statiques. Si le modèle de mobilité à vitesse constante est considéré, la matrice $\mathbf{A}$ dans l'équation (J.26) doit être complétée par la matrice suivante

$$\mathbf{A}_\upsilon = \begin{bmatrix} -\mathbf{C}_\psi(\mathbf{t} \otimes \mathbf{1}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{S}_\psi(\mathbf{t} \otimes \mathbf{1}) \\ -\mathbf{S}_\psi(\mathbf{t} \otimes \mathbf{1}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{C}_\psi(\mathbf{t} \otimes \mathbf{1}) \\ \mathbf{C}_\phi\mathbf{1} & \mathbf{S}_\phi\mathbf{1} \end{bmatrix} \tag{J.31}$$

Pour obtenir

$$\underbrace{\left[\begin{array}{c|c} \mathbf{A} & \mathbf{A}_{\upsilon} \\ \mathbf{O} & \end{array}\right]}_{=\mathbf{A}'} \underbrace{\left[\begin{array}{c} \mathbf{p} \\ \upsilon_x \\ \upsilon_y \end{array}\right]}_{=\mathbf{p}'} = \underbrace{\left[\begin{array}{c} \mathbf{b} \\ \mathbf{F}_d\mathbf{1} \end{array}\right]}_{=\mathbf{b}'}. \tag{J.32}$$

La solution WLS est alors

$$\hat{\mathbf{p}}'_{WLS} = (\mathbf{A}'^t\mathbf{C}_{\mathbf{b}'}^{-1}\mathbf{A}')^{-1}\mathbf{A}'^t\mathbf{C}_{\mathbf{b}'}^{-1}\mathbf{b}'. \tag{J.33}$$

J.4.2 Estimation LS pour ToA/AoD/DS localisation

Contrairement à la 1^{ere} methode où nous avons tiré une solution LS fondée sur la connaissance de tous les 4 sous-ensembles de LDP, dans la 2^e méthode nous ne considérons pas la connaissance des AoA. Nous le faisons afin de tenir compte des nombreux scénarios réalistes dans lesquels les AoA estimations ne sont pas disponibles ou ne sont pas fiable. Dans cette méthode $K = 3$ et $\boldsymbol{\theta} = [\mathbf{d}^t, \boldsymbol{\psi}^t, \mathbf{f}_d^t]^t$. Ensuite, nous décrivons le concept de base de l'algorithme et nous fournissons la solution finale. Le calcul de la solution peut être trouvé dans la section 4.3. Dans la 1^{ere} étape de cette méthode, les paramètres de nuisance en coordonnées polaires sont estimés. Au lieu de considérer $\mathbf{p}_{nui} = [\mathbf{x}_s^t, \mathbf{y}_s^t]^t$, nous considérons $\mathbf{p}'_{nui} = [\mathbf{d}_{bs}^t, \mathbf{p}_{\upsilon,\alpha}^t]^t$, où $\mathbf{p}_{\upsilon,\alpha}^t$ est un vecteur de paramètres inconnus qui dépendent de la vitesse (ou de la vitesse et de l'accélération) du MT. Cette étape est basée sur l'observation suivante: La DS peut être écrit comme un rapport d'une quantité sur $d_{mts_{ij}}$ et la dérivée partielle de cette quantité en fonction du temps ne dépend pas de coordonnées inconnues. Le vecteur estimé des paramètres de nuisance est donné par

$$\hat{\mathbf{p}}'_{nui} = (\mathbf{Z}^t(\mathbf{C}_{\mathbf{w}})^{-1}\mathbf{Z})^{-1}\mathbf{Z}^t(\mathbf{C}_{\mathbf{w}})^{-1}\mathbf{w} \quad . \tag{J.34}$$

En outre, il peut être démontré que

$$\hat{\mathbf{d}}_{bss} = (\mathbf{Z}_f^t\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{P}\mathbf{Z}_f)^{-1}\mathbf{Z}_f^t\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{P}\mathbf{w}. \tag{J.35}$$

où nous avons introduit

$$\mathbf{Z} = [\mathbf{R}_t\mathbf{F}_d(\mathbf{1}\otimes\mathbf{I}), -(dt)^n\mathbf{V}] \tag{J.36}$$

$$\mathbf{w} = \mathbf{R}_t(\mathbf{f}_d \odot \mathbf{d}) \tag{J.37}$$

$$\mathbf{C}_{\mathbf{w}} = \mathbf{R}_t(\sigma_{f_d}^2\mathbf{d}\mathbf{d}^t \odot \mathbf{I} + \sigma_d^2\mathbf{f}_d\mathbf{f}_d^t \odot \mathbf{I} + \sigma_{f_d}^2\sigma_d^2\mathbf{I})\mathbf{R}_t^t \tag{J.38}$$

$$\mathbf{Z}_f = \mathbf{R}_t\mathbf{F}_d(\mathbf{1}\otimes\mathbf{I}) \tag{J.39}$$

$$\mathbf{P} = \mathbf{I} - \mathbf{V}(\mathbf{V}^t\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{V})^{-1}\mathbf{V}^t\mathbf{C}_{\mathbf{w}}^{-1}. \tag{J.40}$$

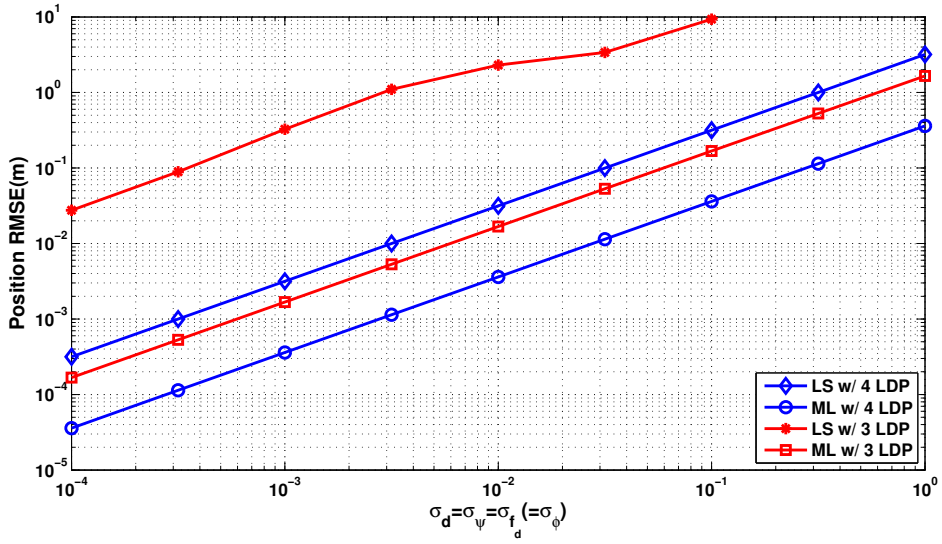

Figure J.8: Position RMSE vs SNR for various methods

$\mathbf{R}_t$ représente les matrices à "différence de temps", données par l'equation (4.24) et $\mathbf{V}$ est donnée dans le tableau 4.1. Dans la 2^e étape, ces estimations sont utilisées afin de formuler un problème de localisation de la ToA et afin d'estimer les paramètres d'intérêt $\mathbf{p}_{int} = [\mathbf{p}_0, \boldsymbol{v}_0, \boldsymbol{\alpha}]^t$ en utilisant une extension de la méthode de lignes-de-position (LoP). La solution est simplement donnée par

$$\hat{\mathbf{p}}_{int} = (\mathbf{A}^t \mathbf{C}_\mathbf{b}^{-1} \mathbf{A})^{-1} \mathbf{A}^t \mathbf{C}_\mathbf{b}^{-1} \mathbf{b} \tag{J.41}$$

où la matrice $\mathbf{A}$ est définie dans l'équation (G.20) et le vecteur $\mathbf{b}$ est défini dans l'équation (G.9).

J.4.3 Résultats de la simulation

Dans ce qui suit, nous évaluons et nous comparons les performances des 2 méthodes ($K = 4$ et $K = 3$) pour un scénario dans lequel le MT est situé à $\{x_0, y_0\} = \{25, 25\}$ et se déplaçe avec la vitesse $\{u_x, u_y\} = \{6, 9\} m/sec$ à l'intérieur d'une micro-cellule (milieu urbain) et réçoit un signal de 4 BS. Toutes les composantes du signal reçu se propagent dans un environnement NLoS. Les coordonnées de la BS et des diffuseurs correspondants sont donnés dans le tableau 4.6 et l'environnement est illustré à la figure 4.2. Le temps d'observation total est $1sec$, au cours de laquelle $N_t = 10$ échantillions de

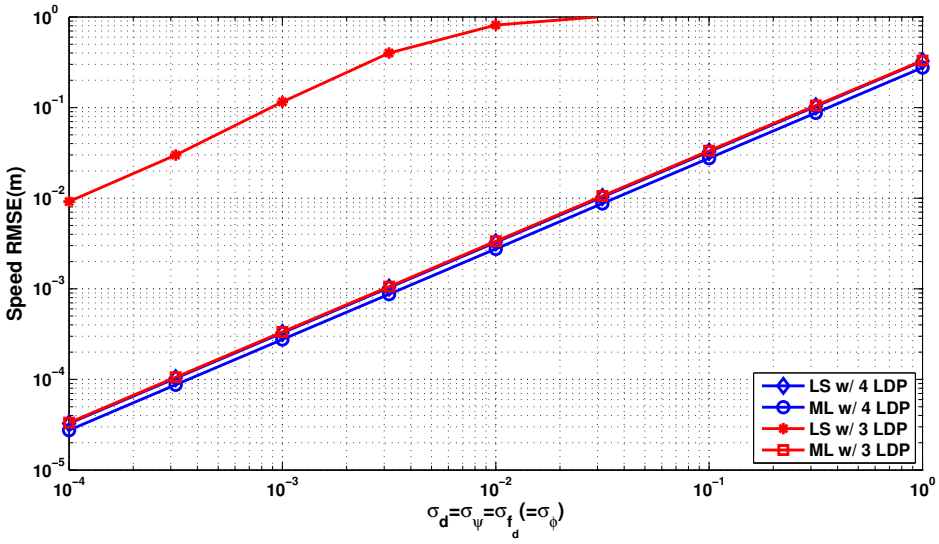

Figure J.9: Speed RMSE vs SNR for various methods

chaque LDP et pour chaque trajet sont estimés et utilisés dans le processus de localisation. Nous effectuons N expériences indépendantes et nous moyennons les résultats. Dans les figures J.8 et J.9 nous traçons les RMSE de la position et de la vitesse du MT respectivement données par

$$RMSE_p = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{x}_{0,k} - x_0)^2 + (\hat{y}_{0,k} - x_0)^2} \tag{J.42}$$

$$RMSE_{sp} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{v}_{x,k} - v_x)^2 + (\hat{v}_{y,k} - v_y)^2} \tag{J.43}$$

en fonction de σ de toutes les estimations LDP (nous considèrons que σ est pareille pour tous les LDP seulement pour la démonstration). Nous pouvons observer que, lorsque seulement 3 LDP sont disponibles, l'estimation LS de la vitesse et de la position est précise seulement pour le SNR élevé (défini comme $\frac{1}{\sigma^2}$). C'est parce que cette méthode dépend de la variation dans le temps des LPD et est très sensible au bruit. Cela ne semble pas provoquer le même problème pour les estimations ML, même si la performance est légèrement moins que la 1^{ere} méthode, comme prévu.

J.5 Estimation directe de la position pour les systmes MIMO-OFDM

Comme son nom l'indique, contrairement aux traditionnelles méthodes de localisation à 2 étapes , la DLE est une méthode qui traite les échantillons du signal reçu et estimate de la position du MT (et éventuellement d'autres paramètres) directement, sans explicitement estimer les LDP. La performance des approches à 2 étapes, a été prouvée à converger vers le CRB pour un SNR élevé et un nombre suffisant d'échantillons de données. Dans les systèmes de communication sans fil, le SNR élevé n'est pas toujours garanti. En outre, si le canal varie rapidement, le nombre d'échantillons de données qui peuvent être utilisés dans le processus d'estimation est très limité. La DLE est une méthode qui permet de localiser efficacement dans ces circonstances. Dans le chapitre 5, nous avons proposé et mis en oeuvre les systèmes MIMO-OFDM qui fonctionnent dans des environnements strictement NLoS ou à trajets multiples. En outre, nous supposons que le MT communique uniquement avec 1 BS.

La DLE devient possible si une application unique inversible existe entre les paramètres d'intérêt (MT position, vitesse etc) et les LPD. Afin d'exprimer toutes les LDP dont la matrice du canal MIMO dépend, à savoir les AoA, les AoD, les délais et les DS, en fonction des coordonnées du MT. Alors, une représentation géométrique appropriée de l'environnement est nécessaire. À cette fin, nous avons fondé notre approche sur le SBM, que nous avons également employé dans les méthodes de localisation à 2 étapes présentées dans les chapitres précédents.

D'abord, nous supposons que le signal se propage dans un environnement strictement NLoS qui change à cause du mouvement du MT, et que chaque composante de signal a rebondi exactement une fois. Le modèle décrit dans la section 1.5.2 peut être appliqué ici. Nous sommes intéressés à estimer conjointement les MT coordonnées au moment t_0 , x_0 et y_0 et les composantes à vitesse constante υ_x et υ_y directement des matrices du signal reçu $\mathbf{Y}_{kl}$. Le modèle de mobilité avec une accélération constante n'est pas considérée ici. Soit $\mathbb{S}_{\mathbf{Y}} = \{\mathbf{Y}_{11}, \ldots, \mathbf{Y}_{N_f N_t}\}$ l'ensemble de matrices du signal reçu. Pour mettre en œuvre l'estimation ML, nous définissons et calculons la log-vraisemblance

$$\mathcal{L} \triangleq \mathcal{L}(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}) = \ln(f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p})) = -\ln(\det(\mathbf{C}_{\mathbf{y}|\mathbf{p}})) - \mathbf{y}^{\dagger}\mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1}\mathbf{y}. \tag{J.44}$$

où $\mathbf{y} = [\mathbf{y}_{11}^t, \ldots, \mathbf{y}_{N_f N_t}^t]^t$ et $\mathbf{C}_{\mathbf{y}|\mathbf{p}}$ est donnée par l'équation (5.11).

En présence d'une composante LoS, la matrice du canal MIMO est

donnée comme une somme de 2 composantes. Dans cette formulation du problème nous avons besoin d'estimer $\mathbf{p}_{int}$ en présence des paramètres de nuisance $\mathbf{p}_{nuis} = [\mathbf{x}_s^t, \mathbf{y}_s^t, \theta]^t$, donc notre objectif est d'estimer le $(2N_s+5)\times 1$ vecteur $\mathbf{p} = [\mathbf{p}_{int}^t, \mathbf{p}_{nuis}^t]^t$. Nous pouvons appliquer directement le résultat sur le vecteur aléatoire centré $\mathbf{y}' = \mathbf{y} - \mathbf{m}_\mathbf{y}$ pour obtenir la nouvelle log-vraisemblance

$$\mathcal{L} = -\ln(\det(\mathbf{C}_{\mathbf{y}|\mathbf{p}})) - (\mathbf{y}-\mathbf{m}_\mathbf{y})^\dagger \mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1}(\mathbf{y}-\mathbf{m}_\mathbf{y}). \tag{J.45}$$

Pour les deux environnements (NLoS et à trajets multiples), l'estimation ML de $\mathbf{p}$ est donnée par la maximisation de la log-vraisemblance correspondante, c'est à dire

$$\widehat{\mathbf{p}} = \underset{\mathbf{p}}{\text{argmax}}\{\mathcal{L}\} \ . \tag{J.46}$$

Étant donné que dans un environnement à trajets multiples, la moyenne conditionnelle et la covariance conditionnelle sur le r.h.s. de (J.45) dépendent des paramètres inconnus, en utilisant la définition de la FIM donnée par la première égalité de l'équation (J.21), nous aboutissons à une somme de 2 termes pour chaque entrée. La solution pour l' entrée i', j' est donnée par

$$J_{i'j'} = tr\Big\{\mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1}\frac{\partial \mathbf{C}_{\mathbf{y}|\mathbf{p}}}{\partial p_{i'}}\mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1}\frac{\partial \mathbf{C}_{\mathbf{y}|\mathbf{p}}}{\partial p_{j'}}\Big\} + 2Re\Big\{\frac{\partial \mathbf{m}_\mathbf{y}^\dagger}{\partial p_{i'}}\mathbf{C}_{\mathbf{y}|\mathbf{p}}^{-1}\frac{\partial \mathbf{m}_\mathbf{y}}{\partial p_{j'}}\Big\}. \tag{J.47}$$

Si un environnement strictement NLoS est considéré, le deuxième terme de r.h.s. de (J.47) est égal à zéro. Pour calculer le CRB, nous construisons les dérivées partielles de la moyenne conditionnelle et de la covariance en utilisant les equations (5.26) et (5.21) et nous substituons le résultat dans (J.47).

J.5.1 Résultats de la simulation

Dans cette section, nous calculons et nous traçons le CRB pour trois environnement différents: Un environnement LoS (l'information de composants NLoS n' est pas disponible), un environnement de propagation à trajets multiples avec 2 trajets NLoS et un trajet LoS, et un environnement strictement NLoS avec 3 trajets. Les coordonnées de la BS, du MT et des diffuseurs considérés, correspondent à une pico-cellule qui s'adapte au modèle elliptique et sont donnés dans le tableau 5.1. Les valeurs des paramètres prises en considération se trouvent dans la section 5.5. Dans les figures J.10 et J.11, nous traçons les RMSE de la position et de la vitesse, respectivement, pour

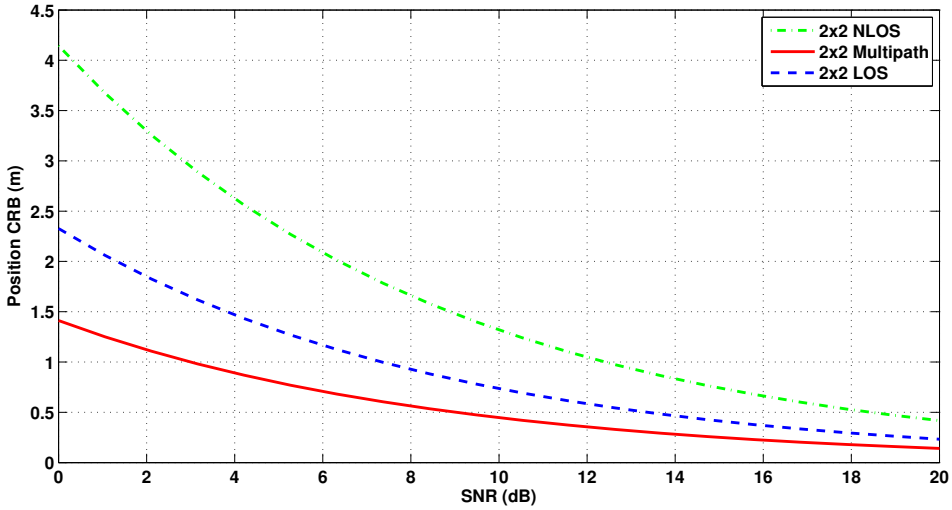

Figure J.10: Position CRB vs SNR, various environments

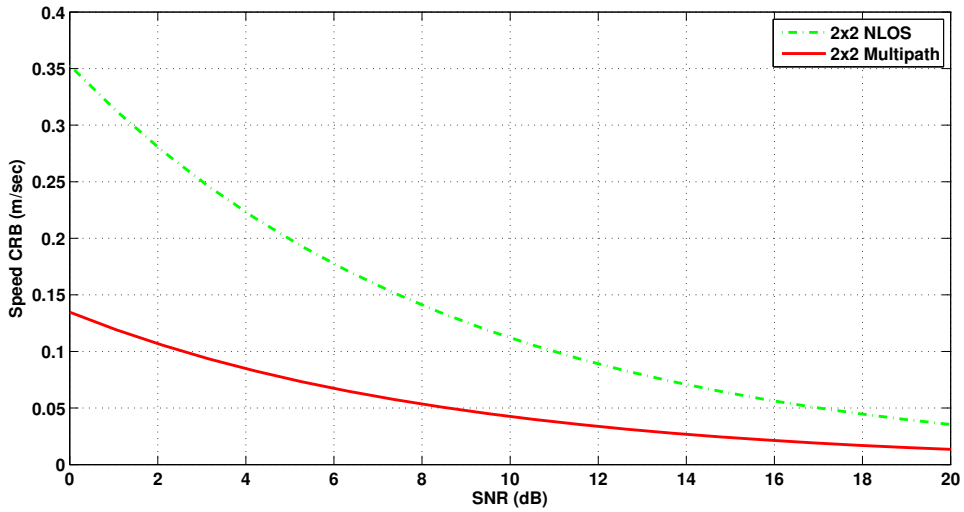

Figure J.11: Speed CRB vs SNR, various environments

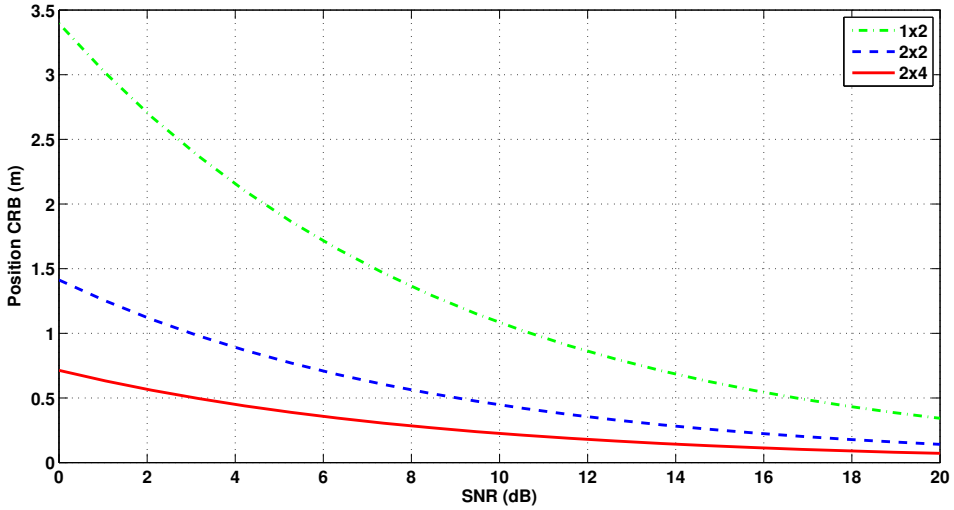

Figure J.12: Position CRB vs SNR, various systems

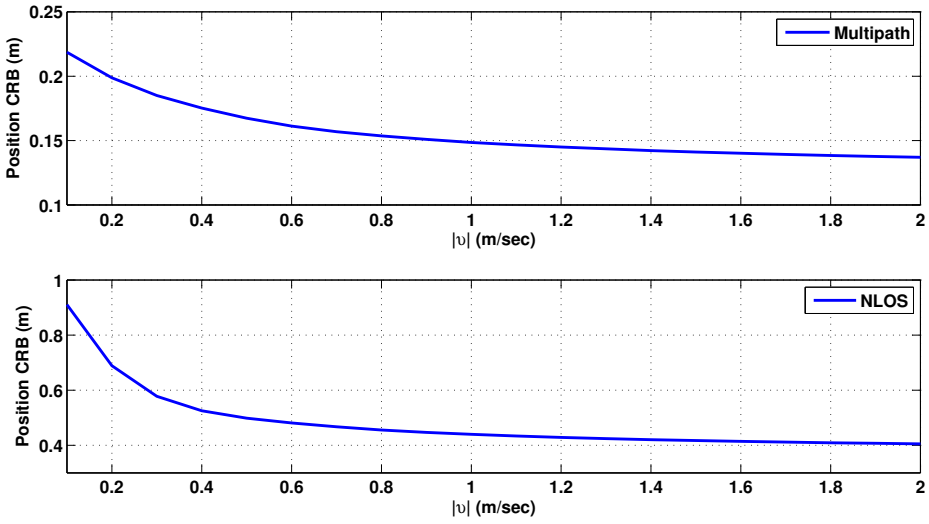

Figure J.13: Position CRB vs Speed Magnitude

un estimateur efficace, en fonction de la SNR pour un système 2×2 MIMO . Le SNR est défini comme:

$$SNR = 10\log_{10}\left(\frac{E\{tr(\mathbf{H}\mathbf{X}\mathbf{X}^{\dagger}\mathbf{H}^{\dagger})\}}{E\{tr(\mathbf{N}\mathbf{N}^{\dagger})\}}\right) = 10\log_{10}\left(\frac{\sigma_{\gamma}^{2}}{\sigma^{2}}\right). \tag{J.48}$$

Les matrices inconnues sont définies dans l'équation (5.32). Les RMSE de la position et de la vitesse sont définis comme

$$RMSE_p = \sqrt{\sigma_{\widetilde{x_0}}^2 + \sigma_{\widetilde{y_0}}^2} = \sqrt{tr([\mathbf{J}^{-1}]_{(1:2,1:2)})} \tag{J.49}$$

$$RMSE_{sp} = \sqrt{\sigma_{\widetilde{v_x}}^2 + \sigma_{\widetilde{v_y}}^2} = \sqrt{tr([\mathbf{J}^{-1}]_{(3:4,3:4)})} \tag{J.50}$$

Nous notons que l'erreur d'estimation est très faible même pour un environnement strictement NLoS. En outre, si les composantes NLoS du signal sont pris en compte avec la composante LoS, la RMSE de la position est significativement réduite (par exemple, 40% à $10dB$) et l'estimation de la vitesse devient possible. Dans la figure J.12, l'effet d'augmentation du nombre d'antennes sur la précision de position est représenté, pour l'environnement de propagation à trajets multiples seulement. Pour un système MISO, la $RMSE_p < 1m$ pour un $SNR > 11dB$, tandis que le système 2×2 peut obtenir la même précision avec un SNR de $3dB$. Enfin, dans la figure J.13, la RMSE de la position en fonction de la vitesse de MT est tracée. C'est évident que le mouvement de la MT a un impact énorme sur la précision de localisation, notamment pour l'environnement NLoS où l'erreur est réduite de plus de 50% lorsque la vitesse est augmentée jusqu'à $2m/s$.

J.6 Suivi du MT pour les systèmes MIMO-OFDM

L'algorithme présenté ici peut estimer uniquement la vitesse, donc la caractérisation "tracking". Il utilise le DDM dans sa version plus générique, comme décrit dans la section 1.5.1. L'avantage principal de l'utilisation de la DDM est qu'aucune hypothèse sur l'environnement de diffusion (nombre de rebonds) est faite. Le désavantage principal est les nombreux nouveaux paramètres de nuisance apparaîssent dans cette représentation du canal.

Soient $\mathbf{p}_{int}$ les paramètres d'intérêt et $\mathbf{p}_{nuis}$ le vecteur de tous les autres paramètres inconnus dans la matrice du canal $\mathbf{H}$. Ces deux vecteurs sont explicitement définis par la suite. Pour le moment, il suffit de constater que $\mathbf{p}_{nuis} = [\boldsymbol{\theta}, \tilde{\mathbf{p}}_{nuis}]$, où $\boldsymbol{\theta} = vec(\boldsymbol{\Theta})$ et $\boldsymbol{\Theta}$, avec toutes les autres matrices

composant chaque $\mathbf{H}(f_k, t_l)$, sont définis dans la section 1.5.1. Selon une estimation ML,

$$\hat{\mathbf{p}}_{int} = \underset{\mathbf{p}_{int}}{\operatorname{argmax}} f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) \tag{J.51}$$

où $f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int})$ est la densité de toutes les matrices du signal reçu, conditionnées uniquement par les paramètres d'intérêt. Nous pouvons obtenir cette densité en marginalisant la densité de tous les paramètres de nuisance, nous obtenons

$$f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) = \int_{\mathcal{A}} f(\mathbb{S}_{\mathbf{Y}}|\mathbb{S}_{\mathbf{H}}) \cdot f(\mathbf{p}_{nuis}) d\mathbf{p}_{nuis} \tag{J.52}$$

Toutes les paramètres, sauf $\boldsymbol{\theta}$, sont distribués uniformément et donc nous pouvons écrire

$$f(\mathbf{p}_{nuis}) = O(s_r, s_t) e^{-\boldsymbol{\theta}^\dagger \boldsymbol{\theta}} \tag{J.53}$$

et nous pouvons avancer en marginalisant le vecteur gaussien $\boldsymbol{\theta}$. Après unlong calcul, nous obtenons

$$f(\mathbb{S}_{\mathbf{Y}}|\mathbf{p}_{int}) = \int_{\tilde{\mathcal{A}}} O'(s_r, s_t) \det((\mathbf{C}_G^* \mathbf{C}_G^t + \sigma^2 \mathbf{I})^{-1}) e^{-(\mathbf{y}^\dagger (\mathbf{C}_G^t \mathbf{C}_G^* + \sigma^2 \mathbf{I})^{-1} \mathbf{y})} d\tilde{\mathbf{p}}_{nuis} \tag{J.54}$$

En outre, si l'on considère que $N_f = 1$, le r.h.s. de l'équation (J.54) ne dépend plus de $\boldsymbol{\tau}$. Cela n'est pas surprenant, car si $N_f = 1$, on peut remplacer $(\boldsymbol{\Theta}_{s_r \times s_t} \odot \mathbf{D}_{s_r \times s_t}(f))$ avec une nouvelle matrice de $\boldsymbol{\Theta}'_{s_r \times s_t}$ qui a la même distribution. L'algorithme proposé pour un système MIMO est résumé ci-dessous

- Mettre $\mathbf{p}_{int} = [sr, \boldsymbol{\phi}]$, $\tilde{\mathbf{p}}_{nuis} = [s_t, \boldsymbol{\psi}, \mathbf{p}^r, \mathbf{p}^t]$ et utiliser (J.54) pour estimer les AoA, en utilisant une seule observation $N_f = N_t = 1$.
- Mettre $\mathbf{p}_{int} = [st, \boldsymbol{\psi}]$, $\tilde{\mathbf{p}}_{nuis} = [s_r, \boldsymbol{\phi}, \mathbf{p}^r, \mathbf{p}^t]$ et utiliser (J.54) pour estimer les AoD, en utilisant une seule observation, $N_f = N_t = 1$.
- Mettre $\mathbf{p}_{int} = [\upsilon_r, \omega_r]$, $\tilde{\mathbf{p}}_{nuis} = [\upsilon_t, \alpha_t, \mathbf{p}^r, \mathbf{p}^t, \boldsymbol{\tau}]$ et utiliser (J.54) pour estimer les paramètres de notre véritable intérêt, en utilisant toutes les observations.

J.6.1 Résultats de la simulation

Afin de calculer la valeur de l'intégrale multidimensionnelle (J.54) et évaluer la performance de l'algorithme de suivi, des simulations de Monte Carlo ont été réalisées. Normalement 100 itérations sont suffisantes pour l'algorithme

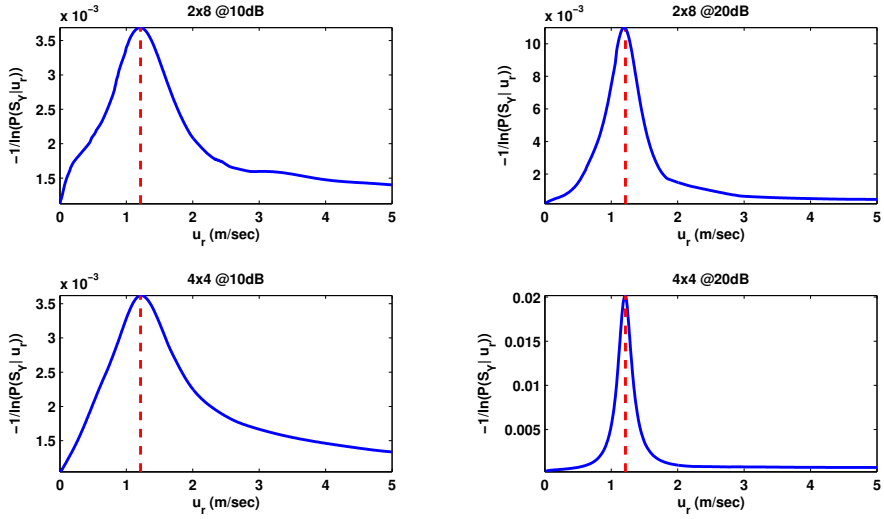

Figure J.14: Log-likelihood for speed magnitude estimation at various SNR

de converger vers la vraie densité. Afin de rendre nos graphiques plus claires, nous avons tracé les 1 dimensions normalisées log-vraisemblances $-\frac{1}{\ln f(\mathbb{S}_{\mathbf{Y}}|\upsilon_r)}$et $-\frac{1}{\ln f(\mathbb{S}_{\mathbf{Y}}|\omega_r)}$ en fonction de υ_r et ω_r, respectivement. La ligne verticale en pointillés sur les figures représente la vraie valeur du paramètre estimé (υ_r ou ω_r).

Dans les figures J.14 et J.15, nous montrons que dans un environnement bruyant (SNR $= 10dB$), υ_r et ω_r peuvent être estimés correctement dans un 4×4 et 2×8 système. D'autre part, dans un système 2×2 ou 2×4, les paramètres d'intérêt sont mal-estimés à faible SNR. Ces résultats indiquent que, lorsque un modèle complex comme le DDM -et pas sa version simplifiée qui suppose un seul rebondit et des chemins distinctes-, est utilisé, il faut chercher des solutions qui utilisent des systèmes MIMO, qui fonctionnent à un SNR élevé et sont équipées avec un grand nombre d' antennes aux deux extrémités.

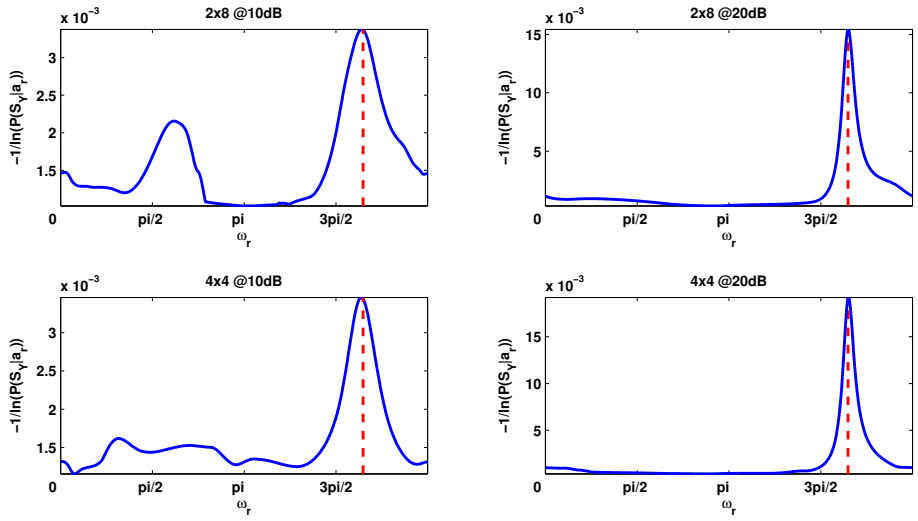

Figure J.15: Log-likelihood for speed direction estimation at various SNR

Bibliography

[1] W. G. Figel, N. H. Shepherd, and W. F. Trammel, "Vehicle Location by a Signal Attenuation Method," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 18, no. 3, pp. 245 – 251, 1969.

[2] FCC, "Revision of the Commission's Rules to Ensure Compatibility with Enhanced 911 Emergency Calling Systems," CC Docket 94 - 102, 1994.

[3] J. H. Reed, K. J. Krizman, B. D. Woerner, and T. S. Rappaport, "An Overview of the Challenges and Progress in Meeting the E-911 Requirement for Location Service," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 30 – 37, 1998.

[4] T. S. Rappaport, J. H. Reed, and B. Woerner, "Position Location Using Wireless Communications on Highways of the Future," *IEEE Commun. Mag.*, vol. 34, no. 10, pp. 33 – 41, 1996.

[5] I. Jami, M. Ali, and R. Ormondroyd, "Comparison of Methods of Locating and Tracking Cellular Mobiles," in *Proc. IEE Colloquium on Novel Methods of Location and Tracking of Cellular Mobiles and Their System Applications.*

[6] B. T. Fang, "Simple Solutions for Hyperbolic and Related Position Fixes," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 5, pp. 748 – 753, 1990.

[7] J. S. Abel and J. O. Smith, "Source Range and Depth Estimation from Multipath Range Difference Measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 8, pp. 1157 – 1165, 1989.

[8] W. H. Foy, "Position-location Solutions by Taylor-Series Estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 12, no. 2, pp. 187 – 194, 1976.

[9] Y. Chan and K. Ho, "A Simple and Efficient Estimator for Hyperbolic Location," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 1905 – 1915, 1994.

[10] H. C. Schau and A. Z. Robinson, "Passive Source Localization Employing Intersecting Spherical Surfaces from Time-of-Arrival Differences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 8, pp. 1223 – 1225, 1987.

[11] F. Cesbron and R. Arnott, "Locating GSM Mobiles Using Antenna Array," *Electronics Letters*, vol. 34, no. 16, pp. 1539 – 1540, 1998.

[12] L. Cong and W. Zhuang, "Hybrid TDOA/AOA Mobile User Location for Wideband CDMA Cellular Systems," *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 439 – 447, 2002.

[13] H. Laitinen, S. Juurakko, T. Lahti, R. Korhonen, and J. Lahteenmaki, "Experimental Evaluation of Location Methods Based on Signal-Strength Measurements," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 287 – 296, 2007.

[14] M. Hellebrandt and R. Mathar, "Location Tracking of Mobiles in Cellular Radio Networks," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1558 – 1562, 1999.

[15] D. Catrein, M. Hellebrandt, R. Mathar, and M. P. Serrano, "Location Tracking of Mobiles: a Smart Filtering Method and its Use in Practice," in *Proc. 59th IEEE Vehicular Technology Conference*, vol. 5, May 2004, pp. 2677 – 2681.

[16] T. Roos, P. Myllymaki, and H. Tirri, "A Statistical Modeling Approach to Location Estimation," vol. 1, no. 1, pp. 59 – 69, 2002.

[17] A. Weiss, "On the Accuracy of a Cellular Location System Based on RSS Measurements," *IEEE Trans. Veh. Technol.*, vol. 52, no. 6, pp. 1508 – 1518, 2003.

[18] O. Hilsenrath and M. Wax, "Radio Transmitter Location Finding for Wireless Communication Network Service and Management," U.S. Patent 6 026 304, Feb. 15, 2000.

[19] H. Laitinen, J. Lahteenmaki, and T. Nordstrom, "Database Correlation Method for GSM Location," in *Proc. 53rd IEEE Vehicular Technology Conference*, vol. 4, May 2001, pp. 2504 – 2508.

[20] M. Triki, D. T. M. Slock, V. Rigal, and P. Francois, "Mobile Terminal Positioning via Power Delay Profile Fingerprinting: Reproducible Validation Simulations," in *Proc. 64th IEEE Vehicular Technology Conference*, Sep. 2006.

[21] K. J. Krizman, T. E. Biedka, and T. S. Rappaport, "Wireless Position Location: Fundamentals, Implementation Strategies, and Sources of Error," in *Proc. 47th IEEE Vehicular Technology Conference*, vol. 2, May 1997, pp. 919 – 923.

[22] J. J. Caffery and G. L. Stuber, "Overview of Radiolocation in CDMA Cellular Systems," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 38 – 45, 1998.

[23] M. Spirito, "On the Accuracy of Cellular Mobile Station Location Estimation," *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 674 – 685, 2001.

[24] R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276 – 280, 1986.

[25] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984 – 995, 1989.

[26] L. Dumont, M. Fattouche, and G. Morrison, "Super-Resolution of Multipath Channels in a Spread Spectrum Location System," *Electronics Letters*, vol. 30, no. 19, pp. 1583 – 1584, 1994.

[27] H. Saarnisaari, "TLS-ESPRIT in a Time Delay Estimation," in *Proc. 47th IEEE Vehicular Technology Conference*, May 1997.

[28] Z. Kostic, M. Sezan, and E. Titlebaum, "Estimation of the Parameters of a Multipath Channel Using Set-Theoretic Deconvolution," *IEEE Trans. Commun.*, vol. 40, no. 6, pp. 1449 – 1452, 1992.

[29] W. Gardner and C.-K. Chen, "Signal-selective time-difference-of-arrival estimation for passive location of man-made signal sources in highly corruptive environments, part I: Theory and method," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1168 – 1184, 1992.

[30] ——, "Signal-selective time-difference-of-arrival estimation for passive location of man-made signal sources in highly corruptive environments, part II: Algorithms and performance," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1185 – 1197, 1992.

[31] Y. Qi, H. Kobayashi, and H. Suda, "On Time-of-Arrival Positioning in a Multipath Environment," *IEEE Trans. Veh. Technol.*, vol. 55, no. 5, pp. 1516 – 1526, 2006.

[32] L. Xiong, "A Selective Model to Suppress NLOS Signals in Angle-of-Arrival (AOA) Location Estimation," in *Proc. 9th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications.*

[33] Y.-T. Chan, W.-Y. Tsui, H. So, and P.-C. Ching, "Time-of-Arrival Based Localization Under NLOS Conditions," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 17 – 24, 2006.

[34] P.-C. Chen, "A Non-Line-of-Sight Error Mitigation Algorithm in Location Estimation," in *Proc. IEEE Wireless Communications and Networking Conference.*

[35] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of Wireless Geolocation in a Non-Line-of-Sight Environment," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 672 – 681, 2006.

[36] S. Venkatraman, J. J. Caffery, and H.-R. You, "A Novel ToA Location Algorithm Using LoS Range Estimation for NLoS Environments," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1515 – 1524, 2004.

[37] S. Al-Jazzar, M. Ghogho, and D. McLernon, "A Joint TOA/AOA Constrained Minimization Method for Locating Wireless Devices in Non-Line-of-Sight Environment," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 468 – 472, 2009.

[38] H. Miao, K. Yu, and M. J. Juntti, "Positioning for NLOS Propagation: Algorithm Derivations and Cramer-Rao Bounds," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2568 – 2580, 2007.

[39] S. Al-Jazzar and J. J. Caffery, "ML and Bayesian TOA Location Estimators for NLOS Environments," in *Proc. 56th IEEE Vehicular Technology Conference*, vol. 2, Sep. 2002, pp. 1178– 1181.

[40] X. Wang, Z. Wang, and B. O'Dea, "A TOA-based Location Algorithm Reducing the Errors due to Non-Line-of-Sight (NLOS) Propagation," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 112 – 116, 2003.

[41] H. Miao, K. Yu, and M. J. Juntti, "Positioning for NLOS Propagation: Algorithm Derivations and Cramer-Rao Bounds," in *Proc. IEEE 2006 International conference on Acoustics, Speech and Signal Processing.*

[42] M. Debbah and R. R. Muller, "Capacity Complying MIMO Channel Models," in *Proc. 37th Annual Asilomar Conference on Signals, Systems and Computers.*

[43] ——, "MIMO Channel Modeling and the Principle of Maximum Entropy," *IEEE Trans. Inform. Theory*, vol. 51, no. 5, pp. 1667 – 1690, 2005.

[44] R. R. Muller, "A Random Matrix Model of Communication Via Antenna Arrays," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2495 – 2506, 2002.

[45] E. Jaynes, *Probability Theory: The logic of science.* Cambridge University Press, 2003.

[46] D. Basu, "On the Elimination of Nuisance Parameters," *JSTOR*, vol. 72, no. 358, pp. 355 – 366, 1977.

[47] J. C. Spall and J. P. Garner, "Parameter Identification for State-Space Models with Nuisance Parameters," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, no. 6, pp. 992 – 998, 1990.

[48] J. O. Berger, B. Liseo, and R. L. Wolpert, "Integrated Likelihood Methods for Eliminating Nuisance Parameters," *JSTOR*, vol. 14, no. 1, pp. 1 – 28, 1999.

[49] K. Papakonstantinou and D. Slock, "Identifiability and Performance Concerns in Location Estimation," 2009, proc. IEEE International Conference on Acoustics, Speech and Signal Processing 2009.

[50] T. J. Rothenberg, "Identification in Parametric Models," *Econometrica*, vol. 39, no. 3, pp. 577 – 591, 1971.

[51] M. Haardt and J. Nossek, "Unitary ESPRIT: How to Obtain Increased Accuracy with a Reduced Computational Burden," *IEEE Trans. Signal Processing*, vol. 43, no. 5, pp. 1232 – 1242, 1995.

[52] ——, "Simultaneous Schur Decomposition of Several Nonsymmetric Matrices to Achieve Automatic Pairing in Multidimensional Harmonic Retrieval Problems," *IEEE Trans. Signal Processing*, vol. 46, no. 1, pp. 161–169, 1998.

[53] A. van der Veen, M. Vanderveen, and A. Paulraj, "Joint Angle and Delay Estimation Using Shift-Invariance Techniques," *IEEE Trans. Signal Processing*, vol. 46, no. 2, pp. 405–418, 1998.

[54] H. Miao, K. Yu, and M. J. Juntti, "2-D Unitary ESPRIT Based Joint AOA and AOD Estimation for MIMO System," in *Proc. IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2006.

[55] A. Swindlehurst, B. Ottersten, R. Roy, and T. Kailath, "Multiple Invariance ESPRIT," *IEEE Trans. Signal Processing*, vol. 40, no. 4, pp. 867 – 881, 1992.

[56] D. J. Torrieri, "Statistical Theory of Passive Location Systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 20, no. 2, pp. 183 – 198, 1984.

[57] N. Levanon, "Lowest GDOP in 2-D Scenarios," *IEE Radar, Sonar and Navigation*, vol. 147, no. 3, pp. 149 – 155, 2000.

[58] C. Fritsche and A. Klein, "Cramer-Rao Lower Bounds for Hybrid Localization of Mobile Terminals," in *Proc. 5th Workshop on Positioning, Navigation and Communication*, Mar. 2008.

[59] M. Laaraiedh, S. Avrillon, and B. Uguen, "Hybrid Data Fusion Techniques for Localization in UWB Networks," in *Proc. 6th Workshop on Positioning, Navigation and Communication*, Mar. 2009.

[60] K. Papakonstantinou and D. Slock, "Hybrid Localization Algorithm based on time-variation of Location Dependent Parameters," in *to be submitted*.

[61] J. C. Liberti and T. S. Rappaport, "A Geometrically Based Model for Line-of-Sight Multipath Radio Channels ," in *Proc. 42th IEEE Vehicular Technology Conference*.

[62] P. Petrus, J. H. Reed, and T. S. Rappaport, "Geometrically Based Statistical Channel Model for Macrocellular Mobile Environments ," in *Proc. Global Telecommunications Conference*.

[63] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified Channel Model for Mobile Radio Systems with Smart Antennas," *IEE Radar, Sonar and Navigation*, vol. 145, no. 1, pp. 32 – 41, 1998.

[64] R. B. Ertel and J. H. Reed, "Angle and Time of Arrival Statistics for Circular and Elliptical Scattering Models," *IEEE Trans. Commun.*, vol. 17, no. 11, pp. 1829 – 1840, 1999.

[65] L. Jiang and S. Y. Tan, "Geometrically Based Statistical Channel Models for Outdoor and Indoor Propagation Environments," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3587 – 3592, 2007.

[66] K. Papakonstantinou and D. Slock, "NLOS Mobile Terminal Position and Speed Estimation," in *Proc. 3rd International Symposium on Communications, Control and Signal Processing*, Mar. 2008.

[67] ——, "Hybrid TOA/AOD/Doppler-Shift Localization Algorithm for NLOS Environments," in *Proc. IEEE 20th Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, 2009.

[68] J. J. Caffery, "A new Approach to the Geometry of ToA Location," in *Proc. 52nd IEEE Vehicular Technology Conference*, vol. 4, Sep. 2000, pp. 1943–1949.

[69] A. Amar and A. J. Weiss, "Direct Position Determination of Multiple Radio Signals," in *Proc. IEEE 2004 International conference on Acoustics, Speech and Signal Processing.*

[70] ——, "New Asymptotic Results on Two fundamental Approaches to Mobile Terminal Location," in *Proc. 3rd International Symposium on Communications, Control and Signal Processing*, Mar. 2008.

[71] ——, "Advances in Direct Position Determination," in *Proc. Sensor Array and Multichannel Signal Processing Workshop*, Jul. 2004.

[72] K. Papakonstantinou, M. Debbah, and D. Slock, "MIMO Mobile Terminal Tracking Using Bayesian Probability Estimation," in *Proc. 42th Annual Asilomar Conference on Signals, Systems and Computers*, Nov. 2007.

[73] H. L. V. Trees, *Optimum Array Processing.* John Wiley and Sons, 2002.

[74] I. E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585 – 595, 1999.

[75] X. Ma, L. Yang, and G. B. Giannakis, "Optimal Training for MIMO Frequency-Selective Fading Channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453 – 466, 2005.

[76] H. Lutkepohl, *Handbook of Matrices*. John Wiley and Sons, 1996.

[77] A. Lee, "Centrohermitian and Skew-Centrohermitian Matrices," *Linear Algebra and its applications*, vol. 29.