HAL
open science

Fingerprinting based Localization of Mobile Terminals

Turgut Mustafa Öktem

To cite this version:

Turgut Mustafa Öktem. Fingerprinting based Localization of Mobile Terminals. Signal and Image processing. Télécom ParisTech, 2011. English. NNT : . pastel-00670218

HAL Id: pastel-00670218

https://pastel.hal.science/pastel-00670218v1

Submitted on 14 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

TELECOM ParisTech

2012-ENST-00xx

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Communication et Electronique »

présentée et soutenue publiquement par

Turgut Mustafa ÖKTEM

le 27 Juin 2011

Localisation de Terminaux Mobiles par Exploitation d'Empreintes

Directeur de thèse : **Dirk SLOCK**

T H È S E

Jury

M. Bernard FLEURY, Prof., Department of Electronic Systems, Aalborg University	Rapporteur
M. Bernard UGUEN, Prof., IETR, Université de Rennes 1	Rapporteur
M. Philippe CIBLAT, Prof., Communications and Electronics Dep., Télécom ParisTech	Examinateur
M. Luc DENEIRE, Prof., Laboratoire I3S, Université de Nice Sophia-Antipolis	Examinateur

TELECOM ParisTech
école de l'Institut Télécom - membre de ParisTech

Abstract

This dissertation is a collection of our approaches to solve the problem of mobile terminal (MT) location estimation. Main focus is on fingerprinting based localization methods which are suitable for multipath and NLoS environments. This feature is one of the many advantages of fingerprinting methods over "traditional" geometrical localization methods, i.e., they exploit multipath and NLoS conditions instead of trying to mitigate them. Hence a curse for "traditional" geometrical localization methods turns into a blessing for fingerprinting methods. This characteristic property of fingerprinting algorithms makes them a promising solution for MT localization problem. Being inspired from this, we steered our research direction towards this field and we can categorize our studies into the following main groups:

- development of new, high precision fingerprinting-based localization algorithms which get use of an additional dimension (Doppler dimension),
- derivations of Cramér-Rao bounds (CRBs) to obtain performance limits for fingerprinting based localization systems, for the estimation of location-dependent parameters (LDPs), as well as the position of the MT,
- investigation of identifiability concerns of the MT position under different circumstances and path amplitude modeling,
- pairwise error probability (PEP) analysis for power delay profile fingerprinting (PDP-F) methods for different cost functions and path amplitude modeling,
- MT tracking based on adaptive Kalman filtering (KF).

This thesis consists of 7 chapters, including introduction and conclusion. We pay special attention to have a meaningful ordering of the chapters.

Chapter 1 is the introduction where we present the state of the art localization systems and then introduce the basics of fingerprinting based localization systems. This chapter serves two purposes: informing readers about fundamentals of localization systems who might be unfamiliar to the subject, and presenting some material which we will get use of throughout the text.

In chapter 2, two new fingerprinting algorithms are introduced, both exploiting the mobility of the MT to increase localization accuracy. Comprehensive derivations and explanations are provided to have an easy to understand context.

Chapter 3 can be considered as the core of our studies. CRBs for the estimation of LDPs and the MT position are derived under various different conditions, path amplitude modelings for PDP-F. Impact of the network geometry on the estimation of MT position is clearly stated. Additionally, local identifiability issues of the MT position are extensively investigated.

Chapter 4 is an extension of our work in chapter 3 to power delay Doppler profile-fingerprinting (PDDP-F) method. Improvements arising with the integration of Doppler shifts of the paths are stated in terms of CRBs.

In chapter 5, we deal with a new problem in the field of localization. PEP analysis is a well-known concept in digital communication, e.g., calculating the probability of error when a vector of symbols $\mathbf{s_m}$ is transmitted but another vector of symbols $\mathbf{s_n}$ is detected at the receiver. The same concept is imported in the field of localization for PDP-F methods utilizing various cost functions.

Chapter 6 deals with MT tracking based on adaptive KF. It is in the context of a project we were involved. Different mobility models are presented and applied to the measurement data. Their performances are compared in terms of the position prediction error.

In the final chapter, there is a brief summary of our conclusions we have obtained during all localization research we have conducted. We try to justify our results. Furthermore, we make a list of ideas about future research subjects and the extension of our work.

Acknowledgements

It was not an easy decision for me whether to make a PhD or not after finishing my masters. What influenced me the most was the joy I got from doing research with my master supervisor Prof. Alper Tunga Erdogan. Having the chance to work with him made me love research and as a result, I found myself in France pursuing a PhD degree.

For the past three and a half years that I was working at Eurecom, I had the chance to work in a very warm environment. I had many friends from many nationalities. I guess this was the best part of working in Eurecom, to be able to meet people from all over the world. There are some friends that I especially want to mention here. Melek comes the first among them to thank to. Whenever we (the Turkish community) needed help, she was always there to help us. To be honest, without friends, I can't imagine that I could complete this long and hard journey. I had so many good memories with them. Especially my friends whom I shared my office: Lorenzo, Antony and Samir. Even though Lorenzo was the newest in my office, I spent much more time with him than anybody else. The cheerful moments we had in Nice while eating our favorite pizza aubergine was priceless. I am also grateful to Daniel who was always very kind, helpful to me. He was like an elder brother for me. I also would like to acknowledge my other friends, namely: Najam, Erhan, Umer, Ufuk and Nesli.

I would also like to thank all my jury members, especially the reporters of my thesis who has spent their time to evaluate my thesis and have given me invaluable feedback to improve it. Thank you very much Prof. Bernard Fleury and Prof. Bernard Uguen.

Also, I must thank to my supervisor Prof. Dirk Slock for his guidance and help during my PhD period.

It would be unfair if I did not mention my dear Nataliya for her endless support and love for me. For the times when I was completely down, she always found a way to encourage me. Thanks for being so kind, so nice, so lovely.

I know that, without the love of my family, I would not be able to accomplish any of my achievements in my life. Hence, my deepest gratitudes go to them, to my parents, Bahri and Esen and to my brothers, Fatih and Sinan. I feel very lucky to be a member of this family. And you, my dear father, you could not have the chance to see that I became a doctor. You have always been the person who inspired me the most in my life with your everlasting energy, cheer, honesty and strong character. Although it is a negligible present compared to what you have given to me, I dedicate my PhD to you.

Contents

Abstract i
Acknowledgements iii
Contents v
List of Figures viii
Acronyms xii
Notation xv

1 Introduction **1**
1.1 Fundamentals of Geometrical Localization Methods 2
1.1.1 Angle of Arrival Methods 3
1.1.2 Time of Arrival Methods 4
1.1.3 Time Difference of Arrival Methods 5
1.1.4 Received Signal Strength Methods 6
1.1.5 Hybrid Methods 6
1.2 Fingerprinting Methods 7
1.3 Identifiability Issues and Performance Bounds 10

2 Power Delay Doppler Profile Fingerprinting **13**
2.1 Introduction 13
2.2 Frequency-Domain PDDP-F 13
2.2.1 Obtaining PDDP from Ray Tracing Data 14
2.2.2 Obtaining PDDP from Measurement Data and The Fingerprinting Operation 19
2.2.3 Simulation Results 21
2.3 Time-Domain PDDP-F 23
2.3.1 Simulation Results 27
2.4 Conclusion 29

3 Performance Analysis of Power Delay Profile Fingerprinting **31**
3.1 Introduction 31

3.2 CRB Analysis of PDP-F 31
3.2.1 Rayleigh Fading Case 33
3.2.2 Deterministic Path Amplitude Case 50
3.2.3 Rician Fading Case 58
3.3 Discussion 61
3.3.1 Localization Performance for Deterministic Path Amplitudes and Rayleigh Fading Cases 63
3.4 Conclusion 72

4 Performance Analysis of Power Delay Doppler Profile Fingerprinting 77
4.1 Introduction 77
4.2 CRB Analysis of PDDP-F 77
4.2.1 Rayleigh Fading Case 79
4.2.2 Deterministic Path Amplitude Case 90
4.3 Conclusion 98

5 Pairwise Error Probability Analysis for Power Delay Profile Fingerprinting 101
5.1 Introduction 101
5.2 PEP of the LS Technique for Deterministic Path Amplitudes 101
5.2.1 Simulation Results 106
5.3 PEP of the GML Technique for Rayleigh Fading 107
5.3.1 Simulation Results 114
5.4 Conclusion 115

6 Mobile Terminal Tracking based on Kalman Filtering 119
6.1 Introduction 119
6.2 Adaptive Kalman Filtering based Tracking 119
6.2.1 Adaptive Kalman Filtering Approaches 119
6.2.2 State-Space Models for Position Tracking 127
6.2.3 Adaptive EM-KF with Fixed-Lag Smoothing for Position Tracking 129
6.2.4 Non-Linear Measurements 129
6.3 Fitting of State Space Mobility Models to M3 Measurements 130
6.4 Conclusion 134

7 Conclusions and Future Work 137

Appendices 141

A CRB of the Position Vector for Non-overlapping Case 143

B Distribution of the Summation of Non-identical Exponential RVs 145

C Résumé Étendu en Français 149

C.1 Abstract en français 149
C.2 Contributions et Cadre de cette Thèse 150
C.3 Résumé du Chapitre 1 153
C.3.1 Questions d'identifiabilité et Bornes de Performance . 155
C.4 Résumé du Chapitre 2 156
C.4.1 Domaine de Fréquence PDDP-F 156
C.4.2 Domaine Temporel PDDP-F 159
C.5 Résumé du Chapitre 3 162
C.5.1 Conclusions 167
C.6 Résumé du Chapitre 4 168
C.6.1 Conclusions 169
C.7 Résumé du Chapitre 5 169
C.7.1 PEP de la Technique LS pour des Amplitudes de Chemin Déterministes 170
C.7.2 PEP de la Technique de GML pour Rayleigh Fading . 170
C.7.3 Conclusions 172
C.8 Résumé du Chapitre 6 174
C.9 Futures travaux 175

List of Figures

1.1 AoA-based localization with two BSs (triangulation). 3
1.2 ToA-based localization with three BSs (trilateration). 4
1.3 TDoA-based localization with three BSs. 5
1.4 Hybrid AoA/ToA-based localization with one BS. 7
1.5 A multipath environment. 8

2.1 MIMO multipath propagation parameters 14
2.2 Power Delay Angle of Arrival Profile 20
2.3 Power Delay Doppler Profile 21
2.4 Power Delay Doppler Profile with pulse shape and windowing effects included 22
2.5 Performance comparison of deterministic PDP-F and frequency-domain PDDP-F algorithms. 24
2.6 Performance results of time-domain PD(D)P-F as a function of n. 28
2.7 Time domain PDDP-F for $n = 2$ with and without speed vector estimation. 29

3.1 RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 1$ MHz. 40
3.2 RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 10$ MHz. 41
3.3 RMSE of $\hat{\sigma_i^2}$ as a function of σ_i^2 for $W = 1$ MHz. 42
3.4 RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 1$ MHz, $\sigma_j^2 = 2$. . 48
3.5 RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 1$ MHz, $\sigma_j^2 = 20$. 49
3.6 RMSE of $\hat{\sigma_i^2}$ as a function of σ_i^2 for $W = 1$ MHz, $\sigma_j^2 = 2$. . . 50
3.7 RMSE of $\hat{\sigma_i^2}$ as a function of σ_i^2 for $W = 1$ MHz, $\sigma_j^2 = 20$. . . 51
3.8 BS-MT geometry. 62
3.9 RMSE of $\hat{\mathbf{r}}$ as a function of SNR_1 and SNR_2 for $\psi_2 - \psi_1 = \pi/2$ for the deterministic case. 65
3.10 RMSE of $\hat{\mathbf{r}}$ as a function of SNR_1 and SNR_2 for $\psi_2 - \psi_1 = \pi/2$ for Rayleigh fading case. 66

3.11 RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 10dB$ and $\psi_2 - \psi_1 = \pi/4$ for the deterministic case. 73
3.12 RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 20dB$ and $\psi_2 - \psi_1 = \pi/4$ for the deterministic case. 74
3.13 RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 10dB$ and $\psi_2 - \psi_1 = \pi/2$ for the deterministic case. 75
3.14 RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 20dB$ and $\psi_2 - \psi_1 = \pi/2$ for the deterministic case. 75
3.15 RMSE of $\hat{\mathbf{r}}$ as a function of $\psi_2 - \psi_1$, $SNR_1 = 10dB$, $SNR_2 = 0dB$ for the Rayleigh fading case. 76
3.16 RMSE of $\hat{\mathbf{r}}$ as a function of $\psi_2 - \psi_1$, $SNR_1 = 10dB$, $SNR_2 = 10dB$ for the Rayleigh fading case. 76

4.1 RMSE of $\hat{f}_i$ as a function of SNR_i. 97
4.2 RMSE of $\hat{\phi}_i$ as a function of SNR_i. 98

5.1 Pairwise error probability for $\alpha = 1.1$. 107
5.2 Pairwise error probability for $\alpha = 1.2$. 108
5.3 Pairwise error probability for $\alpha = 1.4$. 109
5.4 Pairwise error probability for $\beta = 1.1$. 115
5.5 Pairwise error probability for $\beta = 1.2$. 116
5.6 Pairwise error probability for $\beta = 1.4$. 117

6.1 A sample of an aerial photo with trajectory superposed. . . . 133
6.2 A sample of an aerial photo with trajectory superposed. . . . 133
6.3 A sample of an aerial photo with trajectory superposed. . . . 134
6.4 Position prediction error histogram for the temporally white, spatially colored noise acceleration model. 135
6.5 Position prediction error histogram for the spatio-temporally white acceleration model. 136
6.6 Position prediction error histogram for the AR(1) (Markov) velocity model. 136

C.1 Un environnement multipath. 154
C.2 Localisation basée sur ToA avec trois BSs (trilateration). . . 155
C.3 MIMO multipath paramètres de propagation 156
C.4 Power Delay Angle of Arrival Profile 158
C.5 Power Delay Doppler Profile 158
C.6 Power Delay Doppler Profile avec la forme des impulsions et le fenêtrage effets inclus 159

C.7 Comparaison des performances du PDP-F déterministe and domaine de fréquence PDDP-F. 160
C.8 Résultat de performance de domaine temporel PD(D)P-F en fonction de n. 162
C.9 RMSE de $\hat{\tau}_i$ en fonction de SNR_i pour $W = 1$ MHz, $\sigma_j^2 = 2$. 163
C.10 RMSE de $\hat{\tau}_i$ en fonction de SNR_i pour $W = 1$ MHz, $\sigma_j^2 = 20$. 164
C.11 RMSE de $\hat{\sigma_i^2}$ en fonction de σ_i^2 pour $W = 1$ MHz, $\sigma_j^2 = 2$. . . 164
C.12 RMSE de $\hat{\sigma_i^2}$ en fonction de σ_i^2 pour $W = 1$ MHz, $\sigma_j^2 = 20$. . 165
C.13 La géométrie de la BS-MT. 166
C.14 RMSE de $\hat{\mathbf{r}}$ en fonction de SNR_1 et SNR_2 pour $\psi_2 - \psi_1 = \pi/2$ pour la cas deterministic. 166
C.15 RMSE de $\hat{\mathbf{r}}$ en fonction de SNR_1 et SNR_2 pour $\psi_2 - \psi_1 = \pi/2$ pour la cas Rayleigh fading. 167
C.16 PEP pour $\alpha = 1.1$. 171
C.17 PEP pour $\alpha = 1.2$. 171
C.18 PEP pour $\alpha = 1.4$. 172
C.19 PEP pour $\beta = 1.1$. 173
C.20 PEP pour $\beta = 1.2$. 173
C.21 PEP pour $\beta = 1.4$. 174
C.22 Histogramme erreur de position de prédiction pour le blanc temporellement, le bruit spatialement coloré modèle d'accélération.175
C.23 Histogramme erreur de position de prédiction pour l'accélération spatio-temporellement blanc modèle. 176
C.24 Histogramme erreur de position de prédiction pour AR(1) (Markov) modèle de vitesse. 176

Acronyms

Here, we list the main acronyms used in this document.

AoA	Angle(s) of Arrival.
AoD	Angle(s) of Departure.
AGC	Automatic Gain Control.
AR	Autoregressive.
AWGN	Additive White Gaussian Noise.
BE	Bayesian Estimation.
BS	Base Station.
CIR	Channel Impulse Response.
CLT	Central Limit Theorem.
CRB	Cramér-Rao Bound.
DLE	Direct Location Estimation.
DO	Delay Offset.
EKF	Extended Kalman Filter.
EM	Expectation-Maximization.
FCC	(U.S.) Federal Communications Commission.
FIM	Fisher Information Matrix.
FO	Frequency Offset.
GDoP	Geometric Dilution of Precision.
GML	Gaussian Maximum Likelihood.
i.i.d.	independent and identically distributed.
ISD	Itakura-Saito Distance.
KF	Kalman Filter.
LDP	Location-Dependent Parameter.
LF	Location Fingerprinting.
LMMSE	Linear Minimum Mean Squared Error.
LoS	Line-of-Sight.
LS	Least Squares.

MAP	Maximum A Posteriori.
MIMO	Multiple Input Multiple Output.
MISO	Multiple Input Single Output.
ML	Maximum Likelihood.
MMSE	Minimum Mean Squared Error.
MPC	Multipath Component.
MSE	Mean Square Error.
MT	Mobile Terminal.
NLoS	Non Line-of-Sight.
pdf	probability density function.
PD	Positive Definite.
PDAoAP	Power Delay Angle of Arrival Profile.
PDDP	Power Delay Doppler Profile.
PDDP-F	Power Delay Doppler Profile-Fingerprinting.
PDDSP	Power Delay Doppler Space Profile.
PDP	Power Delay Profile.
PDP-F	Power Delay Profile-Fingerprinting.
PEP	Pairwise Error Probability.
PSD	Positive Semi-Definite.
r.h.s.	right hand side.
RMSE	Root Mean Squared Error.
RPEM	Recursive Prediction Error Method.
RSS	Received Signal Strength.
SIMO	Single Input Multiple Output.
SNR	Signal-to-Noise Ratio.
SISO	Single-Input Single-Output.
TDoA	Time Difference of Arrival.
ToA	Time of Arrival.
UWB	Ultra-Wideband.
w.r.t.	with respect to.
2D (3D, 4D)	two (three, four) dimensional.

Notation

Throughout this document, upper case and lower case boldface symbols will represent matrices and column vectors respectively.

$\mathrm{tr}\,\{\cdot\}$	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 $\mathbf{a}$.
$\lVert\mathbf{A}\rVert_F$	Frobenius norm of matrix $\mathbf{A}$.
$\hat{(\cdot)}$	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 $\mathbf{A}$.
$\mathbf{A}^H$	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 $\mathbf{A}$.
$\mathbf{A} \geq \mathbf{B}$	means that $\mathbf{A} - \mathbf{B}$ is non-negative definite.
$f(\cdot)$	pdf of the continuous random variable in parenthesis.
$\Pr(\cdot)$	Probability of the event in parenthesis.
$\mathbb{E}\{\cdot\}$	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
$\max, \min$	Maximum and minimum.
$\sim$	Distributed according to.
$\odot$	Hadamard Product.
$\otimes$	Kronecker Product.

Chapter 1

Introduction

For the past few years, there has been a high interest in mobile positioning systems from both academic and industrial world [1, 2]. The primary motivation for the development of mobile positioning systems was due to the mandatory requirement of E-911 service by the U.S. Federal Communications Commission (FCC) [3, 4]. Although the starting was because of security-emergency need, later it has found various applications in many fields. Position information of the mobile is indeed very useful and can be evaluated in many ways. For example, with the position information of the MT, it is possible to make beamforming in the direction of the mobile to decrease the interference between the users in the cell, increase the range and throughput of the system and so on. Also intra and inter-system handoffs can be handled more properly. Therefore it is not surprising that the problem of MT localization attracts so much interest from researchers. Research in the field of localization in a more general sense can be categorized under the "parameter estimation" topic which is widely popular for a considerably long time.

Due to the difficulty of the localization problem under realistic propagation conditions, researchers first try to estimate meaningful parameters that can in turn be used for the localization process. Those parameters are known as location-dependent parameters (LDPs). As the name suggests, these LDPs can be expressed as a function of the MT position via the exploitation of geometrical relations. Hence localization of the MT can be

carried out afterwards by solving a system of equations. The most commonly used LDPs are the angle of arrival (AoA), time of arrival (ToA), time difference of arrival (TDoA), and the received signal strength (RSS) [5]. Hybrid localization methods use a combination of two or more of these LDPs to increase the accuracy further.

"Traditional" geometrical localization methods are by far the most well-known methods in the field of localization. These methods have a two step localization process. In the 1st step, LDPs are estimated in an adequate number of base stations (BSs). Afterwards, by using these LDPs, location of the MT is estimated by using geometric relations. However the biggest drawback of these methods is that they require line-of-sight (LoS) signals with the BSs. Moreover localization of the MT is not possible with only one BS if only one type of LDP is used. Although they are easy to use, these factors limit their applicability. We will briefly introduce them soon in the text.

Due to the many drawbacks of the "traditional" geometrical localization methods, research was steered towards another direction. As a result, fingerprinting-based localization methods have arisen which are also the main subject of our research. The main advantage of fingerprinting methods is that localization is possible even with only one BS. Moreover they do not suffer from non line-of-sight (NLoS) conditions, instead they exploit it. This was our main motivation to conduct research in this field. Implementation of fingerprinting-based localization methods is completely different than the geometrical methods. As the name suggests, they are based on matching "fingerprints" obtained from the BS-MT link with the entries of an already existing database consisting of "fingerprints" of discrete points over the area of interest. These "fingerprints" are basically LDPs or functions of LDPs. Wide applicability of fingerprinting methods in any environment (indoor and outdoor localization, dense urban environments, NLoS conditions) make them a good solution for the problem of MT localization. As they form the core of our research, we will elaborate them in detail later.

1.1 Fundamentals of Geometrical Localization Methods

This section is devoted to introducing the basics of traditional geometrical methods. For interested readers, [6] is a very good reference for this section. As we expressed before, localization process is performed in two steps for these methods. In the 1st step, LDPs are estimated by using sufficient num-

ber of BSs. Number of BSs required depends on the type of LDPs utilized in the system. Identifiability of the MT location is the highest priority (depends also on the dimension of localization, e.g. 2D or 3D). Another factor (or demand) is the high localization accuracy. In order to satisfy these two constraints, number of BSs required might be high depending on the LDPs used. In the 2nd step, a system of equations are solved with these estimated LDPs to estimate the location of the MT with respect to the coordinates of the BSs. We will now present a very brief summary of the most well-known ones of these methods.

1.1.1 Angle of Arrival Methods

The easiest AoA estimation is accomplished via the calculation of phase difference of the waves on different antenna elements that they impinge on. Hence, multiple antennas are required for this purpose at the receiver. Also the antenna array response must be known (antenna array response is also a function of the spacing between the antenna elements). In figure 1.1, we provide a simple sketch of the problem.

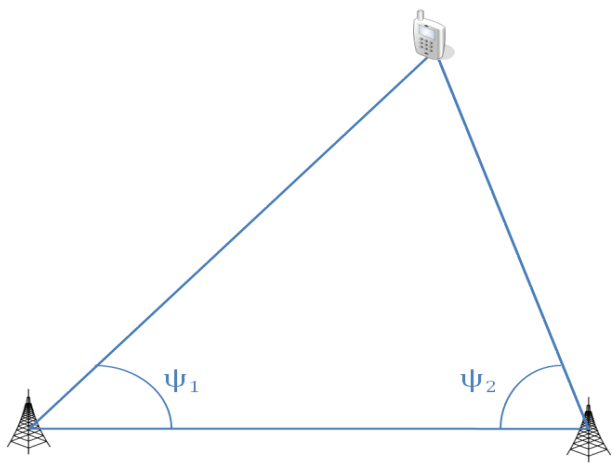

Figure 1.1: AoA-based localization with two BSs (triangulation).

It is clearly seen that at least two BSs are required to uniquely determine the MT location in 2D. With the estimated AoAs, location of the MT is calculated in terms of the coordinates of the BSs and AoAs via simple system of equations via the well-known technique called triangulation. To increase the localization accuracy, more BSs might be utilized and in that case MT location can be computed via least squares (LS) technique. Even though

it looks like a very elegant localization method, it has some drawbacks. Intuitively, by just looking at figure 1.1, we can understand the reason. Localization is accomplished via the estimated AoAs. Therefore small errors in the AoA estimates might lead to a high position estimation error when the MT is far from the BSs. Position estimation error increases when the MT moves further from the BSs.

1.1.2 Time of Arrival Methods

Estimating the arrival times of the LoS paths in each BS-MT link and then converting these ToA estimates to distance estimates via the speed of propagation is the basic principle of ToA-based localization methods. In general, ToA estimation is performed by means of correlation techniques or matched filtering [7]. After obtaining the ToA estimates, and hence the distance estimates, localization of the MT can be done via a well-known technique known as trilateration as shown in figure 1.2.

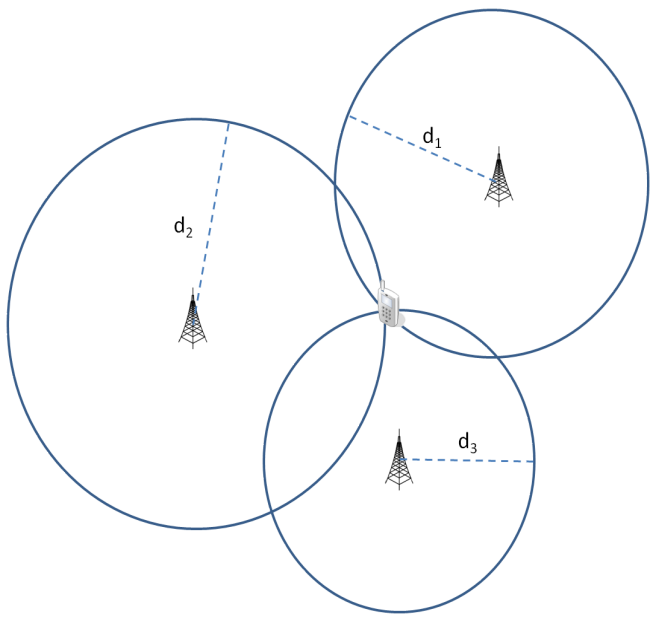

Figure 1.2: ToA-based localization with three BSs (trilateration).

For 2D localization, MT location estimate is obtained by the intersection of circles (for 3D, it will be an intersection of spheres) and at least three BSs are required for a unique position estimation. Major drawback of ToA-based localization systems is that they require strict time synchronization among

all BSs and the MT. Otherwise unknown delay offsets (DOs), unsynchronized BSs might seriously degrade the localization accuracy. Localization performance of ToA-based techniques in the perfectly synchronized case can be found in [8]. What we mean by localization performance is the Cramér-Rao bound (CRB) for the estimation of MT location. We will introduce CRBs soon in the local identifiability section.

1.1.3 Time Difference of Arrival Methods

Due to the strict synchronization requirement of ToA methods, a more flexible method in terms of synchronization was developed. For TDoA-based localization methods, synchronization among BSs is enough [9]. This is because in this technique, the difference between the ToA of the signal at two different BSs is utilized instead of the absolute arrival times. Below, we provide a simple sketch of the problem.

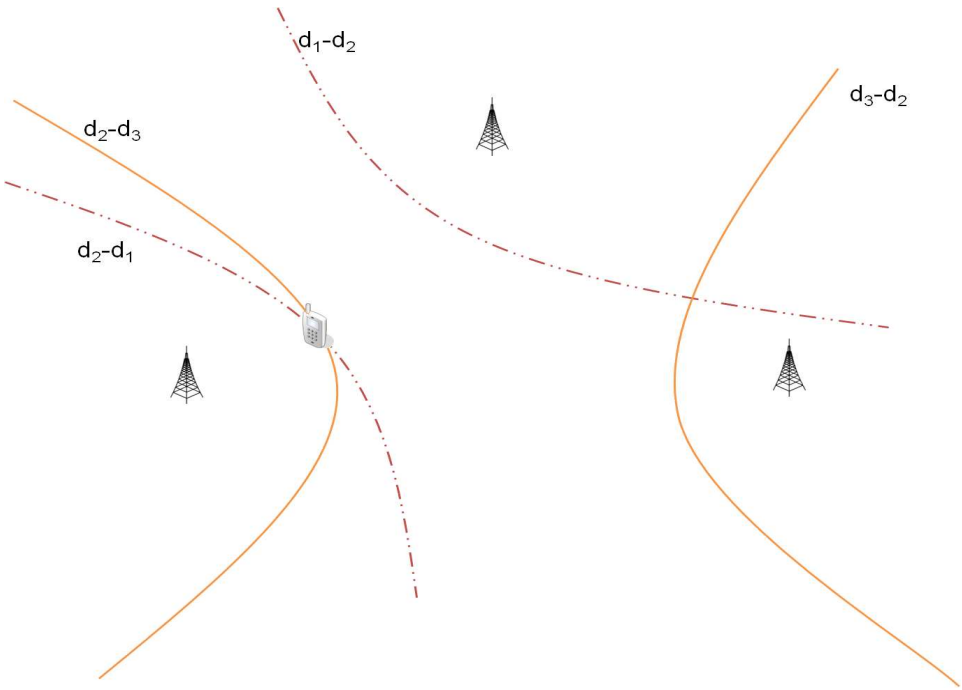

Figure 1.3: TDoA-based localization with three BSs.

For three BSs case, there are two TDoA estimates at hand, and each TDoA estimate defines a hyperbola passing through the MT with foci at the BSs. Estimate of the MT position is obtained by their intersection as seen in

figure 1.3. However it is not as easy as ToA-based localization since now non-linear equations need to be solved [10]. A unique position estimate can be obtained under geometric regularity conditions with at least three BSs for 2D localization.

TDoA estimates might be obtained either from ToA estimates, or from cross-correlations of the received signals at the two BSs and by calculating the delay corresponding to the highest cross-correlation output.

Note that relaxation of the synchronization requirement comes with a cost in terms of performance degradation with respect to the ToA-based localization techniques for the same number of BSs. This is reasonable because available data is decreased by one for TDoA (2 TDoA estimates can be obtained from 3 ToA estimates). A comprehensive study about the performance comparisons of ToA and TDoA-based localization systems can be found in [11].

1.1.4 Received Signal Strength Methods

RSS-based localization methods were first introduced by [12]. The most attractive feature of RSS-based methods is the availability of RSS measurements in almost all systems as it is easy to obtain. However RSS-based localization methods have low accuracy. The accuracy of distance estimates obtained from RSS data deteriorates as the distance between MT and BS increases [13]. Therefore it is not favorable to use RSS data standalone to perform localization. However, as we just mentioned, RSS data is widely available. Therefore it is better to take advantage of it by so-called hybrid data fusion techniques. As the name indicates, these techniques utilize different kinds of LDPs, fuse them to obtain a higher localization accuracy. For example in the same system configuration, fusing RSS and ToA data gives better results than the systems using only ToA or RSS data standalone [14].

1.1.5 Hybrid Methods

We just talked about hybrid methods. In hybrid schemes, a combination of LDPs are estimated, and localization is performed by using them all together. There are various hybrid schemes, such as ToA/RSS, ToA/AoA, etc. For example the hybrid ToA/AoA system can estimate the position of the MT uniquely (in 2D) using only one BS as seen in figure 1.5 [15].
Besides increasing the localization accuracy, hybrid schemes also improve the identifiability of the MT position as we have just seen in the ToA/AoA system where localization is possible with only one BS. Therefore they have

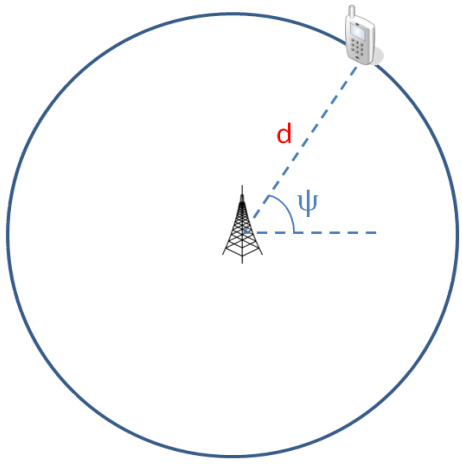

Figure 1.4: Hybrid AoA/ToA-based localization with one BS.

multiple advantages over single-LDP geometrical techniques.

1.2 Fingerprinting Methods

As mentioned before, traditional geometrical techniques perform poorly in challenging environments, such as in dense urban and indoor environments. Moreover they are designed to work in LoS conditions, and LoS condition between all BS-MT links might not be satisfied simultaneously. For example, for a ToA-based localization system in 2D, at least three BSs must be present, and they must all have LoS condition with the MT as we have seen before. Due to this reason, some approaches were proposed to guarantee the algorithm to work properly also in NLoS conditions. Nájar et al. [16] proposed a novel idea. During LoS condition, by estimating the ToA of the LoS and a NLoS path, the time offset (bias) between the two ToAs is calculated. When LoS condition is no longer present, the bias is then subtracted from the ToA of the NLoS path to estimate ToA of the LoS path. In general by adding a Kalman filtering stage, the accuracy of position estimate can be improved. The use of Kalman filter (KF) allows the tracking of the position trajectory, the velocity of the mobile and ToA bias caused by multipaths. In figure 1.5, we illustrate a typical multipath environment.

As can be deduced from the examples, these geometric localization techniques work robust in LoS conditions and try to minimize the effects of the NLoS conditions. However its effect cannot be eliminated completely, lead-

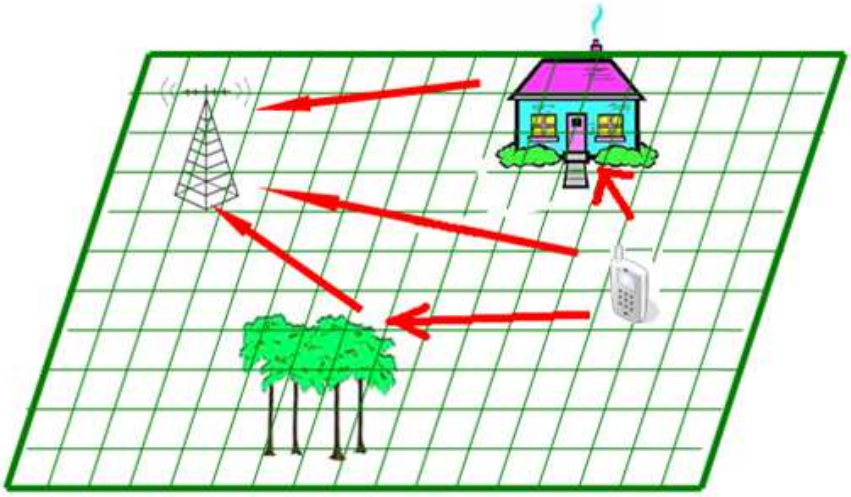

Figure 1.5: A multipath environment.

ing to some irreducible errors in the position estimates for the techniques relying on LoS paths. Therefore new techniques have been developed which try not to eliminate NLoS effects, but instead get advantage of it. Hence fingerprinting techniques were born. The main idea is to store the LDPs or some functions of LDPs, such as channel impulse responses (CIRs) or power delay profiles (PDPs) of the discrete locations in the coverage area of the BS in a database [17–20]. Compared to the distance or angle-related parameters in the "traditional" geometric approaches, PDPs or CIRs can give considerably more information about the MT position. Moreover, simple fingerprinting approaches also exist which store easy to obtain LDPs, such as RSS information in the database and they can also achieve good localization accuracies [21]. Position estimation of the MT is carried out via a correlation or likelihood kind of operation which compares the measured parameter with the entries stored in the database. We will now summarize the basic concepts of general fingerprinting-based localization methods in detail in two steps. First step is creating and maintaining the database. This is the off-line part. Second step is matching the received signal fingerprint to an entry in the database which is performed on-line.

- Creating and maintaining the database:
  - Choose suitable fingerprints which must be unique for every discrete location in the database. There is a trade-off here between

complexity and performance. Selecting a lot of very complicated, detailed fingerprints will end up in a database which will be difficult to maintain. Localization accuracy might be better, but on the other hand matching operation will also last long in such a case. Hence choosing fingerprints with as few parameters as possible and at the same time achieving a good localization accuracy is the target of many researchers.

  - Run (ray tracing or ray launching) simulations or perform measurement campaigns over the area of interest to obtain the unique fingerprints corresponding to a specific position. Then these fingerprints should be stored in a large database. Every time a localization task is carried out, this database will be used as a reference.
  - Update the database periodically or when it is necessary. The propagation environment might not stay the same all the time, especially in dense urban areas. There can be new constructions in the field and many other similar factors which can significantly change the environment. Hence, it is necessary to have an up-to-date database.

- Matching operation:
  - Evaluate the signature of the signal received at the BS and extract the fingerprint out of it.
  - Compare the fingerprint with the ones in the database to find the best match. The matching operation is usually based on maximizing correlation or likelihood between them. The location resulting in the best match is considered as the location estimate.

Fingerprinting techniques are classified among direct location estimation (DLE) techniques, i.e., position estimation is performed in one step. If we remember the "traditional" geometrical techniques, they perform localization in two steps.

Location fingerprinting (LF) technique (introduced by U.S. Wireless Corp. of San Ramon, Calif.) relies on signal structure characteristics [1, 22–24]. By using multipath propagation pattern, LF creates a signature unique to a given location. For LF, it is enough to have only one BS-MT link (multiple BSs are not required) to determine the location of the mobile. Ahonen and Eskelinen suggest using the measured PDPs in the database for fingerprints, because amplitudes and delays of the multipath

components (MPCs) can create a unique position dependent signature [25]. In [26], authors provide deterministic and Bayesian methods for power delay profile fingerprinting (PDP-F) based localization. This work was important for us, since we were inspired from it while developing new fingerprinting algorithms.

For ultra-wideband (UWB) systems employing ToA algorithms, localization accuracy is quite good due to the very high bandwidth utilized in the system [27, 28]. Number of resolvable paths of the multipath channel defined in [29] is quite large in these systems. On the other hand its range is quite limited. Therefore it is mainly used in indoor localization applications. Outdoor localization which is also the scope of our studies, suffers from considerably lower bandwidth used in the system which degrades the resolvability of the MPCs. Therefore, we develop new fingerprinting algorithms exploiting new dimensions which increase the resolvability of paths. They will be introduced in chapter 2.

1.3 Identifiability Issues and Performance Bounds

In this section, we will talk about the identifiability of the MT position, i.e., the ability to estimate it. Identifiability analysis can be divided into two, namely local identifiability and global identifiability analysis. Global identifiability concept is easy to understand. For example, for 2D localization for a ToA-based localization system, three BSs are enough to have a unique position estimate as we have already seen in figure 1.2. However local identifiability of MT position is a bit more tricky. Local identifiability is a similar issue in the sense that the position of the mobile must be uniquely identified around a local neighborhood of the MT. Hence if only signals from two BSs are available (still for the 2D ToA system), the intersection of two circles will result in two possible candidates for the MT position. In this case it is clear that there is no global identifiability. However local identifiability is present. For a TDoA system in 2D, local identifiability cannot be obtained with two BSs as can be inferred from figure 1.3. This is due to the lack of synchronization. When we examine the DO issue for fingerprinting-based localization systems in the upcoming chapters, it will be more clear after analytical derivations. To summarize, no global identifiability in the presence of local identifiability means that there are discrete (not continuous) ambiguities left. For example, with one BS, there is neither local identifiability nor global identifiability for ToA systems. As it is obvious, if there is no local identifiability, it also implies that global identifiability is also not

present. No local identifiability means that there are continuous ambiguities left. This concept is important for us as we will extensively investigate the local identifiability conditions of MT position for various cases later. The reason that we explore identifiability analysis with performance bounds is that they are related to each other via Fisher Information Matrix (FIM).

CRB sets the limit on the covariance matrix of unbiased estimators. For an unbiased estimator $\hat{\mathbf{r}}$ of $\mathbf{r}$, the correlation matrix of the parameter estimation error $\widetilde{\mathbf{r}}$ is bounded below by the inverse of the Fisher Information Matrix (FIM) $\mathbf{J_r}$ as shown below:

$$R_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}} = \mathbb{E}\{(\hat{\mathbf{r}} - \mathbf{r})(\hat{\mathbf{r}} - \mathbf{r})^T\} \geq \mathbf{J_r}^{-1} \tag{1.1}$$

where the FIM is given by:

$$\mathbf{J_r} = \mathbb{E}\Big\{ \left(\frac{\partial \mathcal{LL}}{\partial \mathbf{r}}\right) \left(\frac{\partial \mathcal{LL}}{\partial \mathbf{r}}\right)^T \Big\} \tag{1.2}$$

where $\mathcal{LL}$ is the log-likelihood of the measurement data which depends on the parameter vector $\mathbf{r}$ given by:

$$\mathcal{LL} = \ln\, p\,(\,\mathbf{z}|\,\mathbf{r}\,), \tag{1.3}$$

$p\,(\,\mathbf{z}|\,\mathbf{r}\,)$ being the p.d.f of the measurement data $\mathbf{z}$ conditioned upon the parameter vector $\mathbf{r}$. We will now introduce an important theorem about local identifiability of the unknown parameter vector $\mathbf{r}$:

Theorem 1. *Let* $\mathbf{r}^0$ *be a regular point of* $\mathbf{J_r}(\mathbf{r})$*. Then* $\mathbf{r}^0$ *is locally identifiable if and only if* $\mathbf{J_r}(\mathbf{r}^0)$ *is non-singular [30],*

and a point $\mathbf{r}^0$ is said to be a regular point of the matrix $\mathbf{J_r}$ if there exists an open neighborhood of $\mathbf{r}^0$ in which $\mathbf{J_r}$ has constant rank. This theorem tells us that unknown parameters become identifiable when the FIM evaluated at the true values is nonsingular. This theorem will be used quite frequently in the further chapters.

Chapter 2

Power Delay Doppler Profile Fingerprinting

2.1 Introduction

In this chapter, we are going to introduce two new fingerprinting algorithms, namely frequency-domain PDDP-F and time-domain PDDP-F. By default, all fingerprinting-based localization systems need databases. As we expressed in chapter 1, they can be constructed off-line either by ray tracing or by ray launching simulation methods over the geographical area of interest [31, 32]. The area is divided into several discrete sections, each section having a unique fingerprint. We will begin with the frequency-domain PDDP-F algorithm and after that we will introduce the time-domain PDDP-F algorithm.

2.2 Frequency-Domain PDDP-F

We will start with the frequency-domain PDDP-F algorithm. There are two main steps that we are going to mention soon to understand and apply the frequency-domain PDDP-F. First one is constructing the PDDP from the ray tracing data, and the other one is constructing PDDP from the measurement data.

2.2.1 Obtaining PDDP from Ray Tracing Data

Here in this section, we will first introduce a formulation of PDDP for a general MIMO channel, then simplify the case for the SISO case which we are interested in this chapter. Finally we will explain step by step how to apply the formulation to a given ray tracing data.

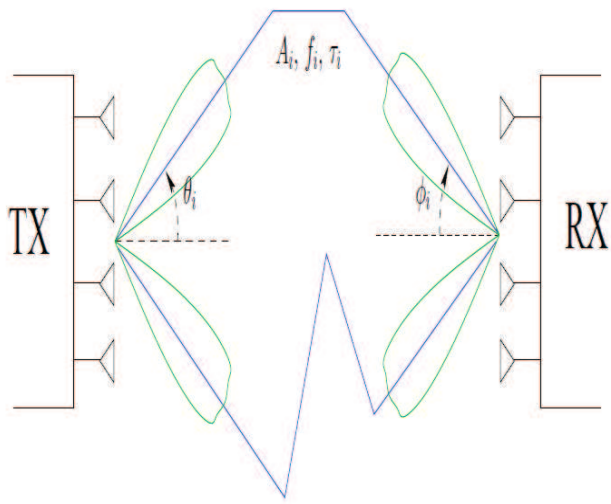

Figure 2.1: MIMO multipath propagation parameters

Consider a specular wireless MIMO channel model with multiple (N_t) transmit and (N_r) receive antennas. The time-varying channel impulse response is:

$$\mathbf{h}(\tau, t) = \sum_{i=1}^{Np} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{a}_R(\phi_i)\, \mathbf{a}_T^T(\theta_i)\, p(\tau - \tau_i) \tag{2.1}$$

where $\mathbf{h}$ is rank 1 in 3 dimensions. The N_p pathwise contributions are characterized by these additional parameters:

- $p(t)$: convolution of the transmit and receive filters (pulse shape)
- f_i: Doppler shift

- $\tau_i(t)$: delay
- $A_i(t)$: complex attenuation coefficient (amplitude and phase of the ray)
- θ_i: angle of departure (AoD)
- ϕ_i: angle of arrival (AoA)
- $\mathbf{a}_R(.)$, $\mathbf{a}_T(.)$: (Rx/Tx side) antenna array response (if only a single antenna is present on one side or the other, then the corresponding $\mathbf{a}(.) = 1$)

We shall assume here 2D propagation, an extension to 3D is immediate. Note: in case the Tx & Rx array responses are unknown, one should instead consider a parameterization of the following form:

$$\mathbf{h}(\tau, t) = \sum_{i=1}^{N_p} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{a}_{R,i}\, \mathbf{a}_{T,i}^T\, p(\tau - \tau_i) \tag{2.2}$$

with $\mathbf{a}_{R,i}$, $\mathbf{a}_{T,i}$ unknown vectors. Note that also the pulse shape may need to be adjusted to measurements.

The channel impulse response in (2.1) results in fact from the propagation channel

$$\begin{array}{l} c(\tau, t, \phi, \theta, v, \phi_v) = \\ \sum_{i=1}^{N_p} A_i(t)\, e^{j2\pi\, f_i\, t}\, \delta(\phi - \phi_i)\, \delta(\theta - \theta_i)\, \delta(\tau - \tau_i) \end{array} \tag{2.3}$$

where we shall assume the channel evolution over a short time period t so that the AoA ϕ, the AoD θ, the path delay τ_i and even the complex path amplitude A_i can be considered as constant.

Any Doppler shift in the propagation channel is actually assumed to be due to the mobility of the mobile terminal (any mobility in the environment would have to be captured by $A_i(t)$). Assume the terminal speed vector to have a magnitude μ and an orientation ϕ_μ (if $\phi_i = \phi_\mu$, then incoming wave and speed vector are aligned, but are evolving in opposite directions). Mobility of the terminal leads to a Doppler shift for path i as follows:

$$f_i = \cos(\phi_i - \phi_\mu)\, \mu\, / \lambda \tag{2.4}$$

where λ is the carrier wavelength. The channel impulse response in (2.1) is the convolution of the propagation channel with the system elements:

$$\mathbf{h}(\tau, t) = c(\tau, t, \phi, \theta, v, \phi_v) * p(\tau) * \mathbf{a}_R(\phi) * \mathbf{a}_T^T(\theta) \ . \tag{2.5}$$

Consider now sampling the impulse response with a sampling period τ_s leading to N_τ samples and then vectorizing it:

$$\underbrace{\underline{\mathbf{h}}(t)}_{N\times 1} = \begin{bmatrix} \underline{\mathbf{h}}(\tau_s, t) \\ \underline{\mathbf{h}}(2\tau_s, t) \\ \vdots \\ \underline{\mathbf{h}}(N_\tau \tau_s, t) \end{bmatrix} = \sum_{i=1}^{N_p} A_i(t)\, e^{j2\pi\, f_i\, t}\, \underline{\mathbf{h}}_i \tag{2.6}$$

where $\underline{\mathbf{h}}(\tau_s, t)$ is the vectorized version of the N_r x N_t channel for the first delay element at time t and

$$\underline{\mathbf{h}}_i = \underline{p}(\tau_i) \otimes \mathbf{a}_T(\theta_i) \otimes \mathbf{a}_R(\phi_i)\ ,\ \ \underline{p}(\tau) = \begin{bmatrix} p(\tau_s - \tau) \\ p(2\tau_s - \tau) \\ \vdots \\ p(N_\tau \tau_s - \tau) \end{bmatrix} \tag{2.7}$$

where $N = N_t N_r N_\tau$ = # TX antennas x # RX antennas x delay spread, and $\otimes$ denotes the Kronecker product: for two matrices A and B, we get the block matrix $A \otimes B = [a_{ij} B]$.

In case of a SISO channel, the sampled CIR is simply:

$$\underline{\mathbf{h}}(t) = \underbrace{\left[\mathbf{p}_{\tau_1} \cdots \mathbf{p}_{\tau_{N_p}}\right]}_{\mathbf{P}_\tau} \underbrace{\begin{bmatrix} A_1(t)\, e^{j2\pi\, f_1\, t} \\ \vdots \\ A_{N_p}(t)\, e^{j2\pi\, f_{N_p}\, t} \end{bmatrix}}_{\mathbf{b}(t)}, \tag{2.8}$$

where $\mathbf{p}_{\tau_i}$ is the complex pulse delayed by τ_i samples. Two possible models can now be considered for the path amplitudes:

- Gaussian model: $A_i(t)$ Gaussian, characterized by a power (variance)
- deterministic model: $A_i(t)$ deterministic unknowns

We shall consider here the Gaussian case (other random models could be considered also, at least for the introduction of the profiles). We are now ready to introduce the Power Delay Doppler Space Profile (PDDSP). At the propagation level we get

$$\begin{aligned} &\mathrm{PDDSP}_c(\tau, f, \phi, \theta, v, \phi_v) \\ &= \int \mathbb{E}\, c(\tau, t_1 + t, ...) c^*(\tau, t_1, ...)\, e^{-j2\pi f t} dt \\ &= \textstyle\sum_{i=1}^{N_p} \sigma_i^2\, \delta(\tau - \tau_i)\, \delta(f - f_i)\, \delta(\phi - \phi_i)\, \delta(\theta - \theta_i)\, . \end{aligned} \tag{2.9}$$

where $\sigma_i^2 = \mathbb{E}|A_i|^2$, and the expectation is at least over the (independent and uniformly distributed) random phases in the A_i, and possibly over the amplitudes also if they are not deterministic. At the channel response level, we get

$$\begin{aligned}
&\mathrm{PDDSP}_h(\tau, f) \\
&= \int \mathbb{E}\, \underline{\mathbf{h}}(\tau, t_1 + t) \underline{\mathbf{h}}^H(\tau, t_1)\, e^{-j2\pi f t} dt \\
&= \textstyle\sum_{i=1}^{N_p} \sigma_i^2\, |p(\tau - \tau_i)|^2\, \delta(f - f_i)\, \mathbf{a}_i \mathbf{a}_i^H \\
&= \textstyle\sum_{i=1}^{N_p} \sigma_i^2\, |p(\tau - \tau_i)|^2\, \delta(f - f_i)\, R_T(\theta_i) \otimes R_R(\phi_i)
\end{aligned} \tag{2.10}$$

where $\mathbf{a}_i = \mathbf{a}_T(\theta_i) \otimes \mathbf{a}_R(\phi_i)$ and we introduced the spatial covariances

$$R_T(\theta_i) = \mathbf{a}_T(\theta_i)\mathbf{a}_T^H(\theta_i)\ , \quad R_R(\phi_i) = \mathbf{a}_R(\phi_i)\mathbf{a}_R^H(\phi_i)\ . \tag{2.11}$$

In the case of a SISO channel, we get the Power Delay Doppler Profile (PDDP)

$$\mathrm{PDDP}_h(\tau, f) = \sum_{i=1}^{N_p} \sigma_i^2\, |p(\tau - \tau_i)|^2\, \delta(f - f_i)\ . \tag{2.12}$$

Now the formulation is complete and we explain how to apply the above formulation to the ray tracing data. The construction of the PDDP is explained step-by-step below:

1. First we create the 2D delay-Doppler profile by only taking the rays into account (pulse shape and windowing effects not included yet). Delay, Doppler and power information of each ray is known. One thing to keep in mind is that delay and Doppler domains should be discretized properly according to the parameters in the measurement data.

   The discretization in delay domain which we call $\Delta\tau$ is fixed, and equal to the sampling duration τ_s like in the measurement data defined before. For the Doppler domain discretization, channel estimations are carried out every t_s seconds for the measurement data, so the highest Doppler frequency + frequency offset (FO) that can be observed is in the range $[-f_s/2, f_s/2]$ where $f_s = 1/t_s$. The FFT length we use in the measurement data is denoted by N_{FFT}, so the Doppler discretization in the measurement data is given by f_s/N_{FFT}. Therefore we use the same for the ray tracing data which we call $\Delta\Omega$.

   The FO, that we mentioned above is due to the difference between the carrier frequencies of the local oscillators of the transmitter and receiver. Normally in a uniform scattering environment, the mean Doppler spread $\bar{\Omega}$ given by:

$$\bar{\Omega} = \frac{\int_{\Psi} \Omega \; D(\Omega) \; d\Omega}{\int_{\Psi} D(\Omega) \; d\Omega} \tag{2.13}$$

would be close to 0, where $D(\Omega)$ is the Doppler spectrum and Ψ denotes the range of possible Doppler shifts. So if there is FO, there will be a nonzero mean Doppler spread $\bar{\Omega}$. Unless the FO is too high, leading to aliasing, it is not a problem for our algorithm, also the DO is not a problem either as we will see soon.

2. Next step is summing up the power of the rays which are in the same grid. Here we directly sum up their individual powers, as we compute the expected average power by the following formula in the grid (we are averaging over the random phases assuming uniform distribution over $[0, 2\pi]$):

$$\mathbb{E}_{\varphi_{g,1} \ldots \varphi_{g,r}} |(A_{g,1} e^{j\varphi_{g,1}} + \cdots + A_{g,r} e^{j\varphi_{g,r}})|^2 = \sum_{i=1}^{r} A_{g,i}^2 \tag{2.14}$$

where r is the number of rays in the grid, $A_{g,i}$ is the magnitude of the ith ray in the grid. It is also easy to see that mean of the rays is equal to 0, where we will use this fact in the time-domain PDDP-F section.

3. Last step is to include the effects of pulse shape in the delay domain, and the windowing in the Doppler domain. This is an easy process. It is enough to make linear convolution in the delay domain of the PDDP matrix with the absolute squared pulse shape.

   And the other operation as we mentioned before is for windowing effect in the measurement data. We have M channel estimates per point. And we choose a certain number of consecutive channel estimates (N_{Window}) among them for computing Fourier Transforms. This process is called windowing, and should be included in the ray tracing PDDP.

   Also it is important to choose a suitable window and a window length N_{Window}. Although Rectangular window has the narrowest main lobe for a given N_{Window}, its side lobes are not negligible leading to spectral leakage [33]. Therefore we decided to choose the Hamming window whose side lobes decay much faster than the Rectangular window. Hamming window is given as:

$$w[k] = 0.54 - 0.46 \cos\left(\frac{2\pi k}{N_{Window} - 1}\right), \qquad 0 \leq k \leq N_{Window} - 1.$$

To include the effect of windowing, it is enough to make cyclic convolution in the Doppler domain of the PDDP matrix with the absolute squared DFT of the window.

It will be more instructive to summarize the whole process with illustrative figures. With the AoA, delay and power information of the rays, we can construct the Power Delay Angle of Arrival Profile (PDAoAP) as shown in figure 2.2. Then the next step is to form the PDDP with the help of the speed vector. With the fusion of information from the AoA of the rays and the information from the speed vector (speed magnitude and direction of motion), Doppler shift of each ray can be computed. Hence PDDP can be constructed as in figure 2.3. The last thing to do is to include the pulse shape and windowing effects. After all, we end up with the final PDDP as shown in figure 2.4.

2.2.2 Obtaining PDDP from Measurement Data and The Fingerprinting Operation

Obtaining the PDDP from the measurement data is easier. N_{Window} consecutive channel estimates are chosen with the Hamming window explained before. Each channel estimate is a vector of length N_τ. For each delay element among N_τ, absolute squared DFT is computed with respect to the time variable to see its variation in time (Doppler information). Consequently this gives the 2D delay-Doppler profile of the measurement data. One thing to pay attention is that all hardware related effects must be removed from the channel estimates e.g. automatic gain control (AGC) and others to make a true comparison.

After computing ray tracing and measurement data PDDPs, next step is to check the similarity between these matrix profiles. There are consecutive channel measurements, obtained in the BS, and the objective is to see which ray tracing data in the database will give the highest match to this measurement. The cost function is defined as the similarity between the matrices. Among the K element database, the one corresponding to the position of the $\hat{k}^{th}$ ray tracing data will be chosen:

$$\hat{k} = arg \max_{k \in [1,K]} J(PDDP_M, PDDP_{RT_k}) \tag{2.15}$$

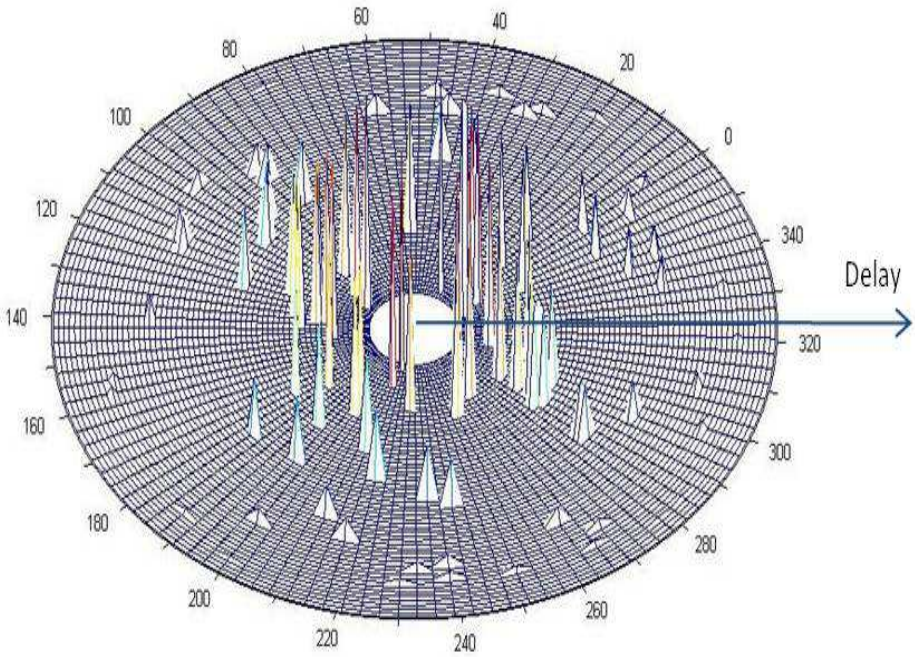

Figure 2.2: Power Delay Angle of Arrival Profile

where J is the likelihood (cost) function, $PDDP_M$ is the PDDP obtained from the measurement data and $PDDP_{RT_k}$ is the PDDP from the kth entry in the ray tracing database. For the likelihood function J, one reasonable candidate is to use the inner product criteria defined for matrices normalized by their norms as below:

$$J(\mathbf{A},\mathbf{B}) = \frac{\mathrm{tr}\,(\mathbf{A}^T\mathbf{B})}{\sqrt{\mathrm{tr}\,(\mathbf{A}^T\mathbf{A})\,\mathrm{tr}\,(\mathbf{B}^T\mathbf{B})}} = \frac{\mathrm{tr}\,(\mathbf{A}^T\mathbf{B})}{\|\mathbf{A}\|_F\,\|\mathbf{B}\|_F}. \tag{2.16}$$

where $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$. But one thing to note is that, before using the above formula, the ray tracing and measurement PDDPs must be perfectly aligned in the delay and Doppler dimensions. Any DO or FO must be handled very precisely. Therefore it is reasonable to choose a computationally effective solution. 2D FFT operation to check the highest correlation between the two matrices normalized by their norms is a very

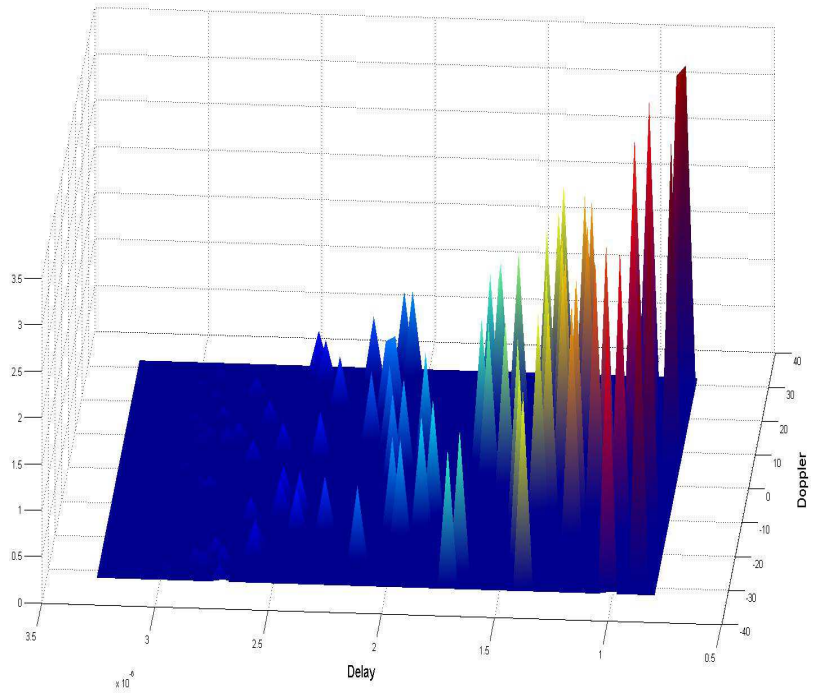

Figure 2.3: Power Delay Doppler Profile

fast operation given by:

$$\frac{IFFT(FFT(\mathbf{A}) \odot conj(FFT(\mathbf{B})))}{\|\mathbf{A}\|_F \, \|\mathbf{B}\|_F} \tag{2.17}$$

where $\odot$ is the Hadamard (element-wise) multiplication and *conj* denotes conjugate. The maximum entry in the resulting matrix is the highest correlation between the two in the perfectly aligned case. Also the position of the maximum entry gives the DO and FO if there is any. Moreover to ensure linear correlation in this operation, proper amount of zero-padding must be introduced in both dimensions.

2.2.3 Simulation Results

Figure 2.5 is the simulation result for the comparison of deterministic PDP-F and frequency-domain PDDP-F over a range of SNR values. In our simu-

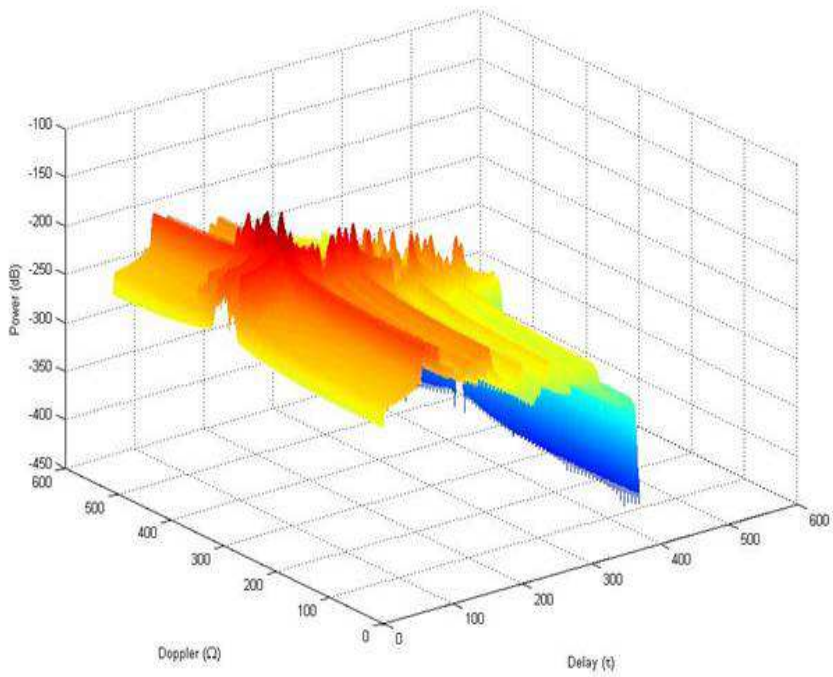

Figure 2.4: Power Delay Doppler Profile with pulse shape and windowing effects included

lation environment, we have created 15 ray tracing points ($K = 15$), and then generated a measurement data from the first ray tracing point by introducing noise. The objective is to see in how many cases the algorithm will match with the first point. The path loss exponent is taken as two, and we generate more than 1000 rays in every iteration to imitate a real RT database as much as possible. There is no spatial relation between these 15 points, we just generate random channel parameters for each of them. For the system parameters, we have a sampling frequency of 9.1429 MHz and a wavelength of 0.4249 m. For the PDDP-F, the size of the 2D FFT is 1024 x 512 when we are computing the correlation. To compute the spectrum, we use an FFT length of 512. The channel estimates are obtained every 4 ms, window length is 200, speed magnitude is chosen randomly between 1-60 km/h and direction angle between 0-359 degrees. For the pulse shape, root-raised-cosine filters are used in the transmitter and in the receiver resulting

in a raised cosine pulse in total with a roll-off factor β of 0.8. Moreover PDDP-F simulations are done in two different ways. In the first case, we assume that speed vector of the mobile is known beforehand and in the second case, we try to estimate the speed vector (its magnitude and direction). As can be expected, the performance of the algorithm when we know the speed vector beforehand is much better than the case when we estimate the speed vector. Estimation of the speed vector is carried out by an exhaustive search method. We know that Doppler spectrum is a function of the speed vector. The speed vector leading to the highest correlation between the RT and measurement spectrums is our estimation. However as we see from the plot, its performance is not that good. The reason might be due to the non-parametric spectrum computation. However as we see in the plot, frequency-domain PDDP-F (when the speed vector is known beforehand) always outperforms the deterministic PDP-F algorithm in all the SNR values. This difference comes from the usage of the additional Doppler dimension which PDP-F cannot benefit from. By deterministic PDP-F, what we mean is that the RT and measurement PDPs are compared via correlation. Computation of the correlation is carried out effectively by using the FFT operation again. However it is now 1D FFT whereas it was 2D FFT for the PDDP-F as we had a 2D profile.

2.3 Time-Domain PDDP-F

As in the frequency-domain version of the algorithm, here we first provide the formulation of the time-domain algorithm, then we explain in detail the application of the algorithm on the ray tracing data. We know that, sampled channel taps might be the superposition of several rays which arrive within the same sampling duration as in a diffuse channel environment:

$$A_l(t) = \sum_{k=1}^{K_l} A_{l,k} e^{j\varphi_{l,k}(t)} \; e^{j2\pi f_{l,k} t}, \tag{2.18}$$

As we explained before, mean of these channel taps is 0, due to averaging over random phases. Also the expected average power is just the summation of the individual powers of the incoming rays. The central limit theorem lets us model these taps as Gaussian random variables. In this section, we propose the time-domain version of the PDDP-F algorithm which exploits the second-order statistics of the channel. We assume that the complex fading vector $\mathbf{b}(t)$, and the additive noise $\mathbf{v}(t)$ are i.i.d. zero-mean Gaussian vector processes, i.e.,

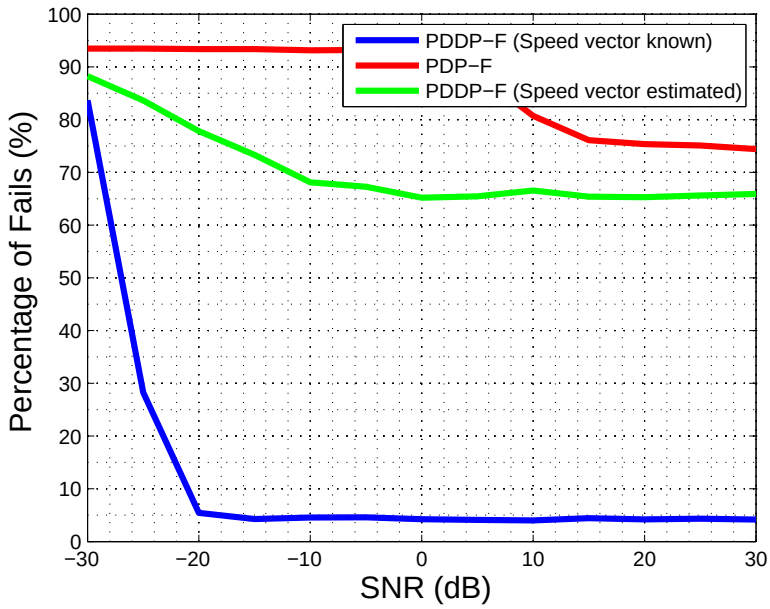

Figure 2.5: Performance comparison of deterministic PDP-F and frequency-domain PDDP-F algorithms.

$$\begin{aligned} \mathbf{b}(t) &\sim \mathcal{N}\left(\mathbf{0}, \mathbf{C_b}\right) \\ \mathbf{v}(t) &\sim \mathcal{N}\left(\mathbf{0}, \sigma_v^2 \mathbf{I}_N\right) \end{aligned} \tag{2.19}$$

where $\mathcal{N}\left(\mathbf{0}, \mathbf{C_b}\right)$ denotes the zero-mean complex Gaussian vector with covariance matrix $\mathbf{C_b}$ (we will soon explain how to derive it), and σ_v^2 is the channel estimation error variance. With the statistical model of (2.19), $\widehat{\mathbf{h}}(t)$ is modeled as an i.i.d. complex Gaussian vector with $\widehat{\mathbf{h}}(t) \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}\right), \mathbf{C}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C_b} \mathbf{P}_\tau{}^H + \sigma_v^2 \mathbf{I}_N$.

With the Gaussian modeling of $\widehat{\mathbf{h}}(t)$, we can propose a Maximum Likelihood solution to the localization problem. Our aim is also to take into account the Doppler variation of the channel. Therefore we stack consecutive $\widehat{\mathbf{h}}(t)$ channel estimates in a vector, instead of taking just one, and compute the covariance matrices based on this. Now consider the channel response at multiple consecutive time instants $t = t_s, 2t_s, \ldots, nt_s$:

$$\underbrace{\underline{\mathbf{h}}}_{nN_\tau N_r N_t \times 1} = \begin{bmatrix} \underline{\mathbf{h}}(t_s) \\ \underline{\mathbf{h}}(2t_s) \\ \vdots \\ \underline{\mathbf{h}}(nt_s) \end{bmatrix} . \tag{2.20}$$

Then we get

$$\underline{\mathbf{h}} = \sum_{i=1}^{N_p} A_i \, \underline{e}(f_i) \otimes \underline{\mathbf{h}}_i \ , \ \ \underline{e}(f) = \begin{bmatrix} e^{j2\pi f t_s} \\ e^{j2\pi f 2t_s} \\ \vdots \\ e^{j2\pi f n t_s} \end{bmatrix} \tag{2.21}$$

We get for the covariance matrix of $\underline{\mathbf{h}}$

$$C_{\underline{\mathbf{hh}}} = \sum_{i=1}^{N_p} \sigma_i^2 \, R_f(f_i) \otimes R_\tau(\tau_i) \otimes R_T(\theta_i) \otimes R_R(\phi_i) \tag{2.22}$$

where

$$R_f(f) = \underline{e}(f)\underline{e}^H(f) \ , \ \ R_\tau(\tau) = \underline{p}(\tau)\underline{p}^H(\tau) \ . \tag{2.23}$$

Note that R_f is Toeplitz. In the case of a SISO channel, we have $C_{\underline{\mathbf{hh}}} = \sum_{i=1}^{N_p} \sigma_i^2 \, R_f(f_i) \otimes R_\tau(\tau_i)$ and the PDDP is related to the diagonal part of this matrix, after taking DFT of the R_f part.

To be more specific for a SISO channel, if there are M channel estimates available, they are divided into $M-n+1$ groups each group having n consecutive channel estimates. For example there can be 2 such groups for $M = 4$ and $n = 3$, i.e. $\widehat{\mathbf{h}}(t_s), \widehat{\mathbf{h}}(2t_s), \widehat{\mathbf{h}}(3t_s)$ for group 1 and $\widehat{\mathbf{h}}(2t_s), \widehat{\mathbf{h}}(3t_s), \widehat{\mathbf{h}}(4t_s)$ for group 2. We stack these vector groups into a longer vector as:

$$\underline{\widehat{\mathbf{h}}}(m) = \begin{bmatrix} \widehat{\mathbf{h}}(m \ t_s) \\ \vdots \\ \widehat{\mathbf{h}}((m+n-1) \ t_s) \end{bmatrix} . \tag{2.24}$$

Now, the Gaussian Log-Likelihood can be constructed with $M-n+1$ vectors as:

$$\mathcal{LL} \propto -(M-n+1) \ln \left(\det \mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}} \right) - \sum_{m=1}^{M-n+1} \underline{\widehat{\mathbf{h}}}(m)^H \mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}{}^{-1} \underline{\widehat{\mathbf{h}}}(m) \tag{2.25}$$

where $\mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}$ is the covariance matrix of $\underline{\widehat{\mathbf{h}}}$. Instead of the usual Maximum Likelihood approaches to estimate the path parameters by maximizing the likelihood with respect to the parameters, the likelihood is evaluated by substituting the position dependent path parameters from the database and hence it provides the likelihood of position. In other words, covariance matrices of the ray tracing database ($\mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}$) are created off-line by the position dependent parameters (using delays, powers, Doppler shifts of the rays), then the likelihood is evaluated with the above formulation for the measurement data. The position giving the highest likelihood is the position estimate of the mobile. (2.25) can be written equivalently as:

$$\mathcal{LL} \propto -\ln\left(\det \mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}\right) - \operatorname{tr}\left\{\mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}{}^{-1}\widehat{\mathbf{C}}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}\right\} \tag{2.26}$$

where $\widehat{\mathbf{C}}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}} = \dfrac{1}{M-n+1}\displaystyle\sum_{m=1}^{M-n+1} \underline{\widehat{\mathbf{h}}}(m)\underline{\widehat{\mathbf{h}}}(m)^H$ is the observed sample covariance matrix which is asymptotically unbiased. (C.11) clearly shows that the log-likelihood is just a covariance matching operation between the measurement covariance matrix and the pre-computed K covariance matrices of the database. One last thing to mention is the derivation of the covariance matrices of the ray tracing database ($\mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}$). We will begin with (2.24) and just present the derivation for $n = 2$. Deriving for any n is straightforward afterwards.

$$\mathbb{E}\left\{\underline{\widehat{\mathbf{h}}}(1)\underline{\widehat{\mathbf{h}}}(1)^H\right\} = \begin{bmatrix} \mathbb{E}\left\{\widehat{\mathbf{h}}_1\widehat{\mathbf{h}}_1^H\right\} & \mathbb{E}\left\{\widehat{\mathbf{h}}_1\widehat{\mathbf{h}}_2^H\right\} \\ \mathbb{E}\left\{\widehat{\mathbf{h}}_2\widehat{\mathbf{h}}_1^H\right\} & \mathbb{E}\left\{\widehat{\mathbf{h}}_2\widehat{\mathbf{h}}_2^H\right\} \end{bmatrix} \tag{2.27}$$

where $\mathbb{E}\left\{\widehat{\mathbf{h}}_1\widehat{\mathbf{h}}_1^H\right\} = \mathbb{E}\left\{\widehat{\mathbf{h}}_2\widehat{\mathbf{h}}_2^H\right\} = \mathbf{C}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}} = \mathbf{P}_\tau\mathbf{C}_\mathbf{b}\mathbf{P}_\tau^H + \sigma_v^2\mathbf{I}_N$, and $\mathbf{C}_\mathbf{b}$ is a diagonal matrix given as:

$$\mathbf{C}_\mathbf{b} = \begin{bmatrix} \sum_{k=1}^{K_1} A_{1,k}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_{k=1}^{K_L} A_{L,k}^2 \end{bmatrix}, \tag{2.28}$$

where we used (2.18) to derive $\mathbf{C}_\mathbf{b}$ ($A_{l,k}$'s are the magnitudes of the rays). As can be seen from the formulation, there is no Doppler information yet because the calculated covariance is for the same time instant. The idea of stacking consecutive channel estimates brings the Doppler information which will be clear while deriving $\mathbb{E}\left\{\widehat{\mathbf{h}}_1\widehat{\mathbf{h}}_2^H\right\}$. We will not make the derivation for

$\mathbb{E}\left\{\widehat{\mathbf{h}}_2\widehat{\mathbf{h}}_1^H\right\}$ because it is just the transpose-conjugate of the other. $\mathbb{E}\left\{\widehat{\mathbf{h}}_1\widehat{\mathbf{h}}_2^H\right\}$ $= \left(\mathbb{E}\left\{\widehat{\mathbf{h}}_2\widehat{\mathbf{h}}_1^H\right\}\right)^H = \mathbf{P}_\tau\mathbf{C}_\mathbf{d}\mathbf{P}_\tau^H$, where $\mathbf{C}_\mathbf{d}$ is derived as:

$$\begin{bmatrix} \sum_{k=1}^{K_1} A_{1,k}^2\, e^{-j2\pi f_{1,k}t_s} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sum_{k=1}^{K_L} A_{L,k}^2\, e^{-j2\pi f_{L,k}t_s} \end{bmatrix}$$

by using (2.18) again. Now, the Doppler contributions of each ray is visible in the covariance matrix. As can be seen, the overall covariance matrix is a function of delays, powers and Doppler shifts of rays. We aim to increase the localization accuracy by incorporating this additional Doppler information.

2.3.1 Simulation Results

Figure 2.6 is the simulation result for the comparison of Bayesian PDP-F described in [26] and the time-domain PDDP-F over a range of SNR values for different n. The simulation environment is the same as for the frequency-domain PDDP-F simulations. We see three curves in the plot where $n = 1$ corresponds to the Bayesian PDP-F case. It is obvious that time-domain PDDP-F outperforms Bayesian PDP-F. Increasing n (number of consecutive channel estimates) also increases the success rate. If we also compare with the frequency-domain PDDP-F algorithm, we see that the time-domain PDDP-F is more robust and success rate is higher for $n \geq 3$. Also one drawback of the frequency-domain PDDP-F is that its non-parametric spectrum might suffer from limited resolution. In these simulations, we assumed that the speed vector is known beforehand.

As was the case in the frequency-domain PDDP-F, simulations for the case when the speed vector is estimated jointly was carried out for $n = 2$. However to reduce the simulation time, we had to reduce the ray tracing points ($K = 5$ now). Apart from that, simulation environment and the system parameters are the same. Joint speed vector estimation was carried out by maximizing the likelihood with respect to the speed vector in every ray tracing point. Hence it is again an exhaustive search over the possible values of the speed magnitude and the direction of motion. The results are demonstrated in figure 2.7. As expected, there is a loss of performance when compared to the case of known speed vector.

The time-domain PDDP-F approach can be seen as an elegant method for localization. Instead of trying to match only the PDPs (diagonal elements of the covariance matrices for the same time instant), the whole covariance

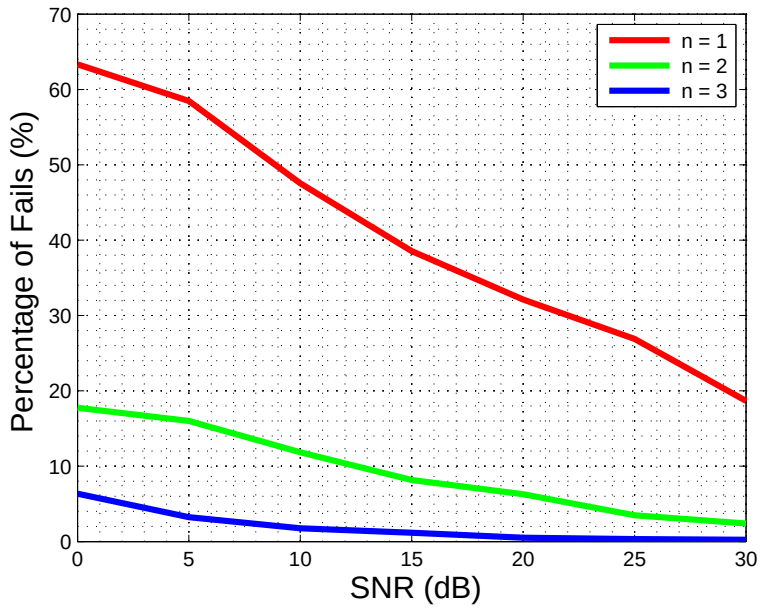

Figure 2.6: Performance results of time-domain PD(D)P-F as a function of n.

matrices are compared, also by taking into account the Doppler information. However, there are some important things to note. First of all, DO and FO problems are also present here. So care must be taken for them. The advantage of the time-domain PDDP-F over its frequency-domain version lies in the weighting matrix (inverse of the covariance matrix) used. If we consider the noiseless case, the weighting matrix tries to balance the weak and strong rays. In other words, strong rays (rays with high power) are weighted by small coefficients whereas weak rays are weighted by higher coefficients. As a result, this gives the advantage of observing even very little details which would be mostly ignored in the frequency-domain case. However if we consider noise, this could also lead to noise amplification when the path contribution is below the noise level.

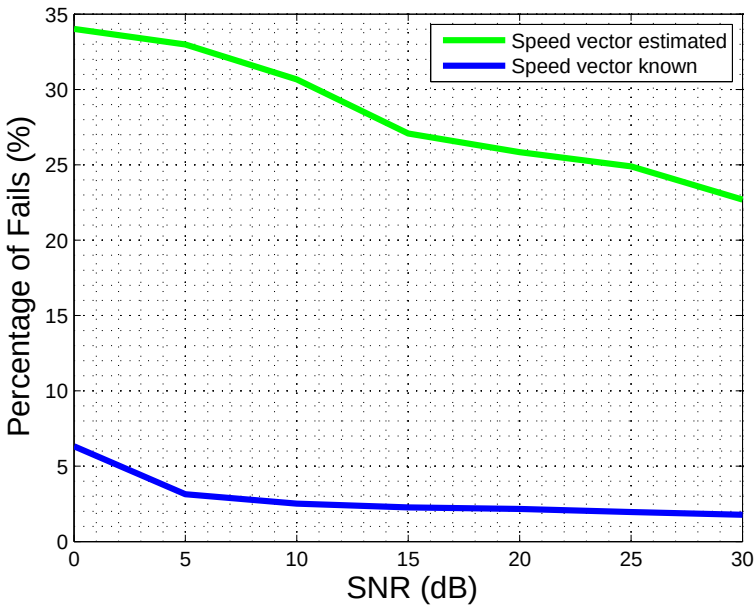

Figure 2.7: Time domain PDDP-F for $n = 2$ with and without speed vector estimation.

2.4 Conclusion

In this chapter, two new fingerprinting methods have been introduced utilizing the time variation of the channel due to MT mobility. Actually we introduced the most general profile, namely PDDSP, of which PDDP is a special case (for SISO channels). First method, frequency-domain PDDP-F resolves the paths not only in delay dimension but also in Doppler dimension. Naturally, with this 2D profile instead of the classical 1D profiles of PDP-F, higher accuracies were achieved. Also the cost function implemented via 2D FFT operations is considerably fast and also increases the robustness of the algorithm against FO and DO. For the systems having small bandwidths, resolvability of rays might be a problem which can cause low localization accuracies. Hence, as we proposed in this chapter, introducing a new dimension to increase the resolvability of rays resulted in significant increase in localization performance. Therefore the Doppler dimension has the effect of increasing the accuracy, and also the identifiability. Having $N_p \geq 2$ paths is sufficient for the identifiability unless they arrive at the same delay-

Doppler grid. For a channel with delay spread N_τ, the probability that two paths arrive in the same sampling duration is $1/N_\tau$, however the probability that both are located in the same delay-Doppler grid is much lower than this for the frequency-domain PDDP-F algorithm. Hence resolution is increased considerably by exploiting the Doppler information.

Second algorithm, which we call time-domain PDDP-F, is based on exploiting the SOS of the time varying channel. It is another high resolution localization technique. The novelty that we proposed was based on stacking consecutive channel estimates. By this way, Doppler information of the paths have become visible in the covariance matrices. We have experienced the performance increase with the increasing n, n being the consecutive channel estimates stacked. Even for $n = 2$, the performance was quite satisfying. However, in this case, FO and/or DO problems (if they are present), must be carefully handled for the method to work properly.

Chapter 3

Performance Analysis of Power Delay Profile Fingerprinting Methods

3.1 Introduction

This chapter deals with performance analysis of PDP-F algorithms. Various path amplitude models, overlapping and non-overlapping pulses, DO issues, impact of network geometry on the localization performance are investigated. Local identifiability and CRB analysis are carried out for each of these cases. Although we have introduced channel model in chapter 2, we present it once more for the completeness of the text. However, we will try to avoid repeating obvious material.

3.2 CRB Analysis of PDP-F

We start by introducing the channel model. As one can expect, PDP is also obtained from the CIR by averaging over phases and also over path amplitudes if they are random. The time varying CIR between the BS and MT can be written as:

$$h(t,\tau) = \sum_{i=1}^{N_p} A_i(t)\ p(\tau - \tau_i(t)) \tag{3.1}$$

After sampling the CIR with a sampling period of τ_s leading to N_τ samples and stacking them in a vector, we obtain:

$$\mathbf{h}(t) = \begin{bmatrix} h(\tau_s, t) \\ h(2\tau_s, t) \\ \vdots \\ h(N_\tau \tau_s, t) \end{bmatrix} = \sum_{i=1}^{N_p} A_i(t)\ \mathbf{p}_{\tau_i} \tag{3.2}$$

where $\mathbf{p}_\tau$ is defined as:

$$\mathbf{p}_\tau = \begin{bmatrix} p(\tau_s - \tau) \\ p(2\tau_s - \tau) \\ \vdots \\ p(N_\tau \tau_s - \tau) \end{bmatrix} \tag{3.3}$$

which is the sampled complex pulse shape vector having a delay equal to the delay of the path in samples and has N nonzero samples. If we write (3.2) in matrix notation and include the channel estimation noise, we obtain the estimated CIR vector as:

$$\hat{\mathbf{h}}(t) = \underbrace{\begin{bmatrix} \mathbf{p}_{\tau_1} \cdots \mathbf{p}_{\tau_{N_p}} \end{bmatrix}}_{\mathbf{P}_\tau} \underbrace{\begin{bmatrix} A_1(t) \\ \vdots \\ A_{N_p}(t) \end{bmatrix}}_{\mathbf{b}(t)} + \mathbf{v}(t), \tag{3.4}$$

As we have stated in chapter 2, two possible models can be considered for the path amplitudes:

- Gaussian model: $A_i(t)$ Gaussian with zero mean, characterized by a power (variance) i.e., $var(A_i) = \sigma_i^2$, which corresponds to Rayleigh fading case (magnitude has Rayleigh distribution)
- deterministic model: $A_i(t)$ deterministic unknowns

We will investigate both of these cases, derive the CRBs and see the conditions of local identifiability of the position vector. The method to be used is the GML based PDP-F for the Rayleigh fading case. To see the details of GML based PDP-F, readers can refer to [26].

We will derive the FIM of LDPs and the position vector $\mathbf{r}$ as well for local identifiability analysis and to obtain CRBs. As we have stated in chapter 1, inverse of the FIM determines the lower bound of the estimation error variance of unbiased estimators which is known as the CRB.

3.2.1 Rayleigh Fading Case

Let θ represent the vector of LDPs for a channel with N_p paths. If we consider the delays and the variances of the complex path amplitudes as LDPs, θ is defined as:

$$\theta = \left[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ \sigma_1^2,\ \sigma_2^2,\ \cdots\ \sigma_{N_p}^2\right]^T \tag{3.5}$$

where τ_i and σ_i^2 represent the delay and the amplitude variance of the ith path respectively. The log-likelihood of the data vector for complex white Gaussian noise is given as:

$$\mathcal{LL} \propto -\ln\left(\det\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}\right)\right) - \left(\hat{\mathbf{h}} - \mu\right)^H \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\left(\hat{\mathbf{h}} - \mu\right) \tag{3.6}$$

From (3.6), the elements of the FIM $\mathbf{J}_\theta$ for a general complex Gaussian scenario is given by [34]

$$[\mathbf{J}_\theta]_{ij} = \operatorname{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_{\mathbf{i}}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_{\mathbf{j}}}\right) + 2\Re\left(\left[\frac{\partial \mu}{\partial \theta_{\mathbf{i}}}\right]^H \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\left[\frac{\partial \mu}{\partial \theta_{\mathbf{j}}}\right]\right). \tag{3.7}$$

Note that we are computing the FIM in the true position. The covariance matrix and the mean vector which were computed off-line according to the parameters of a database entry (each entry in the database corresponds to a different position with unique parameters such as path delays, amplitudes, etc.) belong to the same position of the measured channel estimates. Moreover, also one of our interests is the local identifiability of the position vector $\mathbf{r} = [x, y]$ which denotes the coordinates of the mobile position. Hence there will be a FIM transformation of parameters from θ to $\mathbf{r}$. This transformation can be obtained by the following formula [35]:

$$\mathbf{J}_{\mathbf{r}} = \mathbf{F}\mathbf{J}_\theta\mathbf{F}^H \tag{3.8}$$

where $\mathbf{F} = \frac{\partial \theta}{\partial \mathbf{r}}\Big|_{\mathbf{r}=\mathbf{r}_0}$ ($\mathbf{r}_0 = [x_0, y_0]^T$ being the true position of the mobile) is a $2 \times 2N_p$ matrix which in this case defined as:

$$\mathbf{F} = \left[\begin{array}{cccccc} \frac{\partial \tau_1}{\partial x} & \cdots & \frac{\partial \tau_{N_p}}{\partial x} & \frac{\partial \sigma_1^2}{\partial x} & \cdots & \frac{\partial \sigma_{N_p}^2}{\partial x} \\ \frac{\partial \tau_1}{\partial y} & \cdots & \frac{\partial \tau_{N_p}}{\partial y} & \frac{\partial \sigma_1^2}{\partial y} & \cdots & \frac{\partial \sigma_{N_p}^2}{\partial y} \end{array} \right] \Bigg|_{x=x_0, y=y_0} \tag{3.9}$$

Note that if LDP vector θ is defined differently then there will be a different $\mathbf{F}$ matrix. We will now make some comments on the parameters σ_i^2 and also on the matrix $\mathbf{F}$.

Remark 1. *For the path amplitude variances, they are mostly modeled by distance dependent attenuation which is accompanied by a path-loss coefficient (isotropic model). In that case* $\sigma_i^2 = \frac{k}{\tau_i^\gamma}$ *where* k *is a positive constant depending on the propagation speed of the wave, antenna gains, etc and* γ *is the path-loss coefficient* ($\gamma \geq 2$)*. In such a condition,* σ_i^2 *is just a function of* τ_i*. So only* τ_i *carries position-dependent information. On the other hand we can consider* σ_i^2 *itself as a position-dependent parameter (anisotropic model). For example in a given position it might be a function of the surrounding geography which will cause reflections, refractions and so on. It is obvious that in that case each path will carry two distinct information about position instead of one. For an isotropic model, the variation in path amplitudes becomes smaller in magnitude with the increasing distance due to the exponential decay factor. This can be seen immediately with a simple example. Suppose for a path-loss coefficient of 2* ($\gamma = 2$)*,* $\sigma_i^2(d_i) = \frac{g}{d_i^2}$ *where* g *is a constant and* d_i *denotes the path length. If we make a first order Taylor series expansion around* d_i*, we get* $\sigma_i^2(d_i + \Delta d) \approx \frac{g}{d_i^2} - \frac{2g\Delta d}{d_i^3}$*. So* $\frac{\partial \sigma_i^2}{\partial d} \approx \frac{\sigma_i^2(d_i+\Delta d) - \sigma_i^2(d_i)}{\Delta d} = \frac{-2g}{d_i^3}$*. As it is clear the change in the path amplitude is inversely proportional to the path length with an exponential factor. Consequently we can say that longer paths have less variation in path amplitudes with respect to a position change than shorter paths for the isotropic model, i.e.,* $\left|\frac{\partial \sigma_i^2}{\partial d}\right| < \left|\frac{\partial \sigma_j^2}{\partial d}\right|$ *or* $\frac{\partial \sigma_i^2}{\partial d} > \frac{\partial \sigma_j^2}{\partial d}$ *due to the negative sign for* $i > j$*. In fact with high probability, this inequality still holds for the anisotropic model as well. Also note that by chain rule, we have* $\frac{d\sigma_i^2}{dx} = \frac{d\sigma_i^2}{d\tau_i}\frac{d\tau_i}{dx} = \eta_i \frac{d\tau_i}{dx}$ *where* $\eta_i = -k\gamma\tau_i^{-(\gamma+1)}$ *for the isotropic model. We can say that* $\mathbf{F}$ *is a generic matrix. Hence it is full rank (rank two) with probability* 1 *for the anisotropic case. For the isotropic modeling,* $rank(\mathbf{F}) = \min(2, N_p)$*.*

Proposition 1. *From theorem 1, for local identifiability of* $\mathbf{r}$*,* $\mathbf{J_r}$ *must be full rank (rank 2). For* $\mathbf{J_r}$ *to be full rank, it is required that* $\mathbf{J}_\theta$ *has at least rank 2. Therefore it is never possible to achieve the local identifiability of* $\mathbf{r}$ *for the isotropic case with* $N_p = 1$*.*

3.2.1.1 Anisotropic Path Amplitude Variances

We see that channel estimates have zero mean because $\mathbb{E}\mathbf{b}(t) = \mathbf{0}$ also the noise vector has zero mean. Hence the second term in (3.7) vanishes. The covariance matrix of the channel estimates $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$ can be easily obtained from (5.3) and it is given by $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C}_b \mathbf{P}_\tau^H + \sigma_v^2\, \mathbf{I}$, σ_v^2 being the channel estimation error variance. $\mathbf{C}_b$ is a diagonal matrix with entries $[\sigma_1^2, \sigma_2^2, \cdots, \sigma_{N_p}^2]$. The diagonal structure of $\mathbf{C}_b$ comes from the uncorrelated scattering assumption of the paths. So for the GML technique with Rayleigh fading, the FIM is:

$$[\mathbf{J}_\theta]_{ij} = \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_\mathbf{i}} \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_\mathbf{j}}\right). \tag{3.10}$$

We will investigate a special case of the CIR which will make the derivation of the FIM easier. We will assume that pulse contributions corresponding to different path delays do not overlap with each other. Moreover, this assumption will be valid throughout the whole chapter unless otherwise stated. In other words, pulses for different path delays are orthogonal to each other. This makes the pulse matrix $\mathbf{P}_\tau$ an orthogonal matrix (orthogonal columns), i.e. $\mathbf{P}_\tau^H \mathbf{P}_\tau = e_p \mathbf{I}$ where $e_p = \|\mathbf{p}(\tau)\|^2$ is the pulse energy. After this assumption, we can derive the elements of the FIM by using (3.10). We can explicitly obtain the inverse of the covariance matrix by using Woodbury's matrix identity. By exploiting the orthogonality of the pulse matrix $\mathbf{P}_\tau$, the inverse covariance matrix is obtained as:

$$\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}{}^{-1} = \sigma_v^{-2}\mathbf{I} - \sigma_v^{-2} \sum_{i=1}^{N_p} \frac{\sigma_i^2}{e_p\, \sigma_i^2 + \sigma_v^2} \mathbf{p}_{\tau_i} \mathbf{p}_{\tau_i}^H. \tag{3.11}$$

For the preparation of the computation of the FIM entries, we first compute the partial derivatives of the covariance matrix with respect to the parameters as follows:

$$\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2} = \mathbf{p}_{\tau_i} \mathbf{p}_{\tau_i}^H, \tag{3.12}$$

$$\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_i} = -\sigma_i^2 \left(\mathbf{p}'_{\tau_i} \mathbf{p}_{\tau_i}^H + \mathbf{p}_{\tau_i} \mathbf{p}'^H_{\tau_i}\right) \tag{3.13}$$

where

$$\mathbf{p}'_{\tau} = \begin{bmatrix} p'(\tau_s - \tau) \\ p'(2\tau_s - \tau) \\ \vdots \\ p'(N_\tau \tau_s - \tau) \end{bmatrix} \tag{3.14}$$

and $p'(n\tau_s - \tau)$ being defined as:

$$p'(n\tau_s - \tau) = \left.\frac{dp(t)}{dt}\right|_{t=n\tau_s-\tau}. \tag{3.15}$$

From these partial derivatives, and by using (3.10), (3.11) and the assumption that $\mathbf{p}_{\tau_i}^H \mathbf{p}_{\tau_j} = \delta_{ij} e_p$, we obtain the FIM entries:

$$\begin{aligned} \mathbf{J}_{\tau_i,\tau_i} &= \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{i}}} \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{i}}}\right) && (3.16) \\ &= \sigma_i^4 \sigma_v^{-4} \text{tr}\left((\mathbf{B_i} - c_i \mathbf{C_i})(\mathbf{B_i} - c_i \mathbf{C_i})\right) && (3.17) \\ &= \sigma_i^4 \sigma_v^{-4} \left(\text{tr}\left(\mathbf{B_i}\mathbf{B_i}\right) + c_i^2 \text{ tr}\left(\mathbf{C_i}\mathbf{C_i}\right) - 2c_i \text{ tr}\left(\mathbf{B_i}\mathbf{C_i}\right)\right) && (3.18) \end{aligned}$$

where

$$\begin{aligned} \mathbf{B_i} &= \left(\mathbf{p}'_{\tau_i} \mathbf{p}_{\tau_i}^H + \mathbf{p}_{\tau_i} \mathbf{p}'^H_{\tau_i}\right) && (3.19) \\ \mathbf{C_i} &= \left(\alpha \mathbf{p}_{\tau_i} \mathbf{p}_{\tau_i}^H + e_p \mathbf{p}_{\tau_i} \mathbf{p}'^H_{\tau_i}\right) && (3.20) \\ \alpha &= \mathbf{p}_{\tau_i}^H \mathbf{p}'_{\tau_i} = a + jb && (3.21) \\ c_i &= \frac{\sigma_i^2}{e_p \sigma_i^2 + \sigma_v^2}. && (3.22) \end{aligned}$$

Remark 2. *After a careful inspection, we realize that $a = 0$. To show that, let us write $e_p = \mathbf{p}_\tau^H \mathbf{p}_\tau$. If we take its derivative with respect to τ we obtain:*

$$\frac{\partial e_p}{\partial \tau} = -\left(\mathbf{p}'^H_\tau \mathbf{p}_\tau + \mathbf{p}_\tau^H \mathbf{p}'_\tau\right) = -(\alpha^* + \alpha) = -2a = 0. \tag{3.23}$$

Hence we see that $a = 0$, and $\alpha = jb$. This is the most general case. Obviously if the pulse is real, then b also becomes 0 which results in $\alpha = 0$. Another thing that we should emphasize is that if the pulse is symmetric (for complex or real pulse shape) around its center, this also results in $\alpha = 0$. This is because in this case, the derivative of the pulse is antisymmetric (odd) around its center which makes this inner product 0. To summarize, if the pulse is real or symmetric around its center α becomes 0, otherwise $\alpha = jb$.

If we turn back to the calculation of $\mathbf{J}_{\tau_i,\tau_i}$, after doing the algebra, we obtain the result as follows:

$$\begin{aligned}
\mathbf{J}_{\tau_i,\tau_i} &= \sigma_i^4\sigma_v^{-4}\left(-2b^2+2e_pe_d+2c_ib^2e_p-2c_ie_p^2e_d\right) && (3.24)\\
&= 2\,\sigma_i^4\sigma_v^{-4}(1-e_p\,c_i)(e_p\,e_d-b^2) && (3.25)\\
&= \frac{2\,\sigma_i^4\sigma_v^{-2}(e_p\,e_d-b^2)}{e_p\,\sigma_i^2+\sigma_v^2} && (3.26)
\end{aligned}$$

where $\mathbf{p}'^H_{\tau_i}\mathbf{p}'_{\tau_i}=e_d$.

Remark 3. *The variable e_d that was just defined above gives information about the effective bandwidth of the pulse shape. To see that, let us define the mean square bandwidth of the pulse shape $\overline{F^2}$ which is computed as:*

$$\overline{F^2}=\frac{\int_{-\infty}^{\infty}f^2|P(f)|^2df}{\int_{-\infty}^{\infty}|P(f)|^2df}, \tag{3.27}$$

where $P(f)$ is the Fourier transform of the pulse. By using Parseval's theorem, the following equality is obtained:

$$4\pi^2\overline{F^2}=\frac{\int_0^T|p'(t)|^2dt}{\int_0^T|p(t)|^2dt}=\frac{\int_{-\infty}^{\infty}4\pi^2f^2|P(f)|^2df}{\int_{-\infty}^{\infty}|P(f)|^2df}=\frac{e_d}{e_p}, \tag{3.28}$$

T being the pulse duration and assuming that the sampling frequency satisfies the Nyquist criteria. Therefore we can define W, the effective bandwidth of the pulse as the square root of $\overline{F^2}$, i.e., $W=\sqrt{\overline{F^2}}=\frac{\sqrt{e_d/e_p}}{2\pi}$.

By using the same methodology we continue to calculate the rest of the FIM.

$$\begin{aligned}
\mathbf{J}_{\sigma_i^2,\sigma_i^2} &= \mathrm{tr}\left(\mathbf{C}^{-1}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}\frac{\partial\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial\sigma_\mathbf{i}^2}\mathbf{C}^{-1}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}\frac{\partial\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial\sigma_\mathbf{i}^2}\right) && (3.29)\\
&= \sigma_v^{-4}\left(1-e_pc_i\right)^2\ \mathrm{tr}\left(\mathbf{p}_{\tau_i}\mathbf{p}^H_{\tau_i}\mathbf{p}_{\tau_i}\mathbf{p}^H_{\tau_i}\right) && (3.30)\\
&= \sigma_v^{-4}e_p^2\left(1-e_pc_i\right)^2 && (3.31)\\
&= \left(\frac{e_p}{e_p\sigma_i^2+\sigma_v^2}\right)^2, && (3.32)
\end{aligned}$$

and

$$\mathbf{J}_{\sigma_i^2,\tau_i} = \mathbf{J}_{\tau_i,\sigma_i^2} = \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{i}}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_{\mathbf{i}}^{\mathbf{2}}}\right) \tag{3.33}$$

$$= -2ae_p\sigma_i^2\sigma_v^{-4}\left(1-e_pc_i\right)^2 = 0, \tag{3.34}$$

since $a = 0$. This is an important result. Estimate of the path delay is uncorrelated with the estimate of the path amplitude variance. For $N_p > 1$, we have the cross terms of the FIM for different paths, e.g. $\mathbf{J}_{\tau_{\mathbf{i}},\tau_{\mathbf{j}}}$ or $\mathbf{J}_{\sigma_{\mathbf{i}}^{\mathbf{2}},\tau_{\mathbf{j}}}$ and etc. We still have the non-overlapping pulse assumption. Under this assumption, $\mathbf{J}_{\tau_{\mathbf{i}},\tau_{\mathbf{j}}}$ (for $i \neq j$) is given by:

$$\mathbf{J}_{\tau_i,\tau_j} = \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{i}}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{j}}}\right) \tag{3.35}$$

$$= \sigma_i^2\sigma_j^2\sigma_v^{-4}\,\text{tr}\left((\mathbf{B_i}-c_i\mathbf{C_i})(\mathbf{B_j}-c_j\mathbf{C_j})\right) \tag{3.36}$$

$$= \sigma_i^2\sigma_j^2\sigma_v^{-4}\left(\text{tr}\left(\mathbf{B_iB_j}\right)+c_ic_j\ \text{tr}\left(\mathbf{C_iC_j}\right)-c_i\ \text{tr}\left(\mathbf{C_iB_j}\right)-c_j\ \text{tr}\left(\mathbf{B_iC_j}\right)\right).$$

We observe that overall summation is composed of cross product terms such as $\mathbf{B_iB_j}$ or $\mathbf{B_iC_j}$, etc. These cross products all happen to be the $\mathbf{0}$ matrix. The proof is very simple and comes from the orthogonality of the pulses of different paths. The matrices $\mathbf{B_i}$ and $\mathbf{C_i}$ have N^2 nonzero terms located in the indices $[\tau_i : \tau_i + N - 1, \tau_i : \tau_i + N - 1]$, while the matrices $\mathbf{B_j}$ and $\mathbf{C_j}$ have N^2 nonzero terms located in the indices $[\tau_j : \tau_j + N - 1, \tau_j : \tau_j + N - 1]$. Obviously, these cross products end up in the $\mathbf{0}$ matrix. Therefore $\mathbf{J}_{\tau_{\mathbf{i}},\tau_{\mathbf{j}}} = 0$. With the same reasoning $\mathbf{J}_{\sigma_{\mathbf{i}}^{\mathbf{2}},\sigma_{\mathbf{j}}^{\mathbf{2}}}$ and $\mathbf{J}_{\tau_{\mathbf{i}},\sigma_{\mathbf{j}}^{\mathbf{2}}}$ are all equal to 0 for $i \neq j$.
After having completely derived the FIM for the LDP vector, we can check the conditions for the rank to be at least two to achieve the local identifiability of $\mathbf{r}$. We will first investigate the case when $N_p = 1$. In this case FIM has the following structure:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & 0 \\ 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} \end{bmatrix} \tag{3.37}$$

Obviously to achieve a rank of two, the diagonals of the matrix must be nonzero. As can be seen from (3.32), $\mathbf{J}_{\sigma_1^2,\sigma_1^2}$ is always positive. For $\mathbf{J}_{\tau_1,\tau_1}$, the following condition must hold:

$$e_p\ e_d \neq b^2. \tag{3.38}$$

We can also state in the following form:

$$|\mathbf{p}_\tau^H \mathbf{p}'_\tau|^2 \neq \|\mathbf{p}_\tau\|^2 \, \|\mathbf{p}'_\tau\|^2. \tag{3.39}$$

Note that we have not used τ_1, but instead we just used τ because the statement is independent of the delay. What we observe is that local identifiability of $\mathbf{r}$ depends on the pulse shape and its derivative for $N_p = 1$. By using the Cauchy-Schwarz inequality we have:

$$\|\mathbf{p}_\tau\|^2 \, \|\mathbf{p}'_\tau\|^2 \geq |\mathbf{p}_\tau^H \mathbf{p}'_\tau|^2. \tag{3.40}$$

So unless one vector is a scalar multiple of the other vector (pulse shape and its derivative), the equality never holds resulting in a rank two matrix (full rank in this case). This is an important result because local identifiability of $\mathbf{r}$ can be achieved with only one path. Another thing to emphasize is that if the pulse is real or symmetric, then α and consequently b becomes zero. In this case $\mathbf{J}_{\tau_1,\tau_1}$ is always nonzero. Hence local identifiability of $\mathbf{r}$ is achieved without any constraints in this case. We can easily extend the investigation for $N_p > 1$. Moreover it is also possible to extract the CRBs for the estimation of the elements of the LDP vector θ. For that purpose diagonal entries of the inverse of $\mathbf{J}_\theta$ must be computed. For $N_p > 1$ the FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \mathbf{J}_{\tau_{N_p},\tau_{N_p}} & \vdots & \vdots & \vdots \\ \vdots & \cdots & \cdots & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & \vdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \mathbf{J}_{\sigma_{N_p}^2,\sigma_{N_p}^2} \end{bmatrix}.$$

Note that $\mathbf{J}_\theta$ is a $2N_p \times 2N_p$ diagonal matrix. Therefore computing the CRBs is quite easy and given by:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}_{\tau_i,\tau_i}} = \frac{\sigma_v^2 \, (e_p \, \sigma_i^2 + \sigma_v^2)}{2 \, \sigma_i^4 (e_p \, e_d - b^2)}. \tag{3.41}$$

A first comment on this result is that CRB($\hat{\tau}_i$) decreases with the increasing path power (σ_i^2) and as expected increases with the increasing noise power. Most importantly CRB($\hat{\tau}_i$) decreases with the increasing effective bandwidth of the pulse shape as expected. For a real (or a symmetric) pulse shape, we can equivalently express CRB($\hat{\tau}_i$) as following:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{8\pi^2 W^2 SNR_i}\left(1 + \frac{1}{SNR_i}\right) \tag{3.42}$$

where $SNR_i = \frac{e_p \sigma_i^2}{\sigma_v^2}$ is the signal to noise ratio (SNR) of the ith path. In the figures below, we plot the RMSE (root mean square error) of $\hat{\tau}_i$ which is the square root of CRB($\hat{\tau}_i$) as a function of SNR_i for different bandwidth values. Also we normalize the pulse energy e_p to 1 and use a real pulse shape ($\alpha = 0$). Noise variance σ_v^2 is set to 1.

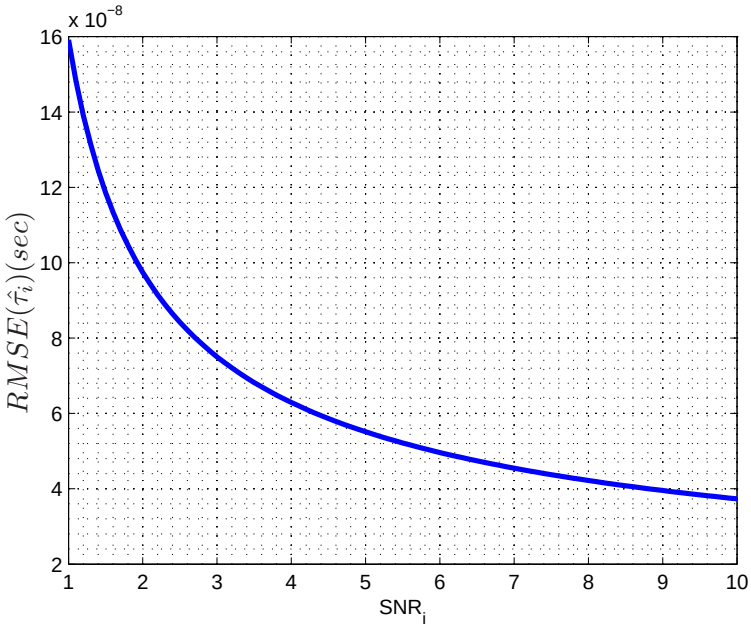

Figure 3.1: RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 1$ MHz.

Figures clearly express the points that we have just emphasized. Having a larger bandwidth or a larger path power simplifies the estimation of the delay.

For the estimation of the path amplitude variance σ_i^2, we obtain the following expression for the CRB:

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \frac{1}{\mathbf{J}_{\sigma_i^2,\sigma_i^2}} = \left(\sigma_i^2 + \sigma_v^2/e_p\right)^2. \tag{3.43}$$

We can express it equivalently as:

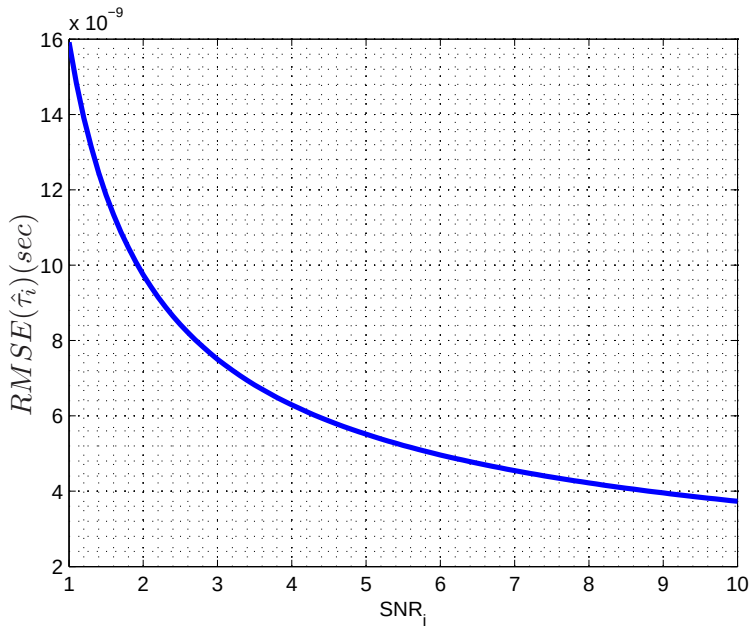

Figure 3.2: RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 10$ MHz.

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \sigma_i^4 \left(1 + \frac{1}{SNR_i}\right)^2. \tag{3.44}$$

Although the expression for CRB($\hat{\sigma_i^2}$) is quite simple, we introduce a plot for it too, with the same parameters that were used for CRB($\hat{\sigma_i^2}$). Only one plot will be given since CRB($\hat{\sigma_i^2}$) is independent of the effective bandwidth of the pulse shape. Another thing to pay attention to is that, unlike the estimation of τ_i, the estimation of σ_i^2 becomes difficult as σ_i^2 increases. This might be interpreted in the way that as the variance of a random variable is increasing, then it becomes more difficult to estimate that variance as expected.

An important comment about these results is that the CRBs for the ith path (CRB($\hat{\tau}_i$) and CRB($\hat{\sigma_i^2}$)) depends only on the parameters of the ith path. This is due to the non-overlapping pulse assumption we have made before. And needless to say that higher σ_v^2 makes the estimation more difficult for all the parameters. As regards local identifiability, we see that local identifiability of the position vector $\mathbf{r}$ is always guaranteed for $N_p > 1$. This is

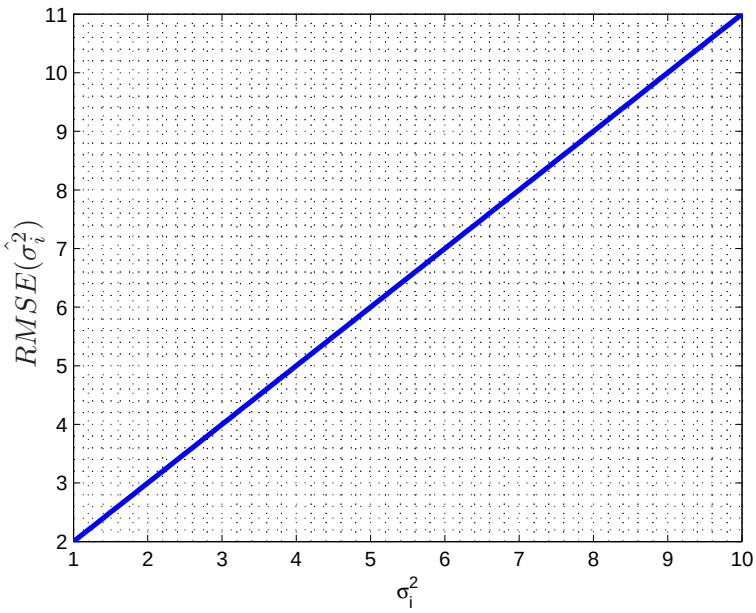

Figure 3.3: RMSE of $\hat{\sigma_i^2}$ as a function of σ_i^2 for $W = 1$ MHz.

because the condition $rank(\mathbf{J}_\theta) \geq N_p$ always holds, hence guaranteeing the local identifiability of $\mathbf{r}$.

3.2.1.1.1 Effect of the Delay Offset Delay synchronization is an important factor in ToA based localization systems. If there is a synchronization error between BS and MT, then there will be an offset value τ_0 in the delays, i.e., $\tau_{i_0} = \tau_i + \tau_0$, where τ_i is the actual delay of the ith path. Hence DO τ_0 is an important nuisance parameter which must be jointly estimated (hence a total of $2N_p + 1$ parameters to estimate). We will try to estimate τ_i from the estimates of τ_{i_0} and τ_0. We still investigate the non-overlapping pulse case again. The partial derivatives are now given as:

$$\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2} = \mathbf{p}_{\tau_{i_0}} \mathbf{p}_{\tau_{i_0}}^H, \tag{3.45}$$

$$\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{i_0}} = -\sigma_i^2 \left(\mathbf{p}'_{\tau_{i_0}} \mathbf{p}_{\tau_{i_0}}^H + \mathbf{p}_{\tau_{i_0}} \mathbf{p}'^H_{\tau_{i_0}} \right), \tag{3.46}$$

$$\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_0} = \sum_{i=1}^{N_p} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{i_0}}. \tag{3.47}$$

After these derivations, entries of the FIM can be calculated. Obviously $\mathbf{J}_{\tau_{i_0},\tau_{i_0}}$ is given by (3.26), and $\mathbf{J}_{\sigma_i^2,\sigma_i^2}$ still by (3.32), and $\mathbf{J}_{\tau_{i_0},\sigma_i^2} = 0$. By (3.47), we obtain: $\mathbf{J}_{\tau_{i_0},\tau_0} = \mathbf{J}_{\tau_{i_0},\tau_{i_0}}$, $\mathbf{J}_{\tau_0,\tau_0} = \sum_{i=1}^{N_p} \mathbf{J}_{\tau_{i_0},\tau_{i_0}}$ and $\mathbf{J}_{\tau_0,\sigma_i^2} = 0$. Hence for $N_p = 1$, the FIM will be:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \\ 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & 0 \\ \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \end{bmatrix}. \tag{3.48}$$

Clearly $rank(\mathbf{J}_\theta) = 2$ as long as $\mathbf{J}_{\tau_{1_0},\tau_{1_0}}$ is nonzero. But surprisingly this is not enough to achieve the local identifiability of $\mathbf{r}$. The reason is due to the matrix $\mathbf{F}$. With the inclusion of τ_0, it now has the structure (for $N_p = 1$):

$$\mathbf{F} = \begin{bmatrix} \frac{\partial \tau_{1_0}}{\partial x} & \frac{\partial \sigma_1^2}{\partial x} & \frac{\partial \tau_0}{\partial x} \\ \frac{\partial \tau_{1_0}}{\partial y} & \frac{\partial \sigma_1^2}{\partial y} & \frac{\partial \tau_0}{\partial y} \end{bmatrix}, \tag{3.49}$$

where

$$\frac{\partial \tau_{1_0}}{\partial x} = \frac{\partial \tau_1}{\partial x}\frac{\partial \tau_{1_0}}{\partial \tau_1} = \frac{\partial \tau_1}{\partial x}, \tag{3.50}$$

$$\frac{\partial \tau_0}{\partial x} = \frac{\partial \tau_1}{\partial x}\frac{\partial \tau_0}{\partial \tau_1} = -\frac{\partial \tau_1}{\partial x}, \tag{3.51}$$

$$\frac{\partial \tau_{1_0}}{\partial y} = \frac{\partial \tau_1}{\partial y}\frac{\partial \tau_{1_0}}{\partial \tau_1} = \frac{\partial \tau_1}{\partial y}, \tag{3.52}$$

$$\frac{\partial \tau_0}{\partial y} = \frac{\partial \tau_1}{\partial y}\frac{\partial \tau_0}{\partial \tau_1} = -\frac{\partial \tau_1}{\partial y}. \tag{3.53}$$

Hence due to this chain rule, the structure of matrix $\mathbf{F}$ becomes:

$$\mathbf{F} = \begin{bmatrix} \mathbf{f1} & \mathbf{f2} & -\mathbf{f1} \end{bmatrix}, \tag{3.54}$$

where $\mathbf{f1}, \mathbf{f2}$ and $-\mathbf{f1}$ are the columns of $\mathbf{F}$. We know from (C.14) that: $\mathbf{J_r} = \mathbf{F}\mathbf{J}_\theta\mathbf{F}^H$, which results in $\mathbf{J_r} = \mathbf{J}_{\sigma_1^2,\sigma_1^2}\mathbf{f2}\,\mathbf{f2}^H$ which is clearly rank 1. Therefore we conclude that local identifiability of $\mathbf{r}$ is not possible for $N_p = 1$. This was an expected result. Because in case of a DO, the delay does not

carry any information about position for $N_p = 1$. Therefore at the end we just have a 1D information coming from the amplitude, which is not sufficient to achieve local identifiability. Also because of the rank deficiency of $\mathbf{J}_\theta$, CRBs cannot be derived. For $N_p > 1$, FIM becomes:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & \cdots & \cdots & \cdots & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \\ 0 & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & \mathbf{J}_{\tau_{N_{p0}},\tau_{N_{p0}}} & \vdots & \vdots & \vdots & \mathbf{J}_{\tau_{N_{p0}},\tau_{N_{p0}}} \\ \vdots & \cdots & \cdots & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & \vdots & \vdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \ddots & 0 & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \mathbf{J}_{\sigma_{N_p}^2,\sigma_{N_p}^2} & 0 \\ \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & \cdots & \mathbf{J}_{\tau_{N_{p0}},\tau_{N_{p0}}} & 0 & \cdots & 0 & \mathbf{J}_{\tau_0,\tau_0} \end{bmatrix}.$$

Rank of $\mathbf{J}_\theta$ is $2N_p$ for a real pulse shape and at most $2N_p$ for any pulse shape as long as $e_p\, e_d > b^2$. It is easy to see the reason for the rank deficiency. The last row (column) is a summation of the first N_p rows (columns). Hence it does not contribute to the rank. But now local identifiability of $\mathbf{r}$ is achieved. However CRBs cannot be obtained again.

3.2.1.1.2 Overlapping Pulses Although our main concern was just to investigate the case of non-overlapping pulses because it is simpler and closed form analytic expressions can be obtained, we now investigate a simple case of 2 overlapping pulses. We define the correlation coefficient ρ, where $|\rho| \in [0, 1]$, as a measure to indicate the overlapping of the pulses. Obviously $|\rho| = 0$ corresponds to the case of non-overlapping situation and $|\rho| = 1$ means that pulses are just on top of each other. The correlation coefficient ρ is given as:

$$\rho = \frac{\mathbf{p}_{\tau_1}^H \mathbf{p}_{\tau_2}}{\|\mathbf{p}_{\tau_1}\| \|\mathbf{p}_{\tau_2}\|} = \frac{\mathbf{p}_{\tau_1}^H \mathbf{p}_{\tau_2}}{e_p}. \tag{3.55}$$

As a result, the matrix $\mathbf{P}_\tau$ is not orthogonal anymore and $\mathbf{P}_\tau^H \mathbf{P}_\tau = e_p\, \mathbf{R}$, where $\mathbf{R}$ is given by:

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}. \tag{3.56}$$

Similarly, we will introduce the correlation coefficient for the derivatives of the pulses: $\mathbf{P'}_\tau^H \mathbf{P'}_\tau = e_d\, \mathbf{N}$ where $\mathbf{N}$ is given as:

$$\mathbf{N} = \begin{bmatrix} 1 & \rho_d \\ \rho_d^* & 1 \end{bmatrix}. \tag{3.57}$$

Clearly $\rho_d \to 1$ when $\rho \to 1$ and $\rho_d \to 0$ when $\rho \to 0$. For this overlapping pulses case, some formulas will change. (3.11) is not valid anymore, and the inverse of the covariance matrix is obtained as:

$$\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}{}^{-1} = \sigma_v^{-2}\,\mathbf{I} - \sigma_v^{-4}\mathbf{P}_\tau \underbrace{\left(\mathbf{C}_b^{-1} + \frac{e_p}{\sigma_v^2}\mathbf{R}\right)^{-1}}_{\mathbf{G}} \mathbf{P}_\tau^H. \tag{3.58}$$

$\mathbf{G}$ is computed easily as:

$$\mathbf{G} = \frac{1}{q}\begin{bmatrix} \sigma_2^{-2} + e_p/\sigma_v^2 & -\rho\, e_p/\sigma_v^2 \\ -\rho^*\, e_p/\sigma_v^2 & \sigma_1^{-2} + e_p/\sigma_v^2 \end{bmatrix}, \tag{3.59}$$

where $q = (e_p/\sigma_v^2 + \sigma_1^{-2})(e_p/\sigma_v^2 + \sigma_2^{-2}) - |\rho|^2\, e_p^2/\sigma_v^4$. Obviously $\mathbf{G}$ is always a nonsingular matrix and q always positive. Partial derivatives of $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$ with respect to the LDPs are still the same as our uncorrelated scattering assumption is still conserved ((3.12), (3.13) still valid). To simplify the equations, we will assume that pulse is real. Therefore ρ and ρ_d also become real ($\rho = \rho^*$, $\rho_d = \rho_d^*$). In addition, we will make the following approximation: $\mathbf{p}_{\tau_i}^H \mathbf{p'}_{\tau_j} \approx 0$ for $i \neq j$. As we remember for real pulses, $\mathbf{p}_{\tau_i}^H \mathbf{p'}_{\tau_i}$ is already 0. With these at hand, entries of the FIM are calculated by using (3.10) again:

$$\begin{aligned}
\mathbf{J}_{\tau_i,\tau_i} &= \operatorname{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_i}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_i}\right) && (3.60)\\
&= 2e_p e_d \sigma_i^4 \sigma_v^{-4}\left(1 - e_p\,\sigma_v^{-2}\,[\mathbf{RGR}]_{i,i}\right) && (3.61)\\
&= 2e_p e_d \sigma_i^2 \sigma_v^{-4}\left(e_p/\sigma_v^2 - e_p\,\rho^2/\sigma_v^2 + \sigma_j^{-2}\right)/q && (3.62)
\end{aligned}$$

where $[\mathbf{RGR}]_{i,i}$ refers to the element in the ith row, i^{th} column position of $\mathbf{RGR}$. At first glimpse, we recognize the differences between the corresponding entry for the non-overlapping pulse case. The correlation coefficient ρ and the power of the other path σ_j^2 now all affect the information obtained from the delay of the ith path. We continue to derive the other entries of the FIM:

$$\begin{aligned}
\mathbf{J}_{\tau_i,\sigma_i^2} &= \operatorname{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_i}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2}\right) && (3.63)\\
&= 0, && (3.64)
\end{aligned}$$

$$\begin{aligned}
\mathbf{J}_{\sigma_i^2,\sigma_i^2} &= \operatorname{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_{\mathbf{i}}^{\mathbf{2}}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_{\mathbf{i}}^{\mathbf{2}}}\right) && (3.65)\\
&= e_p^2\sigma_v^{-4}\left(1 - e_p\,\sigma_v^{-2}\,[\mathbf{RGR}]_{i,i}\right)^2 && (3.66)\\
&= e_p^2\,\sigma_v^{-4}\sigma_i^{-4}\left(e_p/\sigma_v^2 - e_p\,\rho^2/\sigma_v^2 + \sigma_j^{-2}\right)^2/q^2, && (3.67)
\end{aligned}$$

where we see the effects of the jth path again due to the overlapping. We keep on deriving the rest of the entries. In the non-overlapping case, the cross entries of the FIM all happened to be 0, e.g. $\mathbf{J}_{\tau_i,\tau_j} = 0$ for $i \neq j$. However, that is not the case anymore, as it is seen in the derivation below:

$$\begin{aligned}
\mathbf{J}_{\tau_i,\tau_j} &= \operatorname{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{i}}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_{\mathbf{j}}}\right) && (3.68)\\
&= e_p\,e_d\,[\mathbf{N}]_{j,i}\left([\mathbf{R}]_{i,j} - e_p\,\sigma_v^{-2}\,[\mathbf{RGR}]_{i,j}\right)\\
&+ e_p\,e_d\,[\mathbf{N}]_{i,j}\left([\mathbf{R}]_{j,i} - e_p\,\sigma_v^{-2}\,[\mathbf{RGR}]_{j,i}\right) && (3.69)\\
&= 2e_p\,e_d\rho\rho_d\left(1 - e_p\,\sigma_v^{-2}(e_p/\sigma_v^2 - e_p/\sigma_v^2\rho^2 + \sigma_i^{-2} + \sigma_j^{-2})/q\right) && (3.70)\\
&= 2e_p\,e_d\,\rho\,\rho_d\,\sigma_i^{-2}\sigma_j^{-2}/q, && (3.71)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{J}_{\sigma_i^2,\sigma_j^2} &= \operatorname{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_{\mathbf{i}}^{\mathbf{2}}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_{\mathbf{j}}^{\mathbf{2}}}\right) && (3.72)\\
&= e_p^2\sigma_v^{-4}\Big([\mathbf{R}]_{j,i}\,[\mathbf{R}]_{i,j} - e_p\sigma_v^{-2}\,[\mathbf{R}]_{j,i}\,[\mathbf{RGR}]_{i,j}\\
&- e_p\sigma_v^{-2}\,[\mathbf{R}]_{i,j}\,[\mathbf{RGR}]_{j,i} + e_p^2\sigma_v^{-4}\,[\mathbf{RGR}]_{i,j}\,[\mathbf{RGR}]_{j,i}\Big) && (3.73)\\
&= \rho^2 e_p^2\sigma_v^{-4}(1 - e_p\,\sigma_v^{-2}(e_p/\sigma_v^2 - e_p/\sigma_v^2\rho^2 + \sigma_i^{-2} + \sigma_j^{-2})/q)^2 && (3.74)\\
&= \rho^2 e_p^2\sigma_v^{-4}\sigma_i^{-4}\sigma_j^{-4}/q^2. && (3.75)
\end{aligned}$$

where $\mathbf{J}_{\tau_i,\tau_j} = \mathbf{J}_{\tau_j,\tau_i}$ and $\mathbf{J}_{\sigma_i^2,\sigma_j^2} = \mathbf{J}_{\sigma_j^2,\sigma_i^2}$. Obviously both $\mathbf{J}_{\tau_i,\tau_j}$ and $\mathbf{J}_{\sigma_i^2,\sigma_j^2}$ become 0 when ρ is 0. All the FIM entries have been derived, and the FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & \mathbf{J}_{\tau_1,\tau_2} & 0 & 0 \\ \mathbf{J}_{\tau_2,\tau_1} & \mathbf{J}_{\tau_2,\tau_2} & 0 & 0 \\ 0 & 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & \mathbf{J}_{\sigma_1^2,\sigma_2^2} \\ 0 & 0 & \mathbf{J}_{\sigma_2^2,\sigma_1^2} & \mathbf{J}_{\sigma_2^2,\sigma_2^2} \end{bmatrix}. \tag{3.76}$$

FIM is a block diagonal matrix, hence computing its inverse is easy. We just need to compute the inverses of the 2×2 blocks. As a result we obtain the following results for the CRBs:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}_{\tau_i,\tau_i} - \mathbf{J}^2_{\tau_i,\tau_j}/\mathbf{J}_{\tau_j,\tau_j}}, \tag{3.77}$$

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \frac{1}{\mathbf{J}_{\sigma_i^2,\sigma_i^2} - \mathbf{J}^2_{\sigma_i^2,\sigma_j^2}/\mathbf{J}_{\sigma_j^2,\sigma_j^2}}. \tag{3.78}$$

In the following plots, CRBs are shown for different parameters. The values used for e_p and σ_v^2 are the same as the ones used before to make a true comparison between the non-overlapping case and the overlapping case. As expected, the CRB values for the overlapping case is globally higher than the corresponding values of the non-overlapping case. And, obviously, as the correlation between the pulses increases, the CRB values also increase because CRB($\hat{\tau}_i$) depends on both ρ and ρ_d while CRB($\hat{\sigma_i^2}$) depends only on ρ. Another factor affecting the CRBs is now the power of the other path (σ_j^2). It appears as an interference. Hence CRB values increase monotonically with the increasing σ_j^2 for the estimation of both of the parameters. As regards local identifiability, we see that it is achieved for this 2 overlapping pulse case.

3.2.1.2 Isotropic Path Amplitude Variances

As we have already expressed in Remark 1, it is not possible to achieve the local identifiability of $\mathbf{r}$ for $N_p = 1$ due to matrix $\mathbf{F}$ (it will have a rank of 1 for $N_p = 1$). As we will show soon, it will also be clear that $\mathbf{J}_\theta$ is rank deficient for $N_p = 1$. Now we model the path amplitude variances as distance dependent, i.e., $\sigma_i^2 = \frac{k}{\tau_i^\gamma}$. As τ_i and σ_i^2 are coupled now, we will apply chain rule to derive the elements of the FIM. We do not need the entries explicitly for local identifiability analysis. To distinguish the entries from the anisotropic case, we will use the notation $\mathbf{J}'_{\tau_i,\tau_i}$ for example. By using (3.10) again we have:

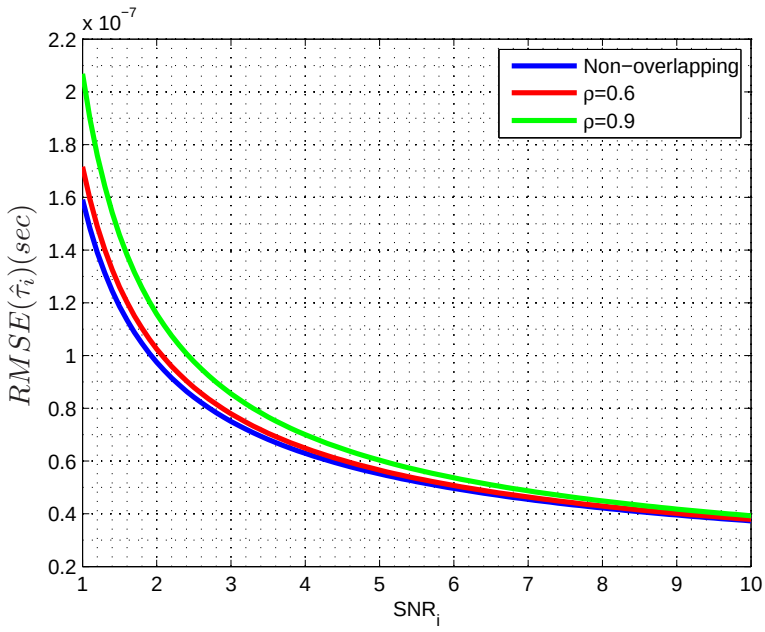

Figure 3.4: RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 1$ MHz, $\sigma_j^2 = 2$.

$$\mathbf{J}'_{\tau_i,\tau_i} = \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}{}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2}\frac{d\sigma_i^2}{d\tau_i}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}{}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2}\frac{d\sigma_i^2}{d\tau_i}\right)$$

where $\frac{d\sigma_i^2}{d\tau_i} = -k\gamma\tau_i^{-(\gamma+1)} = \eta_i$. Hence $\mathbf{J}'_{\tau_i,\tau_i} = \eta_i^2\mathbf{J}'_{\sigma_i^2,\sigma_i^2}$. The partial derivative of $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$ with respect to σ_i^2 is now given as:

$$\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2} = \mathbf{p}_{\tau_i}\mathbf{p}_{\tau_i}^H - \frac{\sigma_i^2}{\eta_i}\left(\mathbf{p}'_{\tau_i}\mathbf{p}_{\tau_i}^H + \mathbf{p}_{\tau_i}\mathbf{p}'^H_{\tau_i}\right), \tag{3.79}$$

and $\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \tau_i} = \eta_i\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \sigma_i^2}$. For this reason we have $\mathbf{J}'_{\sigma_i^2,\tau_i} = \mathbf{J}'_{\tau_i,\sigma_i^2} = \eta_i\,\mathbf{J}'_{\sigma_i^2,\sigma_i^2}$. Therefore for $N_p = 1$ we have:

$$\mathbf{J}_\theta = \mathbf{J}'_{\sigma_1^2,\sigma_1^2}\begin{bmatrix} \eta_1^2 & \eta_1 \\ \eta_1 & 1 \end{bmatrix}, \tag{3.80}$$

which is clearly rank 1. For $N_p > 1$, we will have the same structure of the $2N_p \times 2N_p$ matrix $\mathbf{J}_\theta$ again. However now its rank is always equal to N_p due to the scalar multiplication of the N_p independent rows. Hence for $N_p > 1$, local identifiability of $\mathbf{r}$ can be achieved. Since $\mathbf{J}_\theta$ is always a rank

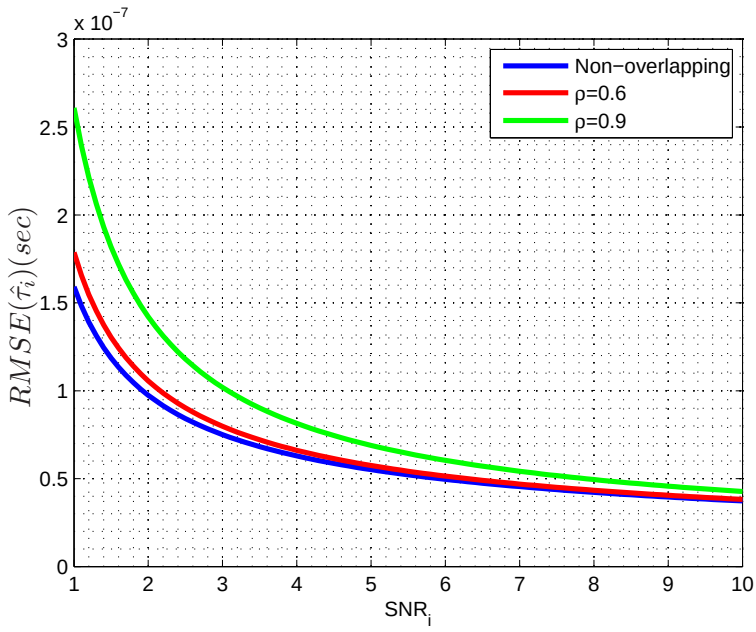

Figure 3.5: RMSE of $\hat{\tau}_i$ as a function of SNR_i for $W = 1$ MHz, $\sigma_j^2 = 20$.

deficient matrix, it is not possible to calculate the CRBs for the LDPs in that configuration. Hence we change the strategy here. The rank deficiency results from the fact that parameters are coupled (σ_i^2 is a function of τ_i). Therefore for the LDP, only delays will be accounted, FIM will consist of only delays, and their CRBs will be calculated. CRBs for the estimation of σ_i^2's will be calculated by the transformation of parameters technique [34]. We obtain easily:

$$\mathbf{J}'_{\tau_i,\tau_i} = \mathbf{J}_{\tau_i,\tau_i} + \eta_i^2 \mathbf{J}_{\sigma_i^2,\sigma_i^2}, \tag{3.81}$$

where $\mathbf{J}_{\tau_i,\tau_i}$ and $\mathbf{J}_{\sigma_i^2,\sigma_i^2}$ are given by (3.26) and (3.32) respectively in the anisotropic case. We see that the information for delay is higher than the anisotropic case, and this is an expected result. The reason is that, now not only delay, but also the path power carries information about the delay and vice versa. The second term in (3.81) is the information coming from the path power which increases the overall information for the delay. The CRB for delays is just the inverse of $\mathbf{J}'_{\tau_i,\tau_i}$, i.e.,

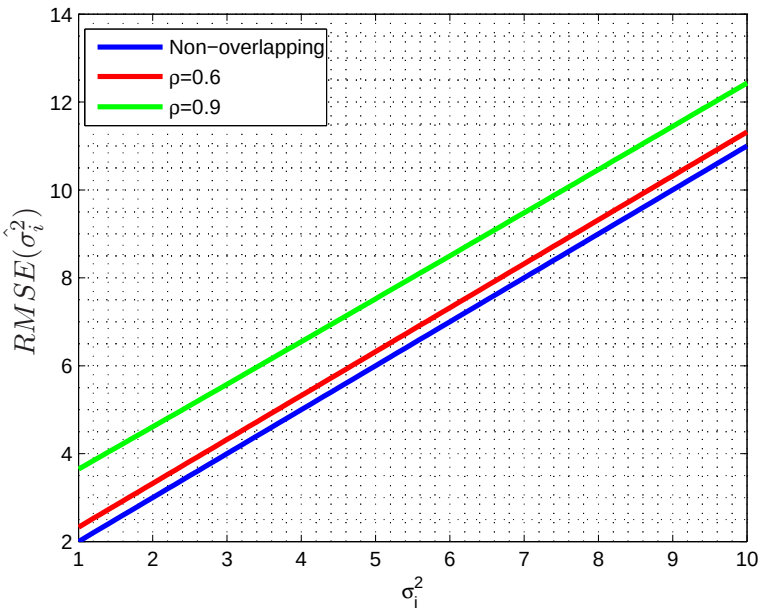

Figure 3.6: RMSE of $\hat{\sigma_i^2}$ as a function of σ_i^2 for $W = 1$ MHz, $\sigma_j^2 = 2$.

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}'_{\tau_i,\tau_i}} = \frac{1}{\mathbf{J}_{\tau_i,\tau_i} + \eta_i^2 \mathbf{J}_{\sigma_i^2,\sigma_i^2}}. \tag{3.82}$$

Hence, estimating delay is easier than the anisotropic case. By the transformation of parameters technique, we obtain the CRB for the amplitude variance:

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \frac{1}{\mathbf{J}_{\sigma_i^2,\sigma_i^2} + \mathbf{J}_{\tau_i,\tau_i}/\eta_i^2}. \tag{3.83}$$

Estimating the path amplitude variance is also easier than the anisotropic case due to the same reason.

3.2.2 Deterministic Path Amplitude Case

We turn back to the channel model in (5.3) and write the complex path amplitude of path i in polar form as $A_i(t) = a_i(t)e^{j\phi_i(t)}$. Now in this section the path magnitudes are modeled as deterministic unknowns, i.e., $a_i(t)$'s. Based on the modeling of the phases, the discussion is divided into two.

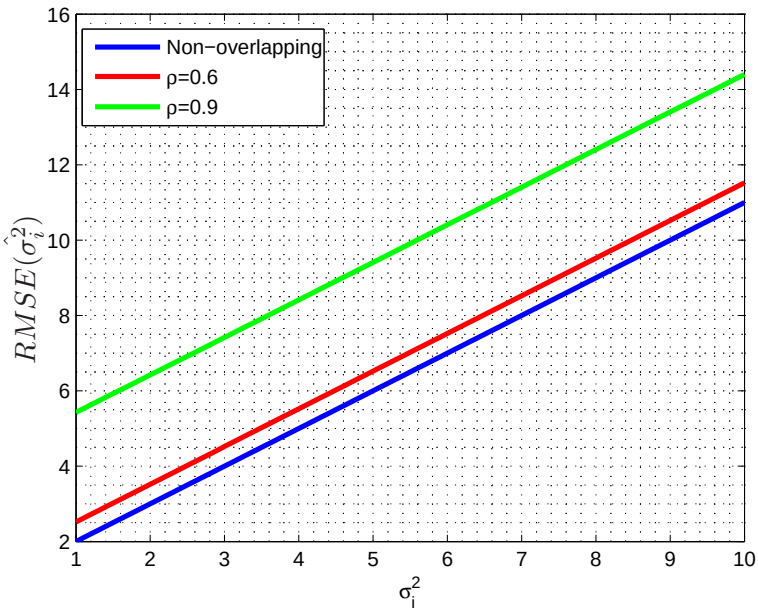

Figure 3.7: RMSE of $\hat{\sigma_i^2}$ as a function of σ_i^2 for $W = 1$ MHz, $\sigma_j^2 = 20$.

The former consists in modeling them as deterministic unknowns and in the latter they are modeled as random variables with uniform distribution over $[0, 2\pi)$. We will explore these two cases separately.

3.2.2.1 Deterministic Modeling of the Phases

Having a different assumption in the path amplitude changes the mean and the covariance matrix of the channel estimates. As we now have deterministic path amplitudes and phases, the mean of the channel estimates is not 0 and given by $\mu = \mathbf{P}_\tau \mathbf{b}(t)$. We also have a different covariance matrix $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \sigma_v^2 \mathbf{I}$. Under these conditions, the computation of the FIM will be completely different. Now we define the LDP vector as:

$$\theta = \left[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ a_1,\ a_2,\ \cdots\ a_{N_p},\ \phi_1,\ \phi_2,\ \cdots\ \phi_{N_p}\right]^T. \tag{3.84}$$

If we check (3.7), it is easy to see that, unlike the Rayleigh fading case, now the first term vanishes because the covariance matrix is not a function of the LDP vector elements, hence its derivatives with respect to these elements

are zero. The second term involving the mean now remains as we have a nonzero mean and also it is a function of the LDP vector elements. So the FIM is given as:

$$[\mathbf{J}_\theta]_{ij} = 2\Re\left(\left[\frac{\partial\mu}{\partial\theta_i}\right]^H \mathbf{C}^{-1}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}\left[\frac{\partial\mu}{\partial\theta_\mathbf{j}}\right]\right). \tag{3.85}$$

3.2.2.1.1 Anisotropic Path Amplitudes In the previous section, we modeled the path amplitudes either isotropically or anisotropically. The similar concept is also applied here, but now for the path magnitudes (a_i's). We model them as a genuine function of position as we have done before. The procedure to follow is the same. We will compute the FIM entries again. Before starting the computation of the FIM entries, we need the following partial derivatives:

$$\frac{\partial\mu}{\partial\tau_i} = -a_i\, e^{j\phi_i}\mathbf{p}'_{\tau_i}, \tag{3.86}$$

$$\frac{\partial\mu}{\partial a_i} = e^{j\phi_i}\mathbf{p}_{\tau_i}, \tag{3.87}$$

and

$$\frac{\partial\mu}{\partial\phi_i} = j\, a_i\, e^{j\phi_i}\mathbf{p}_{\tau_i}. \tag{3.88}$$

With these partial derivatives, the entries of the FIM can be computed as:

$$\mathbf{J}_{\tau_i,\tau_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial\tau_i}\right]\right) \tag{3.89}$$

$$= \frac{2}{\sigma_v^2}\, a_i^2\, e_d, \tag{3.90}$$

$$\mathbf{J}_{\tau_i,a_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial a_i}\right]\right) \tag{3.91}$$

$$= -\frac{2a_i}{\sigma_v^2}\Re\left(\alpha\right) \tag{3.92}$$

$$= -\frac{2a_i}{\sigma_v^2}\, a \tag{3.93}$$

$$= 0, \tag{3.94}$$

$$\mathbf{J}_{a_i,a_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial a_i}\right]^H\left[\frac{\partial\mu}{\partial a_i}\right]\right) \tag{3.95}$$

$$= \frac{2}{\sigma_v^2}\, e_p, \tag{3.96}$$

$$\mathbf{J}_{\phi_i,\phi_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial \phi_i}\right]^H\left[\frac{\partial\mu}{\partial \phi_i}\right]\right) \tag{3.97}$$

$$= \frac{2}{\sigma_v^2}\, a_i^2\, e_p, \tag{3.98}$$

$$\mathbf{J}_{\phi_i,a_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial \phi_i}\right]^H\left[\frac{\partial\mu}{\partial a_i}\right]\right) \tag{3.99}$$

$$= 0, \tag{3.100}$$

and

$$\mathbf{J}_{\phi_i,\tau_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial \phi_i}\right]^H\left[\frac{\partial\mu}{\partial \tau_i}\right]\right) \tag{3.101}$$

$$= -\frac{2}{\sigma_v^2}\, a_i^2\, b, \tag{3.102}$$

where α is defined in (3.22). Note that the correlation between delay and phase disappears for a real (or symmetric) pulse shape. Again we see that stronger paths carry more information for the delay parameter. For $N_p > 1$, we have the cross terms again. However due to the same reasoning (non-overlapping pulse assumption), they are all 0 again. For $N_p = 1$, the FIM is given as follows:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & 0 & \mathbf{J}_{\tau_1,\phi_1} \\ 0 & \mathbf{J}_{a_1,a_1} & 0 \\ \mathbf{J}_{\tau_1,\phi_1} & 0 & \mathbf{J}_{\phi_1,\phi_1} \end{bmatrix} = \frac{2}{\sigma_v^2}\begin{bmatrix} a_1^2\, e_d & 0 & -a_1^2\, b \\ 0 & e_p & 0 \\ -a_1^2\, b & 0 & a_1^2\, e_p \end{bmatrix}. \tag{3.103}$$

Obviously, for any N_p, FIM has a rank of at least $2N_p$. Hence local identifiability of $\mathbf{r}$ is guaranteed. When $e_p\, e_d \neq b^2$, it always has a rank of $3N_p$ (full

rank). From now on, we will assume that the pulse is real (or symmetric) which leads to $b = 0$. In this case $\mathbf{J}_{\tau_i,\phi_i}$ will be 0, resulting in a diagonal FIM. Since now the FIM is diagonal, we do not need to jointly estimate the phases anymore. Therefore the LDP vector will consist of only τ_i's and a_i's now. In that case, CRBs for τ_i and a_i are:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}_{\tau_i,\tau_i}} = \frac{\sigma_v^2}{2\,a_i^2\,e_d} = \frac{1}{8\pi^2 W^2 SNR_i}, \tag{3.104}$$

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{1}{\mathbf{J}_{a_i,a_i}} = \frac{\sigma_v^2}{2\,e_p} = \frac{a_i^2}{2SNR_i}. \tag{3.105}$$

where $SNR_i = \frac{e_p\,a_i^2}{\sigma_v^2}$. We will not provide the plots of these CRBs as we have given the corresponding plots for the Rayleigh fading case. The main difference that we realize is that estimating the delay is easier than in the Rayleigh fading case as expected. And as usual, estimation improves with the decreasing noise power for both of the parameters. Higher bandwidth and higher path power makes the estimation of the delay easier. And having a larger pulse energy (e_p) improves the estimation of the path amplitudes.

3.2.2.1.1.1 Effect of the Delay Offset The same DO issue is revisited here for the deterministic path amplitude case. We begin with the partial derivatives again:

$$\frac{\partial \mu}{\partial \tau_{i_0}} = -a_i\; e^{j\phi_i}\mathbf{p}'_{\tau_{i_0}}, \tag{3.106}$$

$$\frac{\partial \mu}{\partial a_i} = e^{j\phi_i}\mathbf{p}_{\tau_{i_0}}, \tag{3.107}$$

and

$$\frac{\partial \mu}{\partial \tau_0} = \sum_{i=1}^{N_p} \frac{\partial \mu}{\partial \tau_{i_0}}. \tag{3.108}$$

By (3.108), we obtain: $\mathbf{J}_{\tau_{i_0},\tau_0} = \mathbf{J}_{\tau_{i_0},\tau_{i_0}}$, $\mathbf{J}_{\tau_0,\tau_0} = \sum_{i=1}^{N_p} \mathbf{J}_{\tau_{i_0},\tau_{i_0}}$ and $\mathbf{J}_{\tau_0,a_i} = 0$. Hence for $N_p = 1$, the FIM will be:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \\ 0 & \mathbf{J}_{a_1,a_1} & 0 \\ \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \end{bmatrix}. \tag{3.109}$$

The discussion is the same here as in the corresponding DO part in the Rayleigh fading section. Obviously $rank(\mathbf{J}_\theta)$ is always two but due to the same reason that occurred in the Rayleigh fading section, again local identifiability of $\mathbf{r}$ cannot be achieved for $N_p = 1$. This is because of the structure of matrix $\mathbf{F}$, i.e., $\mathbf{F} = [\mathbf{f1}\ \mathbf{f2}\ -\mathbf{f1}]$ which makes $\mathbf{J_r}$ a rank 1 matrix. CRBs cannot be derived due to the rank deficiency and for local identifiability of $\mathbf{r}$, $N_p > 1$ is required.

3.2.2.1.1.2 Overlapping Pulses The same issue as in the Rayleigh fading case will be investigated for the deterministic path amplitudes here. The scenario is the same again, with 2 overlapping and real pulses. Hence there is no need to redefine the parameters ρ and ρ_d once again. The partial derivatives with respect to τ_i and a_i are the same and given by (3.86) and (3.87) respectively. The only difference between the overlapping and non-overlapping cases are the presence of the cross elements for the overlapping pulse case. Obviously $\mathbf{J}_{\tau_i,\tau_i}$, $\mathbf{J}_{\tau_i,a_i}$ and $\mathbf{J}_{a_i,a_i}$ are the same as in the non-overlapping case. The new ones are obtained as:

$$\mathbf{J}_{\tau_i,\tau_j} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial\tau_j}\right]\right) \tag{3.110}$$

$$= \frac{2}{\sigma_v^2}\, a_i\, a_j\, e_d\, \rho_d \cos(\phi_j - \phi_i), \tag{3.111}$$

$$\mathbf{J}_{a_i,a_j} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial a_i}\right]^H\left[\frac{\partial\mu}{\partial a_j}\right]\right) \tag{3.112}$$

$$= \frac{2}{\sigma_v^2}\, e_p\, \rho \cos(\phi_j - \phi_i), \tag{3.113}$$

and $\mathbf{J}_{\tau_i,a_j} = 0$. Evidently, for this overlapping case, FIM entries depend on the phases. Therefore phases must also be jointly estimated. However we will make a simplifying assumption that phases are known beforehand so that they do not need to be estimated. In this case, FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & \mathbf{J}_{\tau_1,\tau_2} & 0 & 0 \\ \mathbf{J}_{\tau_1,\tau_2} & \mathbf{J}_{\tau_2,\tau_2} & 0 & 0 \\ 0 & 0 & \mathbf{J}_{a_1,a_1} & \mathbf{J}_{a_1,a_2} \\ 0 & 0 & \mathbf{J}_{a_1,a_2} & \mathbf{J}_{a_2,a_2} \end{bmatrix}. \tag{3.114}$$

The CRBs are easy to derive:

$$\begin{aligned}
\mathbb{E}(\tau_i - \hat{\tau}_i)^2 &\geq \frac{1}{\mathbf{J}_{\tau_i,\tau_i} - \mathbf{J}^2_{\tau_i,\tau_j}/\mathbf{J}_{\tau_j,\tau_j}} = \frac{\sigma_v^2}{2\,a_i^2\,e_d(1-\rho_d^2\,\cos^2(\phi_j-\phi_i))} \geq \frac{1}{\mathbf{J}_{\tau_i,\tau_i}}, \\
\mathbb{E}(a_i - \hat{a}_i)^2 &\geq \frac{1}{\mathbf{J}_{a_i,a_i} - \mathbf{J}^2_{a_i,a_j}/\mathbf{J}_{a_j,a_j}} = \frac{\sigma_v^2}{2\,e_p(1-\rho^2\,\cos^2(\phi_j-\phi_i))} \geq \frac{1}{\mathbf{J}_{a_i,a_i}}.
\end{aligned}$$

The second inequalities are just given to express that CRBs for this case are higher than the CRBs for the non-overlapping pulse case, as expected. Equality occurs if and only if the phase difference between the pulses is $\phi_j - \phi_i = m\pi/2$, m being an odd integer. In order to have finite CRBs, the denominators must be nonzero (to have an invertible FIM). These inequalities must hold for finite CRBs: $\rho^2\,\cos^2(\phi_j-\phi_i) \neq 1$ and $\rho_d^2\,\cos^2(\phi_j-\phi_i) \neq 1$. The only way for the equalities to be satisfied is when ρ becomes 1 (remember that ρ_d also becomes 1 in that case) and $\phi_j - \phi_i = m\pi$, m being an integer. The interpretation of these results is that when the pulses are exactly on top of each other (so that ρ and ρ_d both become 1) and the phase difference of the paths is an integer multiple of π, CRBs become infinite. Hence estimation of delays and amplitudes become impossible in that case.

3.2.2.1.2 Isotropic Path Amplitudes Now we model path magnitudes as distance dependent and write $a_i = \frac{m}{\tau_i^{\gamma/2}}$ where m is a positive constant and γ is the path-loss coefficient again. We will have a similar discussion that we had in the isotropic modeling in the Rayleigh fading case. Hence to have local identifiability, at least two paths are required again. If the LDP vector is composed of delays and path amplitudes, the FIM will have a rank of N_p due to the coupling between a_i and τ_i. Due to the rank deficiency of the FIM, CRBs cannot be computed in this configuration. Therefore we will apply the same method again. LDP vector will consist of only delays. From that, the CRBs for delays will be computed. And by applying the transformation of parameters technique, the CRBs for the path amplitudes will be derived. The partial derivative with respect to a_i is now given by:

$$\frac{\partial \mu}{\partial a_i} = e^{j\phi_i}\mathbf{p}_{\tau_i} - \frac{1}{\kappa_i} a_i\, e^{j\phi_i}\mathbf{p}'_{\tau_i}, \tag{3.115}$$

where $\kappa_i = \frac{da_i}{d\tau_i} = -\frac{m\gamma}{2}\tau_i^{-(\gamma/2+1)}$. Also $\frac{\partial \mu}{\partial \tau_i} = \kappa_i \frac{\partial \mu}{\partial a_i}$. With these we obtain:

$$\mathbf{J}'_{\tau_i,\tau_i} = \mathbf{J}_{\tau_i,\tau_i} + \kappa_i^2 \mathbf{J}_{a_i,a_i} = \frac{2}{\sigma_v^2}(a_i^2\,e_d + \kappa_i^2 e_p), \tag{3.116}$$

where $\mathbf{J}_{\tau_i,\tau_i}$ and $\mathbf{J}_{a_i,a_i}$ are given in the anisotropic section by (3.90) and (3.96) respectively. Obviously information has increased when compared with its anisotropic counterpart for delay. The CRBs are derived as:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}'_{\tau_i,\tau_i}} = \frac{1}{\mathbf{J}_{\tau_i,\tau_i} + \kappa_i^2 \mathbf{J}_{a_i,a_i}}, \tag{3.117}$$

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{1}{\mathbf{J}_{a_i,a_i} + \mathbf{J}_{\tau_i,\tau_i}/\kappa_i^2}. \tag{3.118}$$

Hence estimating both the delay and the path amplitude is easier than the anisotropic case, as expected.

3.2.2.2 Uniform Random Modeling of the Phases

Now, the second alternative for modeling the phases is investigated. Instead of modeling them as deterministic unknowns, they are now modeled as uniform random variables over $[0, 2\pi)$, i.e., $f_\Phi(\phi) = \frac{1}{2\pi}$. In fact this is a more appropriate model for the phases. In this case, it is easy to see that the mean and the covariance matrix of the channel estimates will be different, i.e., $\mu = \mathbf{0}$ and $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C}_b \mathbf{P}_\tau^H + \sigma_v^2 \mathbf{I}$ where $\mathbf{C}_b$ is a diagonal matrix with entries $[a_1^2, a_2^2, \cdots, a_{N_p}^2]$. Interestingly, the structure of the problem becomes very similar to the Rayleigh fading case. Obviously, the entries of the FIM will be calculated by (3.10):

$$[\mathbf{J}_\theta]_{ij} = \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_\mathbf{i}} \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1} \frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_\mathbf{j}}\right). \tag{3.119}$$

If we define the LDP vector as:

$$\theta = \left[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ a_1^2,\ a_2^2,\ \cdots\ a_{N_p}^2\right]^T, \tag{3.120}$$

we will end up with the same results of the Rayleigh fading case as expected. The only difference would be to substitute a_i^2 for σ_i^2 in the end results. For example CRBs for τ_i and a_i^2 would be given by (3.41) and (3.43) respectively with σ_i^2's replaced by a_i^2's for the anisotropic case. For the isotropic case, the results of the corresponding isotropic section can be used with the same transformation. Hence there is no need to repeat the same things here again. By this simple substitution, we can use the results of the Rayleigh fading section directly.

An interesting thing to do is comparing these results with the ones where we have modeled the phases as deterministic unknowns. As one can expect, the corresponding CRBs will be higher now than their counterparts in the

deterministic phase modeling. If we start with the delays we have (for real or symmetric pulses):

$$\frac{1}{8\pi^2 W^2 SNR_i}\left(1+\frac{1}{SNR_i}\right) > \frac{1}{8\pi^2 W^2 SNR_i} \tag{3.121}$$

with $SNR_i = \frac{e_p\, a_i^2}{\sigma_v^2}$ as defined previously. For the estimation of the path amplitudes, we now have to apply the transformation of parameters technique to obtain the CRBs for a_i's from the CRBs of a_i^2's. By this way we obtain:

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{(\frac{1}{2a_i})^2}{\left(\frac{e_p}{e_p a_i^2 + \sigma_v^2}\right)^2} = \frac{a_i^2}{4} + \frac{\sigma_v^4}{4a_i^2 e_p^2} + \frac{\sigma_v^2}{2e_p} > \frac{\sigma_v^2}{2e_p}. \tag{3.122}$$

These results clearly show that the CRBs for the uniform modeling of the phases are strictly higher than the ones where the phases are modeled as deterministic unknowns.

3.2.3 Rician Fading Case

In this section, we investigate the case where the path amplitudes have both deterministic (non-fading) and random (fading) components at the same time. Indeed, this is the most general case we have dealt with so far. Rician fading is a combination of Rayleigh fading on top of deterministic path amplitudes. In this case, $A_i(t)$ is modeled as a complex Gaussian random variable having a mean of $a_i e^{j\phi_i}$ and a variance of σ_i^2 around this mean. Similarly, based on the modeling of the phases, we can split the analysis into two parts.

3.2.3.1 Deterministic Modeling of the Phases

In this case, there will be a non-zero mean of the paths. The mean will be the same as given in the deterministic path amplitudes deterministic modeling of the phases section, i.e., $\mu = \mathbf{P}_\tau \mathbf{b}$ where $b_i = a_i e^{j\phi_i}$. The covariance matrix is computed as: $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C}_b \mathbf{P}_\tau^H + \sigma_v^2 \mathbf{I}$, where $\mathbf{C}_b$ is a diagonal matrix with entries $[\sigma_1^2, \sigma_2^2, \cdots, \sigma_{N_p}^2]$. In this case the LDP vector is:

$$\theta = \Big[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ \sigma_1^2,\ \sigma_2^2,\ \cdots\ \sigma_{N_p}^2,\ a_1,\ a_2,\ \cdots\ a_{N_p}\Big]^T. \tag{3.123}$$

For simplicity we will just explore the anisotropic case. To calculate the FIM entries, now it is required to use the full formula given by (3.7) because both

the mean and the covariance matrix are now functions of the LDP vector elements. It is not difficult to realize that $\mathbf{J}_{\tau_i,\sigma_i^2} = \mathbf{J}_{\tau_i,a_i} = \mathbf{J}_{a_i,\sigma_i^2} = 0$. The entries corresponding to different paths are still 0. Hence FIM is a diagonal matrix. Also the expression of $\mathbf{J}_{\sigma_i^2,\sigma_i^2}$ is still given by (3.32). Hence the CRB for the estimation of σ_i^2 is the same and given by (3.43). The other entries are computed as (for a real or symmetric pulse):

$$\mathbf{J}_{\tau_i,\tau_i} = \mathbf{J^R}_{\tau_i,\tau_i} + \mathbf{J^D}_{\tau_i,\tau_i}, \tag{3.124}$$

where $\mathbf{J^R}_{\tau_i,\tau_i}$ stands for the FIM entry calculated for Rayleigh fading given by (3.26), and $\mathbf{J^D}_{\tau_i,\tau_i}$ is the FIM entry calculated in the deterministic path amplitude section given by (3.90). The FIM entry just becomes the summation of the information coming from the Rayleigh and deterministic parts. As the CRB is just the inverse of the FIM entry, we can say that in the case of a real (or symmetric) pulse, estimation of τ_i for Rician fading is easier due to the increased information. For the estimation of a_i, we calculate the FIM entry as:

$$\mathbf{J}_{a_i,a_i} = \mathbf{J^D}_{a_i,a_i} - \frac{2}{\sigma_v^2}\frac{e_p^2\,\sigma_i^2}{e_p\,\sigma_i^2 + \sigma_v^2}, \tag{3.125}$$

where $\mathbf{J^D}_{a_i,a_i}$ represents the FIM entry calculated in the deterministic path amplitude section given by (3.96). CRB for a_i is just the inverse of $\mathbf{J}_{a_i,a_i}$. As the information decreased now, we can conclude that estimation of a_i is more difficult for Rician fading case.

3.2.3.2 Uniform Random Modeling of the Phases

Modeling the phases with uniform distribution has an effect on the mean and variance of the channel estimates for Rician fading. We now have $\mu = \mathbf{0}$ and $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C}_b \mathbf{P}_\tau^H + \sigma_v^2 \mathbf{I}$ where $\mathbf{C}_b$ is a diagonal matrix with entries $[a_1^2 + \sigma_1^2, a_2^2 + \sigma_2^2, \cdots, a_{N_p}^2 + \sigma_{N_p}^2]$. Once more we have the same structure of the Rayleigh fading case. We just investigate the anisotropic case again. Obviously, the entries of the FIM will be computed by (3.10):

$$[\mathbf{J}_\theta]_{ij} = \text{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_\mathbf{i}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-1}\frac{\partial \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}}{\partial \theta_\mathbf{j}}\right). \tag{3.126}$$

If we define the LDP vector as:

$$\theta = \left[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ a_1^2 + \sigma_1^2,\ a_2^2 + \sigma_2^2,\ \cdots\ a_{N_p}^2 + \sigma_{N_p}^2\right]^T, \tag{3.127}$$

we end up with the same results we had in the Rayleigh fading section. The only modification is the substitution of $a_i^2 + \sigma_i^2$ for σ_i^2. But in this way we obtain the CRBs for the elements of vector θ. To obtain the CRBs for vector β ($\beta = f(\theta)$), which is:

$$\beta = \Big[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ \sigma_1^2,\ \sigma_2^2,\ \cdots\ \sigma_{N_p}^2,\ a_1,\ a_2,\ \cdots\ a_{N_p}\Big]^T, \tag{3.128}$$

we get use of the transformation of parameters technique again:

$$\mathbf{J}_\beta{}^{-1} \geq \frac{\partial \mathbf{f}(\theta)}{\partial \theta} \mathbf{J}_\theta{}^{-1} \frac{\partial \mathbf{f}(\theta)}{\partial \theta}^T, \tag{3.129}$$

where $\dfrac{\partial \mathbf{f}(\theta)}{\partial \theta}$ is the $3N_p \times 2N_p$ Jacobian matrix. In this case it is derived as:

$$\frac{\partial \mathbf{f}(\theta)}{\partial \theta} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & k_1 & & & \\ & & & & \ddots & & \\ & & & & & & k_{N_p} \end{bmatrix}, \tag{3.130}$$

where $k_i = \frac{\partial a_i}{\partial (a_i^2 + \sigma_i^2)} = \frac{1}{2a_i}$. Hence, as expected, the CRB for the estimation of τ_i is:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{8\pi^2 W^2 SNR_i} \left(1 + \frac{1}{SNR_i}\right) \tag{3.131}$$

where we now define SNR_i as: $\frac{e_p\,(a_i^2 + \sigma_i^2)}{\sigma_v^2}$. It can be deduced that the CRB of τ_i is smaller than the Rayleigh fading case (where there is no deterministic component) due to the increased SNR_i. Hence estimation of τ_i is easier now. The estimation of σ_i^2 is obtained as:

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \left(a_i^2 + \sigma_i^2 + \sigma_v^2/e_p\right)^2. \tag{3.132}$$

Obviously the CRB for σ_i^2 is higher when compared with the Rayleigh fading case. Finally, the CRB for the estimation of a_i is calculated as:

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{(\frac{1}{2a_i})^2}{\left(\frac{e_p}{e_p(a_i^2+\sigma_i^2)+\sigma_v^2}\right)^2} = \frac{1}{4a_i^2}\left(a_i^2 + \sigma_i^2 + \sigma_v^2/e_p\right)^2, \tag{3.133}$$

which is the highest CRB calculated for a_i so far.

3.3 Discussion

In this section, the purpose is to exploit the results obtained so far to make a performance comparison between fingerprinting-based localization algorithms and some other existing localization algorithms. Up to now, the main focus was on the derivation of CRBs of the LDPs and on the local identifiability issue of the position vector $\mathbf{r}$. Now the main concern is the CRB of the position vector (localization performance of the algorithm). In order to remind, the CRBs for the position vector is obtained via (C.14). So far, we did not investigate in detail the matrix $\mathbf{F}$ present in that equation. Indeed $\mathbf{F}$ signifies the geometry of the network, i.e., the surrounding geography, the relative position of the MT, BS and the scatters and etc. In figure 3.8, a simple sketch of the problem is demonstrated.
In fact this is a simplified figure showing only single bounces from the scatterers. However a path from the BS might pass through multiple bounces before reaching the MT, and also a LoS path might be present. Nevertheless, for the ToA-based localization systems, the change in the delay of a path with respect to a change in the position only depends on the AoA. Therefore, only the position of the last scatterer affects the change of the delay of the path through the AoA. It is easy to see it with an example. Suppose that i^{th} path passes through n_i scatterers to reach the MT (n_i^{th} one being the last scatterer before the MT). The path length l_i is given as:

$$l_i = l_{i_1} + \cdots + l_{i_{n_i+1}}, \tag{3.134}$$

where l_{i_j} is the length of the piece the path traverses on its way from the MT to the BS. So l_i can be rewritten as:

$$l_i = \sqrt{(x_{BS} - x_{s_{i1}})^2 + (y_{BS} - y_{s_{i1}})^2} + \cdots + \sqrt{(x_{s_{in_i}} - x)^2 + (y_{s_{in_i}} - y)^2}. \tag{3.135}$$

where $(x_{s_{ij}}, y_{s_{ij}})$ stands for the position of the jth scatterer of the ith path. The change in delay with respect to a change in position is then written:

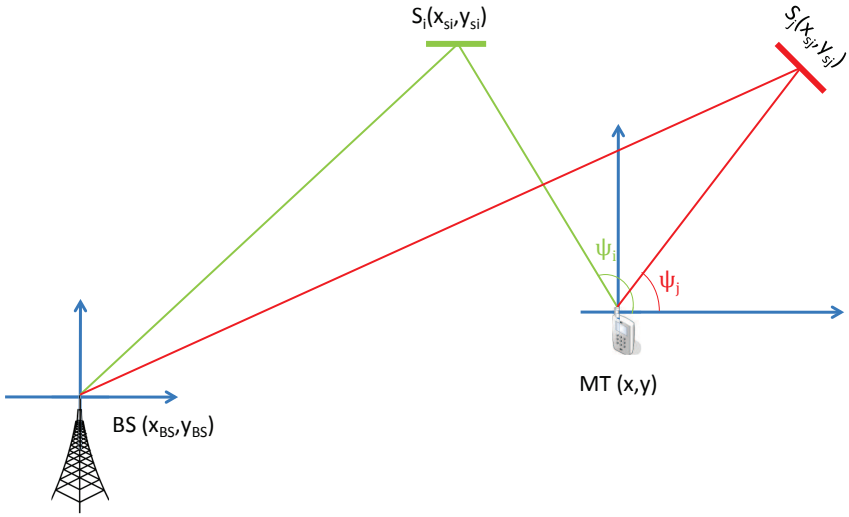

Figure 3.8: BS-MT geometry.

$$\frac{\partial \tau_i}{\partial x} = \frac{1}{c}\frac{\partial l_i}{\partial x} = \frac{1}{c}\frac{x - x_{s_{in_i}}}{\sqrt{(x_{s_{in_i}} - x)^2 + (y_{s_{in_i}} - y)^2}} = -\frac{1}{c}\cos\psi_i \quad (3.136)$$

$$\frac{\partial \tau_i}{\partial y} = \frac{1}{c}\frac{\partial l_i}{\partial y} = \frac{1}{c}\frac{y - y_{s_{in_i}}}{\sqrt{(x_{s_{in_i}} - x)^2 + (y_{s_{in_i}} - y)^2}} = -\frac{1}{c}\sin\psi_i \quad (3.137)$$

where ψ_i is the AoA of the ith path measured with respect to the positive x axis, as shown in figure 3.8 and c is the speed of light. Moreover these equations are also valid for the LoS path since the derivative depends only on AoA. Hence the matrix $\mathbf{F}$ consisting of these kinds of partial derivatives represents the geometry of the network which evidently affects the localization performance of fingerprinting algorithms. This leads us to the well-known concept of Geometric Dilution of Precision (GDoP). For example in indoor positioning, the relative position of the nodes with respect to the mobile has a great impact on the localization performance, which eventually gave rise

to this GDoP concept. The minimum RMSE of position estimate (also the lowest GDoP) for the non-overlapping case and for equal path amplitudes (all path SNR_i's being equal) occurs when $\mathbf{F}$ has orthogonal rows. We will soon see this in detail.

3.3.1 Localization Performance for Deterministic Path Amplitudes and Rayleigh Fading Cases

The purpose of this section is to evaluate the localization performance of fingerprinting algorithms for different path amplitude modeling we have investigated so far. However we will not pass through all the cases we have investigated before. For all the cases, we assume that pulse shape is either real or symmetric.

3.3.1.1 Anisotropic Case with Non-overlapping Pulses

Here we investigate the anisotropic case with non-overlapping pulses first. Only delays will be counted as LDPs. In that case $\mathbf{F}$ will be:

$$\mathbf{F} = -\frac{1}{c}\begin{bmatrix} \cos\psi_1 & \cos\psi_2 & \cdots & \cos\psi_{N_p} \\ \sin\psi_1 & \sin\psi_2 & \cdots & \sin\psi_{N_p} \end{bmatrix}. \tag{3.138}$$

We want to derive the CRB of the position vector $\mathbf{r}$, which is: $\mathbb{E}\|\mathbf{r}-\hat{\mathbf{r}}\|^2 = \mathbb{E}(x-\hat{x})^2 + \mathbb{E}(y-\hat{y})^2$. This is just equal to the trace of the inverse of the position FIM, i.e., $\mathrm{tr}\,(\mathbf{J_r}^{-1})$. We know from matrix theory that, only for a 2×2 matrix:

$$\mathrm{tr}\,(\mathbf{J_r}^{-1}) = \frac{\mathrm{tr}\,(\mathbf{J_r})}{\det \mathbf{J_r}}. \tag{3.139}$$

The term in the numerator is calculated as:

$$\mathrm{tr}\,(\mathbf{J_r}) = \frac{1}{c^2}\sum_{i=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i}, \tag{3.140}$$

and the term in the denominator as:

$$\det \mathbf{J_r} = \frac{1}{c^4}\sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j}\sin^2(\psi_i-\psi_j). \tag{3.141}$$

The proof is given in appendix A. We interpret this result saying that positioning accuracy depends on the differences between the $\binom{N_p}{2}$ AoA couples

of the paths. Indeed this is a reasonable result because differences between AoAs signify the relative position of the scatters with respect to each other and the MT. Therefore having an expression consisting of the differences between the AoA couples makes complete sense. If we had summations of AoAs instead of differences in the overall expression, in that case it would not give any clue about the relative positioning of the scatters.

After the derivation we can now formulate the CRB of the position vector. Indeed we have obtained a generic formula. The term $\mathbf{J}_{\tau_i,\tau_i}$ might correspond either to the Rayleigh fading case or to the deterministic path amplitudes case. We will now present the result for the deterministic path amplitudes with deterministic phases. Hence for $N_p \geq 2$:

$$\mathbb{E}\|\mathbf{r}-\hat{\mathbf{r}}\|^2 = c^2 \frac{\frac{2\,e_d}{\sigma_v^2}\sum_{i=1}^{N_p} a_i^2}{\frac{4\,e_d^2}{\sigma_v^4}\sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} a_i^2\,a_j^2\sin^2(\psi_i-\psi_j)} \tag{3.142}$$

$$= \zeta \frac{\sum_{i=1}^{N_p} SNR_i}{\sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} SNR_i\,SNR_j\sin^2(\psi_i-\psi_j)}, \tag{3.143}$$

with $SNR_i = \frac{e_p\,a_i^2}{\sigma_v^2}$ and $\zeta = \frac{c^2}{8\pi^2\,W^2}$. This is an important result and has the same structure of the equation given in [35]. In that paper, N_p refers to the number of the BSs receiving LoS signals from the MT. But in our case, we have one BS and N_p paths (either all of them NLoS or 1 LoS and the rest NLoS), and demonstrate the CRB of the position estimate. The important thing in our case is that this is the result of a non-overlapping pulse assumption, which means that paths are well separated. We see the effects of the geometry through the differences of the AoAs. The geometry might make the position estimation easier or more difficult (GDoP effect). Even in an extreme case, estimation of the position might be impossible because of the geometry, i.e., when all the pairwise differences between AoAs are either 0 (when all the AoAs are the same) or π radians which nullifies the denominator of (3.143), leading to an infinite CRB. This kind of situation describes a geometry where all the last scatterers of the paths are aligned on the same line (if all the paths are NLoS), or the BS and the last scatterers of the paths all lie on the same line (when there is 1 LoS path and the rest is NLoS) passing through the MT. However the probability of such a situation to take place is extremely small (almost 0). On the other hand, to minimize the RMSE of the position estimate, the pairwise differences between AoAs must be either $\pi/2$ or $3\pi/2$ radians ($\psi_1 - \psi_2 = \pi/2$ or $\psi_1 - \psi_2 = 3\pi/2$) for $N_p = 2$ as clearly seen. Obviously this defines an orthogonal geometry

for $N_p = 2$. Also note that in that case the rows of $\mathbf{F}$ become orthogonal. Below, we provide a plot of this result for $N_p = 2$ and $W = 10$ MHz. We see that RMSE of the position estimate might go well below one meter (indeed 0.4775 meters minimum for $SNR_1 = SNR_2 = 20dB$).

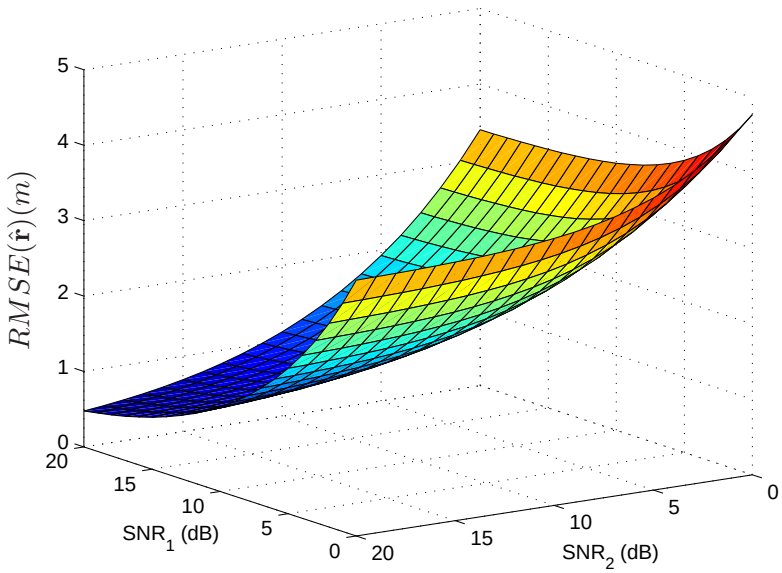

Figure 3.9: RMSE of $\hat{\mathbf{r}}$ as a function of SNR_1 and SNR_2 for $\psi_2 - \psi_1 = \pi/2$ for the deterministic case.

Derivation of the RMSE of the position estimate for the Rayleigh fading case is straightforward. Just need to replace the $\mathbf{J}_{\tau_i,\tau_i}$'s with the corresponding entries from the Rayleigh fading section. We easily obtain the result:

$$\mathbb{E}\|\mathbf{r} - \hat{\mathbf{r}}\|^2 = \zeta \frac{\sum_{i=1}^{N_p} \frac{SNR_i^2}{SNR_i+1}}{\sum_{i=1}^{N_p-1} \sum_{j=i+1}^{N_p} \frac{SNR_i^2\, SNR_j^2}{(SNR_i+1)(SNR_j+1)} \sin^2(\psi_i - \psi_j)}, \quad (3.144)$$

where $SNR_i = \frac{e_p\, \sigma_i^2}{\sigma_v^2}$. We will also provide a plot for it for the same configuration of the deterministic plot. The minimum RMSE of the position estimate is now 0.4798 meters. As expected, RMSE of position estimate is always higher than the deterministic case. This is not surprising as the FIM

entries of the LDPs are lower for Rayleigh fading for the same SNR_i and W values. Indeed we will soon generalize this statement.

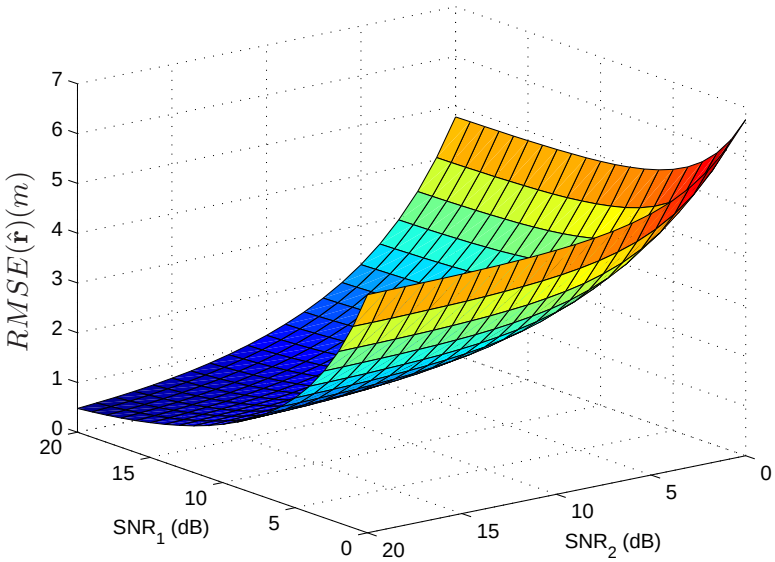

Figure 3.10: RMSE of $\hat{\mathbf{r}}$ as a function of SNR_1 and SNR_2 for $\psi_2 - \psi_1 = \pi/2$ for Rayleigh fading case.

If we also talk about GDoP, the minimum GDoP that can be achieved is $2/\sqrt{N_p}$ in 2D scenarios [36]. The GDoP is given by:

$$\text{GDoP} = \sqrt{\text{tr}\,\{(\mathbf{F}\mathbf{F}^T)^{-1}\}}. \tag{3.145}$$

Also minimum RMSE of position estimate occurs when $\mathbf{F}$ has orthogonal rows for the case of equal path amplitudes (all SNR_i's being equal). For any N_p ($N_p \geq 2$), this can be achieved when the difference between consecutive AoAs is π/N_p radians, i.e., $\psi_i - \psi_{i-1} = \pi/N_p,\ i = 2, \cdots N_p$. In fact, we have obtained this result via simulations. For sure, it is a critical point in the cost function in the denominator of the RMSE expression (when all SNR_i's are equal). However in the second derivative test, the Hessian turned out to be negative semi-definite which made the test inconclusive. Moreover there are multiple global maxima as clearly seen, and the one we propose is one

of them. Intuitively we will make some more comments about these results. We can list them as:

1. An increase in the SNR of any path always improves the position estimation. In fact in a more general statement, we can say that any increase in the information ($\mathbf{J}_{\tau_i,\tau_i}$) of any path (or multiple paths) improves the position estimation if the network geometry is held fixed. This might be due to an increase in the SNR_i's or due to an increase in the bandwidth W of the system.

2. From item 1 we figure out that the accuracy of position estimation for the isotropic case is strictly better than the anisotropic case for the same network geometry. This is because $\mathbf{J'}_{\tau_i,\tau_i} > \mathbf{J}_{\tau_i,\tau_i}$.

3. If we add one additional path (assuming the new path does not disturb the non-overlapping pulse assumption) while keeping the rest of the geometry fixed, this also always improves the position estimation even if the new path has a very low SNR.

4. As expected, the biggest contribution for the position estimation comes from the path with the highest SNR value.

We have carried out the analysis using only delays as LDPs. However as one can expect, using more information would yield a better localization accuracy. We mean that, in addition to the delays, we can also get use of the information coming from the path amplitudes ($\mathbf{J}_{a_i,a_i}$'s). To see the improvement in the localization, suppose first that we use only one of them, without loss of generality $\mathbf{J}_{a_1,a_1}$. Let us call $\mathbf{J}_\mathbf{r}$, the FIM of $\mathbf{r}$ using only delays as LDPs, and $\mathbf{J}'_\mathbf{r}$, the FIM of $\mathbf{r}$ using both delays and the path amplitudes as LDPs. The relation between them can be written as:

$$\mathbf{J}'_\mathbf{r} = \mathbf{J}_\mathbf{r} + \mathbf{J}_{a_1,a_1}\,\mathbf{q}_1\,\mathbf{q}_1^T, \tag{3.146}$$

where we define $\mathbf{q}_1$ as:

$$\mathbf{q}_1 = \begin{bmatrix} \frac{\partial a_1}{\partial x} \\ \frac{\partial a_1}{\partial y} \end{bmatrix}. \tag{3.147}$$

We need the inverses of the FIM's to reach to a conclusion about their localization performances. By using the matrix inversion lemma:

$$\mathbf{J}'^{-1}_\mathbf{r} = \mathbf{J}_\mathbf{r}^{-1} - \frac{\mathbf{J}_{a_1,a_1}\mathbf{J}_\mathbf{r}^{-1}\mathbf{q}_1\,\mathbf{q}_1^T\mathbf{J}_\mathbf{r}^{-1}}{1+\mathbf{J}_{a_1,a_1}\mathbf{q}_1^T\mathbf{J}_\mathbf{r}^{-1}\mathbf{q}_1}. \tag{3.148}$$

Obviously, the second term in the right-hand side of the equation is a positive semidefinite matrix. Therefore we can say that:

$$\mathbf{J_r}^{-1} \geq \mathbf{J'_r}^{-1}, \tag{3.149}$$

which means that, $\mathbf{J_r}^{-1}$ is a more positive definite matrix than $\mathbf{J'_r}^{-1}$ ($\mathbf{J_r}^{-1} - \mathbf{J'_r}^{-1} \geq \mathbf{0}$). Moreover we can also write:

$$\text{tr}\left(\mathbf{J_r}^{-1}\right) \geq \text{tr}\left(\mathbf{J'_r}^{-1}\right). \tag{3.150}$$

The interpretation of this result is that covariance matrix of the position error when only delays are counted is greater than the covariance matrix of the position error when both delays and one of the path amplitudes are counted. From (3.148) we can also see that, when $\mathbf{J}_{a_1,a_1}$ increases, localization accuracy also increases, as expected. Extending this result to any number of (greater than one) $\mathbf{J}_{a_i,a_i}$'s being incorporated in the analysis is straightforward. The procedure to follow is similar. Let us assume that p of the $\mathbf{J}_{a_i,a_i}$'s will be used in addition to the N_p delays being used where $1 < p \leq N_p$. Without loss of generality, we can assume that the indexes begin from $i = 1$ and end in $i = p$. In this case:

$$\mathbf{J'_r} = \mathbf{J_r} + \mathbf{\Gamma} \tag{3.151}$$

where

$$\mathbf{\Gamma} = \sum_{i=1}^{p} \mathbf{J}_{a_i,a_i}\, \mathbf{q}_i\, \mathbf{q}_i^T, \tag{3.152}$$

with $\mathbf{q}_i$ being defined the same way as $\mathbf{q}_1$ before. If we use the matrix inversion lemma again, we obtain:

$$\mathbf{J'_r}^{-1} = \mathbf{J_r}^{-1} - \mathbf{J_r}^{-1}\left(\mathbf{\Gamma}^{-1} + \mathbf{J_r}^{-1}\right)^{-1}\mathbf{J_r}^{-1}. \tag{3.153}$$

The second term in the r.h.s. of the equation is a positive definite matrix. Hence, once more we can write the following:

$$\mathbf{J_r}^{-1} \geq \mathbf{J'_r}^{-1}, \tag{3.154}$$

whose interpretation was already made before.

3.3.1.2 Isotropic Case with Non-overlapping Pulses

In this section, we show the derivation only for deterministic case. It is straightforward to extend to the Rayleigh fading case. However, the comments we will make on the results will be valid for both of the cases. In the 2nd item in the list of the previous section, we already mentioned that isotropic case results in a better position accuracy than the anisotropic case when both are utilizing only delays. Now we will investigate the isotropic case which utilizes only path amplitudes (a_i's) for positioning. In this case, $\mathbf{J}_\theta$ is a diagonal matrix with $\mathbf{J'}_{a_i,a_i}$'s on its diagonal. Remember that we have: $\frac{\partial a_i}{\partial x} = \kappa_i \frac{\partial \tau_i}{\partial x}$. Now in this case $\mathbf{F}$ becomes:

$$\mathbf{F} = -\frac{1}{c}\begin{bmatrix} \kappa_1 \cos\psi_1 & \kappa_2 \cos\psi_2 & \cdots & \kappa_{N_p} \cos\psi_{N_p} \\ \kappa_1 \sin\psi_1 & \kappa_2 \sin\psi_2 & \cdots & \kappa_{N_p} \sin\psi_{N_p} \end{bmatrix}. \tag{3.155}$$

Then we can write $c^2\,\mathbf{J_r} =$

$$\begin{bmatrix} \cos\psi_1 & \cdots & \cos\psi_{N_p} \\ \sin\psi_1 & \cdots & \sin\psi_{N_p} \end{bmatrix} \begin{bmatrix} \kappa_1^2 \mathbf{J'}_{a_1,a_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \kappa_{N_p}^2 \mathbf{J'}_{a_{N_p},a_{N_p}} \end{bmatrix} \begin{bmatrix} \cos\psi_1 & \sin\psi_1 \\ \vdots & \vdots \\ \cos\psi_{N_p} & \sin\psi_{N_p} \end{bmatrix}. \tag{3.156}$$

We see that we end up with the same structure as in the previous section (the matrix $\mathbf{F}$ of the previous section), but now on the diagonals in the middle matrix there are the terms $\kappa_i^2 \mathbf{J'}_{a_i,a_i}$'s. We already know that, $\kappa_i^2 \mathbf{J'}_{a_i,a_i} = \mathbf{J'}_{\tau_i,\tau_i}$. Therefore this is the same result as if we are performing localization with only $\mathbf{J'}_{\tau_i,\tau_i}$'s. As we mentioned in the previous section, this always gives a better positioning accuracy than its anisotropic counterpart. The comments we have made about GDoP before are still valid. The important remark to be mentioned here is that localization accuracy enhances as the path-loss coefficient γ increases.

3.3.1.3 Anisotropic Case for Non-overlapping Pulses with LoS and NLoS Paths

We will now talk about another interesting and a quite frequent scenario in wireless communications. In this situation, there is a direct LoS path between BS and MT and the rest of the paths are NLoS. It is more reasonable to model the LoS path with a deterministic path amplitude and a deterministic phase. For the NLoS paths, Rayleigh fading modeling is more realistic. Consequently $\mathbf{J}_{\tau_1,\tau_1}$ is given by (3.90) and $\mathbf{J}_{\tau_i,\tau_i}$'s are given

by (3.26) for $i \in [2, N_p]$. Derivation of $\mathbb{E}\|\mathbf{r} - \hat{\mathbf{r}}\|^2$ is again carried out via equations (3.139), (3.140), (3.141).

3.3.1.4 Anisotropic Case with 2 Consecutive Pulses Overlapping

In this section, a more realistic scenario will be investigated, in which two consecutive pulses are considered to be overlapping with each other. Indeed this is a scenario very likely to happen. Once again, we take only the delays into account as being the LDPs. In this case, FIM of the LDPs is not diagonal, but instead tridiagonal as seen below:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & \mathbf{J}_{\tau_1,\tau_2} & & & \\ \mathbf{J}_{\tau_1,\tau_2} & \mathbf{J}_{\tau_2,\tau_2} & \mathbf{J}_{\tau_2,\tau_3} & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \mathbf{J}_{\tau_{N_p-1},\tau_{N_p}} \\ & & & \mathbf{J}_{\tau_{N_p-1},\tau_{N_p}} & \mathbf{J}_{\tau_{N_p},\tau_{N_p}} \end{bmatrix}. \tag{3.157}$$

To derive the CRB of the position estimate, we will utilize (3.139) again. However calculating the determinant of $\mathbf{J}_\mathbf{r}$ is more challenging now. The difficulty arises from the non-diagonal structure of $\mathbf{J}_\theta$. On the other hand it is easier to derive the trace and it is given by:

$$\mathrm{tr}\left(\mathbf{J}_\mathbf{r}\right) = \frac{1}{c^2}\left(\sum_{i=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i} + 2\sum_{i=1}^{N_p-1} \mathbf{J}_{\tau_i,\tau_{i+1}} \cos(\psi_{i+1} - \psi_i)\right). \tag{3.158}$$

The procedure to calculate the trace is similar to the one before. We just write $\mathbf{J}_\mathbf{r}$ directly as the sum of 2×2 matrices again, and the rest is easy to see. Calculating the determinant is more difficult due to the terms coming from the side diagonals of $\mathbf{J}_\theta$. However, in a similar manner as we calculated the determinant for the non-overlapping case, we now derive it as follows:

$$\begin{aligned}
\det \mathbf{J_r} &= \frac{1}{c^4}\sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j}\sin^2(\psi_i-\psi_j) \\
&+ \frac{2}{c^4}\sum_{i=1}^{N_p}\sum_{j=1}^{N_p-1} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_{j+1}}\sin(\psi_{j+1}-\psi_i)\,\sin(\psi_j-\psi_i) \\
&+ \frac{1}{c^4}\sum_{i=1}^{N_p-1}\sum_{j=1}^{N_p-1} \mathbf{J}_{\tau_i,\tau_{i+1}}\mathbf{J}_{\tau_j,\tau_{j+1}}\left(\sin(\psi_{j+1}-\psi_i)\,\sin(\psi_j-\psi_{i+1})+\right. \\
&\quad \left.\sin(\psi_j-\psi_i)\,\sin(\psi_{j+1}-\psi_{i+1})\right). \qquad (3.159)
\end{aligned}$$

Just to remind, we have from (3.139):

$$\mathbb{E}\|\mathbf{r}-\hat{\mathbf{r}}\|^2 = \frac{\mathrm{tr}\,(\mathbf{J_r})}{\det \mathbf{J_r}}.$$

Hence plugging equations (3.158), (3.159) into the expression above, we can obtain the CRB for position estimate. The CRB expression we just derived is generic like the ones given by equations (3.140), (3.141) which is valid for all non-overlapping cases for the cases where two consecutive pulses overlap. Hence it is valid both for deterministic path amplitudes and Rayleigh fading cases. However for the deterministic case with deterministic phases, we should make the assumption that we know the phases beforehand to be able to use this CRB formula.

3.3.1.5 Anisotropic Case with 2 Overlapping Pulses

The anisotropic case with 2 overlapping pulses is a special case of the section which was just investigated (two consecutive pulses overlapping) for $N_p = 2$. We first consider the deterministic path amplitudes case with deterministic phases. For the Rayleigh fading case, we will just provide the plots of the results. As we expressed before, we assume to know the phases beforehand for the deterministic case. In that case, $\mathbf{J}_\theta$ is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & \mathbf{J}_{\tau_1,\tau_2} \\ \mathbf{J}_{\tau_1,\tau_2} & \mathbf{J}_{\tau_2,\tau_2} \end{bmatrix}. \qquad (3.160)$$

The matrix $\mathbf{F}$ is still given by (3.138) for $N_p = 2$. By using equations (3.158), (3.159) we obtain the expression of $\mathbb{E}\|\mathbf{r}-\hat{\mathbf{r}}\|^2$:

$$
\begin{aligned}
&= c^2 \frac{\mathbf{J}_{\tau_1,\tau_1} + \mathbf{J}_{\tau_2,\tau_2} + 2\,\mathbf{J}_{\tau_1,\tau_2}\cos(\psi_2-\psi_1)}{(\mathbf{J}_{\tau_1,\tau_1}\mathbf{J}_{\tau_2,\tau_2} - \mathbf{J}^2_{\tau_1,\tau_2})\sin^2(\psi_2-\psi_1)} \qquad (3.161)\\
&= \zeta \frac{SNR_1 + SNR_2 + 2\sqrt{SNR_1\,SNR_2}\rho_d\cos(\phi_2-\phi_1)\,\cos(\psi_2-\psi_1)}{SNR_1\,SNR_2(1-\rho_d^2\cos^2(\phi_2-\phi_1))\sin^2(\psi_2-\psi_1)}.
\end{aligned}
$$

Indeed (3.161) is a generic result. We will also exploit it for the Rayleigh fading case by just substituing the corresponding FIM entries of the Rayleigh fading section. We recognize that if the phase difference between the two paths is $\pi/2$ or $3\pi/2$, then we obtain exactly the same positioning accuracy as in the non-overlapping pulse case (for $N_p = 2$) in the same network geometry. Surprisingly, the positioning accuracy of this overlapping case might be even better than its non-overlapping counterpart for some values of the phase difference, SNR's and correlation coefficient. Therefore we cannot make a certain conclusion about which one is better in every situation. In the figures below, we plot the RMSE of the position estimate as a function of the phase difference. Bandwidth is set to 10 MHz ($W = 10$ MHz) for all the plots. We observe the results that we expected to see. Clearly, in general, the non-overlapping case performs better than the overlapping case, and increasing correlation between the pulses makes the estimation more difficult.

Now, by using (3.161), we will try to see what happens in the same situation for the Rayleigh fading case. We present two plots to make a comparison between overlapping and non-overlapping Rayleigh fading cases for the same bandwidth ($W = 10$ MHz). As we do not have any phases in this situation, the plots are obtained with respect to the $\psi_2 - \psi_1$. Clearly, the estimation of position becomes impossible when $\psi_2 - \psi_1 \to 0$. Differently from the deterministic path amplitudes case, now overlapping case always results in a higher RMSE of position estimate than its non-overlapping counterpart.

3.4 Conclusion

This chapter was the core of our research where we explored performance bounds (CRBs) for PDP-F. Many different cases for path amplitude modeling, phase modeling, local identifiability, overlapping and non-overlapping pulses were investigated. Also the notion of isotropic/anisotropic path amplitude modeling was introduced. We have found out many important results on various subjects. We can briefly list them as follows:

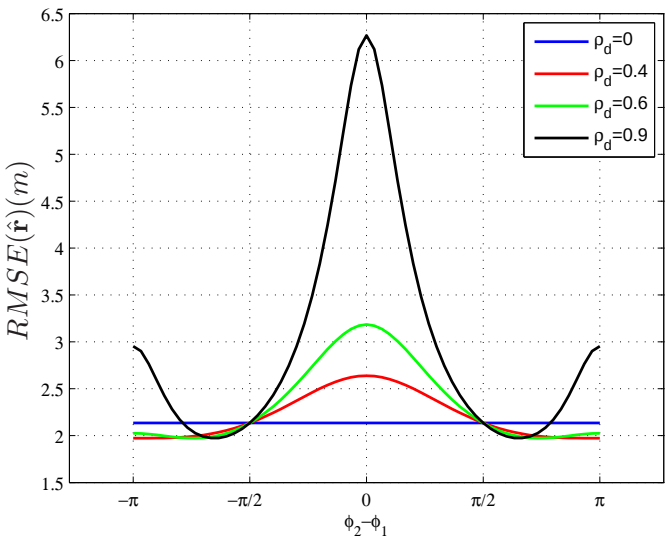

Figure 3.11: RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 10dB$ and $\psi_2 - \psi_1 = \pi/4$ for the deterministic case.

- Local identifiability of the position vector $\mathbf{r}$ might be achieved even with one path for the anisotropic case. Surprisingly pulse shape plays an important role in local identifiability. For the isotropic case, at least two paths are required.

- Estimation of path delays improves with SNR and effective bandwidth of the pulse shape W. On the other hand, estimation of path amplitudes is independent of W. Also estimating the same LDP under the Rayleigh fading modeling is more difficult than the deterministic path amplitude with deterministic phases modeling.

- In case of DO, at least two paths are required for the local identifiability of $\mathbf{r}$ even for the anisotropic case.

- Closed-form CRB expressions for the estimation of LDPs under the overlapping pulse assumption has also been demonstrated. We have seen that interfering pulses hampers the estimation of the LDPs. If a ToA based localization system relying on the LoS path is utilized, and if the LoS pulse overlaps with some other pulses, then the localization accuracy might be substantially degraded depending on the

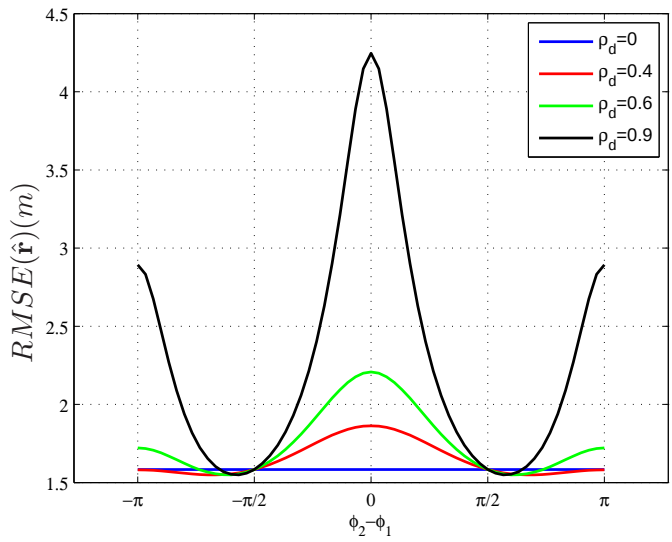

Figure 3.12: RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 20dB$ and $\psi_2 - \psi_1 = \pi/4$ for the deterministic case.

overlapping ratio.

- Closed-form expressions of the CRB for the estimation of $\mathbf{r}$ have been derived for both deterministic and Rayleigh fading cases. We have directly seen the positive effects of W and SNR's of the individual paths. Also impact of the surrounding geometry (GDoP) has also been explicitly shown. The geometry conditions enhancing or degrading the localization performance were explained. Also the special case of two overlapping pulses has been demonstrated. For the Rayleigh fading case, it always degrades the localization performance. However for the deterministic case, it might improve the localization performance for some specific values of the phase difference, overlapping ratios, etc.

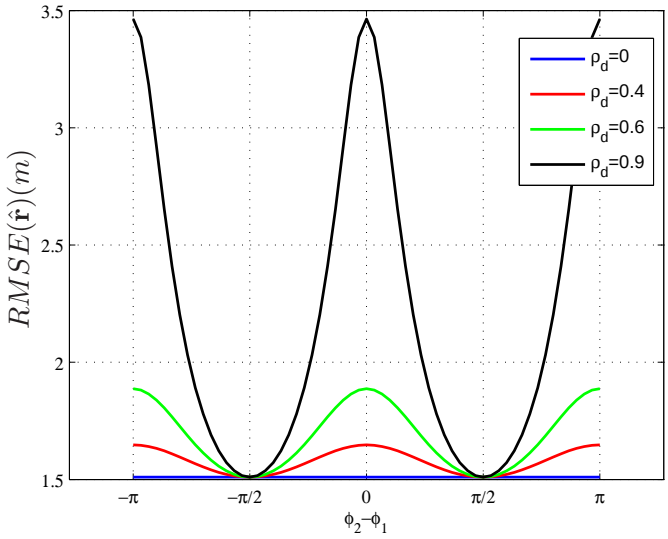

Figure 3.13: RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 10dB$ and $\psi_2 - \psi_1 = \pi/2$ for the deterministic case.

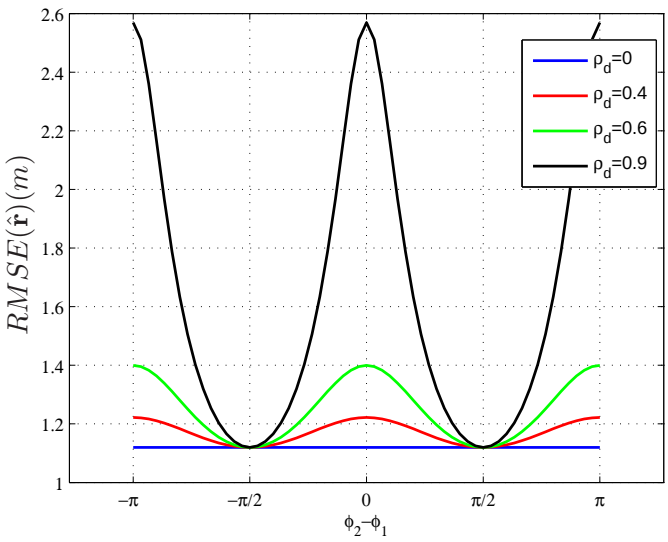

Figure 3.14: RMSE of $\hat{\mathbf{r}}$ as a function of $\phi_2 - \phi_1$, $SNR_1 = 10dB$, $SNR_2 = 20dB$ and $\psi_2 - \psi_1 = \pi/2$ for the deterministic case.

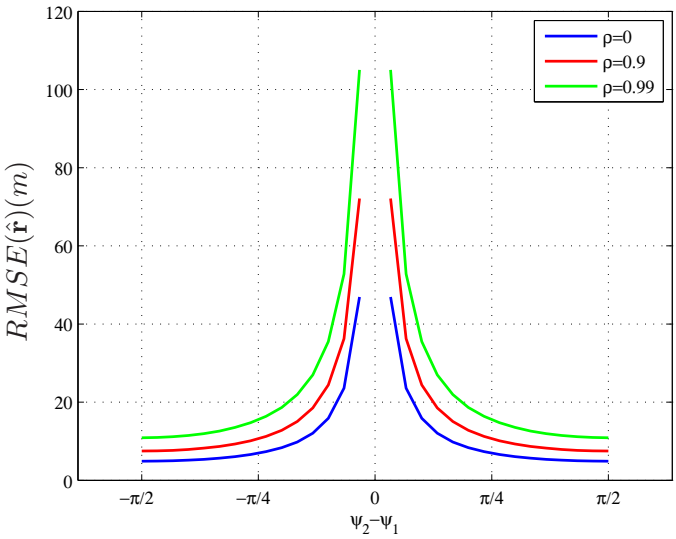

Figure 3.15: RMSE of $\hat{\mathbf{r}}$ as a function of $\psi_2 - \psi_1$, $SNR_1 = 10dB$, $SNR_2 = 0dB$ for the Rayleigh fading case.

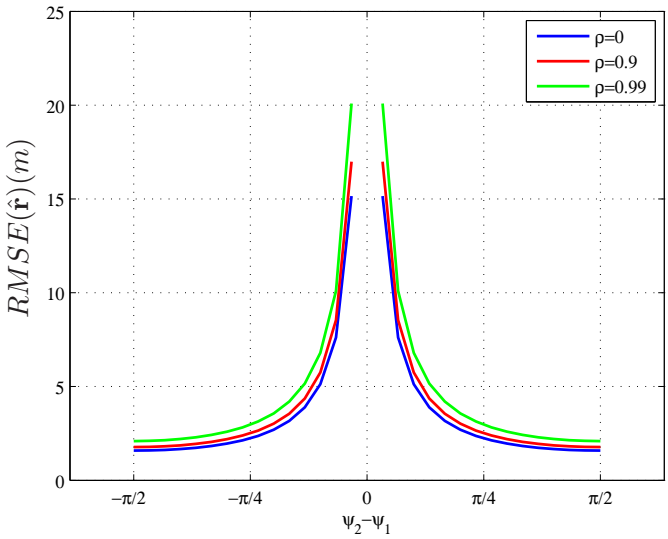

Figure 3.16: RMSE of $\hat{\mathbf{r}}$ as a function of $\psi_2 - \psi_1$, $SNR_1 = 10dB$, $SNR_2 = 10dB$ for the Rayleigh fading case.

Chapter 4

Performance Analysis of Power Delay Doppler Profile Fingerprinting Methods

4.1 Introduction

In this chapter, the aim is to extend the analysis that we have performed for PDP-F to the PDDP-F method. We aim to see the differences on the estimation of the same LDPs for PDDP-F algorithm. Also the CRBs of the estimation of new LDPs are among our targets. We are also interested in observing the local identifiability analysis of the MT position after the inclusion of Doppler shifts of the paths. Since the concepts are very similar to the concepts in chapter 3, we will follow the same methodology. However, we will try to avoid repeating the same things unless they are important.

4.2 CRB Analysis of PDDP-F

We have to revisit the channel model to include the effects of time variation. Now, as the name implies, we have another parameter in the CIR which models the time variation of the channel, which is the Doppler shift of each ray. This new model is presented as:

$$h(t,\tau) = \sum_{i=1}^{N_p} A_i(t)\ e^{j2\pi f_i t}\ p(\tau - \tau_i(t)). \tag{4.1}$$

We are familiar with the other parameters from the previous chapter, except for the Doppler parameter. The Doppler shift of each ray is given by:

$$f_i = \nu \cos(\psi_i - \psi_\nu)/\lambda \tag{4.2}$$

where ν is the magnitude of the velocity, λ is the carrier wavelength, ψ_i and ψ_ν denote the orientation of the incoming wave and the speed vector respectively. Let us sample the CIR and then stack it in a vector. We obtain the following:

$$\mathbf{h}(t) = \begin{bmatrix} h(\tau_s, t) \\ h(2\tau_s, t) \\ \vdots \\ h(N_\tau \tau_s, t) \end{bmatrix} = \sum_{i=1}^{N_p} A_i(t)\ e^{j2\pi f_i t}\ \mathbf{p}_{\tau_i} \tag{4.3}$$

where we implicitly assume that all paths are resolvable, which was also the case in the PDP-F analysis. If we write (4.3) in matrix notation and include the channel estimation noise, we obtain the estimated CIR vector:

$$\hat{\mathbf{h}}(t) = \underbrace{\begin{bmatrix} \mathbf{p}_{\tau_1} \cdots \mathbf{p}_{\tau_{N_p}} \end{bmatrix}}_{\mathbf{P}_\tau} \underbrace{\begin{bmatrix} A_1(t)\ e^{j2\pi f_1 t} \\ \vdots \\ A_{N_p}(t)\ e^{j2\pi f_{N_p} t} \end{bmatrix}}_{\mathbf{b}(t)} + \mathbf{v}(t), \tag{4.4}$$

We will also assume, like in chapter 3, that pulses from different paths are non-overlapping. In other words we assume that $\mathbf{P}_\tau$ is an orthogonal matrix again. For the path amplitudes, there are again two possible models:

- Gaussian model: $A_i(t)$ Gaussian with zero mean, characterized by a power (variance) i.e. $var(A_i) = \sigma_i^2$,
- deterministic model: $A_i(t)$ deterministic unknowns.

The GML-based PDDP-F localization technique utilizing the Gaussian model is explained in detail in [37]. However, for completeness, we will sum up some important concepts.

4.2.1 Rayleigh Fading Case

Let θ represent the vector of location dependent parameters (LDP) with N_p paths. If we consider the delays, the variances of the complex path amplitudes and Doppler shifts as LDPs, θ is defined as:

$$\theta = \left[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ \sigma_1^2,\ \sigma_2^2,\ \cdots\ \sigma_{N_p}^2,\ f_1,\ f_2,\ \cdots\ f_{N_p}\right]^T. \tag{4.5}$$

As we know before, the elements of the FIM for the parameter vector θ is given by (3.7). For the computation of the FIM, we need to derive the covariance matrix of the channel estimates. For the PDDP-F, the covariance matrix is a bit different and includes the effects of the Doppler shifts. Before deriving it we need to introduce some preliminary information of PDDP-F which can be found in [37]. The main idea is to stack consecutive channel estimates in time in a vector. By this way, the Doppler effect can be seen. Now consider the channel response at multiple consecutive time instants $t = t_s, 2t_s, \ldots, nt_s$:

$$\underbrace{\underline{\mathbf{h}}}_{nN_\tau \times 1} = \begin{bmatrix} \mathbf{h}(t_s) \\ \mathbf{h}(2t_s) \\ \vdots \\ \mathbf{h}(nt_s) \end{bmatrix} \tag{4.6}$$

Then we get

$$\underline{\mathbf{h}} = \sum_{i=1}^{N_p} A_i\, \mathbf{e}(f_i) \otimes \mathbf{p}_{\tau_i}\ ,\ \ \mathbf{e}(f) = \begin{bmatrix} e^{j2\pi f t_s} \\ e^{j2\pi f 2t_s} \\ \vdots \\ e^{j2\pi f n t_s} \end{bmatrix} \tag{4.7}$$

We get for the covariance matrix of $\underline{\mathbf{h}}$

$$C_{\underline{\mathbf{h}}\underline{\mathbf{h}}} = \sum_{i=1}^{N_p} \sigma_i^2\, \mathbf{R}_f(f_i) \otimes \mathbf{R}_\tau(\tau_i) \tag{4.8}$$

where

$$\mathbf{R}_f(f) = \mathbf{e}(f)\mathbf{e}^H(f)\ ,\ \ \mathbf{R}_\tau(\tau) = \mathbf{p}_\tau \mathbf{p}_\tau^H. \tag{4.9}$$

We will illustrate this derivation for $n = 2$, and then do the CRB analysis for this case. Extension for any n is similar, but becomes computationally more complex as n increases. For $n = 2$ we have:

$$\underline{\hat{\mathbf{h}}} = \begin{bmatrix} \hat{\mathbf{h}}(t_s) \\ \hat{\mathbf{h}}(2t_s) \end{bmatrix}. \tag{4.10}$$

The covariance matrix is:

$$\mathbb{E}\left\{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}^H\right\} = \mathbf{C}_{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}} = \begin{bmatrix} \mathbb{E}\left\{\hat{\mathbf{h}}(t_s)\hat{\mathbf{h}}(t_s)^H\right\} & \mathbb{E}\left\{\hat{\mathbf{h}}(t_s)\hat{\mathbf{h}}(2t_s)^H\right\} \\ \mathbb{E}\left\{\hat{\mathbf{h}}(2t_s)\hat{\mathbf{h}}(t_s)^H\right\} & \mathbb{E}\left\{\hat{\mathbf{h}}(2t_s)\hat{\mathbf{h}}(2t_s)^H\right\} \end{bmatrix} \tag{4.11}$$

where $\mathbb{E}\left\{\hat{\mathbf{h}}(t_s)\hat{\mathbf{h}}(t_s)^H\right\} = \mathbb{E}\left\{\hat{\mathbf{h}}(2t_s)\hat{\mathbf{h}}(2t_s)^H\right\} = \mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}} = \mathbf{P}_\tau\mathbf{C}_b\mathbf{P}_\tau^H + \sigma_v^2\,\mathbf{I}$, and $\mathbf{C}_b$ is a diagonal matrix with entries $[\sigma_1^2, \sigma_2^2, \cdots, \sigma_{N_p}^2]$. $\mathbb{E}\left\{\hat{\mathbf{h}}(t_s)\hat{\mathbf{h}}(2t_s)^H\right\} = \left(\mathbb{E}\left\{\hat{\mathbf{h}}(2t_s)\hat{\mathbf{h}}(t_s)^H\right\}\right)^H = \mathbf{P}_\tau\mathbf{C}_d\mathbf{P}_\tau^H$ where $\mathbf{C}_d$ is given as:

$$\mathbf{C}_d = \begin{bmatrix} \sigma_1^2 e^{-j2\pi f_1 t_s} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{N_p}^2 e^{-j2\pi f_{N_p} t_s} \end{bmatrix}. \tag{4.12}$$

With these derivations we can write the covariance matrix as:

$$\mathbf{C}_{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}} = \underbrace{\begin{bmatrix} \mathbf{P}_\tau & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_\tau \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} \mathbf{C}_b & \mathbf{C}_d \\ \mathbf{C}_d^H & \mathbf{C}_b \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} \mathbf{P}_\tau^H & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_\tau^H \end{bmatrix}}_{\mathbf{P^H}} + \sigma_v^2\,\mathbf{I}. \tag{4.13}$$

Now with the same reasoning, for the GML-based PDDP-F, the second term in (3.7) vanishes due to the zero mean. So the entries of the FIM are given by (3.10) again. Just for reminding we write it again,

$$[\mathbf{J}_\theta]_{ij} = \operatorname{tr}\left(\mathbf{C}_{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}}^{-1}\frac{\partial\mathbf{C}_{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}}}{\partial\theta_i}\mathbf{C}_{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}}^{-1}\frac{\partial\mathbf{C}_{\underline{\hat{\mathbf{h}}}\underline{\hat{\mathbf{h}}}}}{\partial\theta_j}\right). \tag{4.14}$$

Although we have derived the covariance matrix, we also need its inverse explicitly. By using Woodbury's matrix identity again, and exploiting the orthogonality of matrix $\mathbf{P}$ the inverse is obtained as:

$$\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1} = \sigma_v^{-2}\mathbf{I} - \sigma_v^{-4}\mathbf{P}\underbrace{\left[\frac{\sigma_v^2}{e_p}\mathbf{I} - \frac{\sigma_v^4}{e_p^2}\left(\mathbf{D} + \frac{\sigma_v^2}{e_p}\mathbf{I}\right)^{-1}\right]}_{\mathbf{G}}\mathbf{P}^H. \tag{4.15}$$

In fact we made a simple trick here, and used the Woodbury's matrix identity twice to compute the inverse of the covariance matrix. The reason to do this is to avoid ending up in an equation involving the inverse of matrix $\mathbf{D}$. Because $\mathbf{D}$ is a singular positive semi-definite matrix. To see that we simply check its determinant:

$$\det(\mathbf{D}) = \det(\mathbf{C_b}\,\mathbf{C_b} - \mathbf{C_d}\,\mathbf{C_d^H}). \tag{4.16}$$

The reason why we can write the determinant in the form above is due to the commutative property of the diagonal matrices. We realize that moreover, $\mathbf{C_d}\,\mathbf{C_d^H} = \mathbf{C_b^2}$, which makes the determinant zero resulting in a singular matrix $\mathbf{D}$. Hence in (4.15), there is no term involving the inverse of $\mathbf{D}$ directly. The inverse results after adding a scaled identity matrix to $\mathbf{D}$, which guarantees invertibility.

4.2.1.1 Anisotropic Path Amplitude Variances

We start with the anisotropic case again. For the computation of the FIM entries, partial derivatives of the covariance matrix with respect to the parameters are required:

$$\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \sigma_i^2} = \mathbf{P}\begin{bmatrix} \mathbf{e}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i \end{bmatrix}\underbrace{\begin{bmatrix} 1 & z_i \\ z_i^* & 1 \end{bmatrix}}_{\mathbf{Z_i}}\begin{bmatrix} \mathbf{e}_i^T & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i^T \end{bmatrix}\mathbf{P}^H, \tag{4.17}$$

$$\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial f_i} = \mathbf{P}\begin{bmatrix} \mathbf{e}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i \end{bmatrix}\underbrace{\begin{bmatrix} 0 & s_i \\ s_i^* & 0 \end{bmatrix}}_{\mathbf{S_i}}\begin{bmatrix} \mathbf{e}_i^T & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i^T \end{bmatrix}\mathbf{P}^H, \tag{4.18}$$

$$\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \tau_i} = -\mathbf{P}'\begin{bmatrix} \mathbf{e}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i \end{bmatrix}\underbrace{\begin{bmatrix} [\mathbf{C}_b]_{i,i} & [\mathbf{C}_d]_{i,i} \\ [\mathbf{C}_d^H]_{i,i} & [\mathbf{C}_b]_{i,i} \end{bmatrix}}_{\mathbf{D_i}}\begin{bmatrix} \mathbf{e}_i^T & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i^T \end{bmatrix}\mathbf{P}^H \tag{4.19}$$

$$-\mathbf{P}\begin{bmatrix} \mathbf{e}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i \end{bmatrix}\begin{bmatrix} [\mathbf{C}_b]_{i,i} & [\mathbf{C}_d]_{i,i} \\ [\mathbf{C}_d^H]_{i,i} & [\mathbf{C}_b]_{i,i} \end{bmatrix}\begin{bmatrix} \mathbf{e}_i^T & \mathbf{0} \\ \mathbf{0} & \mathbf{e}_i^T \end{bmatrix}\mathbf{P}'^H, \tag{4.20}$$

where $z_i = e^{-j2\pi f_i t_s}$, $s_i = -j2\pi t_s \sigma_i^2 z_i$, $\mathbf{D}_i = \sigma_i^2 \mathbf{Z}_i$ and $\mathbf{e}_i$ is the $N_p \times 1$ elementary vector having 1 in its ith position. After having calculated these partial derivatives, we can begin deriving the elements of the FIM.

$$\begin{aligned} \mathbf{J}_{\tau_i,\tau_i} &= \operatorname{tr}\left(\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1} \frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \tau_i} \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1} \frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \tau_i} \right) && (4.21) \\ &= 2\sigma_v^{-4} \left(e_p\, e_d - b^2\right) \left(\operatorname{tr}\left(\mathbf{D}_i \mathbf{D}_i\right) - \sigma_v^{-2} e_p \operatorname{tr}\left(\mathbf{G}_i \mathbf{D}_i \mathbf{D}_i\right)\right) && (4.22) \end{aligned}$$

where $\mathbf{G}_i$ is defined similarly as $\mathbf{D}_i$. To obtain the matrix $\mathbf{G}_i$, we will get use of the block matrix inversion lemma:

$$\begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Z} & \mathbf{T} \end{bmatrix}^{-1} = \begin{bmatrix} \left(\mathbf{X} - \mathbf{Y}\mathbf{T}^{-1}\mathbf{Z}\right)^{-1} & -\mathbf{X}^{-1}\mathbf{Y}\left(\mathbf{T} - \mathbf{Z}\mathbf{X}^{-1}\mathbf{Y}\right)^{-1} \\ -\mathbf{T}^{-1}\mathbf{Z}\left(\mathbf{X} - \mathbf{Y}\mathbf{T}^{-1}\mathbf{Z}\right)^{-1} & \left(\mathbf{T} - \mathbf{Z}\mathbf{X}^{-1}\mathbf{Y}\right)^{-1} \end{bmatrix}. \tag{4.23}$$

By using this block matrix inversion lemma and replacing the matrix $\mathbf{D}$ defined in (4.13), we obtain $\mathbf{G}_i$ as follows:

$$\mathbf{G}_i = \begin{bmatrix} \sigma_i^2\sigma_v^2/(2e_p\sigma_i^2 + \sigma_v^2) & \sigma_i^2\sigma_v^2 z_i/(2e_p\sigma_i^2 + \sigma_v^2) \\ \sigma_i^2\sigma_v^2 z_i^*/(2e_p\sigma_i^2 + \sigma_v^2) & \sigma_i^2\sigma_v^2/(2e_p\sigma_i^2 + \sigma_v^2) \end{bmatrix} = \gamma_i \mathbf{Z}_i, \tag{4.24}$$

with $\gamma_i = \dfrac{\sigma_i^2\sigma_v^2}{2e_p\sigma_i^2 + \sigma_v^2}$. Hence with $\mathbf{G}_i$ and $\mathbf{D}_i$, we can calculate:

$$\begin{aligned} \operatorname{tr}\left(\mathbf{D}_i\mathbf{D}_i\right) &= 4\sigma_i^4, && (4.25) \\ \operatorname{tr}\left(\mathbf{G}_i\mathbf{D}_i\mathbf{D}_i\right) &= 8\gamma_i\sigma_i^4. && (4.26) \end{aligned}$$

All these at hand, we obtain:

$$\mathbf{J}_{\tau_i,\tau_i} = \frac{4\sigma_i^4\sigma_v^{-2}(e_p e_d - b^2)}{e_p\,\sigma_i^2 + \sigma_v^2/2}. \tag{4.27}$$

Check the similarity between the corresponding entry for the PDP-F case in (3.26). As expected, the information has increased since we are now effectively using two channel estimates. Information is now higher than twice its value in the PDP-F section. In fact if we replaced the channel estimation error variance with $\sigma_v^2/2$ in (3.26), we would have obtained the same result above. Hence this reminds us the noise averaging effect. We continue the computation of the other entries of the FIM.

$$
\begin{aligned}
\mathbf{J}_{\tau_i,\sigma_i^2} &= \operatorname{tr}\left(\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \tau_i}\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \sigma_i^2}\right) && (4.28)\\
&= -2\,a\,e_p\,\sigma_v^{-4}\,(\operatorname{tr}(\mathbf{D}_i\mathbf{Z}_i) && (4.29)\\
&- \sigma_v^{-2}e_p\ \operatorname{tr}(\mathbf{D}_i\mathbf{G}_i\mathbf{Z}_i)\\
&- \sigma_v^{-2}e_p\ \operatorname{tr}(\mathbf{G}_i\mathbf{D}_i\mathbf{Z}_i)\\
&+ \sigma_v^{-4}e_p^2\ \operatorname{tr}(\mathbf{G}_i\mathbf{D}_i\mathbf{G}_i\mathbf{Z}_i)\,)\\
&= 0, && (4.30)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{J}_{\tau_i,f_i} &= \operatorname{tr}\left(\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \tau_i}\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial f_i}\right) && (4.31)\\
&= -2\,a\,e_p\,\sigma_v^{-4}\,(\operatorname{tr}(\mathbf{D}_i\mathbf{S}_i) && (4.32)\\
&- \sigma_v^{-2}e_p\ \operatorname{tr}(\mathbf{D}_i\mathbf{G}_i\mathbf{S}_i)\\
&- \sigma_v^{-2}e_p\ \operatorname{tr}(\mathbf{G}_i\mathbf{D}_i\mathbf{S}_i)\\
&+ \sigma_v^{-4}e_p^2\ \operatorname{tr}(\mathbf{G}_i\mathbf{D}_i\mathbf{G}_i\mathbf{S}_i)\,)\\
&= 0 && (4.33)
\end{aligned}
$$

since $a = 0$. This is an important result. Estimate of the path delay is uncorrelated with the estimate of the path Doppler shift. We go on with the other entries of the FIM.

$$
\begin{aligned}
\mathbf{J}_{\sigma_i^2,\sigma_i^2} &= \operatorname{tr}\left(\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \sigma_i^2}\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial \sigma_i^2}\right) && (4.34)\\
&= \sigma_v^{-4}e_p^2\ (\operatorname{tr}(\mathbf{Z}_i\mathbf{Z}_i) && (4.35)\\
&- e_p\sigma_v^{-2}\ \operatorname{tr}(\mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i)\\
&- e_p\sigma_v^{-2}\ \operatorname{tr}(\mathbf{G}_i\mathbf{Z}_i\mathbf{Z}_i)\\
&+ e_p^2\sigma_v^{-4}\ \operatorname{tr}(\mathbf{G}_i\mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i)\,)
\end{aligned}
$$

where

$$\begin{aligned}
\text{tr}\ (\mathbf{Z}_i\mathbf{Z}_i) &= 4, && (4.36)\\
\text{tr}\ (\mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i) &= 8\gamma_i, && (4.37)\\
\text{tr}\ (\mathbf{G}_i\mathbf{Z}_i\mathbf{Z}_i) &= 8\gamma_i, && (4.38)\\
\text{tr}\ (\mathbf{G}_i\mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i) &= 16\gamma_i^2. && (4.39)
\end{aligned}$$

Hence we get:

$$\mathbf{J}_{\sigma_i^2,\sigma_i^2} = \left(\frac{e_p}{e_p\,\sigma_i^2 + \sigma_v^2/2}\right)^2. \tag{4.40}$$

Once more, the reader should compare it with its corresponding entry in (3.32). Obviously substituting $\sigma_v^2/2$ for the channel estimation error variance would end up in the same result above. For the remaining entries of the FIM, we continue to derive:

$$\begin{aligned}
\mathbf{J}_{\sigma_i^2,f_i} &= \text{tr}\left(\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial\sigma_i^2}\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\frac{\partial\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}}{\partial f_i}\right) && (4.41)\\
&= \sigma_v^{-4}e_p^2\ (\,\text{tr}\ (\mathbf{Z}_i\mathbf{S}_i) && (4.42)\\
&- e_p\sigma_v^{-2}\ \text{tr}\ (\mathbf{Z}_i\mathbf{G}_i\mathbf{S}_i)\\
&- e_p\sigma_v^{-2}\ \text{tr}\ (\mathbf{G}_i\mathbf{Z}_i\mathbf{S}_i)\\
&+ e_p^2\sigma_v^{-4}\ \text{tr}\ (\mathbf{G}_i\mathbf{Z}_i\mathbf{G}_i\mathbf{S}_i)\ )
\end{aligned}$$

where

$$\begin{aligned}
\text{tr}\ (\mathbf{Z}_i\mathbf{S}_i) &= 0, && (4.43)\\
\text{tr}\ (\mathbf{Z}_i\mathbf{G}_i\mathbf{S}_i) &= 0, && (4.44)\\
\text{tr}\ (\mathbf{G}_i\mathbf{Z}_i\mathbf{S}_i) &= 0, && (4.45)\\
\text{tr}\ (\mathbf{G}_i\mathbf{Z}_i\mathbf{G}_i\mathbf{S}_i) &= 0, && (4.46)
\end{aligned}$$

resulting in $\mathbf{J}_{\sigma_i^2,f_i} = 0$. We obtained a similar result for the delay and Doppler estimation. Now the same situation occurs for the path amplitude. Hence, the path amplitude variance estimate is uncorrelated with the Doppler shift estimate. There is one remaining entry of the FIM to be computed which we derive as following:

$$\begin{aligned}
\mathbf{J}_{f_i,f_i} &= \text{tr}\left(\mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}}{\partial f_i}\mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}^{-1}\frac{\partial \mathbf{C}_{\underline{\widehat{\mathbf{h}}}\underline{\widehat{\mathbf{h}}}}}{\partial f_i}\right) && (4.47)\\
&= \sigma_v^{-4}e_p^2\ (\text{tr}\ (\mathbf{S}_i\mathbf{S}_i) && (4.48)\\
&- e_p\sigma_v^{-2}\ \text{tr}\ (\mathbf{S}_i\mathbf{G}_i\mathbf{S}_i)\\
&- e_p\sigma_v^{-2}\ \text{tr}\ (\mathbf{G}_i\mathbf{S}_i\mathbf{S}_i)\\
&+ e_p^2\sigma_v^{-4}\ \text{tr}\ (\mathbf{G}_i\mathbf{S}_i\mathbf{G}_i\mathbf{S}_i)\ )
\end{aligned}$$

where

$$\begin{aligned}
\text{tr}\ (\mathbf{S}_i\mathbf{S}_i) &= 8\pi^2t_s^2\sigma_i^4, && (4.49)\\
\text{tr}\ (\mathbf{S}_i\mathbf{G}_i\mathbf{S}_i) &= 8\gamma_i\pi^2t_s^2\sigma_i^4, && (4.50)\\
\text{tr}\ (\mathbf{G}_i\mathbf{S}_i\mathbf{S}_i) &= 8\gamma_i\pi^2t_s^2\sigma_i^4, && (4.51)\\
\text{tr}\ (\mathbf{G}_i\mathbf{S}_i\mathbf{G}_i\mathbf{S}_i) &= 0, && (4.52)
\end{aligned}$$

which results in

$$\mathbf{J}_{f_i,f_i} = \frac{8\pi^2t_s^2\,e_p^2\,\sigma_i^4\sigma_v^{-2}}{2\,e_p\,\sigma_i^2+\sigma_v^2}. \tag{4.53}$$

Due to the non-overlapping pulse assumption, the terms corresponding to different paths are null, e.g., $\mathbf{J}_{f_i,f_j}=0$ or $\mathbf{J}_{\tau_i,\sigma_j^2}=0$ for $i\neq j$. The derivation of the FIM entries has been completed. We can now check the conditions for local identifiability and also derive the CRBs. First we investigate the $N_p=1$ case again. In this case FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & 0 & 0\\ 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & 0\\ 0 & 0 & \mathbf{J}_{f_1,f_1}\end{bmatrix} \tag{4.54}$$

Clearly local identifiability of $\mathbf{r}$ is achieved. When $e_p\,e_d\neq b^2$, $\mathbf{J}_\theta$ is full rank for any N_p. For a real (or symmetric) pulse shape, it is always full rank. It is easy to derive the CRBs because of the diagonal nature of the FIM. We just need the inverses of the FIM entries. Note that, for the estimation of the delays, having a real (or symmetric) pulse makes the estimation easier. In this case, the CRB for the estimation of τ_i is given as:

$$\mathbb{E}(\tau_i-\hat{\tau}_i)^2 \geq \frac{1}{16\pi^2W^2SNR_i}\left(1+\frac{1}{2\,SNR_i}\right) \tag{4.55}$$

where $SNR_i = \frac{e_p \sigma_i^2}{\sigma_v^2}$. As we mentioned before, estimating the delay is now easier than it was for the PDP-F case. The CRBs of the other parameters are derived as:

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \sigma_i^4 \left(1 + \frac{1}{2\,SNR_i}\right)^2 \tag{4.56}$$

$$\mathbb{E}(f_i - \hat{f}_i)^2 \geq \frac{1}{4\pi^2\, t_s^2\, SNR_i}\left(1 + \frac{1}{2\,SNR_i}\right). \tag{4.57}$$

It is noteworthy to check the similarity between the CRBs of delay and the Doppler parameter. The delay estimation improves with the increasing bandwidth (W), while the Doppler estimation improves when t_s increases. This is a reasonable result because when t_s increases, then the phase change will also increase. Therefore it will be easier to extract the Doppler shift from the change of phase. But there is one important thing to pay attention here. If t_s increases too much or in other words the channel estimates are obtained less frequently, then this may lead to aliasing for the estimation of f_i, because the highest Doppler shift that can be estimated is bounded by $0.5/t_s$.

4.2.1.1.1 Effect of the Delay Offset It is worthwhile to investigate the effects of the DO also for PDDP-F method. The reason to explore this is that now we expect to achieve local identifiability with only 1 path. This is because we now have an additional parameter, namely f_i, which also carries position dependent information. We know that in case of a DO, delay does not carry any information about position. However with σ_1^2 and f_1, we expect to have local identifiability with $N_p = 1$. We will not repeat the calculations that we have carried out in the PDP-F method because we just need some little modifications and additions. Instead we will provide a brief explanation of the situation now. The LDP vector to be estimated is (for $N_p = 1$):

$$\theta = \left[\tau_{1_0},\ \sigma_1^2,\ f_1,\ \tau_0\right].^T \tag{4.58}$$

Obviously $\mathbf{J}_{\tau_{i_0},\tau_{i_0}}$ is given by (4.27), and $\mathbf{J}_{\sigma_i^2,\sigma_i^2}$ still by (4.40), and $\mathbf{J}_{\tau_{i_0},\sigma_i^2} = 0$. We also have: $\mathbf{J}_{\tau_{i_0},\tau_0} = \mathbf{J}_{\tau_{i_0},\tau_{i_0}}$, $\mathbf{J}_{\tau_0,\tau_0} = \sum_{i=1}^{N_p} \mathbf{J}_{\tau_{i_0},\tau_{i_0}}$ and $\mathbf{J}_{\tau_0,\sigma_i^2} = 0$. Hence for $N_p = 1$, the FIM will be:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \\ 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & 0 & 0 \\ 0 & 0 & \mathbf{J}_{f_1,f_1} & 0 \\ \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} \end{bmatrix}. \tag{4.59}$$

For matrix $\mathbf{F}$, it has the following structure (for $N_p = 1$):

$$\mathbf{F} = \begin{bmatrix} \mathbf{f1} & \mathbf{f2} & \mathbf{f3} & -\mathbf{f1} \end{bmatrix}. \tag{4.60}$$

Consequently, all these lead to

$$\mathbf{J_r} = \mathbf{F}\mathbf{J}_\theta\mathbf{F}^T = \mathbf{J}_{\sigma_1^2,\sigma_1^2}\mathbf{f2}\,\mathbf{f2}^T + \mathbf{J}_{f_1,f_1}\mathbf{f3}\,\mathbf{f3}^T \tag{4.61}$$

which is clearly rank two. Hence, with $N_p = 1$, local identifiability of $\mathbf{r}$ can be achieved even in the case of a DO for PDDP-F method. As mentioned before, this is due to the Doppler parameter which is not present in the PDP-F method.

4.2.1.1.2 Effect of the Frequency Offset In communication systems, when there is a mismatch between the frequencies of the transmitter and receiver oscillators (when the receiver oscillator frequency is not exactly the same as the carrier frequency), another defect arises. This is known as the frequency offset (FO) problem which we also mentioned in chapter 2. This is analogous to the DO problem in the sense that it now adds an offset f_0 to the Doppler shifts, i.e., $f_{i_0} = f_i + f_0$. We will pursue a similar approach that we did for the DO case. For $N_p = 1$, the LDP vector is:

$$\theta = \left[\tau_1,\ \sigma_1^2,\ f_{1_0},\ f_0\right].^T \tag{4.62}$$

Obviously $\mathbf{J}_{\tau_i,\tau_i}$ is still given by (4.27), and $\mathbf{J}_{\sigma_i^2,\sigma_i^2}$ still by (4.40). We also have: $\mathbf{J}_{f_{i_0},f_0} = \mathbf{J}_{f_{i_0},f_{i_0}}$, $\mathbf{J}_{f_0,f_0} = \sum_{i=1}^{N_p} \mathbf{J}_{f_{i_0},f_{i_0}}$. Hence for $N_p = 1$, the FIM will be:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & 0 & 0 & 0 \\ 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & 0 & 0 \\ 0 & 0 & \mathbf{J}_{f_{1_0},f_{1_0}} & \mathbf{J}_{f_{1_0},f_{1_0}} \\ 0 & 0 & \mathbf{J}_{f_{1_0},f_{1_0}} & \mathbf{J}_{f_{1_0},f_{1_0}} \end{bmatrix}. \tag{4.63}$$

The matrix $\mathbf{F}$ has the following structure, for $N_p = 1$:

$$\mathbf{F} = \left[\begin{array}{cccc} \mathbf{f1} & \mathbf{f2} & \mathbf{f3} & -\mathbf{f3} \end{array} \right]. \tag{4.64}$$

Consequently,

$$\mathbf{J_r} = \mathbf{F}\mathbf{J}_\theta\mathbf{F}^T = \mathbf{J}_{\tau_1,\tau_1}\mathbf{f1}\,\mathbf{f1}^T + \mathbf{J}_{\sigma_1^2,\sigma_1^2}\mathbf{f2}\,\mathbf{f2}^T \tag{4.65}$$

which is obviously rank two. Hence, with $N_p = 1$, local identifiability of $\mathbf{r}$ can be achieved even in the case of a FO. It is clearly seen that all the information comes from delay and path amplitudes. Due to the FO, Doppler shift does not carry any information about position.

4.2.1.1.3 Effect of Delay and Frequency Offset Simultaneously

Naturally, there is the possibility that both DO and FO problems take place at the same time. To investigate this case, we will get use of the results of the previous delay and FO sections. For $N_p = 1$, the LDP vector is:

$$\theta = \left[\tau_{1_0},\ \sigma_1^2,\ f_{1_0},\ \tau_0,\ f_0\right].^T \tag{4.66}$$

And for $N_p = 1$, the FIM will be:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 \\ 0 & \mathbf{J}_{\sigma_1^2,\sigma_1^2} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{J}_{f_{1_0},f_{1_0}} & 0 & \mathbf{J}_{f_{1_0},f_{1_0}} \\ \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 & 0 & \mathbf{J}_{\tau_{1_0},\tau_{1_0}} & 0 \\ 0 & 0 & \mathbf{J}_{f_{1_0},f_{1_0}} & 0 & \mathbf{J}_{f_{1_0},f_{1_0}} \end{bmatrix}. \tag{4.67}$$

Additionally, the matrix $\mathbf{F}$, has the following structure for $N_p = 1$:

$$\mathbf{F} = \left[\begin{array}{ccccc} \mathbf{f1} & \mathbf{f2} & \mathbf{f3} & -\mathbf{f1} & -\mathbf{f3} \end{array} \right]. \tag{4.68}$$

In this case, we obtain:

$$\mathbf{J_r} = \mathbf{F}\mathbf{J}_\theta\mathbf{F}^T = \mathbf{J}_{\sigma_1^2,\sigma_1^2}\mathbf{f2}\,\mathbf{f2}^T, \tag{4.69}$$

which is rank one. This is an expected result. Having offsets both in delay and frequency resulted in a lack of knowledge from them. In this case, information about position only comes from the path amplitude which eventually resulted in a rank one FIM for the position vector $\mathbf{r}$. Hence for $N_p = 1$, it is not possible to achieve the local identifiability of $\mathbf{r}$ when we have both DO and FO simultaneously.

4.2.1.2 Isotropic Path Amplitude Variances

The same phenomenon that was investigated for the PDP-F method is now examined for PDDP-F method. The path amplitude variances are modeled as distance dependent, i.e., $\sigma_i^2 = \frac{k}{\tau_i^\gamma}$. We will not repeat the same procedure here. But the difference from the PDP-F method is now the inclusion of the Doppler parameter into the LDP vector. For the LDP, we have only delays and Doppler shifts. For $N_p = 1$, FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J'}_{\tau_1,\tau_1} & 0 \\ 0 & \mathbf{J'}_{f_1,f_1} \end{bmatrix}. \tag{4.70}$$

It is easy to see that $\mathbf{J'}_{f_i,f_i}$ is the same as in the anisotropic case, i.e., $\mathbf{J'}_{f_i,f_i} = \mathbf{J}_{f_i,f_i}$. For the delays, we obtain the same relationship as we had in the PDP-F case:

$$\mathbf{J'}_{\tau_i,\tau_i} = \mathbf{J}_{\tau_i,\tau_i} + \eta_i^2 \mathbf{J}_{\sigma_i^2,\sigma_i^2}. \tag{4.71}$$

Clearly $rank(\mathbf{J}_\theta) = 2$ for $N_p = 1$. There is now the possibility of achieving the local identifiability of $\mathbf{r}$ with only one path. Remember that it was not possible to do that with one path for PDP-F method. However we should check matrix $\mathbf{F}$ with the inclusion of the partial derivatives of f_i's. For $N_p = 1$:

$$\mathbf{F} = \begin{bmatrix} -\frac{1}{c}\cos\psi_1 & \frac{\partial f_1}{\partial x} \\ -\frac{1}{c}\sin\psi_1 & \frac{\partial f_1}{\partial y} \end{bmatrix}. \tag{4.72}$$

where we have:

$$\frac{\partial f_i}{\partial x} = -\frac{\nu}{\lambda}\sin(\psi_i - \psi_\nu)\frac{\partial \psi_i}{\partial x} \tag{4.73}$$

$$\frac{\partial f_i}{\partial y} = -\frac{\nu}{\lambda}\sin(\psi_i - \psi_\nu)\frac{\partial \psi_i}{\partial y} \tag{4.74}$$

where ψ_i which is defined in (3.136) or (3.137), is a function of the MT's coordinates. After substituting ψ_i, we obtain:

$$\frac{\partial \psi_i}{\partial x} / \frac{\partial \psi_i}{\partial y} = -\frac{y - y_{s_{in_i}}}{x - x_{s_{in_i}}} \tag{4.75}$$

whereas

$$\cos\psi_i/\sin\psi_i = \frac{x - x_{s_{in_i}}}{y - y_{s_{in_i}}}, \tag{4.76}$$

which guarantees the full rankness of $\mathbf{F}$ (last two ratios given above cannot be equal to each other). Hence for $N_p = 1$, it is possible to achieve the local identifiability of the position vector $\mathbf{r}$. This result is different from what we have obtained in the corresponding section of PDP-F method. Now local identifiability is possible even with one path. The reason is that now each path carries two distinct information instead of one. One information comes from the Doppler shift of the path, while the other information from the delay (or the amplitude variance as they are coupled). CRBs are easy to derive. The CRB for the estimation of f_i is still given by (4.57). For the delay we have:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}'_{\tau_i,\tau_i}} = \frac{1}{\mathbf{J}_{\tau_i,\tau_i} + \eta_i^2 \mathbf{J}_{\sigma_i^2,\sigma_i^2}}, \tag{4.77}$$

and the CRB for the amplitude variance is:

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \frac{1}{\mathbf{J}_{\sigma_i^2,\sigma_i^2} + \mathbf{J}_{\tau_i,\tau_i}/\eta_i^2} \tag{4.78}$$

by using the transformation of parameters technique.

4.2.2 Deterministic Path Amplitude Case

Now the second alternative for modeling the path amplitudes is investigated as we have also done in the PDP-F case. We revisit (4.4) for the channel model and write the complex path amplitude of path i in as $A_i(t) = a_i(t)e^{j\phi_i(t)}$. We will only deal with the case where the phases are considered as deterministic unknowns. We will investigate the case of real (or symmetric) pulses, which makes $b = 0$. As we will see soon, we now have to include the phases into the LDPs, because there will be correlation between the phases and the Doppler shifts. Moreover, the results will be given for any n ($n > 1$). Now, in this situation, our LDP vector is:

$$\theta = \left[\tau_1,\ \tau_2,\ \cdots\ \tau_{N_p},\ a_1,\ a_2,\ \cdots\ a_{N_p},\ f_1,\ f_2,\ \cdots\ f_{N_p},\ \phi_1,\ \phi_2,\ \cdots\ \phi_{N_p}\right].^T \tag{4.79}$$

Because of the deterministic path amplitudes, the mean of the channel estimates is not zero: $\mu = \mathbf{P}\begin{bmatrix} \mathbf{b}(0) \\ \vdots \\ \mathbf{b}((n-1)t_s) \end{bmatrix}$. It is important to say that size of $\mathbf{P}$ and $\mathbf{P}'$ is changed accordingly $(nN_\tau \times nN_p)$ and all the related matrices, e.g. $\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}$ is now an $nN_\tau \times nN_\tau$ matrix and is given by: $\mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}} = \sigma_v^2\mathbf{I}$. Under these circumstances, FIM is given as:

$$[\mathbf{J}_\theta]_{ij} = 2\Re\left(\left[\frac{\partial\mu}{\partial\theta_i}\right]^H \mathbf{C}_{\underline{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}}^{-1}\left[\frac{\partial\mu}{\partial\theta_j}\right]\right). \tag{4.80}$$

We again consider the case where the pulses from different paths are non-overlapping.

4.2.2.1 Anisotropic Path Amplitudes

Before deriving the FIM entries, one needs the partial derivatives:

$$\frac{\partial\mu}{\partial\tau_i} = -\mathbf{P}'\begin{bmatrix} \mathbf{e}_i & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{e}_i \end{bmatrix}\begin{bmatrix} a_i\ e^{j\phi_i} \\ \vdots \\ a_i\ e^{j\phi_i}\ e^{j2\pi(n-1)f_it_s} \end{bmatrix}, \tag{4.81}$$

$$\frac{\partial\mu}{\partial a_i} = \mathbf{P}\begin{bmatrix} \mathbf{e}_i & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{e}_i \end{bmatrix}\begin{bmatrix} e^{j\phi_i} \\ \vdots \\ e^{j\phi_i}\ e^{j2\pi(n-1)f_it_s} \end{bmatrix}, \tag{4.82}$$

$$\frac{\partial\mu}{\partial f_i} = \mathbf{P}\begin{bmatrix} \mathbf{e}_i & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{e}_i \end{bmatrix}\begin{bmatrix} 0 \\ \vdots \\ j2\pi(n-1)\ t_s\ a_i\ e^{j\phi_i}\ e^{j2\pi f_it_s} \end{bmatrix}. \tag{4.83}$$

After computing the partials, FIM entries can be derived as follows:

$$\mathbf{J}_{\tau_i,\tau_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial\tau_i}\right]\right) \tag{4.84}$$

$$= \frac{2}{\sigma_v^2}\, n\, a_i^2\, e_d, \tag{4.85}$$

$$\begin{aligned}
\mathbf{J}_{\tau_i,a_i} &= \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial a_i}\right]\right) && (4.86)\\
&= -\frac{2}{\sigma_v^2}\,n\,a_i\,a && (4.87)\\
&= 0, && (4.88)
\end{aligned}$$

$$\begin{aligned}
\mathbf{J}_{\tau_i,f_i} &= \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial f_i}\right]\right) && (4.89)\\
&= -\frac{2}{\sigma_v^2}\,n\,(n-1)\,a_i^2\,b\,\pi\,t_s && (4.90)\\
&= 0. && (4.91)
\end{aligned}$$

This is an interesting result. For the Rayleigh fading case, there was no correlation between the estimates of τ_i and f_i for any pulse shape, but here, because of the real (or symmetric) pulse assumption, it is null. The correlation depends on the pulse. We derive the other entries of the FIM:

$$\begin{aligned}
\mathbf{J}_{\tau_i,\phi_i} &= \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\tau_i}\right]^H\left[\frac{\partial\mu}{\partial\phi_i}\right]\right) && (4.92)\\
&= -\frac{2}{\sigma_v^2}\,n\,a_i^2\,b && (4.93)\\
&= 0, && (4.94)
\end{aligned}$$

due to the same reason.

$$\begin{aligned}
\mathbf{J}_{a_i,a_i} &= \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial a_i}\right]^H\left[\frac{\partial\mu}{\partial a_i}\right]\right) && (4.95)\\
&= \frac{2}{\sigma_v^2}\,n\,e_p, && (4.96)
\end{aligned}$$

$$\begin{aligned}
\mathbf{J}_{a_i,f_i} &= \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial a_i}\right]^H\left[\frac{\partial\mu}{\partial f_i}\right]\right) && (4.97)\\
&= 0, && (4.98)
\end{aligned}$$

$$\mathbf{J}_{a_i,\phi_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial a_i}\right]^H\left[\frac{\partial\mu}{\partial\phi_i}\right]\right) \tag{4.99}$$

$$= 0, \tag{4.100}$$

$$\mathbf{J}_{f_i,f_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial f_i}\right]^H\left[\frac{\partial\mu}{\partial f_i}\right]\right) \tag{4.101}$$

$$= \frac{4}{3\,\sigma_v^2}\,n\,(n-1)\,(2\,n-1)\,a_i^2\,e_p\,\pi^2\,t_s^2, \tag{4.102}$$

$$\mathbf{J}_{f_i,\phi_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial f_i}\right]^H\left[\frac{\partial\mu}{\partial\phi_i}\right]\right) \tag{4.103}$$

$$= \frac{2}{\sigma_v^2}\,n\,(n-1)\,a_i^2\,e_p\,\pi\,t_s, \tag{4.104}$$

$$\mathbf{J}_{\phi_i,\phi_i} = \frac{2}{\sigma_v^2}\Re\left(\left[\frac{\partial\mu}{\partial\phi_i}\right]^H\left[\frac{\partial\mu}{\partial phi_i}\right]\right) \tag{4.105}$$

$$= \frac{2}{\sigma_v^2}\,n\,a_i^2\,e_p. \tag{4.106}$$

We have completed the derivation of the FIM. We know that the cross entries corresponding to different paths are null. For $N_p = 1$, FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & 0 & 0 & 0 \\ 0 & \mathbf{J}_{a_1,a_1} & 0 & 0 \\ 0 & 0 & \mathbf{J}_{f_1,f_1} & \mathbf{J}_{f_1,\phi_1} \\ 0 & 0 & \mathbf{J}_{f_1,\phi_1} & \mathbf{J}_{\phi_1,\phi_1} \end{bmatrix}. \tag{4.107}$$

This matrix is always full rank, which guarantees the local identifiability of $\mathbf{r}$. To derive the CBRs for the LDPs for any $N_p > 1$, we first check the structure of the FIM for $N_p > 1$:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \mathbf{J}_{a_1,a_1} & & & & & & \\ & & & \ddots & & & & & \\ & & & & \mathbf{J}_{f_1,f_1} & & & \mathbf{J}_{f_1,\phi_1} & & \\ & & & & & \ddots & & & \ddots & \\ & & & & & & \mathbf{J}_{f_{N_p},f_{N_p}} & & & \mathbf{J}_{f_{N_p},\phi_{N_p}} \\ & & & & \mathbf{J}_{f_1,\phi_1} & & & \mathbf{J}_{\phi_1,\phi_1} & & \\ & & & & & \ddots & & & \ddots & \\ & & & & & & \mathbf{J}_{f_{N_p},\phi_{N_p}} & & & \mathbf{J}_{\phi_{N_p},\phi_{N_p}} \end{bmatrix}.$$

To derive the CRBs for the estimation of LDPs, we need to compute the diagonal entries of the inverse of $\mathbf{J}_\theta$. For this purpose, we will use Cramer's rule.

$$\left[\mathbf{J}_\theta{}^{-1}\right]_{i,i} = \frac{\det\left(\tilde{\mathbf{J}}_{\theta[ii]}\right)}{\det\left(\mathbf{J}_\theta\right)}, \tag{4.108}$$

where $\tilde{\mathbf{J}}_{\theta[ii]}$ is the matrix obtained by deleting the ith row and i^{th} column of $\mathbf{J}_\theta$. Due to the special structure of $\mathbf{J}_\theta$ we can easily make this calculation. We might compute its determinant either via Laplace expansion in a recursive manner or via making elementary row and column swapping operations, to obtain a matrix whose determinant is easy to compute. We prefer the second one. We know that swapping two rows (or two columns) has the effect of multiplying the determinant by -1. We will make row swapping N_p times, and also column swapping N_p times. So $2N_p$ swapping operations leaves the determinant unchanged. After these operations we obtain the following matrix:

$$\mathbf{J}'_\theta = \begin{bmatrix} \mathbf{J}_{\tau_1,\tau_1} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \mathbf{J}_{a_1,a_1} & & & & & & \\ & & & \ddots & & & & & \\ & & & & \mathbf{J}_{f_1,f_1} & \mathbf{J}_{f_1,\phi_1} & & & \\ & & & & \mathbf{J}_{f_1,\phi_1} & \mathbf{J}_{\phi_1,\phi_1} & & & \\ & & & & & & \ddots & & \\ & & & & & & & \ddots & \\ & & & & & & & \mathbf{J}_{f_{N_p},f_{N_p}} & \mathbf{J}_{f_{N_p},\phi_{N_p}} \\ & & & & & & & \mathbf{J}_{f_{N_p},\phi_{N_p}} & \mathbf{J}_{\phi_{N_p},\phi_{N_p}} \end{bmatrix},$$

whose determinant can be computed easily. As clearly seen, the submatrix obtained by selecting the first $2N_p$ rows and $2N_p$ columns is a diagonal matrix, and the rest is the block diagonal part. Now we can apply the Cramer's rule. Computing the CRBs for τ_i's and a_i's is easier, because they just appear as a single factor in the determinant. When that row and column are deleted, then only this single factor will disappear. Hence the ratio will be the inverse of the single factor, e.g., $\det(\mathbf{J}_\theta) = \mathbf{J}_{\tau_i,\tau_i} \det\left(\tilde{\mathbf{J}}_{\theta[ii]}\right)$. Hence we have:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J}_{\tau_i,\tau_i}} = \frac{\sigma_v^2}{2\,n\,a_i^2\,e_d} \tag{4.109}$$

$$= \frac{1}{8\,n\,\pi^2 W^2 SNR_i}, \tag{4.110}$$

where $SNR_i = \frac{e_p\,a_i^2}{\sigma_v^2}$. Similarly, we derive:

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{1}{\mathbf{J}_{a_i,a_i}} = \frac{\sigma_v^2}{2\,n\,e_p} \tag{4.111}$$

$$= \frac{a_i^2}{2\,nSNR_i}. \tag{4.112}$$

By using the Cramer's rule again, we derive the CRBs for the estimation of f_i's and ϕ_i's. For example, for the CRB of f_i, when we delete the row and column where $\mathbf{J}_{f_i,f_i}$ is located, only $\mathbf{J}_{\phi_i,\phi_i}$ will remain from the 2×2 block. Hence, the ratio in the Cramer's rule will be:

$$\mathbb{E}(f_i - \hat{f}_i)^2 \geq \frac{1}{\mathbf{J}_{f_i,f_i} - \mathbf{J}^2_{f_i,\phi_i}/\mathbf{J}_{\phi_i,\phi_i}} = \frac{3}{2\,n\,(n^2-1)\,\pi^2\,t_s^2\,SNR_i}, \tag{4.113}$$

and similarly for ϕ_i:

$$\mathbb{E}(\phi_i - \hat{\phi}_i)^2 \geq \frac{1}{\mathbf{J}_{\phi_i,\phi_i} - \mathbf{J}^2_{f_i,\phi_i}/\mathbf{J}_{f_i,f_i}} = \frac{2\,n-1}{n\,(n+1)\,SNR_i}. \tag{4.114}$$

In order to prevent aliasing, we must still impose f_i's are below $0.5/t_s$. When we replace $n = 2$, and compare it with the corresponding CRB of the Rayleigh fading case, we see that estimating f_i's is easier than the Rayleigh fading case. Naturally we see that it is impossible to estimate the Doppler shift with just one channel estimate (when $n = 1$). This is also parallel to the Rayleigh fading case. In that case, the effects of the Doppler shifts are visible in the covariance matrix of the channel estimates only when $n \geq 2$. Another comment is that with the increase of n, estimating Doppler shift becomes easier than delay, amplitude or phase estimation. This is because the CRB for the others are inversely proportional to n, while for the Doppler shift it is inversely proportional to n^3. In figures 4.1 and 4.2, we provide the CRB plots for the estimation of f_i and ϕ_i respectively for $n = 4$. We use $t_s = 10$ ms, which is reasonable for a channel with not so high Doppler spread (coherence time is not very short).

4.2.2.2 Isotropic Path Amplitudes

As we have done before, we now model the path amplitudes as $a_i = \frac{m}{\tau_i^{\gamma/2}}$ again. We have $\frac{da_i}{d\tau_i} = \kappa_i = -\frac{m\gamma}{2}\tau_i^{-(\gamma/2+1)}$. The real pulse assumption is still valid. LDP vector will consist of delays, Doppler shifts and the phases. By applying the transformation of parameters technique, the CRBs for the path amplitudes will be derived from the CRBs of the delays as carried out before. For $N_p = 1$, FIM is:

$$\mathbf{J}_\theta = \begin{bmatrix} \mathbf{J'}_{\tau_1,\tau_1} & 0 & 0 \\ 0 & \mathbf{J'}_{f_1,f_1} & \mathbf{J'}_{f_1,\phi_1} \\ 0 & \mathbf{J'}_{f_1,\phi_1} & \mathbf{J'}_{\phi_1,\phi_1} \end{bmatrix}. \tag{4.115}$$

Obviously, the FIM entries for Doppler shifts and phases are the same as in the anisotropic case. Hence the CRBs for the estimation of Doppler shifts and phases are still given by Equations (4.113) and (4.114) respectively. For

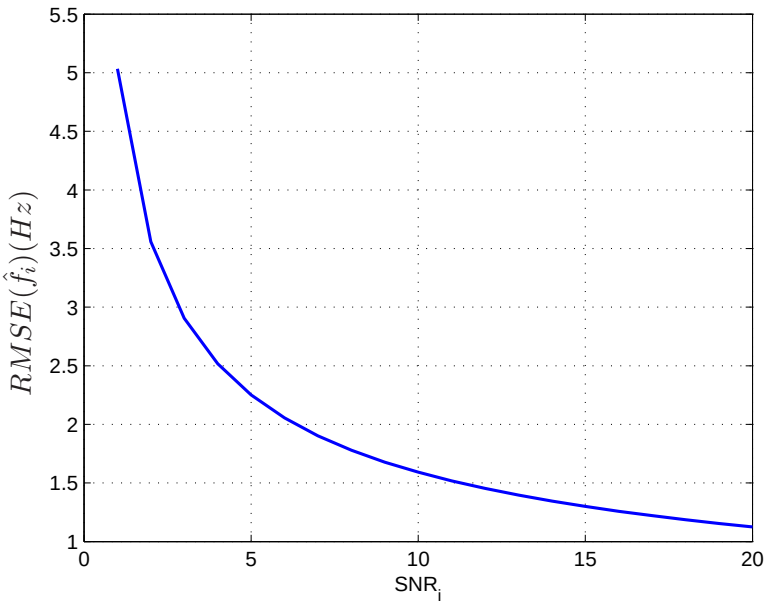

Figure 4.1: RMSE of $\hat{f}_i$ as a function of SNR_i.

the delays and path amplitudes, we first need the following partial derivative to compute the FIM entries:

$$\frac{\partial \mu}{\partial a_i} = \mathbf{P} \begin{bmatrix} \mathbf{e}_i & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{e}_i \end{bmatrix} \begin{bmatrix} e^{j\phi_i} \\ \vdots \\ e^{j\phi_i}\ e^{j2\pi(n-1)f_i t_s} \end{bmatrix} \tag{4.116}$$

$$-\frac{1}{\kappa_i}\ \mathbf{P}' \begin{bmatrix} \mathbf{e}_i & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{e}_i \end{bmatrix} \begin{bmatrix} a_i\ e^{j\phi_i} \\ \vdots \\ a_i\ e^{j\phi_i}\ e^{j2\pi(n-1)f_i t_s} \end{bmatrix}. \tag{4.117}$$

Also $\frac{\partial \mu}{\partial \tau_i} = \kappa_i \frac{\partial \mu}{\partial a_i}$. By exploiting these, we obtain:

$$\mathbf{J}'_{\tau_i,\tau_i} = \mathbf{J}_{\tau_i,\tau_i} + \kappa_i^2 \mathbf{J}_{a_i,a_i} = \frac{2\,n}{\sigma_v^2}(a_i^2\, e_d + \kappa_i^2 e_p), \tag{4.118}$$

where $\mathbf{J}_{\tau_i,\tau_i}$ and $\mathbf{J}_{a_i,a_i}$ are the anisotropic FIM entries. The CRBs are derived as:

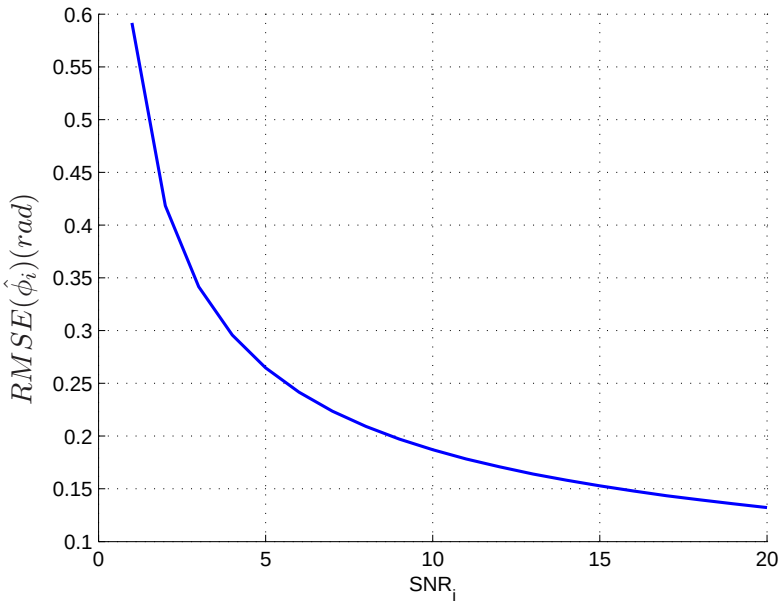

Figure 4.2: RMSE of $\hat{\phi}_i$ as a function of SNR_i.

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{\mathbf{J'}_{\tau_i,\tau_i}} = \frac{1}{\mathbf{J}_{\tau_i,\tau_i} + \kappa_i^2 \mathbf{J}_{a_i,a_i}}, \tag{4.119}$$

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{1}{\mathbf{J}_{a_i,a_i} + \mathbf{J}_{\tau_i,\tau_i}/\kappa_i^2}. \tag{4.120}$$

Hence, estimating both the delay and the path amplitude is easier than the anisotropic case, as expected.

4.3 Conclusion

We focused mainly on the CRBs of the LDPs. As one can expect, estimation performances of the same LDPs are improved when compared with the PDP-F method. This is due to the fact that we are effectively using more channel estimates (hence more data). Another remarkable conclusion is that, mobility of the MT, hence the integration of the Doppler shifts into the scenario enhanced the local identifiability of the MT position. This is because Doppler shifts also provide information about the position which

in turn improves the local identifiability of MT position, e.g. identifiability might be achieved even with one path for isotropic modeling. Readers can ask why we did not mention the localization performance of the PDDP-F. In fact, we implicitly talked about it in chapter 3. There, we have provided generic formulas to see the CRB of the position estimation as a function of the information obtained from the delays. The only thing to be done here is to substitute the new information values obtained for PDDP-F algorithm. We just did not want to repeat the same things here. Needless to say, localization performance of PDDP-F is strictly superior to PDP-F algorithm due to the higher information. However, as it is usually the case, there is a trade-off between complexity and performance. In addition to DO, FO problem might also be present in PDDP-F. Unless both impairments occur simultaneously, local identifiability might still be achieved with only one path.

Chapter 5

Pairwise Error Probability Analysis for Power Delay Profile Fingerprinting Methods

5.1 Introduction

We now deal with a new problem in the field of localization. PEP performance of PDP-F algorithm will be studied. In fact, PEP analysis is a well-known concept in digital communications, e.g., calculating the probability of error when a vector of symbols $\mathbf{s_m}$ is transmitted but another vector of symbols $\mathbf{s_n}$ is detected at the receiver [38]. Under different path amplitude modelings and cost functions, we will try to obtain closed-form results for PEP.

5.2 PEP of the LS Technique for Deterministic Path Amplitudes

The specular SISO CIR is given as:

$$h(t,\tau) = \sum_{i=1}^{N_p} A_i(t)\ p(\tau - \tau_i(t)). \tag{5.1}$$

We already know the parameters from previous chapters. We can write the complex path amplitude of path i in polar form as $A_i(t) = a_i(t)e^{j\phi_i(t)}$. Let us now consider sampling the CIR with a sampling period of τ_s leading to N_τ samples and stacking them in a vector as follows:

$$\mathbf{h}(t) = \begin{bmatrix} h(\tau_s, t) \\ h(2\tau_s, t) \\ \vdots \\ h(N_\tau \tau_s, t) \end{bmatrix} = \sum_{i=1}^{N_p} A_i(t)\ \mathbf{p}_{\tau_i}, \tag{5.2}$$

where $\mathbf{p}_\tau$ is defined in (3.3). If we write (5.2) in matrix notation and include the channel estimation noise, we obtain the estimated CIR vector as:

$$\hat{\mathbf{h}}(t) = \underbrace{\begin{bmatrix} \mathbf{p}_{\tau_1} \cdots \mathbf{p}_{\tau_{N_p}} \end{bmatrix}}_{\mathbf{P}_\tau} \underbrace{\begin{bmatrix} A_1(t) \\ \vdots \\ A_{N_p}(t) \end{bmatrix}}_{\mathbf{a}(t)} + \mathbf{v}(t). \tag{5.3}$$

where $\mathbf{v}(t)$ is the complex additive white Gaussian noise vector with covariance matrix $\sigma_v^2\,\mathbf{I}$. For the techniques relying on PDPs, PDP estimates are classically obtained by averaging the magnitude squared version of the CIR over a time duration. However, if the mobile moves rapidly and/or some paths are not resolvable (due to the limited bandwidth of the pulse-shape $p(t)$, path contributions can overlap), the averaging gives a poor PDP estimation, and then a poor location accuracy. We estimate the PDP vector as:

$$\widehat{\mathbf{PDP}} = \frac{1}{T} \sum_{t=1}^{T} \left| \hat{\mathbf{h}}(t) \right|^2 \tag{5.4}$$

where T is the number of channel observations. There is one thing that needs to be clarified that the absolute squaring operation is element-wise. Hence the resulting PDP estimate is another vector having the same length as the channel estimates. For the path amplitudes, we consider two possibilities:

- deterministic model: $A_i(t)$ deterministic unknowns

- Gaussian model: $A_i(t)$ Gaussian with zero mean, characterized by a power (variance) i.e. $var(A_i) = \sigma_i^2$, which corresponds to Rayleigh fading for the magnitudes.

We will consider the first case now where the path amplitudes are considered as deterministic unknowns. The PEP can be defined as follows when the LS criteria is the cost function:

$$\text{PEP} = \Pr\left(||\widehat{\mathbf{PDP}} - \mathbf{PDP}_F|| < ||\widehat{\mathbf{PDP}} - \mathbf{PDP}_T||\right), \tag{5.5}$$

where $\widehat{\mathbf{PDP}}$ is the estimated PDP vector defined in (5.4), $\mathbf{PDP}_T$ is the true PDP vector which is computed off-line from the stored database and $\mathbf{PDP}_F$ is the PDP vector to be erroneously detected. Every position in the database (Ray Tracing database or any other pre-computed database) is distinguishable from each other, e.g. they have either different number of paths, or path delays or amplitudes (variances) are different. Hence there are unique entries in the database so that fingerprinting can work correctly. $\mathbf{PDP}_T$ and $\mathbf{PDP}_F$ are given as:

$$\mathbf{PDP}_T = \mathbf{PDP}_{true} + \sigma_v^2\mathbf{1}, \quad \mathbf{PDP}_F = \mathbf{PDP}_{false} + \sigma_v^2\mathbf{1}, \tag{5.6}$$

where $\mathbf{1}$ is a vector of all 1's which is added to include the effects of the noise. $\mathbf{PDP}_{true}$ and $\mathbf{PDP}_{false}$ are computed with delays and amplitudes of paths in the database by averaging over random phases. They are given as:

$$\mathbf{PDP}_{true} = \sum_{i=1}^{N_p} a_i^2 \left|\mathbf{p}_{\tau_i}\right|^2, \quad \mathbf{PDP}_{false} = \sum_{i=1}^{L} b_i^2 \left|\mathbf{p}_{\zeta_i}\right|^2, \tag{5.7}$$

Based on (5.4), $\widehat{\mathbf{PDP}}$ can be calculated as:

$$\begin{aligned}\widehat{\mathbf{PDP}} &= \frac{1}{T}\sum_{t=1}^{T}\Big(|\mathbf{v}(t)|^2 + 2(\Re\mathbf{h}(t)\odot\Re\mathbf{v}(t) + \Im\mathbf{h}(t)\odot\Im\mathbf{v}(t))\Big) \\ &+ \mathbf{PDP}_{true},\end{aligned} \tag{5.8}$$

where $\odot$ stands for the Hadamard element-wise multiplication. In this case, it is used to multiply the corresponding elements of the real and imaginary parts of the noise and channel vectors. We will make an important simplification in (5.8) and assume that the terms in the innermost parentheses tend to go to $\mathbf{0}$. In other words they will be replaced by their expected values as

the noise is a zero mean process. In fact it is a reasonable assumption when the number of observations T is high. Hence $\widehat{\mathbf{PDP}}$ is now approximated as:

$$\widehat{\mathbf{PDP}} \approx \mathbf{PDP}_{true} + \frac{1}{T}\sum_{t=1}^{T} |\mathbf{v}(t)|^2 . \tag{5.9}$$

After these calculations and definitions we can turn back to the PEP formulation. In fact PEP can also be stated equivalently as:

$$\text{PEP} = \Pr\left(||\widehat{\mathbf{PDP}} - \mathbf{PDP}_F||^2 < ||\widehat{\mathbf{PDP}} - \mathbf{PDP}_T||^2\right). \tag{5.10}$$

This equivalent formulation is easier to deal with. For simplicity of notation, let us call $\widehat{\mathbf{PDP}}$ as $\mathbf{x}$, $\mathbf{PDP}_{false}$ as $\mathbf{y}$ and $\mathbf{PDP}_{true}$ as $\mathbf{z}$. Then PEP becomes:

$$\Pr\left(\mathbf{x}^T(\mathbf{z}-\mathbf{y}) < \frac{(\mathbf{z}+\sigma_{\mathbf{v}}^{\mathbf{2}}\mathbf{1})^T(\mathbf{z}+\sigma_{\mathbf{v}}^{\mathbf{2}}\mathbf{1}) - (\mathbf{y}+\sigma_{\mathbf{v}}^{\mathbf{2}}\mathbf{1})^T(\mathbf{y}+\sigma_{\mathbf{v}}^{\mathbf{2}}\mathbf{1})}{2}\right). \tag{5.11}$$

If we check (5.9), we immediately recognize that first term of the equation is deterministic while the second term is the random part. If we do the algebra, we can reorganize (5.11) as:

$$\text{PEP} = \Pr\left((\mathbf{z}-\mathbf{y})^T\sum_{t=1}^{T}|\mathbf{v}(t)|^2 < \frac{T}{2}(2k_2 - k_1 - k_3 + 2\,M)\right) \tag{5.12}$$

where $k_1 = \mathbf{z}^T\mathbf{z}$, $k_2 = \mathbf{y}^T\mathbf{z}$, $k_3 = \mathbf{y}^T\mathbf{y}$ and $M = \sigma_v^2\mathbf{1}^T(\mathbf{z}-\mathbf{y})$. Here k_2 is an important parameter which gives information about the overlapping between the vectors. As it is clear, it is always non-negative. It can be 0 if and only if the vectors do not overlap with each other at all. Mathematical formulation of PEP is almost complete. When we explore (5.12), it is a summation of random variables on the left hand side. We can divide the analysis for each turn of T. Let us call the random variable as W_i for the ith loop. So the left hand side as a result becomes a random variable W which is $W = \sum_{i=1}^{T} W_i$. However finding the distribution of W is not easy as we will see later. Therefore we will just compute the distribution of a W_i (mean and variance of W_1 without loss of generality). And then we will call the central limit theorem (CLT) for W as all the W_i's are identically distributed. Remember that $\mathbf{v}(t)$ is a complex white Gaussian noise vector. Hence each element of the $|\mathbf{v}(t)|^2$ vector is composed of sums of squares

of two zero mean Gaussian random variables. It is well known that this leads to the exponential distribution with mean $1/\lambda = \sigma_v^2$, i.e., $f_M(m) = \lambda e^{-m\lambda}$, $m \geq 0$ [39]. Therefore W_1 will be a summation of exponential random variables. However they all have different parameters (different λ's) because of the multiplication by the vector $(\mathbf{z} - \mathbf{y})^T$ which makes the calculation of the overall distribution more difficult. In other words it would be a summation of independent but not identically distributed exponential random variables. If all had the same parameters, we know that this leads to the Erlang distribution [40]. The distribution of W_1 which is a summation of K exponential random variables with means $1/\lambda_i$'s is derived as (see appendix B for the derivation):

$$f_{W_1}(u) = \left(\prod_{i=1}^{K} \lambda_i\right) \left(\sum_{j=1}^{K} \frac{e^{-\lambda_j u}}{\prod\limits_{\substack{l=1 \\ l \neq j}}^{K} (\lambda_l - \lambda_j)}\right). \tag{5.13}$$

However deriving the distribution of W which is a summation of T of these random variables (W_i's) would be very difficult. We can also compute the probability for $T = 1$ with the obtained derivation. However in that case the assumption that we have done in (5.9) will be disturbed. Due to these reasons, we will call the CLT for these T (T being large) i.i.d. random variables as we mentioned before. Before applying the CLT, we have to know the mean and variance of W_i's. By using (5.12), we determine the mean and variance of W_i's as follows:

$$\mu_{W_i} = M, \tag{5.14}$$

$$\sigma_{W_i}^2 = \sigma_v^4 \left(k_1 + k_3 - 2k_2\right). \tag{5.15}$$

By CLT, $\dfrac{W - T\mu_{W_i}}{\sigma_{W_i}\sqrt{T}}$ will tend to have a standard normal distribution ($\mathcal{N}(0,1)$) when T is large. Hence PEP can be reformulated as:

$$\begin{aligned}
\text{PEP} &= \Pr\left(\frac{W - T\mu_{W_i}}{\sigma_{W_i}\sqrt{T}} < \frac{\sqrt{T}}{2\sigma_{W_i}}(2k_2 - k_1 - k_3)\right) \\
&= Q\left(\frac{\sqrt{T}}{2\sigma_v^2}\sqrt{k_1 + k_3 - 2k_2}\right) \qquad (5.16) \\
&= Q\left(\frac{\sqrt{T}}{2\sigma_v^2}||\mathbf{z} - \mathbf{y}||\right). \qquad (5.17)
\end{aligned}$$

And by using the Chernoff bound for the Q function, we can bound the PEP as:

$$\text{PEP} \leq \frac{1}{2}e^{-\frac{T}{8\sigma_v^4}||\mathbf{z}-\mathbf{y}||^2}. \qquad (5.18)$$

We see that PEP decreases when the norm of the difference between the true and false PDPs increase. In fact it is a reasonable result. When they become more and more apart from each other, one can expect that it will be less likely to confuse the true PDP with the false one. The interesting thing is that we reached this result after the approximation given by (5.9) and by the use of the CLT.

5.2.1 Simulation Results

Now, the aim is to justify the theoretical analysis developed so far with simulations. For the simulations, we have used $T = 50$, sampling frequency f_s $(= 1/\tau_s)$ equal to 9.14 MHz, and also we assumed $\mathbf{P}_\zeta = \mathbf{P}_\tau$. Below we present three plots. The SNR calculation is carried out with respect to the CIR of the false position's entry. Hence σ_v^2 is calculated accordingly. For these three plots, the path amplitude coefficients of the true PDP vector are just a multiple of the false PDP vector. Let us call this factor as α. Hence when α increases, the separation between the true and false PDP increases, i.e., $||\mathbf{z} - \mathbf{y}||$ becomes higher. In accordance with the theoretical analysis, PEP is decreasing with the increasing α as can be seen in the plots.

What we see in the plots is that simulations and theoretical analysis are very close to each other in all SNR ranges. This shows that our approximations and the application of the CLT are quite accurate. Also effect of α is clearly seen in the plots. For $\alpha = 1.1$, PEP approaches to 0 around an SNR

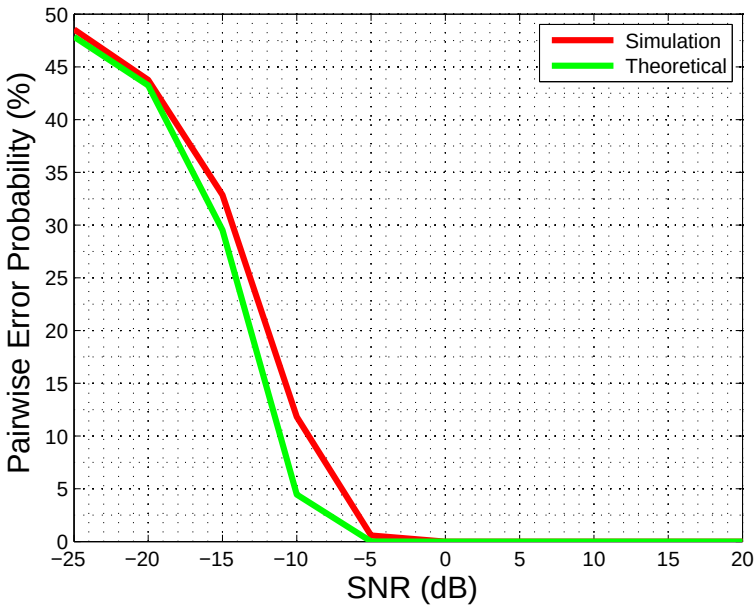

Figure 5.1: Pairwise error probability for $\alpha = 1.1$.

of -5dB. However for $\alpha = 1.4$, PEP approaches to 0 much before, around -15dB. Another comment that we can infer from the plots is that at very low SNR, PEP is almost equal to 0.5 which means that due to the very high noise, PDP's cannot be differentiated much and hence selecting the true or false PDP is dominated mostly by noise which results in an equal probability for true or the false one to be selected.

5.3 PEP of the GML Technique for Rayleigh Fading

In this part, we investigate the PEP analysis for the GML based PDP-F technique. We also have a different assumption for the complex path amplitudes $A_i(t)$. Instead of modeling them as deterministic unknowns, we now model them as complex Gaussian random variables (Rayleigh distribution for the magnitudes). For a complete description of this PDP-F method, readers can refer to [26]. The channel model that we have proposed in the previous section is still valid and given by (5.3). We now assume that pulses

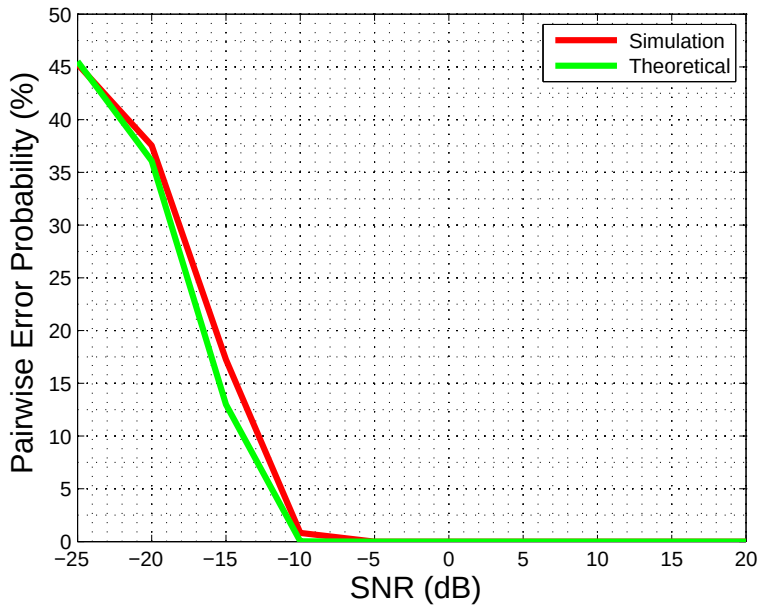

Figure 5.2: Pairwise error probability for $\alpha = 1.2$.

from different paths are non-overlapping ($\mathbf{P}_\tau$ is an orthogonal matrix) to simplify the analysis which is a reasonable assumption in high bandwidth systems. The matching criteria is based on Gaussian log-likelihood. Hence formulation of the PEP is such that the probability that the log-likelihood performed in the true position is lower than the log-likelihood in the false position which results in the false position to be selected. If we have multiple channel estimates, the log-likelihood can be expressed as:

$$\mathcal{LL} \propto -\ln\left(\det\left(\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}\right)\right) - \operatorname{tr}\left(\hat{\mathbf{C}}\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}^{-\mathbf{1}}\right) \tag{5.19}$$

where $\hat{\mathbf{C}} = \frac{1}{T}\sum_{i=1}^{T}\left(\hat{\mathbf{h}}_\mathbf{i} - \mu\right)\left(\hat{\mathbf{h}}_\mathbf{i} - \mu\right)^H$ is the sample covariance matrix obtained from channel estimates. Since the complex path amplitudes $A_i(t)$ and the noise have both zero mean, channel estimates have also zero mean, i.e., $\mu = \mathbf{0}$. For simplicity of notation, let us call $\mathbf{C}_{\hat{\mathbf{h}}\hat{\mathbf{h}}}$ as $\mathbf{C_T}$ which denotes the covariance matrix calculated with the true positions' entries. By using (5.3), we have $\mathbf{C_T} = \mathbf{P}_\tau\mathbf{C_a}\mathbf{P}_\tau^\mathbf{H} + \sigma_\mathbf{v}^\mathbf{2}\mathbf{I}$ where $\mathbf{C}_a$ is a diagonal matrix having $\left[\sigma_{a_1}^2, \sigma_{a_2}^2, \cdots, \sigma_{a_{N_p}}^2\right]$ on its diagonal ($var(A_i) = \sigma_{a_i}^2$). We also introduce $\mathbf{C_F}$

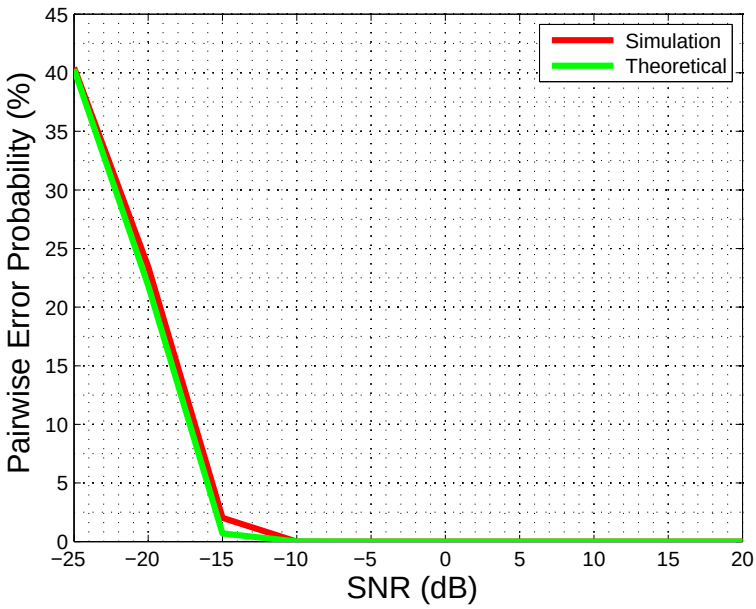

Figure 5.3: Pairwise error probability for $\alpha = 1.4$.

for the covariance matrix computed with the false positions' entries as $\mathbf{C_F} = \mathbf{P}_\zeta \mathbf{C_b} \mathbf{P}_\zeta^{\mathbf{H}} + \sigma_{\mathbf{v}}^{\mathbf{2}} \mathbf{I}$ where $\mathbf{C}_b$ is a diagonal matrix having $\left[\sigma_{b_1}^2, \sigma_{b_2}^2, \cdots, \sigma_{b_L}^2\right]$ on its diagonal and $\mathbf{P}_\zeta$ is defined similarly as $\mathbf{P}_\tau$. After giving the necessary information, we can state PEP as:

$$\text{PEP} = \Pr\left(\mathcal{L}\mathcal{L}_T < \mathcal{L}\mathcal{L}_F\right). \tag{5.20}$$

We know that there are many scenarios to investigate. However we will try to explore the scenario where the error probability is more likely to occur. And also the scenario proposed will also simplify the analysis.

Scenario: The scenario can be summarized as follows:

1. Number of paths are equal, i.e., $L = N_p$.

2. Path delays are equal, i.e., $\tau_i = \zeta_i \ \forall i$.

3. There is no delay synchronization error.

Under these assumptions $\mathbf{P}_\zeta = \mathbf{P}_\tau$. We see that the only differences between the true and the false positions' parameters are the path amplitude variances. By using (5.19) we can restate PEP as:

$$\text{PEP} = \Pr\left(\ln\left(\det \mathbf{C_T}/\det \mathbf{C_F}\right) > \text{tr}\left(\hat{\mathbf{C}}\mathbf{C_F^{-1}}\right) - \text{tr}\left(\hat{\mathbf{C}}\mathbf{C_T^{-1}}\right)\right). \quad (5.21)$$

Under the assumption that $\mathbf{P}_\tau$ being an orthogonal matrix, the determinants can be easily calculated.

$$\begin{aligned}\det\left(\mathbf{C}_T\right) &= \sigma_v^{2N_\tau} \det\left(\mathbf{I} + \frac{1}{\sigma_v^2}\mathbf{P}_\tau \mathbf{C}_a \mathbf{P}_\tau^H\right) \\ &= \sigma_v^{2N_\tau} \prod_{i=1}^{N_p}\left(1 + \frac{e_p \sigma_{a_i}^2}{\sigma_v^2}\right), \end{aligned} \quad (5.22)$$

where we have used the Sylvester's determinant theorem, $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$ and $e_p = S$ is the pulse energy. One thing to note is that the determinant does not depend on the path delays when the pulses are non-overlapping. Similarly we obtain $\det\left(\mathbf{C}_F\right)$. Hence left hand side of (5.21) is:

$$\ln\left(\det \mathbf{C_T}/\det \mathbf{C_F}\right) = \sum_{i=1}^{N_p} \ln\left(\frac{\sigma_v^2 + e_p\sigma_{a_i}^2}{\sigma_v^2 + e_p\sigma_{b_i}^2}\right) = f_1. \quad (5.23)$$

For the inversion of $\mathbf{C}_T$ and $\mathbf{C}_F$ we will use the Woodbury's matrix inversion lemma. We get:

$$\mathbf{C}_T^{-1} = \sigma_v^{-2}(\mathbf{I} - \mathbf{P}_\tau \mathbf{D}_a \mathbf{P}_\tau^H) \quad (5.24)$$
$$\mathbf{C}_F^{-1} = \sigma_v^{-2}(\mathbf{I} - \mathbf{P}_\tau \mathbf{D}_b \mathbf{P}_\tau^H) \quad (5.25)$$

where $\mathbf{D}_a$ and $\mathbf{D}_b$ are diagonal matrices having $[\mathbf{D}_a]_{ii} = \dfrac{\sigma_{a_i}^2}{\sigma_v^2 + e_p\sigma_{a_i}^2}$ and $[\mathbf{D}_b]_{ii} = \dfrac{\sigma_{b_i}^2}{\sigma_v^2 + e_p\sigma_{b_i}^2}$ on their diagonals respectively. By using (5.3) we can write $\hat{\mathbf{C}}$ as:

$$\begin{aligned}
\hat{\mathbf{C}} &= \frac{1}{T}\sum_{i=1}^{T}\hat{\mathbf{h}}_{\mathbf{i}}\hat{\mathbf{h}}_{\mathbf{i}}^{H} = \frac{1}{T}\sum_{i=1}^{T}(\mathbf{P}_\tau\mathbf{a}_i+\mathbf{v}_i)(\mathbf{P}_\tau\mathbf{a}_i+\mathbf{v}_i)^H \\
&= \frac{1}{T}\sum_{i=1}^{T}\left(\mathbf{P}_\tau\mathbf{a}_i\mathbf{a}_i^H\mathbf{P}_\tau^H+\mathbf{P}_\tau\mathbf{a}_i\mathbf{v}_i^H+\mathbf{v}_i\mathbf{a}_i^H\mathbf{P}_\tau^H+\mathbf{v}_i\mathbf{v}_i^H\right) \\
&\approx \frac{1}{T}\sum_{i=1}^{T}\left(\mathbf{P}_\tau\mathbf{a}_i\mathbf{a}_i^H\mathbf{P}_\tau^H+\mathbf{v}_i\mathbf{v}_i^H\right) \qquad (5.26)
\end{aligned}$$

where in the last equation we have made an approximation based on the fact that noise samples and channel coefficients are uncorrelated zero mean Gaussian random variables. Hence for large T we replaced them with their expectations resulting in $\mathbf{0}$. With these at hand, the trace functions can be evaluated by using (5.24), (5.25) and (5.26). By exploiting the properties of the trace function and also the orthogonality of $\mathbf{P}_\tau$ we obtain:

$$\begin{aligned}
\mathrm{tr}\left(\hat{\mathbf{C}}\mathbf{C}_{\mathbf{T}}^{-\mathbf{1}}\right) = \frac{\sigma_v^{-2}}{T}\Bigg[&e_p\ \mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{a}_i\mathbf{a}_i^H\right)+\mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{v}_i\mathbf{v}_i^H\right) \\
&-e_p^2\ \mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{a}_i\mathbf{a}_i^H\mathbf{D}_a\right)-\mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{v}_i\mathbf{v}_i^H\mathbf{P}_\tau\mathbf{D}_a\mathbf{P}_\tau^H\right)\Bigg].
\end{aligned}$$

We assume that random variables are uncorrelated in time. It is well known that distribution remains the same under orthonormal transformations. Therefore we realize that $\mathbf{v}_i$ and $\mathbf{w}_i = \frac{1}{\sqrt{e_p}}\mathbf{P}_\tau^H\mathbf{v}_i$ have the same distribution ($\mathcal{N}(0,\sigma_v^2\mathbf{I})$). However size of the vector changes (size of $\mathbf{I}$ also changes). By this transformation we rewrite the above equation:

$$\begin{aligned}
\mathrm{tr}\left(\hat{\mathbf{C}}\mathbf{C}_{\mathbf{T}}^{-\mathbf{1}}\right) = \frac{\sigma_v^{-2}}{T}\Bigg[&e_p\ \mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{a}_i\mathbf{a}_i^H\right)+\mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{v}_i\mathbf{v}_i^H\right) \\
&-e_p^2\ \mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{a}_i\mathbf{a}_i^H\mathbf{D}_a\right)-e_p\ \mathrm{tr}\left(\sum_{i=1}^{T}\mathbf{w}_i\mathbf{w}_i^H\mathbf{D}_a\right)\Bigg].
\end{aligned}$$

Similarly we can derive $\mathrm{tr}\left(\hat{\mathbf{C}}\mathbf{C}_{\mathbf{F}}^{-\mathbf{1}}\right)$. The term we need in (5.21) is:

$$\mathrm{tr}\left(\hat{\mathbf{C}}\mathbf{C}_{\mathbf{F}}^{-\mathbf{1}}-\hat{\mathbf{C}}\mathbf{C}_{\mathbf{T}}^{-\mathbf{1}}\right) = \frac{e_p\,\sigma_v^{-2}}{T}\left(e_p\sum_{i=1}^{T}\mathbf{a}_i^H\mathbf{D}\mathbf{a}_i+\sum_{i=1}^{T}\mathbf{w}_i^H\mathbf{D}\mathbf{w}_i\right). \qquad (5.27)$$

where $\mathbf{D} = \mathbf{D}_a - \mathbf{D}_b$ being another diagonal matrix. Each element of $\mathbf{w}_k$ and $\mathbf{a}_k$ are complex Gaussian random variables with mean 0. For $\mathbf{w}_k$, every entry has the variance σ_v^2 while i^{th} entry of $\mathbf{a}_k$ has a variance of $\sigma_{a_i}^2$. The matrix $\mathbf{D}$ being diagonal simplifies the analysis substantially. It prevents the coupling of the cross elements of the vectors. Therefore (5.27) represents a summation of squares of Gaussian random variables weighted by $\mathbf{D}$. Since $\mathbf{a}_k^H \mathbf{D} \mathbf{a}_k = \sum_{j=1}^{N_p} [\mathbf{D}]_{jj} |a_{kj}|^2$ and $\mathbf{w}_k^H \mathbf{D} \mathbf{w}_k = \sum_{j=1}^{N_p} [\mathbf{D}]_{jj} |w_{kj}|^2$, each loop of (by loop any of the T iterations is meant) (5.27) is composed of summation of non-identically distributed exponential random variables. One important thing to mention is that in order to consider it as a summation we implicitly assume that $\mathbf{D}$ has all positive elements on its diagonal meaning that $\sigma_{a_i}^2 > \sigma_{b_i}^2 \forall i$. In the previous section that distribution was calculated and given by (5.13). As we have done in the previous section, let us call this distribution as W_i, and let $W = \sum_{i=1}^{T} W_i$ (all W_i's identically distributed). Since the derivation of the distribution of summation of T of them (W) will be difficult, we will call the CLT again for T being large. Before that we need the mean and variance of W_i which is calculated as:

$$\mu_{W_i} = \frac{e_p}{T} \sum_{i=1}^{N_p} [\mathbf{D}]_{ii} \left(e_p \, \sigma_v^{-2} \sigma_{a_i}^2 + 1 \right) = \frac{e_p}{T} \, f_2, \tag{5.28}$$

$$\sigma_{W_i}^2 = \frac{e_p^2}{T^2} \sum_{i=1}^{N_p} [\mathbf{D}]_{ii}^2 \left(e_p^2 \, \sigma_v^{-4} \sigma_{a_i}^4 + 1 \right) = \frac{e_p^2}{T^2} \, f_3. \tag{5.29}$$

We know that $\dfrac{W - T\mu_{W_i}}{\sigma_{W_i}\sqrt{T}}$ will tend to have a standard normal distribution ($\mathcal{N}(0,1)$) when T is large. Hence PEP can be reformulated as:

$$\text{PEP} = \Pr\left(\frac{W - T\mu_{W_i}}{\sigma_{W_i}\sqrt{T}} < \frac{\sqrt{T}}{e_p \sqrt{f_3}} \left(f_1 - e_p \, f_2 \right) \right) \tag{5.30}$$

$$= Q\left(\frac{\sqrt{T}}{e_p \sqrt{f_3}} \left(e_p \, f_2 - f_1 \right) \right). \tag{5.31}$$

And we can use the Chernoff bound for the Q function bounding the PEP as:

$$\text{PEP} \leq \frac{1}{2} e^{-\frac{T}{2\, e_p^2 \, f_3} (e_p \, f_2 - f_1)^2}. \tag{5.32}$$

In fact, (5.31) is a specific case where we imposed constraints on the path delays. In the general ergodic case where there are no constraints on the path delays, using the CLT, we get for the PEP:

$$\text{PEP} = Q\left(\frac{\text{tr}\,\{\mathbf{C_T C_F^{-1}} - \mathbf{I}\} - \ln\det(\mathbf{C_T C_F^{-1}})}{\sqrt{\frac{1}{T}\,\text{tr}\,\{(\mathbf{C_T C_F^{-1}} - \mathbf{I})^2\}}}\right) \tag{5.33}$$

from which we see that a mismatch in every path contributes separately to decreasing the PEP when the path delays are well separated (the numerator of the argument of the Q function is a form of the Itakura-Saito distance between covariance matrices). We will now show the derivation of this formula for the general case (when we do not impose any constraints on the delays). We can write (5.21) equivalently as:

$$\text{PEP} = \Pr\left(\ln\det(\mathbf{C}_T\mathbf{C}_F^{-1}) > \text{tr}\,(\hat{\mathbf{C}}\mathbf{A})\right) \tag{5.34}$$

where $\mathbf{A} = \mathbf{C_F^{-1}} - \mathbf{C_T^{-1}}$. Then:

$$\text{PEP} = \Pr\left(T\,\ln\det(\mathbf{C}_T\mathbf{C}_F^{-1}) > \sum_{i=1}^{T}\hat{\mathbf{h_i}}^H\mathbf{A}\hat{\mathbf{h_i}}\right). \tag{5.35}$$

Let us call $\hat{\mathbf{h_i}}^H\mathbf{A}\hat{\mathbf{h_i}} = x_i$. Before using the CLT, mean and variance of x_i is required. Mean is easy to derive and given as:

$$\mu_{x_i} = \text{tr}\,(\mathbf{A}\mathbf{C}_T) = \text{tr}\,\{\mathbf{C_T C_F^{-1}} - \mathbf{I}\}. \tag{5.36}$$

For the variance we know that $\sigma_{x_i}^2 = \mathbb{E}x_i^2 - (\mu_{x_i})^2$. In order to derive $\mathbb{E}x_i^2$, we will use the following identity for zero mean complex Gaussian vectors [41]:

$$\mathbb{E}\{\hat{\mathbf{h_i}}^H\mathbf{A}\hat{\mathbf{h_i}}\hat{\mathbf{h_i}}^H\mathbf{A}\hat{\mathbf{h_i}}\} = \text{tr}\,(\mathbf{A}\mathbf{C}_T\mathbf{A}\mathbf{C}_T) + (\,\text{tr}\,(\mathbf{A}\mathbf{C}_T))^2\,. \tag{5.37}$$

Therefore:

$$\sigma_{x_i}^2 = \text{tr}\,(\mathbf{A}\mathbf{C}_T\mathbf{A}\mathbf{C}_T) = \text{tr}\,\{(\mathbf{C}_T\mathbf{C}_F^{-1} - \mathbf{I})^2\}, \tag{5.38}$$

so that (5.33) follows via the CLT.

For the non-ergodic case in which the channel $\mathbf{h}$ remains constant in the T estimates $\hat{\mathbf{h_i}}$, the PEP using the CLT becomes:

$$\text{PEP} = \mathbb{E}_{\mathbf{h}}\, Q\left(\frac{\mathbf{h}^H\mathbf{A}\mathbf{h} + \sigma_v^2\,\text{tr}\,(\mathbf{A}) - \ln\det(\mathbf{C_T}\mathbf{C_F^{-1}})}{\frac{1}{\sqrt{T}}\sqrt{\sigma_v^4\,||\mathbf{A}||_F^2 + 2\,\sigma_v^2\mathbf{h}^H\mathbf{A}^2\mathbf{h}}}\right). \tag{5.39}$$

The derivation is similar to the ergodic case. We start with equation (5.35) again. For the non-ergodic case, $\hat{\mathbf{h_i}}$ is not a zero mean vector, it is: $\hat{\mathbf{h_i}} = \mathbf{h} + \mathbf{v}_i$. For the mean of x_i, we obtain it easily:

$$\mu_{x_i} = \mathbf{h}^H\mathbf{A}\mathbf{h} + \sigma_v^2\ \text{tr}\,(\mathbf{A}). \tag{5.40}$$

For the variance $\sigma_{x_i}^2$, we need $\mathbb{E}x_i^2$ again. We will exploit another identity for non-zero mean complex Gaussian vectors:

$$\mathbb{E}\{\hat{\mathbf{h_i}}^H\mathbf{A}\hat{\mathbf{h_i}}\hat{\mathbf{h_i}}^H\mathbf{A}\hat{\mathbf{h_i}}\} = \sigma_v^4\,||\mathbf{A}||_F^2 + (\sigma_v^2\ \text{tr}\,(\mathbf{A}) + \mathbf{h}^H\mathbf{A}\mathbf{h})^2 + 2\,\sigma_v^2\,\mathbf{h}^H\mathbf{A}^2\mathbf{h}. \tag{5.41}$$

Consequently:

$$\sigma_{x_i}^2 = \sigma_v^4\,||\mathbf{A}||_F^2 + 2\,\sigma_v^2\,\mathbf{h}^H\mathbf{A}^2\mathbf{h}, \tag{5.42}$$

so that (5.39) follows via the use of CLT.

5.3.1 Simulation Results

To verify the analytical results, we have performed simulations for the ergodic case. The simulation parameters are the same as in the deterministic section and also $\mathbf{P}_\zeta = \mathbf{P}_\tau$. The SNR calculation is carried out with respect to the CIR of the false position's entry. We provide three plots again. The path amplitude variances of the true position's entries are just a multiple of the false ones, i.e., $\sigma_{a_i}^2 = \beta^2\,\sigma_{b_i}^2$.

Obviously, increasing β has the effect of increasing the distance between the covariance matrices. Therefore, as can also be seen in the simulations, PEP decreases with the increasing β. Moreover, differently from the deterministic case, PEP curves do not all converge to 0 unless $T \to \infty$. Although with the increasing β, they converge closer to 0 in high SNR, some of them (for smaller values of β) stay around non-negligible PEP values. The reason is evident. In the deterministic case, path amplitudes are deterministic. Hence the only random parameter is the noise. As the noise becomes negligible in the high SNR region, PEP decays to almost 0 as expected even for finite T. However in this case, path amplitudes are also random variables.

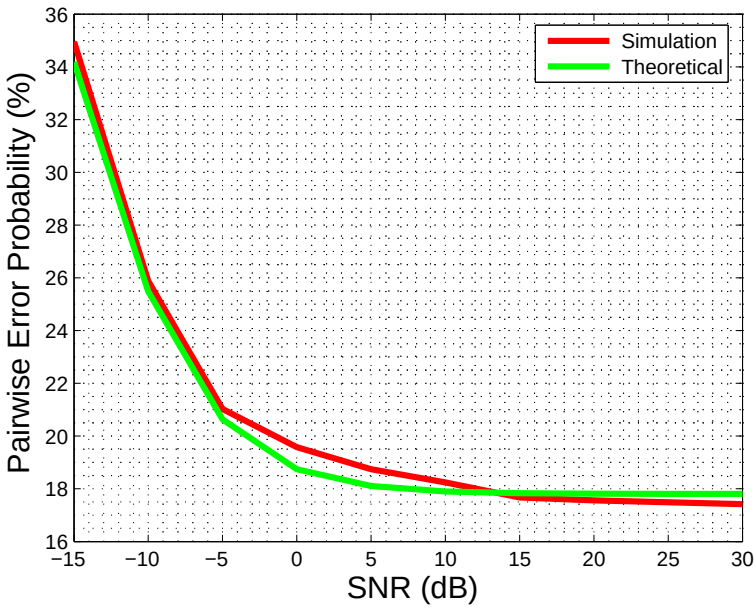

Figure 5.4: Pairwise error probability for $\beta = 1.1$.

Hence in the high SNR region, effect of the noise almost vanishes but on the other hand effects of the randomness of the path amplitudes still remain. Consequently, if the difference between the amplitude variances increases (when β increases), the chance of PEP getting closer to 0 also increases.

5.4 Conclusion

In this chapter, we derived approximate analytic results of PEP for PDP-F for different cost functions and path amplitude modelings. To the best of our knowledge, there has not been any work for the computation of PEP for fingerprinting applications so far. The effects of the pulse shape and other parameters on PEP are explicitly shown. As expected we have shown that the PEP decreases with the increasing T. In the asymptotic case (when $T \to \infty$), PEP goes to 0 for both of the algorithms investigated. Also we have verified the accuracy of the theoretical analysis with the simulations.

For the deterministic path amplitudes case using LS as the cost function, we have observed that PEP decreases with an increase in the difference of

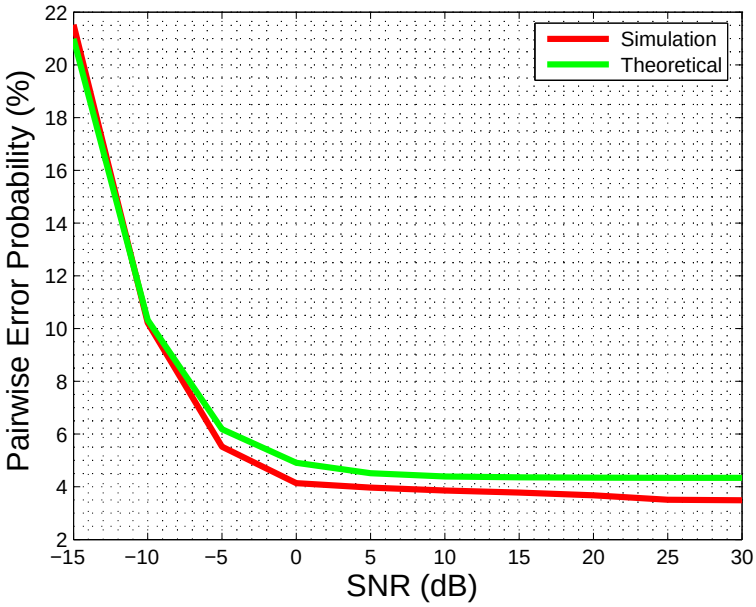

Figure 5.5: Pairwise error probability for $\beta = 1.2$.

the PDP vectors. This was an expected outcome in fact. For the GML based Rayleigh fading case, we have seen the saturation of PEP around a non-zero value in the high SNR regime. This is due to the randomness in the channel coefficients and improves only with the number of channel observations (T). We can say that total mitigation of the saturation is only possible for high T values.

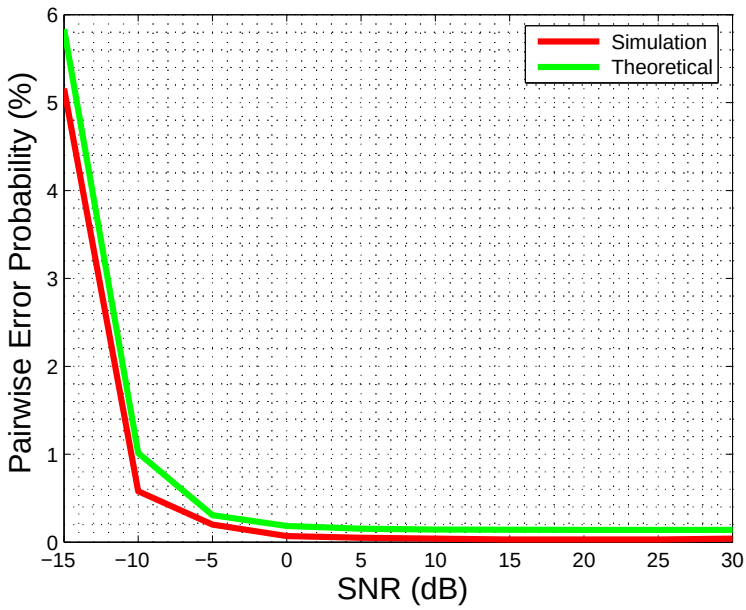

Figure 5.6: Pairwise error probability for $\beta = 1.4$.

Chapter 6

Mobile Terminal Tracking based on Kalman Filtering

6.1 Introduction

We discuss various algorithmic approaches for augmenting the Kalman filter (KF) to an adaptive Kalman filter in which position tracking and acceleration parameter estimation occur simultaneously. The main algorithm proposed is the expectation-maximization-KF (EM-KF), based on fixed-lag smoothing.

6.2 Adaptive Kalman Filtering based Tracking

Even though the Kalman filter is the work horse of position tracking, its use requires the knowledges of various parameters that describe the measurement error variances and the acceleration model. This aspect is often overlooked. In this section we review some known and some not so known methods to adapt these state space model parameters.

6.2.1 Adaptive Kalman Filtering Approaches

We shall first consider the topic of Adaptive Kalman filtering for a general linear state-space model.

Basic Kalman Filter

The KF considers the estimation of a first-order vector autoregressive (Markov) process from linear measurements in white noise. The KF performs this estimation recursively by alternating between filtering (measurement update) and (one step ahead) prediction (time update). An alternative viewpoint is that the Kalman filter recursively generates the innovations of the measurement signal (by a structured Gram-Schmidt approach that decorrelates the consecutive measurements). The KF corresponds to optimal (Minimum Mean Squared Error (MMSE) or Maximum A Posteriori (MAP)) Bayesian estimation of the state sequence if all random sources involved (measurement noise, state noise and state initial conditions) are Gaussian. In the non-Gaussian case, the KF performs Linear MMSE (LMMSE) estimation.

The signal model can be written as

$$\begin{aligned} &\text{state update equation:} \\ \boldsymbol{x}_{k+1} &= \boldsymbol{F}_k\, \boldsymbol{x}_k + \boldsymbol{G}_k\, \boldsymbol{w}_k \\ &\text{measurement equation:} \\ \boldsymbol{y}_k &= \boldsymbol{H}_k\, \boldsymbol{x}_k + \boldsymbol{v}_k \end{aligned} \tag{6.1}$$

for $k = 1, 2, \ldots$, where the initial state $\boldsymbol{x}_0 \sim \mathcal{N}(\hat{\boldsymbol{x}}_0, \boldsymbol{P}_0)$, the measurement noise $\boldsymbol{v}_k \sim \mathcal{N}(0, \boldsymbol{R}_k)$, and the state noise $\boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{Q}_k)$ and all these random quantities are mutually uncorrelated. In the case of time-varying system matrices $\boldsymbol{F}_k$ etc., the form of the equations as they appear in (6.1) is the most logical, with the state update corresponding to a prediction of the state on the basis of the quantities available at time k. In the position tracking application to be considered here, the system matrices are essentially time-invariant, or at most slowly time-varying. In that case the time index k in e.g. $\boldsymbol{F}_k$ refers to the estimate of $\boldsymbol{F}$ available at time k in an adaptive approach. In a logical ordering, a measurement is performed first and then the state gets predicted.

In the following, we introduce the notation $\boldsymbol{y}_{1:k} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\}$. The KF performs Gram-Schmidt orthogonalization (decorrelation) of the measurement variables $\boldsymbol{y}_k$. This is done by computing the LMMSE predictor $\hat{\boldsymbol{y}}_{k|k-1}$ of $\boldsymbol{y}_k$ on the basis of $\boldsymbol{y}_{1:k-1}$, leading to the orthogonalized prediction error (or innovation) $\widetilde{\boldsymbol{y}}_k = \widetilde{\boldsymbol{y}}_{k|k-1} = \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}$. We introduce the correlation matrix notation $R_{\boldsymbol{xy}} = \mathbb{E}\, \boldsymbol{x}\, \boldsymbol{y}^T$ (correlation matrices will usually also be covariance matrices here since the processes $\boldsymbol{y}_k$ and $\boldsymbol{x}_k$ have zero mean and also various estimation errors will have (conditional) zero mean). We denote the covariance matrix $R_{\widetilde{\boldsymbol{y}}_k \widetilde{\boldsymbol{y}}_k} = \boldsymbol{S}_k$. The idea of the innovations approach is

that (linear) estimation in terms of $\boldsymbol{y}_{1:k}$ is equivalent to estimation in terms of $\widetilde{\boldsymbol{y}}_{1:k}$ since one set is obtained from the other by an invertible linear transformation. Now, since the $\widetilde{\boldsymbol{y}}_k$ are decorrelated, estimation in terms of $\widetilde{\boldsymbol{y}}_{1:k}$ simplifies:

$$\hat{\boldsymbol{x}}_{|k} = \sum_{i=1}^{k} R_{\boldsymbol{x}\widetilde{\boldsymbol{y}}_i} \, R^{-1}_{\widetilde{\boldsymbol{y}}_i\widetilde{\boldsymbol{y}}_i} \, \widetilde{\boldsymbol{y}}_i = \boldsymbol{x}_{|k-1} + R_{\boldsymbol{x}\widetilde{\boldsymbol{y}}_k} \, \boldsymbol{S}_k^{-1} \, \widetilde{\boldsymbol{y}}_k \; .$$

This will be used to obtain *predicted* estimates $\hat{\boldsymbol{x}}_{k|k-1}$ with estimation error $\widetilde{\boldsymbol{x}}_{k|k-1} = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}$ with covariance matrix $\boldsymbol{P}_{k|k-1} = R_{\widetilde{\boldsymbol{x}}_{k|k-1}\widetilde{\boldsymbol{x}}_{k|k-1}}$ and also *filtered* estimates $\hat{\boldsymbol{x}}_{k|k}$ with estimation error $\widetilde{\boldsymbol{x}}_{k|k} = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}$ with covariance matrix $\boldsymbol{P}_{k|k} = R_{\widetilde{\boldsymbol{x}}_{k|k}\widetilde{\boldsymbol{x}}_{k|k}}$.

Now exploiting the correlation structure in the signal model, this leads to the following two-step recursive procedure to go from $|k-1$ to $|k$:

Measurement Update

$$\begin{aligned}
\hat{\boldsymbol{y}}_{k|k-1} &= \boldsymbol{H}_k \, \hat{\boldsymbol{x}}_{k|k-1} \\
\widetilde{\boldsymbol{y}}_k &= \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} \\
\boldsymbol{S}_k &= \boldsymbol{H}_k \, \boldsymbol{P}_{k|k-1} \, \boldsymbol{H}_k^T + \boldsymbol{R}_k \\
\boldsymbol{K}_k &= \boldsymbol{P}_{k|k-1} \, \boldsymbol{H}_k^T \, \boldsymbol{S}_k^{-1} \\
\hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \, \widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k|k} &= \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \, \boldsymbol{H}_k \, \boldsymbol{P}_{k|k-1}
\end{aligned} \tag{6.2}$$

Time Update (prediction)

$$\begin{aligned}
\hat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{F}_k \, \hat{\boldsymbol{x}}_{k|k} \\
\boldsymbol{P}_{k+1|k} &= \boldsymbol{F}_k \, \boldsymbol{P}_{k|k} \, \boldsymbol{F}_k^T + \boldsymbol{G}_k \, \boldsymbol{Q}_k \, \boldsymbol{G}_k^T
\end{aligned} \tag{6.3}$$

There are various other ways to formulate these update equations, including performing both steps in one step.

The choice of the initial conditions crucially affects the initial convergence (transient behavior). In the usual case of total absence of prior information on the initial state, one can choose $\hat{\boldsymbol{x}}_0 = 0$, $\boldsymbol{P}_0 = p_0 \, I$ with p_0 a (very) large number. This leads to $\boldsymbol{P}_{1|0} = \boldsymbol{F}_0 \, \boldsymbol{P}_0 \, \boldsymbol{F}_0^T + \boldsymbol{G}_0 \, \boldsymbol{Q}_0 \, \boldsymbol{G}_0^T$, $\hat{\boldsymbol{x}}_{1|0} = \boldsymbol{F}_0 \, \hat{\boldsymbol{x}}_0$.

For numerical stability (in the presence of roundoff errors), it is crucial that the symmetry of the covariance matrices $\boldsymbol{P}_{k|k-1}$, $\boldsymbol{P}_{k|k}$ is maintained throughout the updates (which is not going to be the case with the update of $\boldsymbol{P}_{k|k}$ the way it appears in (6.2)). The easiest way to ensure this is to periodically (e.g. every sample to ease programming) force the symmetry of e.g. $\boldsymbol{P}_{k|k}$ by computing $\boldsymbol{P}_{k|k-1} = \frac{1}{2}(\boldsymbol{P}_{k|k} + \boldsymbol{P}_{k|k}^T)$. Another solution that should work is to update $\boldsymbol{P}_{k|k}$ as $\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \, \boldsymbol{H}_k \, \boldsymbol{P}_{k|k-1}^T = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \, \boldsymbol{S}_k \, \boldsymbol{K}_k^T$.

Extended Kalman Filter (EKF)

For the case of a nonlinear state-space model, the idea of the EKF is to apply the KF to a linearized version of the state-space model, via a first-order Taylor series expansion. So we get

$$\begin{aligned} &\text{state update equation:} \\ \boldsymbol{x}_{k+1} &= f(\boldsymbol{x}_k, \boldsymbol{w}_k) \approx \boldsymbol{F}_k\, \boldsymbol{x}_k + \boldsymbol{G}_k\, \boldsymbol{w}_k \\ &\text{measurement equation:} \\ \boldsymbol{y}_k &= h(\boldsymbol{x}_k) + \boldsymbol{v}_k \approx \boldsymbol{H}_k\, \boldsymbol{x}_k + \boldsymbol{v}_k \end{aligned} \tag{6.4}$$

where

$$\begin{aligned} \boldsymbol{F}_k &= \left.\frac{\partial\, f(\boldsymbol{x}, \boldsymbol{w})}{\partial\, \boldsymbol{x}^T}\right|_{(\boldsymbol{x},\boldsymbol{w})=(\boldsymbol{x}_k,\boldsymbol{w}_k)} \\ \boldsymbol{G}_k &= \left.\frac{\partial\, f(\boldsymbol{x}, \boldsymbol{w})}{\partial\, \boldsymbol{w}^T}\right|_{(\boldsymbol{x},\boldsymbol{w})=(\boldsymbol{x}_k,\boldsymbol{w}_k)} \\ \boldsymbol{H}_k &= \left.\frac{\partial\, h(\boldsymbol{x})}{\partial\, \boldsymbol{x}^T}\right|_{\boldsymbol{x}=\boldsymbol{x}_k} . \end{aligned} \tag{6.5}$$

So, at this point, the basic KF can be applied to the thus obtained approximate linear state-space model.

The EKF approach can be used to adapt some parameters in an otherwise linear state-space model $\boldsymbol{x}_{k+1}^{'} = \boldsymbol{F}^{'}\ \boldsymbol{x}_k^{'} + \boldsymbol{G}^{'}\ \boldsymbol{w}_k$. For instance, consider the case in which one wants to adapt parameters appearing (e.g.) linearly in the matrix $\boldsymbol{F}^{'} = \boldsymbol{F}^{'}(\boldsymbol{\theta})$. One can jointly estimate the unknown constant parameter vector $\boldsymbol{\theta}$ by considering the following state update for them: $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k$. Then one can introduce the augmented state and system matrices

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{x}_k^{'} \\ \boldsymbol{\theta}_k \end{bmatrix}\ ,\ \boldsymbol{F}_k = \begin{bmatrix} \boldsymbol{F}^{'}(\boldsymbol{\theta}_k) & \boldsymbol{C}(\boldsymbol{x}_k^{'}) \\ 0 & I \end{bmatrix}\ ,\ \boldsymbol{G}_k = \begin{bmatrix} \boldsymbol{G}^{'} \\ 0 \end{bmatrix} \tag{6.6}$$

where $\boldsymbol{C}(\boldsymbol{x}_k^{'}) = \dfrac{\partial\, \boldsymbol{F}^{'}(\boldsymbol{\theta})\boldsymbol{x}_k^{'}}{\partial\, \boldsymbol{\theta}^T}$. When running the EKF, the state-dependent system matrices have to be filled with the latest state estimates, so in this case

$$\boldsymbol{F}_k = \begin{bmatrix} \boldsymbol{F}^{'}(\hat{\boldsymbol{\theta}}_{k|k}) & \boldsymbol{C}(\hat{\boldsymbol{x}}_{k|k}^{'}) \\ 0 & I \end{bmatrix} . \tag{6.7}$$

The basic EKF can be improved by correct computation of the covariance matrix $\boldsymbol{P}_{k+1|k}$, accounting for the replacement of $\boldsymbol{x}_k$ by $\hat{\boldsymbol{x}}_{k|k}$ (which in fact

leads to equivalent augmented process noise). The other issue is that the parameters $\boldsymbol{\theta}$ are often not really constant and hence need to be tracked adaptively. This can be done either by introducing some process noise in $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k$ (random walk time evolution) or by introducing exponential weighting (at least for the $\boldsymbol{\theta}$ portion) into the KF updates [42].

This EKF approach allows fairly straightforwardly to estimate parameters in $\boldsymbol{F}_k$, $\boldsymbol{H}_k$, or $\boldsymbol{G}_k$, but much less so in $\boldsymbol{Q}_k$, $\boldsymbol{R}_k$. For adapting (parameters in) $\boldsymbol{Q}$ and $\boldsymbol{R}$, one needs to consider the innovations representation $\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{F}_k\,\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{F}_k\boldsymbol{K}_k\,\widetilde{\boldsymbol{y}}_k$ and consider gradients of the Kalman gain $\boldsymbol{K}_k$ w.r.t. these matrices.

Recursive Prediction Error Method (RPEM-KF)

The RPEM [43], [44] is an adaptive implementation of Maximum Likelihood (ML) parameter estimation. The negative log-likelihood becomes a least-squares criterion in the prediction errors (innovations) and RPEM performs one iteration per sample. Applied to KF, the RPEM can be seen as a more rigorous version of EKF and computes gradients more precisely [45]. Indeed, for the case of a state transition matrix $\boldsymbol{F}_k = \boldsymbol{F}_k(\boldsymbol{\theta})$, the EKF would consider the gradient

$$\frac{\partial \boldsymbol{x}_{k+1}}{\partial \boldsymbol{\theta}^T} = \frac{\partial \boldsymbol{F}_k(\boldsymbol{\theta})\boldsymbol{x}_k}{\partial \boldsymbol{\theta}^T} \tag{6.8}$$

where only the explicit dependence of $\boldsymbol{F}$ on $\boldsymbol{\theta}$ would be considered, whereas the RPEM would consider more correctly

$$\frac{\partial \boldsymbol{x}_{k+1}}{\partial \boldsymbol{\theta}^T} = \frac{\partial \boldsymbol{F}_k(\boldsymbol{\theta})\boldsymbol{x}_k}{\partial \boldsymbol{\theta}^T} + \boldsymbol{F}_k(\boldsymbol{\theta})\,\frac{\partial \boldsymbol{x}_k}{\partial \boldsymbol{\theta}^T}\ . \tag{6.9}$$

RPEM for KF can be found in the references above, but will not be pursued here further. One characteristic of the RPEM is a higher complexity.

Expectation-Maximization (EM-KF)

In EM [46], the parameters are estimated by minimizing expected values of negative loglikelihoods, see e.g. [47] for an application involving KF. For the state update, since $\boldsymbol{G}_k$ is typically a tall matrix, $\boldsymbol{G}_k\,\boldsymbol{w}_k$ has a singular covariance matrix. The state update equation can be rewritten as

$$\boldsymbol{G}_k^+\,\boldsymbol{x}_{k+1} = \boldsymbol{G}_k^+\,\boldsymbol{F}_k\,\boldsymbol{x}_k + \boldsymbol{w}_k \tag{6.10}$$

where $\boldsymbol{G}_k^+ = (\boldsymbol{G}_k^T \boldsymbol{G}_k)^{-1} \boldsymbol{G}_k^T$ is the pseudo-inverse of $\boldsymbol{G}_k$. For the parameters involved in the state update equation, hence the following negative log-likelihood is applicable:

$$\sum_k \{\ln\det(\boldsymbol{Q}_k) + (\boldsymbol{G}_k^+(\boldsymbol{x}_{k+1} - \boldsymbol{F}_k\,\boldsymbol{x}_k))^T \boldsymbol{Q}_k^{-1} (\boldsymbol{G}_k^+(\boldsymbol{x}_{k+1} - \boldsymbol{F}_k\,\boldsymbol{x}_k))\}\ . \quad (6.11)$$

For the parameters involved in the measurement equation, the appropriate log-likelihood is

$$\sum_k \{\ln\det(\boldsymbol{R}_k) + (\boldsymbol{y}_k - \boldsymbol{H}_k\,\boldsymbol{x}_k)^T \boldsymbol{R}_k^{-1} (\boldsymbol{y}_k - \boldsymbol{H}_k\,\boldsymbol{x}_k)\}\ . \quad (6.12)$$

Now the expectation is taken, in principle with the conditional distribution given all data. Hence $\mathbb{E}_{|n}$ involving all data $\boldsymbol{y}_k$ up to the last sample n. This leads to an iterative algorithm with in each iteration a whole fixed-interval smoothing operation. An adaptive version [47], [48] can be obtained by replacing fixed-interval smoothing by fixed-lag smoothing and performing one iteration per time sample. Since the state update equation corresponds to a vector AR(1) model, one may expect (as in [48]) that a lag of 1 should be enough (to guarantee convergence). In [47], complexity is reduced further by suggesting that filtering might be enough. In that case, the (presumably) slowly varying $\hat{\boldsymbol{Q}}_{k+1}$, $\hat{\boldsymbol{F}}_{k+1}$ (for use in the KF at time k+1) get determined by minimizing

$$\sum_{i=1}^{k} \lambda^{k-i}\ \mathbb{E}_{|i}\{\ln\det(\hat{\boldsymbol{Q}}) + (\boldsymbol{G}_i^+(\boldsymbol{x}_{i+1} - \hat{\boldsymbol{F}}\,\boldsymbol{x}_i))^T \hat{\boldsymbol{Q}}^{-1} (\boldsymbol{G}_i^+(\boldsymbol{x}_{i+1} - \hat{\boldsymbol{F}}\,\boldsymbol{x}_i))\} \quad (6.13)$$

w.r.t. $\hat{\boldsymbol{Q}}$, $\hat{\boldsymbol{F}}$ where we introduced an exponential forgetting factor $\lambda \overset{\approx}{<} 1$. (6.13) is equivalent to

$$\gamma_k^{-1} \ln\det(\hat{\boldsymbol{Q}}) + \sum_{i=1}^{k} \lambda^{k-i}\ \operatorname{tr}\{\boldsymbol{G}_i^{+T} \hat{\boldsymbol{Q}}^{-1} \boldsymbol{G}_i^+ \mathbb{E}_{|i}(\boldsymbol{x}_{i+1} - \hat{\boldsymbol{F}}\,\boldsymbol{x}_i)(\boldsymbol{x}_{i+1} - \hat{\boldsymbol{F}}\,\boldsymbol{x}_i)^T\} \quad (6.14)$$

where we introduced $\gamma_k^{-1} = \sum_{i=1}^{k} \lambda^{k-i} = \lambda\,\gamma_{k-1}^{-1} + 1$. γ_k^{-1} behaves initially as $1/k$ but saturates eventually at $\gamma_\infty^{-1} = 1-\lambda$. We shall need

$$\begin{aligned}
\mathbb{E}_{|i}\,\boldsymbol{x}_i\boldsymbol{x}_i^T &= \hat{\boldsymbol{x}}_{i|i}\hat{\boldsymbol{x}}_{i|i}^T + \boldsymbol{P}_{i|i} \\
\mathbb{E}_{|i}\,\boldsymbol{x}_{i+1}\boldsymbol{x}_i^T &= \boldsymbol{F}_i\,\hat{\boldsymbol{x}}_{i|i}\hat{\boldsymbol{x}}_{i|i}^T + \boldsymbol{F}_i\,\boldsymbol{P}_{i|i} \\
\mathbb{E}_{|i}\,\boldsymbol{x}_i\boldsymbol{x}_{i+1}^T &= \hat{\boldsymbol{x}}_{i|i}\hat{\boldsymbol{x}}_{i|i}^T\boldsymbol{F}_i^T + \boldsymbol{P}_{i|i}\,\boldsymbol{F}_i^T \\
\mathbb{E}_{|i}\,\boldsymbol{x}_{i+1}\boldsymbol{x}_{i+1}^T &= \boldsymbol{F}_i\,\hat{\boldsymbol{x}}_{i|i}\hat{\boldsymbol{x}}_{i|i}^T\boldsymbol{F}_i^T + \boldsymbol{P}_{i+1|i} \\
&= \boldsymbol{F}_i\,(\hat{\boldsymbol{x}}_{i|i}\hat{\boldsymbol{x}}_{i|i}^T + \boldsymbol{P}_{i|i})\,\boldsymbol{F}_i^T + \boldsymbol{G}_i\boldsymbol{Q}_i\boldsymbol{G}_i^T\ .
\end{aligned} \quad (6.15)$$

In case of time-invariant $\boldsymbol{G}_k \equiv \boldsymbol{G}$, we can rewrite (6.14) as

$$\ln\det(\hat{\boldsymbol{Q}}) + \operatorname{tr}\{\boldsymbol{G}^{+T}\hat{\boldsymbol{Q}}^{-1}\boldsymbol{G}^{+}(\boldsymbol{M}_k^{11} - \hat{\boldsymbol{F}}\,\boldsymbol{M}_k^{01} - \boldsymbol{M}_k^{10}\,\hat{\boldsymbol{F}}^T + \hat{\boldsymbol{F}}\,\boldsymbol{M}_k^{00}\,\hat{\boldsymbol{F}}^T)\} \quad (6.16)$$

where

$$\begin{aligned}
\boldsymbol{M}_k^{00} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{00} + \gamma_k\,(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k|k}^T + \boldsymbol{P}_{k|k}) \\
\boldsymbol{M}_k^{10} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{10} + \gamma_k\,\boldsymbol{F}_k\,(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k|k}^T + \boldsymbol{P}_{k|k}) \\
\boldsymbol{M}_k^{01} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{01} + \gamma_k\,(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k|k}^T + \boldsymbol{P}_{k|k})\,\boldsymbol{F}_k^T \\
\boldsymbol{M}_k^{11} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{11} + \gamma_k(\hat{\boldsymbol{x}}_{k+1|k}\hat{\boldsymbol{x}}_{k+1|k}^T + \boldsymbol{P}_{k+1|k})\ .
\end{aligned} \quad (6.17)$$

In case of furthermore time-invariant $\boldsymbol{F}_k \equiv \boldsymbol{F}$, $\boldsymbol{Q}_k \equiv \boldsymbol{Q}$, then

$$\begin{aligned}
\boldsymbol{M}_k^{10} &= \boldsymbol{F}\,\boldsymbol{M}_k^{00} \\
\boldsymbol{M}_k^{01} &= \boldsymbol{M}_k^{00}\,\boldsymbol{F}^T \\
\boldsymbol{M}_k^{11} &= \boldsymbol{F}\,\boldsymbol{M}_k^{00}\,\boldsymbol{F}^T + \boldsymbol{G}\,\boldsymbol{Q}\,\boldsymbol{G}^T\ .
\end{aligned} \quad (6.18)$$

As a result, (6.16) can be rewritten as

$$\ln\det(\hat{\boldsymbol{Q}}) + \operatorname{tr}\{\hat{\boldsymbol{Q}}^{-1}\boldsymbol{Q}\} + \operatorname{tr}\{\boldsymbol{G}^{+T}\hat{\boldsymbol{Q}}^{-1}\boldsymbol{G}_i^{+}(\boldsymbol{F}-\hat{\boldsymbol{F}})\,\boldsymbol{M}_k^{00}\,(\boldsymbol{F}-\hat{\boldsymbol{F}})^T)\}\ . \quad (6.19)$$

The optimization of (6.19) now clearly leads to $\hat{\boldsymbol{F}} = \boldsymbol{F}$, $\hat{\boldsymbol{Q}} = \boldsymbol{Q}$, so we just get back the quantities that we use in the KF, without any additional information. Hence, just Kalman filtering in the EM-KF is not enough to adapt the state update parameters. Indeed, it just leads to a snake biting its own tail.

Fixed-Lag Smoothing

Using the innovations approach, we have

$$\hat{\boldsymbol{x}}_{k-1|k} = \hat{\boldsymbol{x}}_{k-1|k-1} + R_{\boldsymbol{x}_{k-1}\widetilde{\boldsymbol{y}}_k}\,\boldsymbol{S}_k^{-1}\,\widetilde{\boldsymbol{y}}_k\ . \quad (6.20)$$

After a few steps, we get the following lag-1 smoothing equations that need to be added to the basic Kalman Filter equations (to be inserted between the Measurement Update and the Time Update)

$$\begin{aligned}
\boldsymbol{K}_{k;1} &= \boldsymbol{P}_{k-1|k-1}\,\boldsymbol{F}_{k-1}^T\,\boldsymbol{H}_k^T \\
\hat{\boldsymbol{x}}_{k-1|k} &= \hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{K}_{k;1}\,\boldsymbol{S}_k^{-1}\,\widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k-1|k} &= \boldsymbol{P}_{k-1|k-1} - \boldsymbol{K}_{k;1}\,\boldsymbol{S}_k^{-1}\,\boldsymbol{K}_{k;1}^T\ .
\end{aligned} \quad (6.21)$$

Adaptive EM-KF with Fixed-Lag Smoothing

Consider now the case in which the state-space model is essentially time-invariant (or slowly time-varying). In that case the time index of the system matrices $\boldsymbol{F}_k$ etc. just reflects at which time the (unknown) system matrices have been adapted. The resulting KF equations with lag-1 smoothing become:

$$\begin{array}{rcl}
\hat{\boldsymbol{y}}_{k|k-1} & = & \boldsymbol{H}_{k-1}\,\hat{\boldsymbol{x}}_{k|k-1} \\
\widetilde{\boldsymbol{y}}_k & = & \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} \\
\boldsymbol{S}_k & = & \boldsymbol{H}_{k-1}\,\boldsymbol{P}_{k|k-1}\,\boldsymbol{H}_{k-1}^T + \boldsymbol{R}_{k-1} \\
\boldsymbol{K}_{k;1} & = & \boldsymbol{P}_{k-1|k-1}\,\boldsymbol{F}_{k-1}^T\,\boldsymbol{H}_{k-1}^T \\
\hat{\boldsymbol{x}}_{k-1|k} & = & \hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{K}_{k;1}\,\boldsymbol{S}_k^{-1}\,\widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k-1|k} & = & \boldsymbol{P}_{k-1|k-1} - \boldsymbol{K}_{k;1}\,\boldsymbol{S}_k^{-1}\,\boldsymbol{K}_{k;1}^T \\
\boldsymbol{K}_k & = & \boldsymbol{P}_{k|k-1}\,\boldsymbol{H}_{k-1}^T\,\boldsymbol{S}_k^{-1} \\
\hat{\boldsymbol{x}}_{k|k} & = & \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\,\widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k|k} & = & \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\,\boldsymbol{H}_{k-1}\,\boldsymbol{P}_{k|k-1} \\
 & & \text{parameter update} \\
\hat{\boldsymbol{x}}_{k+1|k} & = & \boldsymbol{F}_k\,\hat{\boldsymbol{x}}_{k|k} \\
\boldsymbol{P}_{k+1|k} & = & \boldsymbol{F}_k\,\boldsymbol{P}_{k|k}\,\boldsymbol{F}_k^T + \boldsymbol{G}_k\,\boldsymbol{Q}_k\,\boldsymbol{G}_k^T
\end{array} \tag{6.22}$$

So, the system matrices ($\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{Q}$) should be adapted after the smoothing step and before the filtering and prediction steps. We now adapt the system matrices $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{Q}$ from

$$\begin{aligned}
&\ln\det(\hat{\boldsymbol{Q}}) + \operatorname{tr}\{\hat{\boldsymbol{G}}^{+T}\hat{\boldsymbol{Q}}^{-1}\hat{\boldsymbol{G}}^{+}\gamma_k \textstyle\sum_{i=1}^{k}\lambda^{k-i}\,\mathbb{E}_{|i}(\boldsymbol{x}_i - \hat{\boldsymbol{F}}\,\boldsymbol{x}_{i-1})(\boldsymbol{x}_i - \hat{\boldsymbol{F}}\,\boldsymbol{x}_{i-1})^T\} = \\
&\ln\det(\hat{\boldsymbol{Q}}) + \operatorname{tr}\{\hat{\boldsymbol{G}}^{+T}\hat{\boldsymbol{Q}}^{-1}\hat{\boldsymbol{G}}^{+}\,(\boldsymbol{M}_k^{11} - \hat{\boldsymbol{F}}\,\boldsymbol{M}_k^{01} - \boldsymbol{M}_k^{10}\,\hat{\boldsymbol{F}}^T + \hat{\boldsymbol{F}}\,\boldsymbol{M}_k^{00}\,\hat{\boldsymbol{F}}^T)\}
\end{aligned} \tag{6.23}$$

where the matrix definitions become this time

$$\begin{aligned}
\boldsymbol{M}_k^{00} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{00} + \gamma_k\,(\hat{\boldsymbol{x}}_{k-1|k}\hat{\boldsymbol{x}}_{k-1|k}^T + \boldsymbol{P}_{k-1|k}) \\
\boldsymbol{M}_k^{10} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{10} \\
&\quad +\gamma_k\,(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k-1|k}^T + \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k} - \boldsymbol{G}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{G}_{k-1}^T\boldsymbol{H}_{k-1}^T\boldsymbol{S}_k^{-1}\boldsymbol{K}_{k,1}^T) \\
\boldsymbol{M}_k^{01} &= (\boldsymbol{M}_k^{10})^T \\
\boldsymbol{M}_k^{11} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{11} + \gamma_k(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k|k}^T + \boldsymbol{P}_{k|k})\ .
\end{aligned} \tag{6.24}$$

If for example $\boldsymbol{G}$ would be fixed and invertible, minimization of (6.23) w.r.t. $\hat{\boldsymbol{F}}$, $\hat{\boldsymbol{Q}}$ would lead to the following minimizers

$$\begin{array}{rcl}
\boldsymbol{F}_k & = & \boldsymbol{M}_k^{10}\,(\boldsymbol{M}_k^{00})^{-1} \\
\boldsymbol{Q}_k & = & \boldsymbol{G}^{+}(\boldsymbol{M}_k^{11} - \boldsymbol{M}_k^{10}\,(\boldsymbol{M}_k^{00})^{-1}\boldsymbol{M}_k^{01})\boldsymbol{G}^{+T}\ .
\end{array} \tag{6.25}$$

For adapting the parameters in the measurement equation on the other hand, Kalman filtering should be sufficient. So consider

$$\begin{aligned}
&\gamma_k \sum_{i=1}^{k} \lambda^{k-i}\, \mathbb{E}_{|i}\{\ln\det(\hat{\boldsymbol{R}}) + (\boldsymbol{y}_i - \hat{\boldsymbol{H}}\,\boldsymbol{x}_i)^T \hat{\boldsymbol{R}}^{-1}(\boldsymbol{y}_i - \hat{\boldsymbol{H}}\,\boldsymbol{x}_i)\} \\
&= \ln\det(\hat{\boldsymbol{R}}) + \operatorname{tr}\{\hat{\boldsymbol{R}}^{-1}\, \gamma_k \sum_{i=1}^{k} \lambda^{k-i}\, \mathbb{E}_{|i}(\boldsymbol{y}_i - \hat{\boldsymbol{H}}\,\boldsymbol{x}_i)(\boldsymbol{y}_i - \hat{\boldsymbol{H}}\,\boldsymbol{x}_i)^T\} \\
&= \ln\det(\hat{\boldsymbol{R}}) + \operatorname{tr}\{\hat{\boldsymbol{R}}^{-1}\, (\hat{R}_{\boldsymbol{yy},k} - \hat{\boldsymbol{H}}\,\hat{R}_{\boldsymbol{xy},k} - \hat{R}_{\boldsymbol{yx},k}\,\hat{\boldsymbol{H}}^T + \hat{\boldsymbol{H}}\,\hat{R}_{\boldsymbol{xx},k}\hat{\boldsymbol{H}}^T)\}
\end{aligned} \tag{6.26}$$

where

$$\begin{aligned}
\hat{R}_{\boldsymbol{yy},k} &= \gamma_k \sum_{i=1}^{k} \lambda^{k-i}\, \boldsymbol{y}_i\boldsymbol{y}_i^T = (1-\gamma_k)\hat{R}_{\boldsymbol{yy},k-1} + \gamma_k \boldsymbol{y}_k\boldsymbol{y}_k^T \\
\hat{R}_{\boldsymbol{xy},k} &= \gamma_k \sum_{i=1}^{k} \lambda^{k-i}\, \hat{\boldsymbol{x}}_{i|i}\boldsymbol{y}_i^T = (1-\gamma_k)\hat{R}_{\boldsymbol{xy},k-1} + \gamma_k \hat{\boldsymbol{x}}_{k|k}\boldsymbol{y}_k^T \\
\hat{R}_{\boldsymbol{yx},k} &= \hat{R}_{\boldsymbol{xy},k}^T \\
\hat{R}_{\boldsymbol{xx},k} &= \gamma_k \sum_{i=1}^{k} \lambda^{k-i}\, \mathbb{E}_{|i}\boldsymbol{x}_i\boldsymbol{x}_i^T = (1-\gamma_k)\hat{R}_{\boldsymbol{xx},k-1} + \gamma_k(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k|k}^T + \boldsymbol{P}_{k|k})
\end{aligned} \tag{6.27}$$

Minimization of (6.26) w.r.t. $\hat{\boldsymbol{H}}$, $\hat{\boldsymbol{R}}$ yields

$$\begin{aligned}
\boldsymbol{H}_k &= \hat{R}_{\boldsymbol{yx},k}\hat{R}_{\boldsymbol{xx},k}^{-1} \\
\boldsymbol{R}_k &= \hat{R}_{\boldsymbol{yy},k} - \hat{R}_{\boldsymbol{yx},k}\hat{R}_{\boldsymbol{xx},k}^{-1}\hat{R}_{\boldsymbol{xy},k} \quad .
\end{aligned} \tag{6.28}$$

For the initialization, in absence of any side information, one can take $\boldsymbol{M}_0^{00} = 1/p_0\boldsymbol{I}$, $\boldsymbol{M}_0^{10} = \boldsymbol{0}$, $\boldsymbol{M}_0^{11} = \boldsymbol{0}$, $\hat{R}_{\boldsymbol{xx},0} = 1/p_0\boldsymbol{I}$, $\hat{R}_{\boldsymbol{xy},0} = \boldsymbol{0}$, $\hat{R}_{\boldsymbol{yy},0} = \boldsymbol{0}$ where again p_0 is a very large number.

Hybrid EM-EKF

The idea here is to update $\boldsymbol{F}_k$ via EKF and $\boldsymbol{Q}_k$, $\boldsymbol{R}_k$ via EM.

6.2.2 State-Space Models for Position Tracking

White Noise Acceleration

Consider positioning in nD, where $n = 1$, 2 or 3 dimensions. Let $\boldsymbol{p}_k$ be the position at sampling instant k, $\boldsymbol{v}_k$ the velocity (not to be confused with the measurement noise) and $\boldsymbol{a}_k$ the acceleration. In the case of e.g. 3D positioning, $\boldsymbol{p}_k$ is of the form $\boldsymbol{p} = [x\ y\ z]^T$. By simple discretization of the differential equations of motion, we get

$$\begin{aligned}
\boldsymbol{p}_{k+1} &= \boldsymbol{p}_k + \boldsymbol{v}_k \\
\boldsymbol{v}_{k+1} &= \boldsymbol{v}_k + \boldsymbol{a}_k \\
\boldsymbol{a}_k &= \boldsymbol{w}_k \quad .
\end{aligned} \tag{6.29}$$

In the case of modeling the acceleration as (temporally) white noise, the acceleration is the process noise. To simplify the equations, we assume here that the unit of time for velocity and acceleration is the sampling period. The physical speed and acceleration are then $t_s\,\boldsymbol{v}_k$ and $t_s^2\,\boldsymbol{a}_k$ where t_s is the sampling period expressed in seconds, assuming $\boldsymbol{p}_k$ is expressed in meters. We get for the state-space model

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{p}_k \\ \boldsymbol{v}_k \end{bmatrix}, \; \boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{I}_n \\ \boldsymbol{0}_{n,n} & \boldsymbol{I}_n \end{bmatrix}, \; \boldsymbol{G} = \begin{bmatrix} \boldsymbol{0}_{n,n} \\ \boldsymbol{I}_n \end{bmatrix}, \; \boldsymbol{y}_k = \hat{\boldsymbol{p}}_k, \boldsymbol{H} = [\boldsymbol{I}_n \;\; \boldsymbol{0}_{n,n}] \tag{6.30}$$

where $\boldsymbol{0}_{n,m}$ is a $n \times m$ matrix of zeros. We have $\boldsymbol{G}^+ = [\boldsymbol{0}_{n,n} \;\; \boldsymbol{I}_n]$ and $\boldsymbol{a}_k = \boldsymbol{G}^+\boldsymbol{x}_k$. The only unknown system parameter in this case is the acceleration covariance matrix $\boldsymbol{Q}$.

AR(1) (Markov) Acceleration

In this case we assume a first-order autoregressive model for the acceleration $\boldsymbol{a}_{k+1} = \boldsymbol{A}\,\boldsymbol{a}_k + \boldsymbol{w}_k$ where now $\boldsymbol{A}$ and $\boldsymbol{Q}$ are unknown. Note that $\boldsymbol{G}^+\boldsymbol{F} = \boldsymbol{A}\boldsymbol{G}^+$. We get for the state-space model

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{p}_k \\ \boldsymbol{v}_k \\ \boldsymbol{a}_k \end{bmatrix}, \; \boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{I}_n & \boldsymbol{0}_{n,n} \\ \boldsymbol{0}_{n,n} & \boldsymbol{I}_n & \boldsymbol{I}_n \\ \boldsymbol{0}_{n,n} & \boldsymbol{0}_{n,n} & \boldsymbol{A} \end{bmatrix}, \; \boldsymbol{G} = \begin{bmatrix} \boldsymbol{0}_{n,n} \\ \boldsymbol{0}_{n,n} \\ \boldsymbol{I}_n \end{bmatrix}, \; \boldsymbol{H} = [\boldsymbol{I}_n \;\; \boldsymbol{0}_{n,2n}] \; . \tag{6.31}$$

6.2.3 Adaptive EM-KF with Fixed-Lag Smoothing for Position Tracking

White Noise Acceleration

The EM-KF becomes (note that $\boldsymbol{H}\boldsymbol{G} = \boldsymbol{0}$)

$$
\begin{aligned}
\hat{\boldsymbol{y}}_{k|k-1} &= \boldsymbol{H}\,\hat{\boldsymbol{x}}_{k|k-1} \\
\widetilde{\boldsymbol{y}}_k &= \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} \\
\boldsymbol{S}_k &= \boldsymbol{H}\,\boldsymbol{P}_{k|k-1}\,\boldsymbol{H}^T + \boldsymbol{R}_{k-1} \\
\boldsymbol{K}_{k;1} &= \boldsymbol{P}_{k-1|k-1}\,\boldsymbol{F}_{k-1}^T\,\boldsymbol{H}^T \\
\hat{\boldsymbol{x}}_{k-1|k} &= \hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{K}_{k;1}\,\boldsymbol{S}_k^{-1}\,\widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k-1|k} &= \boldsymbol{P}_{k-1|k-1} - \boldsymbol{K}_{k;1}\,\boldsymbol{S}_k^{-1}\,\boldsymbol{K}_{k;1}^T \\
\boldsymbol{K}_k &= \boldsymbol{P}_{k|k-1}\,\boldsymbol{H}^T\,\boldsymbol{S}_k^{-1} \\
\hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\,\widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k|k} &= \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\,\boldsymbol{H}\,\boldsymbol{P}_{k|k-1} \\
\boldsymbol{M}_k^{00} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{00} + \gamma_k\,\boldsymbol{G}^{+}(\hat{\boldsymbol{x}}_{k-1|k}\hat{\boldsymbol{x}}_{k-1|k}^T + \boldsymbol{P}_{k-1|k})\boldsymbol{G}^{+\,T} \\
\boldsymbol{M}_k^{10} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{10} + \gamma_k\,\boldsymbol{G}^{+}(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k-1|k}^T + \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k})\boldsymbol{G}^{+\,T} \\
\boldsymbol{M}_k^{11} &= (1-\gamma_k)\,\boldsymbol{M}_{k-1}^{11} + \gamma_k\,\boldsymbol{G}^{+}(\hat{\boldsymbol{x}}_{k|k}\hat{\boldsymbol{x}}_{k|k}^T + \boldsymbol{P}_{k|k})\boldsymbol{G}^{+\,T} \\
\boldsymbol{Q}_k &= \boldsymbol{M}_k^{11} - \boldsymbol{M}_k^{10} - \boldsymbol{M}_k^{10\,T} + \boldsymbol{M}_k^{00} \\
(\ \boldsymbol{Q}_k &= \tfrac{1}{n}\,\mathrm{tr}\,\{\boldsymbol{Q}_k\}\,\boldsymbol{I}_n\ ) \\
\boldsymbol{R}_k &= (1-\gamma_k)\boldsymbol{R}_{k-1} + \gamma_k\,((\boldsymbol{y}_k - \boldsymbol{H}\hat{\boldsymbol{x}}_{k|k})(\boldsymbol{y}_k - \boldsymbol{H}\hat{\boldsymbol{x}}_{k|k})^T + \boldsymbol{H}\boldsymbol{P}_{k|k}\boldsymbol{H}^T) \\
\hat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{F}_k\,\hat{\boldsymbol{x}}_{k|k} \\
\boldsymbol{P}_{k+1|k} &= \boldsymbol{F}_k\,\boldsymbol{P}_{k|k}\,\boldsymbol{F}_k^T + \boldsymbol{G}\,\boldsymbol{Q}_k\,\boldsymbol{G}^T
\end{aligned}
\tag{6.32}
$$

$\boldsymbol{Q}_k = \frac{1}{n}\,\mathrm{tr}\,\{\boldsymbol{Q}_k\}\,\boldsymbol{I}_n$ gets added in case we want to model the acceleration as also spatially white.

AR(1) (Markov) Acceleration

The only change for the EM-KF in (6.39) is that the update for $\boldsymbol{Q}_k$ gets replaced by

$$
\begin{aligned}
\boldsymbol{A}_k &= \boldsymbol{M}_k^{10}\,(\boldsymbol{M}_k^{00})^{-1} \\
\boldsymbol{Q}_k &= \boldsymbol{M}_k^{11} - \boldsymbol{A}_k(\boldsymbol{M}_k^{10})^T\ .
\end{aligned}
\tag{6.33}
$$

6.2.4 Non-Linear Measurements

In the case the measurement equation is of the form $\boldsymbol{y}_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k$, the EKF needs to be applied as explained earlier. We get

$$
\boldsymbol{y}_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k \;\approx\; \boldsymbol{H}_k\,\boldsymbol{x}_k + \boldsymbol{v}_k \tag{6.34}
$$

where

$$\boldsymbol{H}_k = \left.\frac{\partial\, h(\boldsymbol{x})}{\partial\, \boldsymbol{x}^T}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_{k|k-1}} . \tag{6.35}$$

6.3 Fitting of State Space Mobility Models to M3 Measurements

In this section, by exploiting the algorithms discussed so far, we try different mobility models on the measurement data (GPS position estimates) we have.

Temporally White, Spatially Colored Noise Acceleration

We consider positioning in nD, where $n = 1$, 2 or 3 dimensions. Let $\boldsymbol{p}_k$ be the position at sampling instant k, $\boldsymbol{v}_k$ the velocity (not to be confused with the measurement noise) and $\boldsymbol{a}_k$ the acceleration. In the case of 2D positioning, $\boldsymbol{p}_k$ is of the form $\boldsymbol{p} = [x\ y]^T$. By simple discretization of the differential equations of motion, we get:

$$\begin{aligned} \boldsymbol{p}_{k+1} &= \boldsymbol{p}_k + \boldsymbol{v}_k \\ \boldsymbol{v}_{k+1} &= \boldsymbol{v}_k + \boldsymbol{a}_k \\ \boldsymbol{a}_k &= \boldsymbol{w}_k \ . \end{aligned} \tag{6.36}$$

In the case of modeling the acceleration as (temporally) white noise, the acceleration is the process noise. To simplify the equations, we assume here that the unit of time for velocity and acceleration is the sampling period. The physical speed and acceleration are then $t_s\, \boldsymbol{v}_k$ and $t_s^2\, \boldsymbol{a}_k$ where t_s is the sampling period expressed in seconds, assuming $\boldsymbol{p}_k$ is expressed in meters. We get for the state-space model

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{p}_k \\ \boldsymbol{v}_k \end{bmatrix},\ \boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{I}_2 \\ \boldsymbol{0}_{2,2} & \boldsymbol{I}_2 \end{bmatrix},\ \boldsymbol{G} = \begin{bmatrix} \boldsymbol{0}_{2,2} \\ \boldsymbol{I}_2 \end{bmatrix},\ \boldsymbol{y}_k = \hat{\boldsymbol{p}}_k, \boldsymbol{H} = [\boldsymbol{I}_2\ \ \boldsymbol{0}_{2,2}] \tag{6.37}$$

where $\boldsymbol{0}_{n,m}$ is a $n \times m$ matrix of zeros. The only unknown system parameter in this case is the acceleration covariance matrix $\boldsymbol{Q}$.

Spatio-Temporally White Noise Acceleration

In this case the 2×2 matrix $\boldsymbol{Q}$ is constrained to be of the form $\boldsymbol{Q} = \boldsymbol{q}\, \boldsymbol{I}_2$.

AR(1) (Markov) Velocity

In this case we assume a first-order autoregressive model for the velocity $\boldsymbol{v}_{k+1} = \boldsymbol{A}\, \boldsymbol{v}_k + \boldsymbol{w}_k$ where now $\boldsymbol{A}$ and $\boldsymbol{Q}$ are unknown. We get for the

state-space model

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{p}_k \\ \boldsymbol{v}_k \end{bmatrix}, \ \boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{I}_2 \\ \boldsymbol{0}_{2,2} & \boldsymbol{A} \end{bmatrix}, \ \boldsymbol{G} = \begin{bmatrix} \boldsymbol{0}_{2,2} \\ \boldsymbol{I}_2 \end{bmatrix}, \ \boldsymbol{y}_k = \hat{\boldsymbol{p}}_k, \boldsymbol{H} = [\boldsymbol{I}_2 \ \ \boldsymbol{0}_{2,2}] \tag{6.38}$$

The case of temporally white acceleration is a special case of the AR(1) velocity model with $\boldsymbol{A} = \boldsymbol{I}_2$.

Mobility Model Performance Evaluation via Position Prediction

Although we have seen before that the mobility parameters can be adapted with the EM-Kalman filter, in the case of the M3 measurement campaign with GPS position estimates, we dispose of very good initial position estimates. As a result, the mobility model dynamics can be fitted to the GPS position estimates directly. In order to evaluate the usefulness of the various mobility models, we shall nevertheless take into account the bit of noise on the GPS position estimates, and run the Kalman filter with the estimated mobility model parameters, for each of the three models considered. Then the histogram of the position prediction error $\|\widetilde{\boldsymbol{y}}_k\|$ gets evaluated and compared between models.

The basic Kalman filter becomes:

$$\begin{aligned}
\boldsymbol{F}_k &= \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{I}_2 \\ \boldsymbol{0}_{2,2} & \boldsymbol{A}_k \end{bmatrix} \\
\hat{\boldsymbol{y}}_{k|k-1} &= \boldsymbol{H}\,\hat{\boldsymbol{x}}_{k|k-1} \\
\widetilde{\boldsymbol{y}}_k &= \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1} \\
\boldsymbol{S}_k &= \boldsymbol{H}\,\boldsymbol{P}_{k|k-1}\,\boldsymbol{H}^T + \boldsymbol{R}_{k-1} \\
\boldsymbol{K}_k &= \boldsymbol{P}_{k|k-1}\,\boldsymbol{H}^T\,\boldsymbol{S}_k^{-1} \\
\hat{\boldsymbol{x}}_{k|k} &= \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k\,\widetilde{\boldsymbol{y}}_k \\
\boldsymbol{P}_{k|k} &= \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k\,\boldsymbol{S}_k\,\boldsymbol{K}_k^T \\
\hat{\boldsymbol{x}}_{k+1|k} &= \boldsymbol{F}_k\,\hat{\boldsymbol{x}}_{k|k} \\
\boldsymbol{P}_{k+1|k} &= \boldsymbol{F}_k\,\boldsymbol{P}_{k|k}\,\boldsymbol{F}_k^T + \boldsymbol{G}\,\boldsymbol{Q}_k\,\boldsymbol{G}^T
\end{aligned} \tag{6.39}$$

Initialization: $\hat{\boldsymbol{x}}_{1|0} = [\boldsymbol{y}_1^T \ \boldsymbol{0}_{1,2}]^T$, $\boldsymbol{P}_{1|0} = \text{blockdiag}\{\boldsymbol{R}, 2\boldsymbol{R}\}$. This needs to be augmented (preceded for each sample k) with the following model parameter updates.

Temporally White, Spatially Colored Noise Acceleration

$$\begin{aligned} \boldsymbol{v}_k &= \boldsymbol{y}_k - \boldsymbol{y}_{k-1} \\ \boldsymbol{a}_k &= \boldsymbol{v}_k - \boldsymbol{v}_{k-1} \\ \gamma_k^{-1} &= \lambda\, \gamma_{k-1}^{-1} + 1 \\ \boldsymbol{Q}_k &= (1-\gamma_k)\, \boldsymbol{Q}_{k-1} + \gamma_k\, \boldsymbol{a}_k \boldsymbol{a}_k^T \ . \end{aligned} \tag{6.40}$$

with initializations $\boldsymbol{y}_0 = \boldsymbol{y}_1$, $\boldsymbol{v}_0 = 0$, $\boldsymbol{Q}_0 = \boldsymbol{0}_{2,2}$, $\gamma_0^{-1} = 0$. $\boldsymbol{A}_k = \boldsymbol{I}_2$, $\forall k$.

Spatio-Temporally White Acceleration

$$\begin{aligned} \boldsymbol{v}_k &= \boldsymbol{y}_k - \boldsymbol{y}_{k-1} \\ \boldsymbol{a}_k &= \boldsymbol{v}_k - \boldsymbol{v}_{k-1} \\ \gamma_k^{-1} &= \lambda\, \gamma_{k-1}^{-1} + 1 \\ \boldsymbol{Q}_k &= (1-\gamma_k)\, \boldsymbol{Q}_{k-1} + \tfrac{1}{2}\gamma_k\, \|\boldsymbol{a}_k\|^2\, \boldsymbol{I}_2 \ . \end{aligned} \tag{6.41}$$

with initializations $\boldsymbol{y}_0 = \boldsymbol{y}_1$, $\boldsymbol{v}_0 = 0$, $\boldsymbol{Q}_0 = \boldsymbol{0}_{2,2}$, $\gamma_0^{-1} = 0$. $\boldsymbol{A}_k = \boldsymbol{I}_2$, $\forall k$.

AR(1) (Markov) Velocity

$$\begin{aligned} \boldsymbol{v}_k &= \boldsymbol{y}_k - \boldsymbol{y}_{k-1} \\ \gamma_k^{-1} &= \lambda\, \gamma_{k-1}^{-1} + 1 \\ \boldsymbol{M}_k^{00} &= (1-\gamma_k)\, \boldsymbol{M}_{k-1}^{00} + \gamma_k\, \boldsymbol{v}_{k-1} \boldsymbol{v}_{k-1}^T \\ \boldsymbol{M}_k^{10} &= (1-\gamma_k)\, \boldsymbol{M}_{k-1}^{10} + \gamma_k\, \boldsymbol{v}_k \boldsymbol{v}_{k-1}^T \\ \boldsymbol{M}_k^{11} &= (1-\gamma_k)\, \boldsymbol{M}_{k-1}^{11} + \gamma_k\, \boldsymbol{v}_k \boldsymbol{v}_k^T \\ \boldsymbol{A}_k &= \lambda\, \boldsymbol{A}_{k-1} + (1-\lambda)\, \boldsymbol{M}_k^{10}\, (\boldsymbol{M}_k^{00} + 10^{-2}\ \mathrm{tr}\,\{\boldsymbol{M}_k^{00}\}\, \boldsymbol{I}_2)^{-1} \\ \boldsymbol{Q}_k &= \boldsymbol{M}_k^{11} - \boldsymbol{A}_k (\boldsymbol{M}_k^{10})^T - \boldsymbol{M}_k^{10} \boldsymbol{A}_k^T + \boldsymbol{A}_k \boldsymbol{M}_k^{00} \boldsymbol{A}_k^T \ . \end{aligned} \tag{6.42}$$

with initializations $\boldsymbol{y}_0 = \boldsymbol{y}_1$, $\boldsymbol{v}_0 = 0$, $\boldsymbol{A}_0 = \boldsymbol{I}_2$, $\gamma_0^{-1} = 0$.

Some practical details: for the noise level on the GPS position estimates, we have determined a value by inspection of the estimated trajectories (see trajectory plots) and we have fixed $\boldsymbol{R} = \boldsymbol{I}$ (so a standard deviation of 1m/s on x and y position components). The GPS data for various trajectories are discontinuous due to frame errors in the communication from time to time. These discontinuities lead to significant outliers in the position prediction error and hence are easy to discard.

Examples of trajectories considered in the M3 data

Here we show some of the aerial photos with trajectories superposed (GPS + Kalman filtered for AR(1) (Markov) Velocity model). GPS data is marked with blue '+' while the Kalman filtered data is marked with red 'x' marks.

Figure 6.1: A sample of an aerial photo with trajectory superposed.

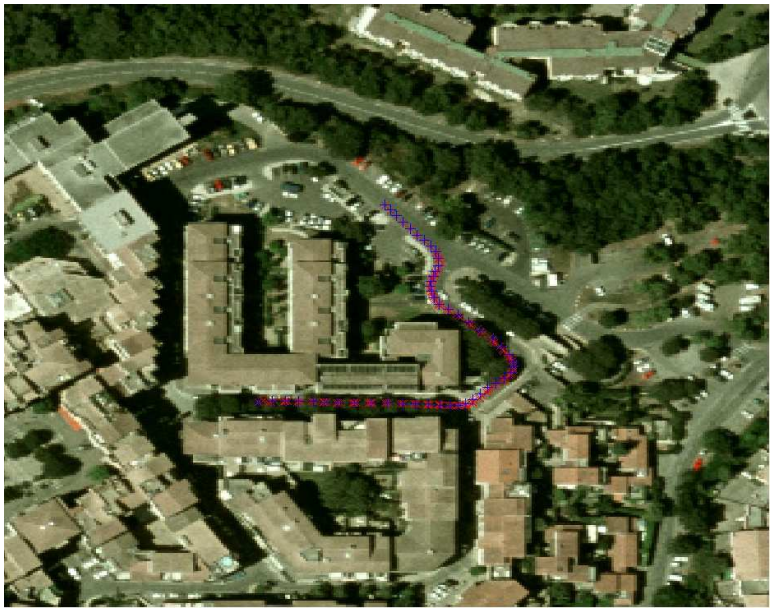

Figure 6.2: A sample of an aerial photo with trajectory superposed.

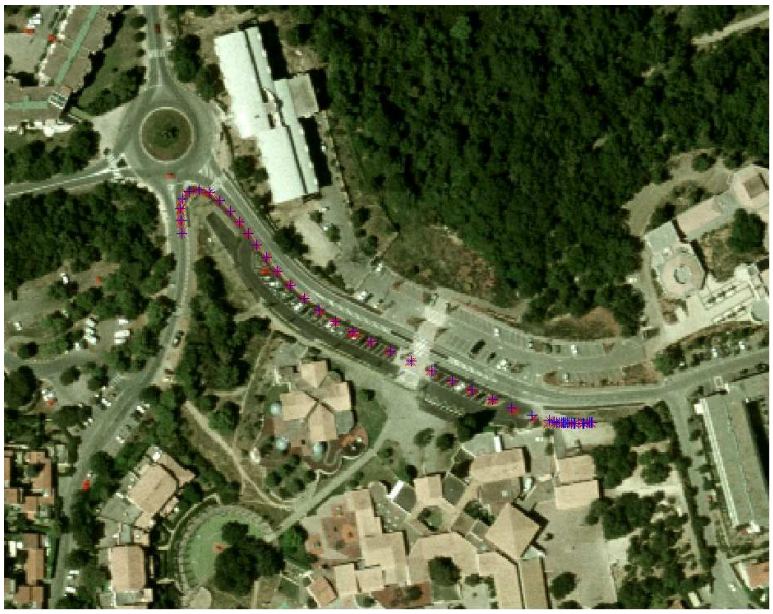

Figure 6.3: A sample of an aerial photo with trajectory superposed.

Histograms of position prediction error for the three mobility models considered

Here we show the histograms of the position prediction errors for all the three mobility models averaged over all the scenarios (driving, walking). According to the plots, we see that the performance of AR(1) (Markov) Velocity model is the best, then comes the Spatio-Temporally White Acceleration model. Temporally White, Spatially Colored Noise Acceleration model performs the worst among them. During all the simulations, the value used for λ was 0.97 and we had a total of 55928 total position samples in our database.

6.4 Conclusion

In this chapter, we have explored adaptive KF approaches, tested different mobility models. In fact, this work was in the context of a project. We had the duty to fit the mobility models to our M3 measurement campaign GPS data. Three different mobility models were tested. The results were surprising. We were expecting to see that temporally white, spatially colored

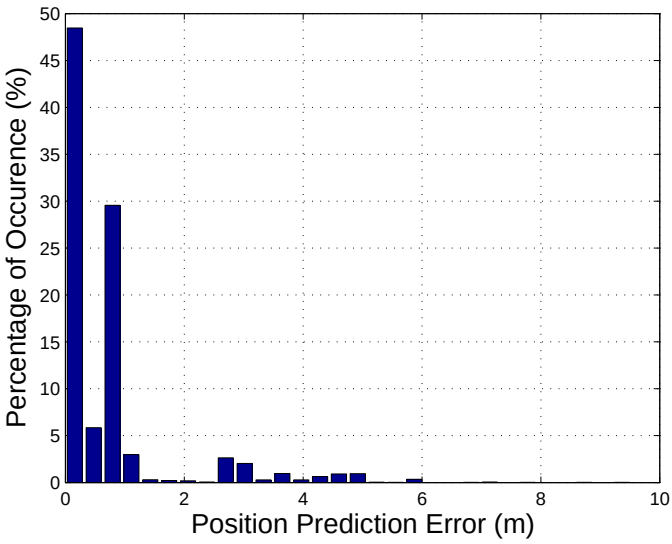

Figure 6.4: Position prediction error histogram for the temporally white, spatially colored noise acceleration model.

noise acceleration model would perform better than the spatio-temporally white acceleration model. However, just the opposite came true. This was a surprising result. Among the models investigated, AR(1) Markov velocity model has achieved the highest performance in terms of position prediction error.

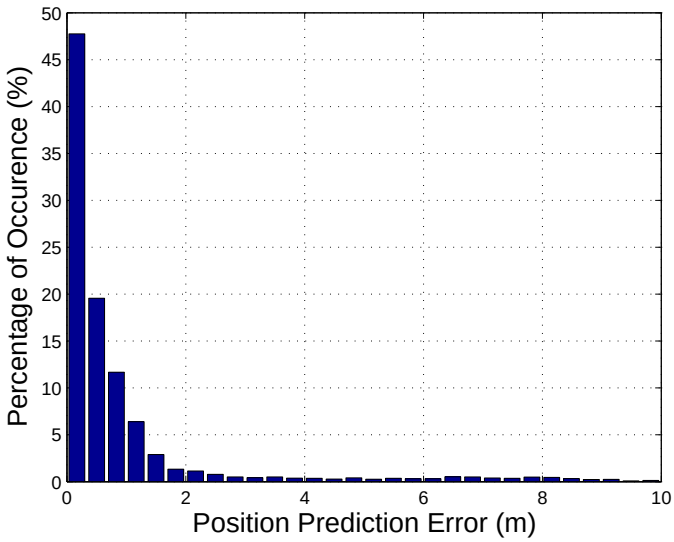

Figure 6.5: Position prediction error histogram for the spatio-temporally white acceleration model.

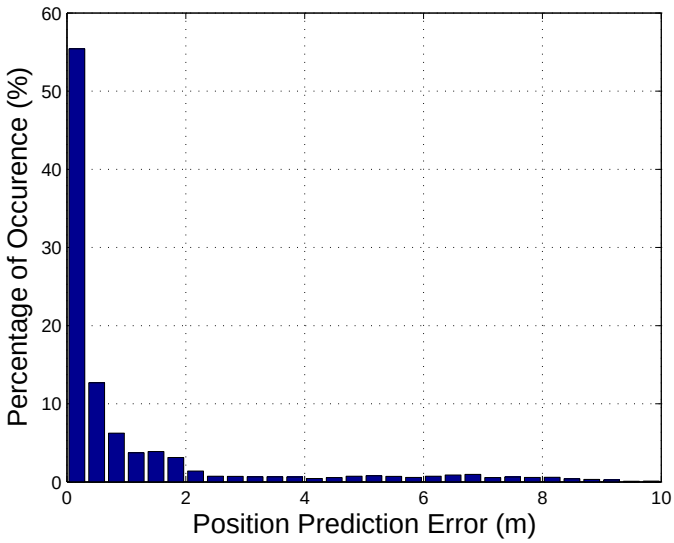

Figure 6.6: Position prediction error histogram for the AR(1) (Markov) velocity model.

Chapter 7

Conclusions and Future Work

Throughout this thesis period, we tried to conduct research in diverse fields. Mainly, our subject of interest was localization of MTs based on fingerprinting methods. In a broader sense, parameter estimation has been our main subject of research. Many results have been obtained on performance bounds of the estimation of parameters, local identifiability issues, new localization algorithms, PEP analysis, etc. Hopefully, the results, findings obtained in this work has shed light on the problems related to our research field.

Our main effort was oriented towards preparing a complete text about the the topics we investigated. For that purpose, we tried to follow a reasonable order. First we introduced two new fingerprinting methods, namely time-domain PDDP-F and frequency-domain PDDP-F. After that we tried to address performance bounds on PDP-F and PDDP-F methods. To the best of our knowledge, performance analysis on fingerprinting based localization methods has not been conducted yet. Therefore we focused on that aspect and presented CRBs on the estimation of LDPs and the position vector for various cases investigated. We figured out many intuitive and comprehensive results on CRBs and also the local identifiability of the position vector $\mathbf{r}$. Next step in our studies was a very interesting subject which is the PEP analysis for PDP-F methods. PEP analysis is a quite well-known topic in the field of digital communications. Receiver designs, detection al-

gorithms are deeply involved with PEP analysis for digital communication channels. Being inspired from this, we tried to convey the same concept to PDP-F based localization methods. This can be perceived as a kind of performance analysis. We managed to derive closed-form results and verified the accuracy of the analytical results with simulations. Final work in this dissertation was about MT tracking based on adaptive KF. In fact this was in the context of a project we were involved in (project WHERE). We investigated three mobility models for EURECOM's M3 measurement campaign and presented the results in the form of position prediction errors.

Although we tried to address a wide variety of issues about fingerprinting based localization, there are for sure many open problems/questions left uninvestigated/unanswered. Some of them listed below might be interesting problems for further research or for other researchers interested in this field:

- For the frequency-domain PDDP-F, window selection is also an important factor effecting the performance of the algorithm. Hence impact of window selection might be further investigated. Both for time-domain and frequency-domain PDDP-F algorithms, joint speed vector estimation case needs to be improved.

- Even though we have introduced the most general PDDSP, we just explored the SISO case afterwards. Hence multi-antenna extensions of the proposed fingerprinting algorithms seem definitely promising solutions to NLoS localization.

- Another important parameter is the pulse shape for the fingerprinting algorithms we have introduced. Therefore its effects on the localization performance can be studied, e.g. different pulse shapes, various roll-off factors for raised cosine filters, etc.

- Studies show that, diffuse channel parameters might have significant importance in channel modeling [49, 50]. In spite of the fact that our fingerprinting algorithms work both in specular and diffuse channel environments, we have only derived the CRBs for the specular channels. Hence it is an open question how the bounds for LDPs, position vector will be effected after the integration of diffuse channel components.

- Deriving the CRBs for PDDSP-F methods might be considered among important future research work. As can be expected, more parameters will be involved in this process. Antenna array responses as a function of AoAs, AoDs will play an important role and at the same time will require new parameters to be jointly estimated.

- We put majority of our effort on exploring the performance bounds of PDP-F algorithms. Therefore, we believe that there are still things to do about CRBs of PDDP-F methods. Mostly, the CRBs for LDPs were derived. Even though we know that after the transformation from LDPs to the position vector $\mathbf{r}$, PDDP-F methods result in strictly better localization performance, it would be nice to see the closed-form results when both delays and Doppler shifts are utilized in the localization process. This requires the integration of the Doppler shifts (f_i's) into the matrix $\mathbf{F}$. We have done this partly in the isotropic path amplitudes variances section. However it requires further investigation, further simplification to obtain intuitive results in the overall CRB expression for the estimation of $\mathbf{r}$.

- We opened a new research path by importing the PEP analysis in PDP-F. However we are just at the beginning. There are still more interesting scenarios to simulate. Moreover for the non-ergodic case in the GML based PDP-F, we could not proceed much. Only the expression was stated. However it must be evaluated (averaging over the channel coefficients). We expect to see interesting results there in terms of diversity coming from delay mismatches between the true and false position's CIRs. Naturally, other interesting future work would be to extend the analysis to PDDP-F methods which seems a bit challenging but not impossible.

- KF based tracking requires deeper thought because we were expecting a different outcome (we were expecting that temporally white spatially colored noise acceleration model would perform better than the spatio-temporally white noise acceleration model) at the end. Therefore the reasons for this surprising result should be discovered.

By far the most crucial item we did not put forth in the preceding list was applying the fingerprinting algorithms to *multi-BS* scenarios. In fact, one of the most appealing features of fingerprinting based localization methods is their ability to locate the MT with just a single BS. This is undoubtedly a very big advantage over classical localization methods which require signals from multiple BSs. Because coordination between BSs, fusion of the information obtained in each BS-MT link require huge effort. However, if there is an already established infrastructure, then why not to exploit it? More interestingly, hybrid approaches might be suggested (already exists between classical geometrical methods such as ToA/RSS, etc.) between fingerprinting and other classical or advanced geometric localization methods

to improve the localization accuracy even further. As a result, one can say that “there is plenty of room for improvement”.

Appendices

Appendix A

CRB of the Position Vector for Non-overlapping Case

We can write $\mathbf{J_r}$ as:

$$\mathbf{J_r} = \frac{1}{c^2} \sum_{i=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i} \begin{bmatrix} \cos^2 \psi_i & \cos\psi_i \sin\psi_i \\ \cos\psi_i \sin\psi_i & \sin^2 \psi_i \end{bmatrix}. \tag{A.1}$$

(3.140) is now easy to see. Proving (3.141) can be done by mathematical induction. It is straightforward to show for $N_p = 2$. Assuming that it holds for any N_p, it is easy to show that it is still valid for $N_p + 1$. However we will directly prove without using induction. If we check (A.1), we can write:

$c^4 \det \mathbf{J_r}$

$$
\begin{aligned}
&= \left(\sum_{i=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i} \cos^2\psi_i\right)\left(\sum_{i=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i} \sin^2\psi_i\right) - \left(\sum_{i=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i} \cos\psi_i \sin\psi_i\right)^2 \\
&= \sum_{i=1}^{N_p}\sum_{j=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \cos\psi_i \sin\psi_j \left(\cos\psi_i \sin\psi_j - \sin\psi_i \cos\psi_j\right) \\
&= \sum_{i=1}^{N_p}\sum_{j=1}^{N_p} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \cos\psi_i \sin\psi_j \sin(\psi_j - \psi_i) \\
&= \sum_{i=1}^{N_p}\sum_{\substack{j=1\\ j\neq i}}^{N_p} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \cos\psi_i \sin\psi_j \sin(\psi_j - \psi_i). \qquad \text{(A.2)}
\end{aligned}
$$

(A.2) stems from the fact that the terms for $i = j$ are null due to the term $\sin(\psi_j - \psi_i)$. Hence $N_p^2 - N_p$ terms are left in the summation. Note that in the $(i, j)^{th}$ turn we have a factor of $\mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \cos\psi_i \sin\psi_j \sin(\psi_j - \psi_i)$ and in the $(j, i)^{th}$ turn there is $\mathbf{J}_{\tau_j,\tau_j}\mathbf{J}_{\tau_i,\tau_i} \cos\psi_j \sin\psi_i \sin(\psi_i - \psi_j)$ in the overall summation. When we add them we have:

$$
\begin{aligned}
& \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \cos\psi_i \sin\psi_j \sin(\psi_j - \psi_i) + \mathbf{J}_{\tau_j,\tau_j}\mathbf{J}_{\tau_i,\tau_i} \cos\psi_j \sin\psi_i \sin(\psi_i - \psi_j) \\
= \; & \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \sin(\psi_i - \psi_j)\left(\cos\psi_j \sin\psi_i - \cos\psi_i \sin\psi_j\right) \\
= \; & \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \sin^2(\psi_i - \psi_j). \qquad \text{(A.3)}
\end{aligned}
$$

Let us plug (A.3) into (A.2); we now have $\binom{N_p}{2} = (N_p^2 - N_p)/2$ terms in the overall summation resulting in:

$$
c^4 \det \mathbf{J_r} = \sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} \mathbf{J}_{\tau_i,\tau_i}\mathbf{J}_{\tau_j,\tau_j} \sin^2(\psi_i - \psi_j), \qquad \text{(A.4)}
$$

which concludes the proof of (3.141).

Appendix B

Distribution of the Summation of Non-identical Exponential RVs

The proof is via mathematical induction. Let us assume that $V_1, V_2, ...V_K$ be independent exponential random variables with parameters $\lambda_1, \lambda_2, ...\lambda_K$ respectively and $W_1 = \sum_{i=1}^{K} V_i$ is the distribution we are interested in, i.e., $f_{W_1}(u)$. We begin to derive this distribution recursively by starting with $K = 2$ and then try to extend it to any K.

$$\begin{aligned} f_{V_1+V_2}(u) &= \int_{-\infty}^{\infty} f_{V_1}(u-v) f_{V_2}(v) dv \\ &= \int_{0}^{u} \lambda_1 e^{-\lambda_1(u-v)} \lambda_2 e^{-\lambda_2 v} dv \\ &= \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2} \left(e^{-\lambda_2 u} - e^{-\lambda_1 u} \right) \\ &= \lambda_1 \lambda_2 \left(\frac{e^{-\lambda_2 u}}{\lambda_1 - \lambda_2} + \frac{e^{-\lambda_1 u}}{\lambda_2 - \lambda_1} \right), \end{aligned} \tag{B.1}$$

where we used convolution to derive the distribution of summation of random variables. Now we have two exponentials for the summation of two

exponential random variables with different λ_i's. Now by using (B.1) next step is to see $f_{V_1+V_2+V_3}(u)$:

$$\begin{aligned}
f_{V_1+V_2+V_3}(u) &= \int_{-\infty}^{\infty} f_{V_3}(u-v) f_{V_1+V_2}(v) dv \\
&= \int_0^u \lambda_3 e^{-\lambda_3(u-v)} \lambda_1 \lambda_2 \left(\frac{e^{-\lambda_2 v}}{\lambda_1 - \lambda_2} + \frac{e^{-\lambda_1 v}}{\lambda_2 - \lambda_1} \right) dv \\
&= \lambda_1 \lambda_2 \lambda_3 \left(\frac{e^{-\lambda_2 u} - e^{-\lambda_3 u}}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)} + \frac{e^{-\lambda_1 u} - e^{-\lambda_3 u}}{(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)} \right) \\
&= \lambda_1 \lambda_2 \lambda_3 \left(\frac{e^{-\lambda_3 u}}{(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)} + \frac{e^{-\lambda_2 u}}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)} \right. \\
&+ \left. \frac{e^{-\lambda_1 u}}{(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)} \right). \qquad \text{(B.2)}
\end{aligned}$$

Because of the integrals involving these exponentials, we will have the similar patterns when we extend the analysis for any number of summations. The distribution has the form given in (5.13). However to prove it formally by induction, we will assume that (5.13) is valid for K summations, and try to derive the distribution for $K+1$ summations. When we begin for the case of K exponential random variables, the coefficient of $e^{-\lambda_k u}$ represented by $c_{\lambda_k}^K$ given by (5.13) is:

$$c_{\lambda_k}^K = \left(\prod_{i=1}^{K} \lambda_i \right) \left(\frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^{K} (\lambda_j - \lambda_k)} \right). \qquad \text{(B.3)}$$

To derive the distribution for the $K+1$ case, the following convolution operation is used:

$$\begin{aligned}
f_{W_1}^{K+1}(u) &= \int_{-\infty}^{\infty} f_{V_{K+1}}(u-v) f_{W_1}^K(v) dv \qquad \text{(B.4)} \\
&= \int_0^u \lambda_{K+1} e^{-\lambda_{K+1}(u-v)} f_{W_1}^K(v) dv \qquad \text{(B.5)}
\end{aligned}$$

where $f_{W_1}^K(u)$ denotes the distribution of the summation of K exponential random variables. It is easy to realize that now the coefficient of $e^{-\lambda_k u}$ has

some additional factors, and we can write it recursively as:

$$c_{\lambda_k}^{K+1} = \frac{\lambda_{K+1}}{\lambda_{K+1} - \lambda_k} c_{\lambda_k}^{K}. \tag{B.6}$$

Hence we can write $c_{\lambda_k}^{K+1}$ as:

$$c_{\lambda_k}^{K+1} = \left(\prod_{i=1}^{K+1} \lambda_i \right) \left(\frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^{K+1} (\lambda_j - \lambda_k)} \right), \tag{B.7}$$

demonstrating that the distribution still has the same form for $K+1$ which concludes the proof by induction.

Appendix C

Résumé Étendu en Français

C.1 Abstract en français

Depuis plusieurs années, le positionnement de terminal mobile reçoit un intérêt particulièrement grand. La motivation principale pour le développement de système de positionnement mobile provient essentiellement dŠune nécessité imposée par le service E-911 de lŠU.S FCC. Bien quŠau départ ils ont été utilisés pour les besoins des systèmes de sécurité dŠurgence, aujourdŠhui ils trouvent des applications dans de nombreux domaine tel que les systèmes cellulaires.

Il existe de nombreux algorithmes développés pour le problème de localisation MT. Les méthodes traditionnelles de localisation géométrique sont conçues pour fonctionner sous les conditions de line-of-sight (LoS). Cependant, les conditions LoS pourrait ne pas être toujours présentes entre la station de base (BS) et le MT. Par conséquent, les techniques de localisation basée sur fingerprinting qui sont également l'objet de cette thèse attire l'attention en raison de leur capacité à travailler aussi en multi trajet et dans des environnements non-line-of-sight (NLoS).

Dans cette thèse, nous introduisons de nouveaux algorithmes de fingerprinting, à savoir lŠalgorithme de power delay Doppler-profile fingerprinting (PDDP-F) qui exploite la mobilité du MT. Le but est d'augmenter la précision de localisation en utilisant la dimension Doppler. Nous étudions également les performances de localisation des algorithmes power delay-profile

fingerprinting (PDP-F) et PDDP-F via la dérivation des bornes de Cramer-Rao (CRBs). L'impact de la géométrie du réseau est également étudié.

Un autre sujet nous nous occupons est lŠanalyse de la probabilité d'erreur par paires (PEP) pour les méthodes PDP-F. Le PEP est une notion bien connue dans les communications numériques, et nous lŠimportons dans le domaine de la localisation pour dériver la probabilité de prendre une décision en faveur d'une mauvaise position.

Le dernier sujet sur lequel nous avons travaillé est le suivi adaptatif de lŠenviron MT en utilisant un filtre de Kalman. Différents modèles de mobilité sont comparés en termes de leurs erreurs de prédiction de position.

C.2 Contributions et Cadre de cette Thèse

Tout au long de cette période de thèse, nous avons essayé de mener des recherches dans divers domaines. Principalement, notre sujet d'intérêt était la localisation des MTs basée sur des méthodes d'empreintes. Dans un sens plus large, l'estimation des paramètres a été notre principal sujet de recherche. De nombreux résultats ont été obtenus sur des bornes de performance de l'estimation des paramètres, des questions d'identifiabilité locales, des algorithmes de localisation de nouvelles, des analyses PEP, etc. Espérons que les résultats obtenus dans ce travail a mis en lumière les problèmes liés à notre domaine de recherche.

Notre effort principal a été orientée vers la préparation d'un texte complet sur les sujets que nous avons étudiés. Pour ce faire, nous avons essayé de suivre un ordre raisonnable. D'abord, nous avons introduit deux nouvelles méthodes d'empreintes, à savoir domaine temporel PDDP-F et domaine de fréquence PDDP-F. Puis, nous avons essayé d'aborder les limites de performance sur PDP-F et PDDP-F méthodes. Pour le meilleur de notre connaissance, l'analyse des performances des méthodes de localisation basés sur les empreintes n'a pas été effectuée pour le moment. Par conséquent, nous nous sommes concentrés sur cet aspect et présentés CRBs sur l'estimation des LDPs et le vecteur de position pour les différents cas étudiés. Nous avons pensé beaucoup de résultats intuitifs et complètes sur les CRBs et aussi la possibilité de locale identifier du vecteur de position $\mathbf{r}$. Prochaine étape dans nos études était un sujet très intéressant qui est l'analyse PEP pour le PDP-F méthodes. Analyse PEP est un sujet bien connu dans le domaine des communications numériques. Conceptions du récepteur, des algorithmes de détection sont profondément impliqués dans l'analyse PEP pour voies de communication numériques. En s'inspirant

de cela, nous avons essayé de transmettre le même concept pour méthodes de localisation basées sur PDP-F. Cela peut être perçu comme une sorte d'analyse de performance. Nous avons réussi à tirer résultats analytiques et vérifié l'exactitude des résultats analytiques avec des simulations. Travail final dans cette thèse était sur le suivi des MT basée sur adaptative KF. En fait, ce fut dans le contexte d'un projet, nous ont été impliqués dans (projet WHERE). Nous avons étudié trois modèles de mobilité pour la campagne de mesure M3 d'EURECOM et a présenté les résultats sous la forme d'erreurs de prédiction de position. Tous ceux-ci nous a expliqué jusqu'ici était un bref résumé de nos réalisations. Maintenant, nous allons essayer de les élaborer en détail.

Dans le chapitre 2, deux méthodes d'empreintes nouvelles ont été introduites en utilisant la variation temporelle de la chaîne en raison de la mobilité de MT. En fait, nous avons introduit le profil le plus général, à savoir PDDSP, dont PDDP est un cas particulier (pour les canaux SISO). Première méthode, domaine de fréquence PDDP-F résout les chemins non seulement dans la dimension de retard, mais aussi dans la dimension Doppler. Naturellement, avec ce profil 2D au lieu de les profils classiques 1D de PDP-F, des exactitudes plus élevées ont été atteints. En outre la fonction de coût mises en IJuvre par l'intermédiaire d'opérations FFT 2D est beaucoup rapide et augmente également la robustesse de l'algorithme contre FO et DO.

Deuxième algorithme, que nous appelons domaine temporel PDDP-F, est basé sur l'exploitation du SOS du canal variant dans le temps. Il est une autre technique de localisation haute résolution. La nouveauté que nous avons proposé a été fondée sur le cumul des estimations de canal consécutifs. De cette manière, l'information Doppler des chemins sont devenus visibles dans les matrices de covariance. Nous avons connu l'augmentation des performances avec le n croissante, n, soit le canal consécutive estimations empilées. Même pour $n = 2$, la performance était tout à fait satisfaisant. Toutefois, dans ce cas, les problèmes FO et/ou DO (si elles sont présentes), doivent être manipulés avec soin pour que la méthode fonctionne correctement. Une partie de ce travail a été publié en:

- Turgut Mustafa Oktem and Dirk T. M. Slock, "**Power delay doppler profile fingerprinting for mobile localization in NLOS**", *in the proceedings of 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Istanbul, Turkey, 2010.

Chapitre 3 était le coeur de notre recherche où nous avons exploré

les limites de performance (CRBs) pour le PDP-F. Beaucoup de cas différents pour la modélisation d'amplitude chemin, la modélisation des phases, l'identifiabilité locale, les impulsions qui se chevauchent et ne se chevauchent pas ont été étudiés. Modèles des chemins d'amplitude isotrope/anisotrope a été introduite aussi. Nous avons découvert de nombreux résultats importants sur des sujets divers. Toutefois, nous avons publié seulement une petite quantité de nos réalisations dans ce chapitre jusqu'ici qui a été publié en:

- Turgut Mustafa Oktem and Dirk T. M. Slock, "**Cramer-Rao bounds for power delay profile fingerprinting based positioning**", *in the proceedings of IEEE 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, 2011.

Nous allons présenter reste de nos réalisations en tant que article journal:

- Turgut Mustafa Oktem and Dirk T. M. Slock, "**Cramer-Rao bounds for power delay profile fingerprinting based positioning**", under preparation.

Chapitre 4 était une extension du travail présenté dans le chapitre 3 pour la méthode PDDP-F. Naturellement, nous voyons l'amélioration dans les CRBs des LDPs puisque nous utilisons plus des estimations de canal maintenant. En outre, il ya aussi l'amélioration de l'identifiabilité de $\mathbf{r}$, par exemple, identifiabilité peut être atteint même avec un chemin pour la modélisation isotrope. Cela est dû à l'exploitation de la nouvelle dimension en provenance de Doppler. En raison de l'augmentation dans les informations obtenues à partir de LDPs, les performances de localisation est sans doute plus élevé que le PDP-F méthode. Cependant, comme c'est généralement le cas, il ya un compromis entre la complexité et la performance. En plus de DO, le problème FO pourrait également être présent dans PDDP-F. Sauf si les deux déficiences se produisent simultanément, identifiabilité locale pourrait encore être atteint avec un seul chemin.

Nous n'avons pas publié les conclusions dans ce chapitre car il est une extension du chapitre précédent. Cependant, les résultats que nous avons obtenus sont originaux et pourrait être considérée comme publiée dans un article de conférence.

Dans le chapitre 5, nous avons étudié un sujet intéressant à propos de l'analyse PEP dans le PDP-F. Nous avons dérivé des résultats analytiques closed-form. Pour le cas de chemin amplitudes déterministe en utilisant LS comme la fonction objectif, nous avons observé que la PEP diminue avec l'augmentation de la différence des vecteurs du PDP. Il s'agissait d'un

résultat attendu, en fait. Pour le cas GML Rayleigh fading, nous avons vu la saturation de la PEP autour d'une valeur non-nulle dans le régime de la haute SNR. Cela est dû à l'aléatoire dans les coefficients du canal et améliore seulement avec le nombre d'observations de canaux (T). Nous pouvons dire que l'atténuation totale de la saturation n'est possible que pour des valeurs élevées de T. Les résultats ont été publiés dans

- Turgut Mustafa Oktem and Dirk T. M. Slock, "**Pairwise error probability analysis for power delay profile fingerprinting based localization**", *in the proceedings of IEEE 73rd Vehicular Technology Conference (VTC)*, Budapest, Hungary, 2011.

Enfin, dans le chapitre 6, dans le contexte de projet WHERE, nous avons effectué un suivi de MT basée sur adaptative KF. Nous avons essayé d'adapter les paramètres de mobilité à nos mesures GPS. Différents modèles de mobilité ont été étudiées. Parmi eux, AR (1) modèle de vitesse de Markov a atteint la plus haute performance en termes d'erreur de prédiction de position. Cependant, temporellement blanc spatialement couleur bruit modèle de accélération moins bons résultats que le modèle d'accélération spatio-temporel du bruit blanc qui était surprenant pour nous. Nous n'avons pas publier ces résultats dans un article, mais plutôt qu'ils ont été présentés dans les éléments livrables du projet de WHERE.

C.3 Résumé du Chapitre 1

Dans ce chapitre d'introduction, nous allons introduire les bases du processus de localisation. La plupart des méthodes bien connues pour la localisation MT sont les techniques que l'on appelle " traditionnelle" géométriques. Ces méthodes fonctionnent généralement la localisation en deux étapes. Dans l'étape 1er, les LDPs sont estimés dans un nombre suffisant de stations de base. Parmi les plus connus LDPs, on peut énumérer AoA, ToA, TDoA, et le RSS [5]. Ensuite, en utilisant les LDPs, l'emplacement de la MT est estimée en utilisant les relations géométriques. Cependant le plus grand inconvénient de ces méthodes est qu'elles nécessitent des signaux LoS avec le BSs.

En raison des nombreux inconvénients des méthodes traditionnelles de localisation "géométriques", la recherche a été dirigée vers une autre direction. En conséquence, les méthodes de localisation basées sur les empreintes ont surgi, qui font également l'objet principal de nos recherches. Le principal avantage des méthodes d'empreintes digitales, c'est que la localisation est

possible même avec un seul BS. En outre, ils ne souffrent pas de conditions NLoS, au contraire, ils l'exploitent. Ce fut notre principale motivation pour mener des recherches dans ce domaine. Dans la figure C.1, nous illustrons un environnement multipath typique.

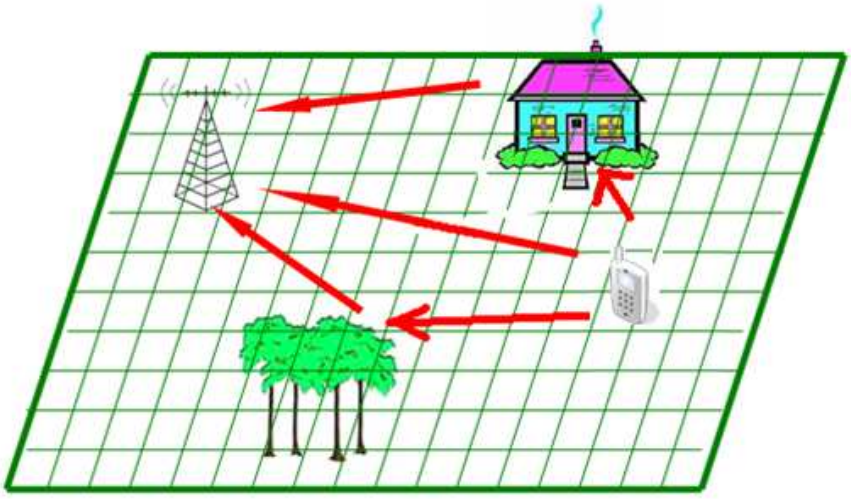

Figure C.1: Un environnement multipath.

Techniques d'empreintes sont classés parmi les techniques de DLE, c'est à dire, l'estimation de position est effectuée en une seule étape. En résumé, les concepts de base de méthodes générales de localisation basés sur les empreintes peuvent être expliquées en deux étapes. La première étape est la création et le maintien d'une base de données des empreintes. Empreintes appropriés doivent être choisis qui doit être unique pour chaque emplacement discret dans la base de données. Ceci peut être réalisé en exécutant (ray tracing ou ray lancement) ou en effectuant des simulations de campagnes de mesures sur la zone d'intérêt pour obtenir des empreintes uniques correspondant à une position spécifique. Ensuite, ces empreintes devraient être stockées dans une base de données grand. C'est la partie off-line. Deuxième étape est de l'empreinte correspondant au signal reçu une entrée dans la base de données qui est effectué en ligne par l'intermédiaire d'une sorte de fonction de likelihood.

C.3.1 Questions d'identifiabilité et Bornes de Performance

Un autre sujet intéressant que nous allons mentionner, c'est au sujet de la identifiabilité de la position MT, c'est à dire, la capacité de l'estimer. Analyse identifiabilité peut être divisé en deux, à savoir l'identifiabilité locale et l'analyse d'identifiabilité globale. Notion d'identifiabilité globale est facile à comprendre. Par exemple, pour de localisation 2D pour un système de localisation basé sur ToA, trois BSs sont suffisantes pour avoir une estimation de la position unique que nous pouvons voir dans la figure suivante:

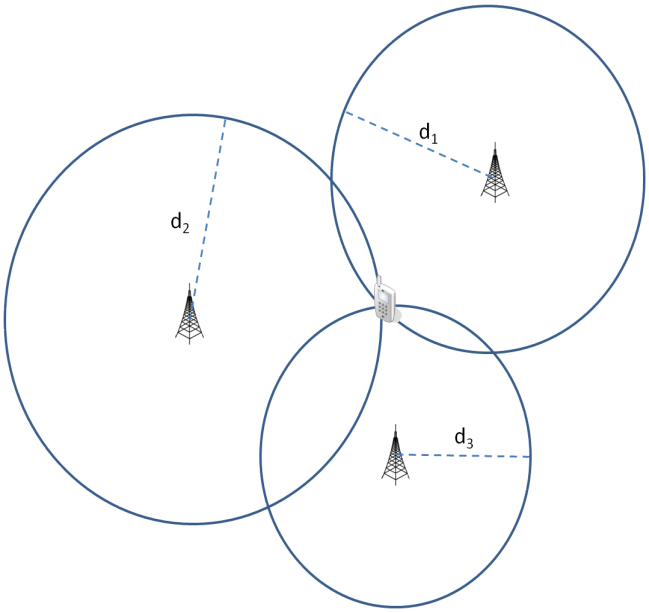

Figure C.2: Localisation basée sur ToA avec trois BSs (trilateration).

Cependant, identifiabilité locale de la position de MT est un peu plus compliqué. Identifiabilité locale est un problème similaire en ce sens que la position du mobile doit être identifié de façon unique autour d'un voisinage local du MT. Par conséquent, si seuls les signaux provenant de deux BSs sont disponibles (toujours pour le 2D ToA système), l'intersection de deux cercles se traduira par deux candidats possibles pour le poste MT. Dans ce cas, il est clair qu'il n'ya pas de identifiabilité globale. Cependant, identifiabilité locale est présente. Il existe un théorème que nous allons utiliser fréquemment dans la thèse qui offre un moyen d'analyse pour faire face à ce problème grâce à l'utilisation de la FIM. Ce théorème nous dit que les paramètres inconnus devenir localement identifiables lorsque la FIM évalué à des vraies valeurs

est inversible.

C.4 Résumé du Chapitre 2

Dans ce chapitre, nous allons introduire deux nouveaux algorithmes, à savoir domaine de fréquence PDDP-F et domaine temporel PDDP-F basé sur d'empreintes. Par défaut, tous les systèmes de localisation basés sur les empreintes besoin de bases de données. Comme nous l'avons exprimé dans le chapitre 1, ils peuvent être construits hors ligne, soit par le ray tracing, ou par des rayons de lancer des méthodes de simulation sur la zone géographique d'intérêt [31, 32]. La zone est divisée en plusieurs sections distinctes, chaque section ayant une empreinte unique. Nous allons commencer par le domaine fréquentiel PDDP-F algorithme et après que nous allons introduire le domaine temporel PDDP-F algorithme.

C.4.1 Domaine de Fréquence PDDP-F

Considérons un canal MIMO sans fil spéculaire illustré ci-dessous avec de multiples (N_t) et de transmettre (N_r) antennes de réception: La réponse

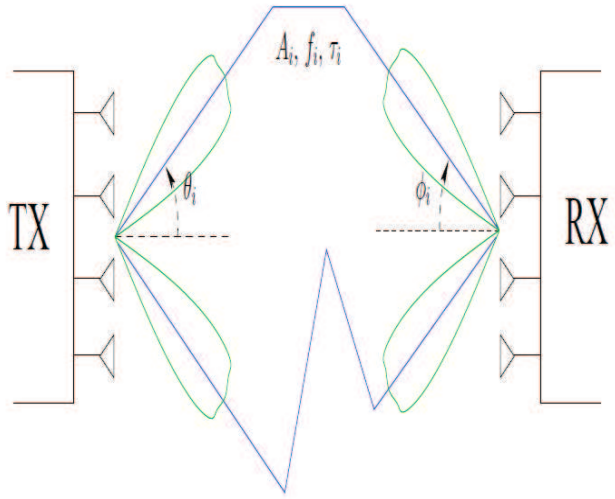

Figure C.3: MIMO multipath paramètres de propagation

varie dans le temps d'impulsion de canal est la suivante:

$$\mathbf{h}(\tau, t) = \sum_{i=1}^{Np} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{a}_R(\phi_i)\, \mathbf{a}_T^T(\theta_i)\, p(\tau - \tau_i) \tag{C.1}$$

Pour ce modèle de canal, nous pouvons maintenant introduire le profil le plus général, qui est la Power Delay Doppler Space Profile (PDDSP) comme suit:

$$\begin{aligned} &\mathrm{PDDSP}_h(\tau, f) \\ &= \int \mathbb{E}\, \underline{\mathbf{h}}(\tau, t_1 + t) \underline{\mathbf{h}}^H(\tau, t_1)\; e^{-j2\pi f t} dt \\ &= \textstyle\sum_{i=1}^{N_p} \sigma_i^2\, |p(\tau - \tau_i)|^2\, \delta(f - f_i)\, \mathbf{a}_i \mathbf{a}_i^H \\ &= \textstyle\sum_{i=1}^{N_p} \sigma_i^2\, |p(\tau - \tau_i)|^2\, \delta(f - f_i)\, R_T(\theta_i) \otimes R_R(\phi_i) \end{aligned} \tag{C.2}$$

où $\mathbf{a}_i = \mathbf{a}_T(\theta_i) \otimes \mathbf{a}_R(\phi_i)$ et nous avons introduit les covariances spatiales:

$$R_T(\theta_i) = \mathbf{a}_T(\theta_i)\mathbf{a}_T^H(\theta_i)\ , \quad R_R(\phi_i) = \mathbf{a}_R(\phi_i)\mathbf{a}_R^H(\phi_i)\ . \tag{C.3}$$

Dans le cas d'un canal SISO, nous obtenons le Power Delay Doppler Profile (PDDP)

$$\mathrm{PDDP}_h(\tau, f) = \sum_{i=1}^{N_p} \sigma_i^2\, |p(\tau - \tau_i)|^2\, \delta(f - f_i)\ . \tag{C.4}$$

Toutefois, les équations ci-dessus ne comprennent pas les effets de fenêtrage. Puisque nous avons une quantité limitée de données, nous avons besoin de prendre des effets de fenêtrage en compte. Ses effets se font sentir à l'image dans le domaine fréquentiel.

Il sera plus instructif de résumer l'ensemble du processus avec les chiffres d'illustration. Avec le AoA, le retard et la puissance des rayons, on peut construire le Power Delay Angle of Arrival Profile (PDAoAP) (PDAoAP) comme le montre la figure C.4. Alors la prochaine étape consiste à former le PDDP avec l'aide du vecteur vitesse. Avec la fusion de l'information à partir du AoA des rayons et les informations à partir du vecteur vitesse (vitesse amplitude et la direction du mouvement), décalage Doppler de chaque rayon peut être calculé. Ainsi PDDP peuvent être construits que dans la figure C.5. La dernière chose à faire est d'inclure la forme des impulsions et des effets de fenêtrage. Après tout, nous nous retrouvons avec le PDDP finale comme le montre la figure C.6.

Pour l'opération de correspondance entre le PDDP et les entrées dans la base de données, nous utilisons le rapide et vigoureuse de l'algorithme FFT 2D:

$$\frac{IFFT(FFT(\mathbf{A}) \odot conj(FFT(\mathbf{B})))}{\|\mathbf{A}\|_F\, \|\mathbf{B}\|_F} \tag{C.5}$$

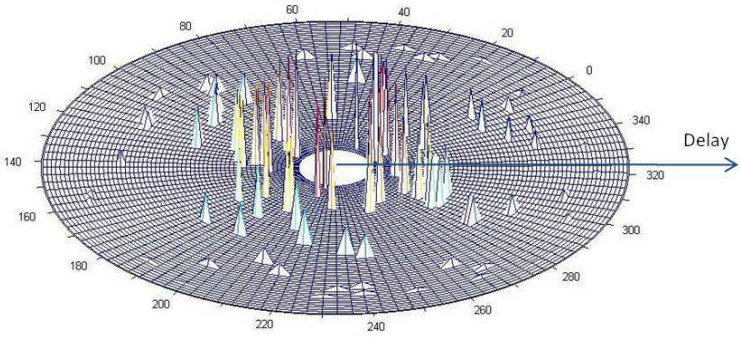

Figure C.4: Power Delay Angle of Arrival Profile

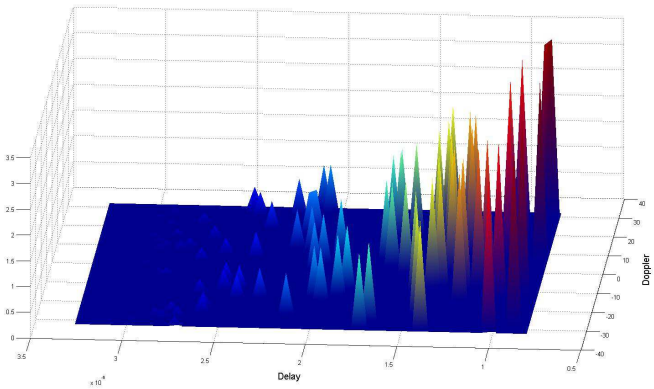

Figure C.5: Power Delay Doppler Profile

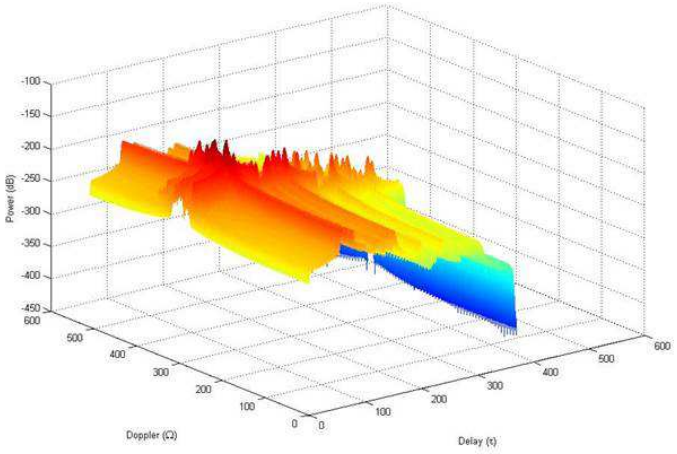

Figure C.6: Power Delay Doppler Profile avec la forme des impulsions et le fenêtrage effets inclus

où $\odot$ est la Hadamard (élément par élément) la multiplication et la *conj* dénote conjugué. L'entrée maximum dans la matrice qui en résulte est la plus forte corrélation entre les deux dans le parfaitement alignés cas.

Pour voir l'efficacité de notre algorithme, nous avons effectué quelques simulations dont les résultats sont présentés dans la figure C.7.

Nous voyons la grande amélioration dans la performance. La raison du succès est le domaine Doppler pris en compte. La capacité de résoudre les rayons en deux dimensions (à la fois en retard et Doppler) a augmenté la capacité de résolution des chemins, et donc la performance de localisation.

C.4.2 Domaine Temporel PDDP-F

Nous savons que, prises de canal échantillonnés pourrait être la superposition de plusieurs rayons qui arrivent pendant la durée d'échantillonnage que dans un environnement canal diffuse. Le théorème central limite nous permet de modéliser ces taps comme des variables random Gaussiennes. Dans cette section, nous proposons la version du domaine temporel de l'algorithme PDDP-F qui exploite les statistiques de second ordre de la chaîne. Nous supposons que le complexe vecteur fading $\mathbf{b}(t)$, et le bruit additif $\mathbf{v}(t)$ sont iid zero-mean processus vecteur Gaussien, à savoir les, les

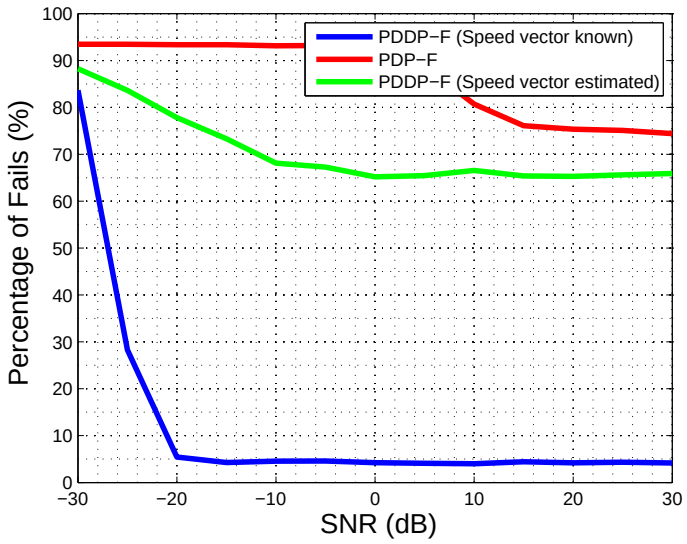

Figure C.7: Comparaison des performances du PDP-F déterministe and domaine de fréquence PDDP-F.

$$\begin{aligned} \mathbf{b}(t) &\sim \mathcal{N}\left(\mathbf{0}, \mathbf{C_b}\right) \\ \mathbf{v}(t) &\sim \mathcal{N}\left(\mathbf{0}, \sigma_v^2 \mathbf{I}_N\right) \end{aligned} \tag{C.6}$$

où $\mathcal{N}\left(\mathbf{0}, \mathbf{C_b}\right)$ désigne la zero-mean vecteur complexe Gaussien de covariance matrice $\mathbf{C_b}$, et σ_v^2 est la variance de l'erreur d'estimation de canal. Avec le modèle statistique de (2.19), $\widehat{\mathbf{h}}(t)$ est modélisé comme un i.i.d. complexe vecteur Gaussien avec $\widehat{\mathbf{h}}(t) \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}\right), \mathbf{C}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}} = \mathbf{P}_\tau \mathbf{C_b} \mathbf{P}_\tau{}^H + \sigma_v^2 \mathbf{I}_N$.
Avec la modélisation Gaussienne de $\widehat{\mathbf{h}}(t)$, nous pouvons vous proposer un solution ML du problème de localisation. Notre objectif est également de prendre en compte la variation Doppler du canal. Par conséquent, nous empiler consécutives $\widehat{\mathbf{h}}(t)$ des estimations de canal en un vecteur, au lieu de prendre un seul, et calculer les matrices de covariance basé sur ceci. Considérons maintenant la réponse du canal au moment consécutive multiples instants $t = t_s, 2t_s, \ldots, nt_s$:

$$\underbrace{\underline{\mathbf{h}}}_{nN_\tau N_r N_t \times 1} = \begin{bmatrix} \underline{\mathbf{h}}(t_s) \\ \underline{\mathbf{h}}(2t_s) \\ \vdots \\ \underline{\mathbf{h}}(nt_s) \end{bmatrix} . \tag{C.7}$$

Alors nous obtenons

$$\underline{\mathbf{h}} = \sum_{i=1}^{N_p} A_i \, \underline{e}(f_i) \otimes \underline{\mathbf{h}}_i \ , \ \ \underline{e}(f) = \begin{bmatrix} e^{j2\pi f t_s} \\ e^{j2\pi f 2t_s} \\ \vdots \\ e^{j2\pi f n t_s} \end{bmatrix} \tag{C.8}$$

Nous obtenons pour la matrice de covariance $\underline{\mathbf{h}}$

$$C_{\underline{\mathbf{hh}}} = \sum_{i=1}^{N_p} \sigma_i^2 \, R_f(f_i) \otimes R_\tau(\tau_i) \otimes R_T(\theta_i) \otimes R_R(\phi_i) \tag{C.9}$$

où

$$R_f(f) = \underline{e}(f)\underline{e}^H(f) \ , \ \ R_\tau(\tau) = \underline{p}(\tau)\underline{p}^H(\tau) \ . \tag{C.10}$$

Notez que R_f est Toeplitz. Dans le cas d'un canal SISO, nous avons $C_{\underline{\mathbf{hh}}} = \sum_{i=1}^{N_p} \sigma_i^2 \, R_f(f_i) \otimes R_\tau(\tau_i)$ et le PDDP est liée à la partie diagonale des cette matrice, après la prise de DFT de la partie R_f. Au lieu de les approches habituelles du ML pour estimer le chemin paramètres de maximiser la probabilité par rapport à la paramètres, la probabilité est évaluée en substituant la position paramètres dépendant du chemin de la base de données et, partant, il fournit la probabilité de la position. En d'autres termes, les matrices de covariance du ray tracing base de données ($\mathbf{C}_{\widehat{\underline{\mathbf{hh}}}}$) sont créés hors-ligne par le positionner les paramètres dépendants (en utilisant les retards, les pouvoirs, l'effet Doppler des rayons), puis la probabilité est évaluée à ce qui précède formulation pour les données de mesure. La position donnant l' vraisemblance la plus élevée est l'estimation de position du mobile. Likelihood peut être exprimée comme suit:

$$\mathcal{LL} \propto -\ln\left(\det \mathbf{C}_{\widehat{\underline{\mathbf{hh}}}}\right) - \ \mathrm{tr}\ \left\{ \mathbf{C}_{\widehat{\underline{\mathbf{hh}}}}{}^{-1} \widehat{\mathbf{C}}_{\widehat{\underline{\mathbf{hh}}}} \right\} \tag{C.11}$$

où $\widehat{\mathbf{C}}_{\widehat{\underline{\mathbf{hh}}}}$ est la covariance de l'échantillon observé matrice.

Une fois encore, nous avons effectué des simulations pour voir la performance de notre algorithme. L'environnement de simulation est le même que pour le domaine fréquentiel PDDP-F simulations. Nous voyons trois

courbes dans la parcelle où $n = 1$ correspond à la bayésienne PDP-F cas. Il est évident que dans le domaine temporel PDDP-F surpasse bayésienne PDP-F. L'augmentation des n (nombre d'estimations de canal consécutifs) augmente également le taux de réussite. Si nous comparons également avec le domaine fréquentiel PDDP-F algorithme, nous voyons que le domaine temporel PDDP-F est plus robuste et le taux de réussite est plus élevé pour $n \geq 3$. En outre un inconvénient de la fréquence-domaine PDDP-F, c'est que sa non-paramétrique du spectre pourrait souffrir de la résolution limitée. Dans ces simulations, nous avons supposé que le vecteur vitesse est connu d'avance.

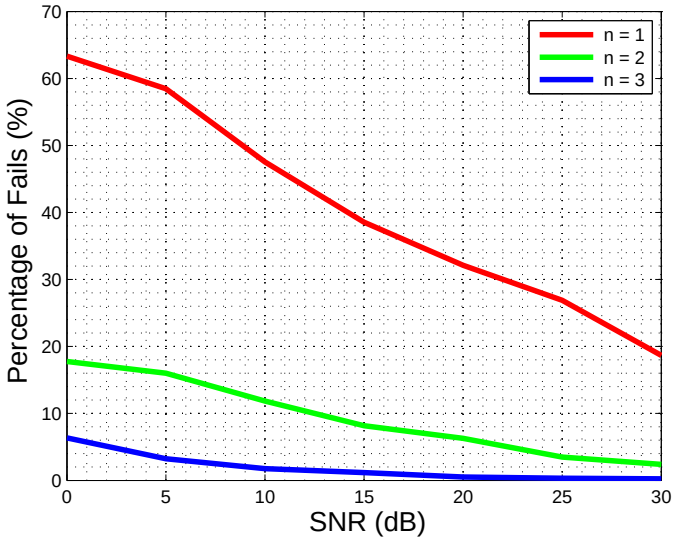

Figure C.8: Résultat de performance de domaine temporel PD(D)P-F en fonction de n.

C.5 Résumé du Chapitre 3

Ce chapitre traite de l'analyse des performances du PDP-F algorithmes. Nous pouvons dire que ce chapitre est le cIJur de la thèse. Divers modèles d'amplitude de chemin, qui se chevauchent et non-chevauchement des impulsions, DO questions, impact de la géométrie du réseau sur la performance de localisation sont étudiées. Identifiabilité locale et l'analyse CRB

sont effectuées pour chacun de ces cas. La plupart de l'analyse s'appuie sur des enquêtes FIM des paramètres à estimer. A la fin, en utilisant la transformation des paramètres, nous obtenons les CRBs du vecteur de position des CRBs des LDPs. Il ya beaucoup de résultats afin de discuter dans ce chapitre. Cependant, nous allons donner un bref résumé de nos conclusions.

Pour le cas de Rayleigh fading anisotrope, on obtient le CRB qui suit pour l'estimation du retard du chemin i^{eme} lorsque les impulsions sont non-chevauchement:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{8\pi^2 W^2 SNR_i}\left(1 + \frac{1}{SNR_i}\right) \tag{C.12}$$

où SNR_i est le SNR de la voie i^{eme} et W est la bande passante effective de l'impulsion. Pour l'estimation de la variance chemin, nous avons:

$$\mathbb{E}(\sigma_i^2 - \hat{\sigma_i^2})^2 \geq \frac{1}{\mathbf{J}_{\sigma_i^2,\sigma_i^2}} = \left(\sigma_i^2 + \sigma_v^2/e_p\right)^2. \tag{C.13}$$

Nous vous fournirons les parcelles de ces résultats dans les chiffres suivants qui incluent également le cas de chevauchement.

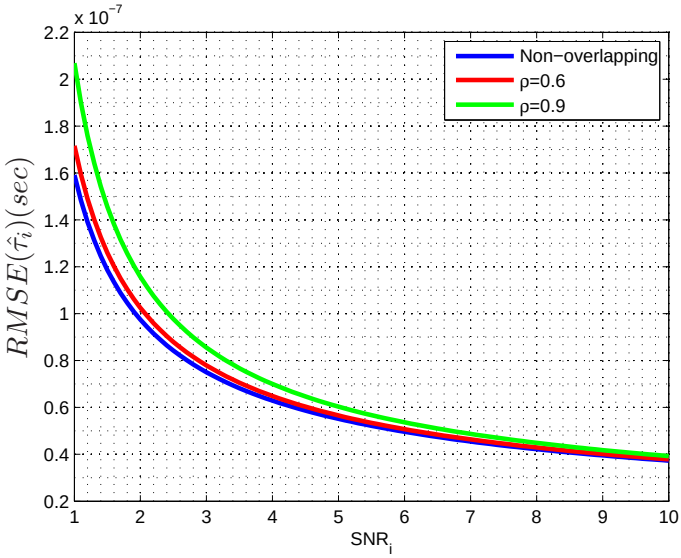

Figure C.9: RMSE de $\hat{\tau}_i$ en fonction de SNR_i pour $W = 1$ MHz, $\sigma_j^2 = 2$.

Il n'est pas difficile de se prononcer sur les chiffres. Augmentation de la bande passante permet l'estimation du retard. En outre, pour le

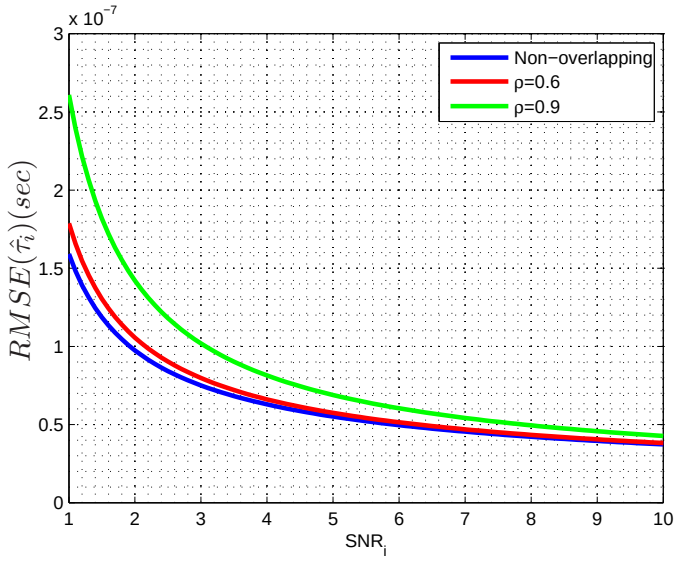

Figure C.10: RMSE de $\hat{\tau}_i$ en fonction de SNR_i pour $W = 1$ MHz, $\sigma_j^2 = 20$.

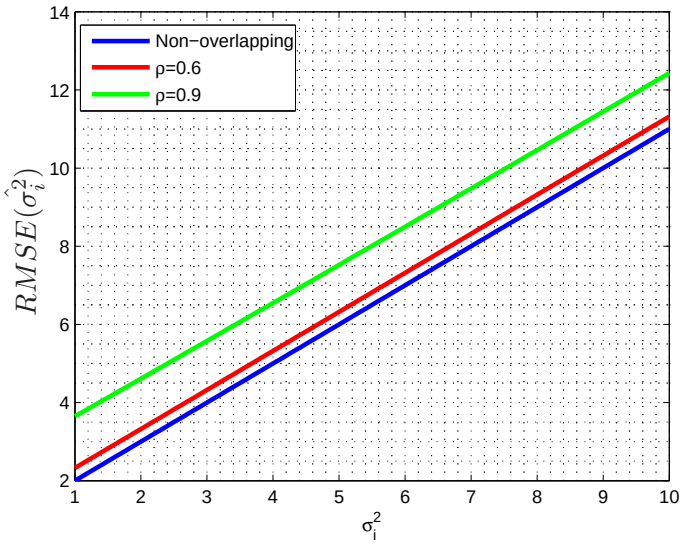

Figure C.11: RMSE de $\hat{\sigma_i^2}$ en fonction de σ_i^2 pour $W = 1$ MHz, $\sigma_j^2 = 2$.

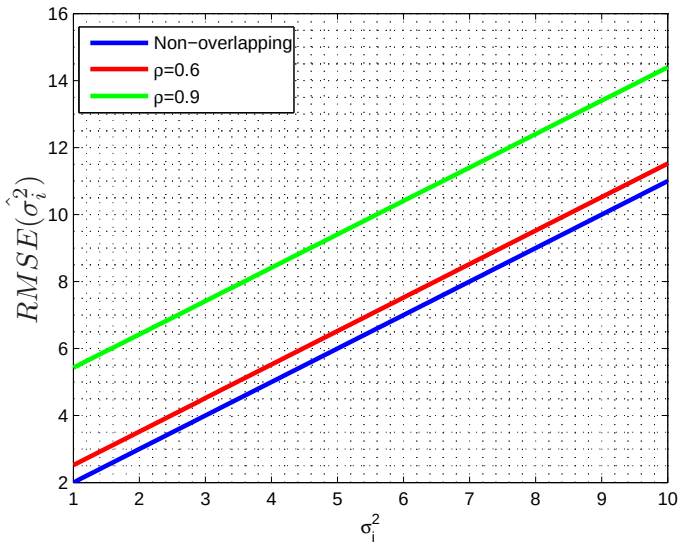

Figure C.12: RMSE de $\hat{\sigma_i^2}$ en fonction de σ_i^2 pour $W = 1$ MHz, $\sigma_j^2 = 20$.

cas de chevauchement, comme l'ingérence de la part des augmentations d'impulsions d'autres, l'estimation devient plus difficile que prévu. Notre intérêt principal est l'obtention du CRB pour l'estimation du vecteur position et étudie également l'analyse identifiabilité locale de la position de la MT. Pour cela, nous allons utiliser la transformation que nous avons mentionné avant, qui peut être obtenu par la formule suivante [35]:

$$\mathbf{J_r} = \mathbf{F}\mathbf{J}_\theta\mathbf{F}^H \tag{C.14}$$

où $\mathbf{F} = \frac{\partial\theta}{\partial\mathbf{r}}\big|_{\mathbf{r}=\mathbf{r}_0}$ ($\mathbf{r}_0 = [x_0, y_0]^T$ est la position réelle du mobile). Pour une configuration générique comme celui ci-dessous, nous allons présenter le résultat pour les amplitudes de chemin déterministes avec des phases déterministes.

Ainsi, pour $N_p \geq 2$ et non-chevauchement des impulsions:

$$\mathbb{E}\|\mathbf{r} - \hat{\mathbf{r}}\|^2 = \zeta \frac{\sum_{i=1}^{N_p} SNR_i}{\sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} SNR_i\, SNR_j \sin^2(\psi_i - \psi_j)}, \tag{C.15}$$

avec $\zeta = \frac{c^2}{8\pi^2 W^2}$. C'est un résultat important et a la même structure de l'équation donnée dans [35]. L'intrigue de ce résultat peut être vu ci-dessous:

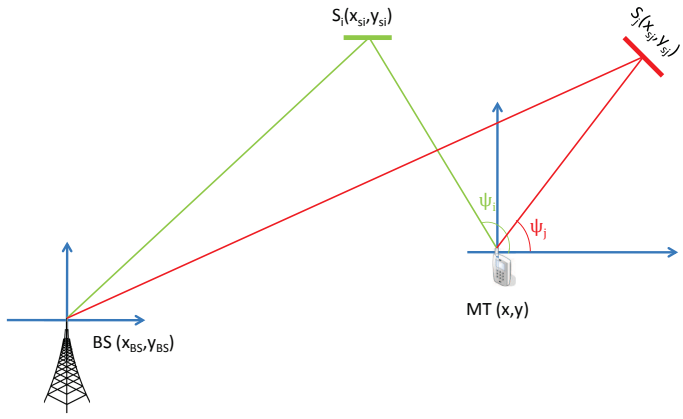

Figure C.13: La géométrie de la BS-MT.

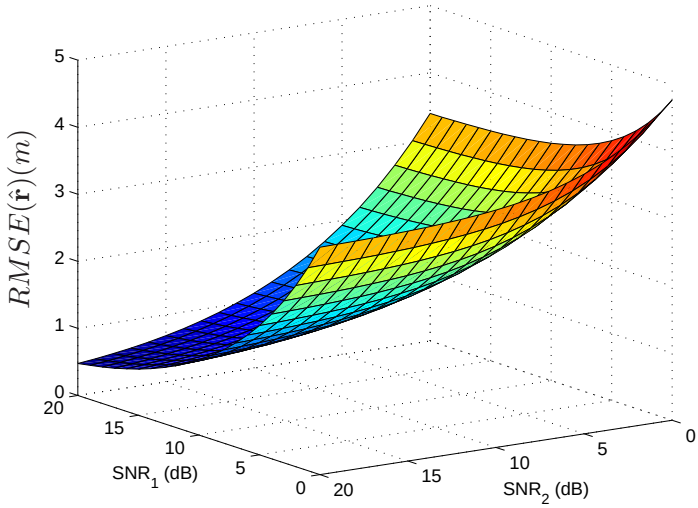

Figure C.14: RMSE de $\hat{\mathbf{r}}$ en fonction de SNR_1 et SNR_2 pour $\psi_2 - \psi_1 = \pi/2$ pour la cas deterministic.

En utilisant la même méthodologie, de dérivation pour le cas de Rayleigh fading est simple:

$$\mathbb{E}\|\mathbf{r}-\hat{\mathbf{r}}\|^2 = \zeta \frac{\sum_{i=1}^{N_p} \frac{SNR_i^2}{SNR_i+1}}{\sum_{i=1}^{N_p-1}\sum_{j=i+1}^{N_p} \frac{SNR_i^2\, SNR_j^2}{(SNR_i+1)(SNR_j+1)} \sin^2(\psi_i-\psi_j)}. \tag{C.16}$$

et son intrigue est donnée dans la figure C.15.

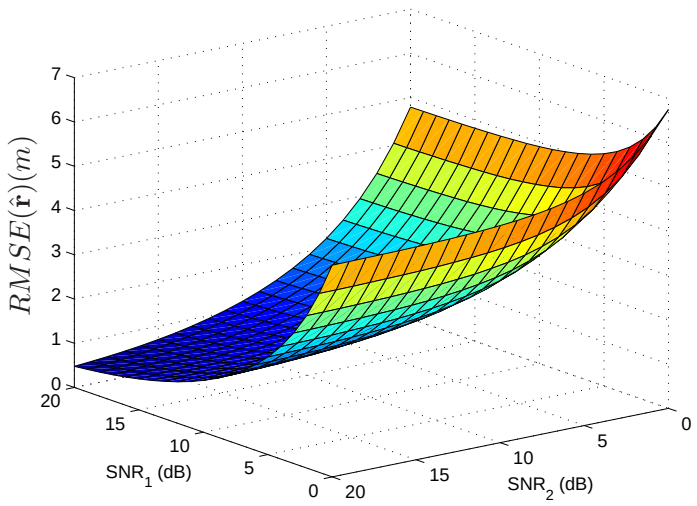

Figure C.15: RMSE de $\hat{\mathbf{r}}$ en fonction de SNR_1 et SNR_2 pour $\psi_2-\psi_1=\pi/2$ pour la cas Rayleigh fading.

C.5.1 Conclusions

Dans le chapitre 3, nous avons découvert de nombreux résultats importants sur des sujets divers. Nous pouvons les énumérer brièvement comme suit:

- identifiabilité locale du vecteur de position $\mathbf{r}$ peut être atteint même avec un chemin pour le cas anisotrope. Étonnamment, forme d'impulsion joue un rôle important dans l'identifiabilité locale. Pour le cas isotrope, au moins deux chemins sont tenus.

- Estimation des retards de voie s'améliore avec SNR et avec la bande passante effective de la forme de l'impulsion W. L'estimation des amplitudes chemin, d'autre part est indépendant de W. En outre l'estimation de la même LDP sous la modélisation évanouissement de Rayleigh est plus difficile que l'amplitude chemin déterministe de la modélisation déterministe phases.

- En cas de DO, au moins deux voies sont nécessaires pour l'identifiabilité locale de $\mathbf{r}$, même pour le cas anisotrope.

- Expressions CRB analytiques pour l'estimation des LDPs pour pulses qui se chevauchent a également été démontrée. Nous avons vu que l'ingérence des impulsions entrave l'estimation des LDPs. Si un système de localisation basée sur ToA s'appuyant sur le chemin LoS est utilisé, et si l'impulsion LoS interfère avec certains autres impulsions, alors la précision de localisation seriez peut-être sensiblement dégradées en fonction du rapport qui se chevauchent.

- Expressions CRB analytiques pour l'estimation de $\mathbf{r}$ ont été tirés à la fois pour les cas déterministe et Rayleigh. Nous avons directement constaté les effets positifs de l 'W et SNR' des trajectoires individuelles. Aussi l'impact de la géométrie environnante (GDOP) a également été mentionnée explicitement. Les conditions de géométrie améliorants ou dégradants de la performance de localisation ont été expliqués. En outre le cas particulier de deux impulsions qui se chevauchent a été démontrée. Pour le cas évanouissement de Rayleigh, il se dégrade toujours la performance de localisation. Toutefois, pour le cas déterministe, il pourrait améliorer la performance de localisation pour certaines valeurs spécifiques de la différence de phase, ratios de chevauchent, etc.

C.6 Résumé du Chapitre 4

Dans ce chapitre, notre objectif est d'étendre l'analyse que nous avons fait dans le chapitre précédent pour PDDP-F. Nous utilisons presque la même méthodologie. Nous ne répéterons pas les mêmes équations à nouveau, au lieu nous allons juste vous présenter nos nouvelles découvertes. Comme nous l'avons expliqué précédemment, pour PDDP-F, nous utilisons plus d'un estimations de canal ($n > 1$). En conséquence, nous nous attendons à une amélioration des estimations. En effet l'analyse FIM il vérifie. Pour le cas déterministe anisotrope où la phase est modélisée comme inconnues déterministes, on obtient les résultats suivants:

$$\mathbb{E}(\tau_i - \hat{\tau}_i)^2 \geq \frac{1}{8\,n\,\pi^2 W^2 SNR_i}, \tag{C.17}$$

De même, nous tirons:

$$\mathbb{E}(a_i - \hat{a}_i)^2 \geq \frac{a_i^2}{2\,nSNR_i}, \tag{C.18}$$

$$\mathbb{E}(f_i - \hat{f}_i)^2 \geq \frac{3}{2\,n\,(n^2-1)\,\pi^2\,t_s^2\,SNR_i}, \tag{C.19}$$

et de même pour ϕ_i:

$$\mathbb{E}(\phi_i - \hat{\phi}_i)^2 \geq \frac{2\,n-1}{n\,(n+1)\,SNR_i}. \tag{C.20}$$

Nous voyons clairement l'effet d'amélioration de la n sur les estimations des LDPs.

C.6.1 Conclusions

Chapitre 4 était une extension du travail présenté dans le chapitre 3 pour la méthode PDDP-F. Naturellement, nous voyons l'amélioration dans les CRBs des LDPs puisque nous utilisons plus des estimations de canal maintenant. En outre, il ya aussi l'amélioration de l'identifiabilité de bfr, par exemple, identifiabilité peut être atteint même avec un chemin pour la modélisation isotrope. Cela est dû à l'exploitation de la nouvelle dimension en provenance de Doppler. En raison de l'augmentation dans les informations obtenues à partir de LDPs, les performances de localisation est sans doute plus élevé que le PDP-F méthode. Cependant, comme c'est généralement le cas, il ya un compromis entre la complexité et la performance. En plus de DO, le problème FO pourrait également être présent dans PDDP-F. Sauf si les deux déficiences se produisent simultanément, identifiabilité locale pourrait encore être atteint avec un seul chemin.

C.7 Résumé du Chapitre 5

Nous avons maintenant affaire à un nouveau problème dans le domaine de la localisation. Dans ce chapitre, les performances des PEP PDP-F algorithme sera étudiée. En fait, l'analyse PEP est un concept bien connu dans la communication numérique, i.e., calcul de la probabilité d'erreur quand un vecteur de symboles $\mathbf{s_m}$ est transmis, mais un autre vecteur de symboles

$\mathbf{s_n}$ est détecté au niveau du récepteur [38]. Sous différentes modélisations d'amplitude de chemin et les fonctions de coût, nous allons essayer d'obtenir des résultats closed form pour la PEP.

C.7.1 PEP de la Technique LS pour des Amplitudes de Chemin Déterministes

Nous utilisons les mêmes modèles de canaux qui ont été utilisées dans les chapitres précédents, et l'analyse est limitée à canaux SISO nouveau. Nous pouvons définir le PEP comme suit lorsque les critères de LS est la fonction de coût pour des amplitudes de chemin déterministes:

$$\text{PEP} = \Pr\left(||\widehat{\mathbf{PDP}} - \mathbf{PDP}_F|| < ||\widehat{\mathbf{PDP}} - \mathbf{PDP}_T||\right), \tag{C.21}$$

où $\widehat{\mathbf{PDP}}$ est le vecteur PDP estimé défini dans (5.4), $\mathbf{PDP}_T$ est le vecteur PDP vrai, qui est calculé hors-ligne à partir de la base de données stockées et $\mathbf{PDP}_F$ est le vecteur PDP à tort détecté. Dans certaines hypothèses simplificatrices et en obtenant un avantage de la CLT (lorsque le nombre d'observations canal T est suffisamment élevée), nous obtenons le résultat suivant intuitive:

$$\text{PEP} = Q\left(\frac{\sqrt{T}}{2\sigma_v^2}||\mathbf{PDP}_{true} - \mathbf{PDP}_{false}||\right). \tag{C.22}$$

où $\mathbf{PDP}_{false}$ and $\mathbf{PDP}_{true}$ sont les versions de débruitées $\mathbf{PDP}_F$ et $\mathbf{PDP}_T$ respectivement. Comme d'habitude, nous vérifions notre analyse des résultats de simulation où α dénote le niveau de séparation entre les vecteurs PDP vrais et faux:

C.7.2 PEP de la Technique de GML pour Rayleigh Fading

Maintenant, nous étudions le cas où les amplitudes de chemin complexes sont modélisés comme des variables random Gaussiennes. Dans ce cas, lorsque la fonction de coût est basée GML, nous pouvons formaliser PEP comme suit:

$$\text{PEP} = \Pr\left(\mathcal{LL}_T < \mathcal{LL}_F\right). \tag{C.23}$$

Cela peut être interprété de telle sorte que la probabilité que le likelihood logarithmique effectué dans la position vraie est inférieure à le likelihood logarithmique dans la position fausse qui conduit à la position de fausse à être sélectionné. Dans le cas général ergodique, en utilisant la CLT, nous obtenons pour le PEP:

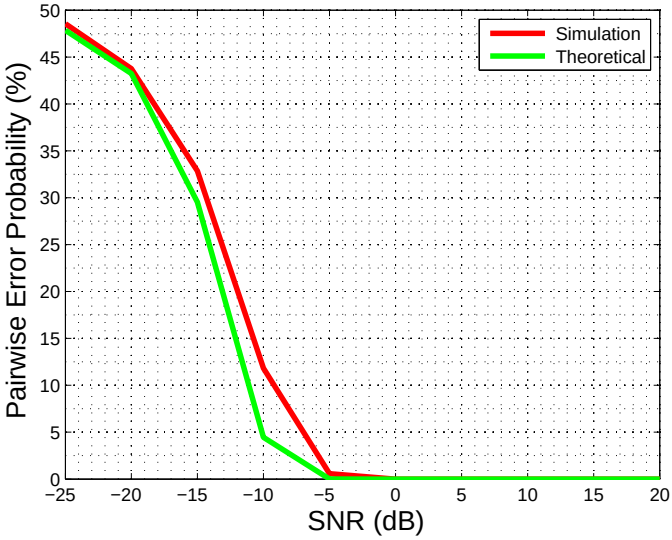

Figure C.16: PEP pour $\alpha = 1.1$.

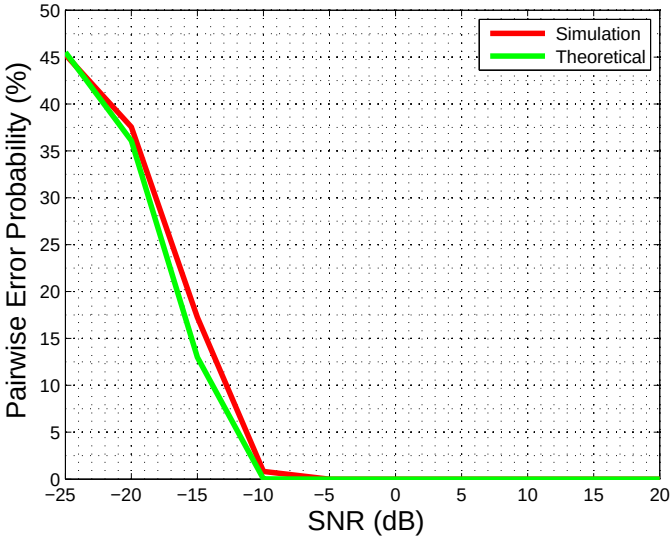

Figure C.17: PEP pour $\alpha = 1.2$.

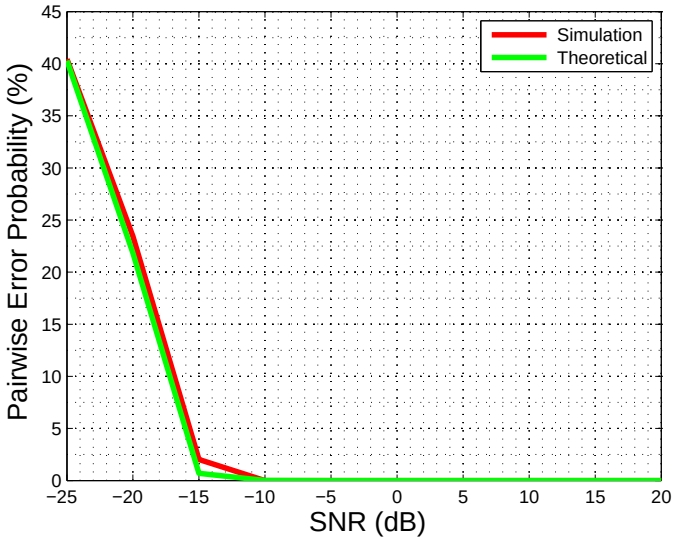

Figure C.18: PEP pour $\alpha = 1.4$.

$$\mathrm{PEP} = Q\left(\frac{\mathrm{tr}\left\{\mathbf{C_T}\mathbf{C_F^{-1}} - \mathbf{I}\right\} - \ln\det(\mathbf{C_T}\mathbf{C_F^{-1}})}{\sqrt{\frac{1}{T}\,\mathrm{tr}\left\{(\mathbf{C_T}\mathbf{C_F^{-1}} - \mathbf{I})^{\mathbf{2}}\right\}}}\right) \tag{C.24}$$

à partir de laquelle nous voyons que un décalage dans chaque chemin séparément contribue à diminuer le PEP lorsque les retards de voie sont bien séparés (le numérateur de l'argument de la fonction Q est une forme de la distance d'Itakura-Saito entre les matrices de covariance).

Pour vérifier les résultats de l'analyse, nous avons effectué des simulations pour le cas ergodique. Les écarts d'amplitude chemin de la véritable position des entrées sont sélectionnées pour être juste un multiple de celles de faux, i.e., $\sigma_{a_i}^2 = \beta^2\,\sigma_{b_i}^2$.

De toute évidence, l'augmentation de β a pour effet d'augmenter la distance entre les matrices de covariance. Par conséquent, comme on peut aussi être vu dans les simulations, PEP diminue avec le croissante β.

C.7.3 Conclusions

Un autre sujet intéressant que nous avons étudié était l'analyse PEP dans le PDP-F. Nous avons dérivé des résultats analytiques closed-form. Pour le

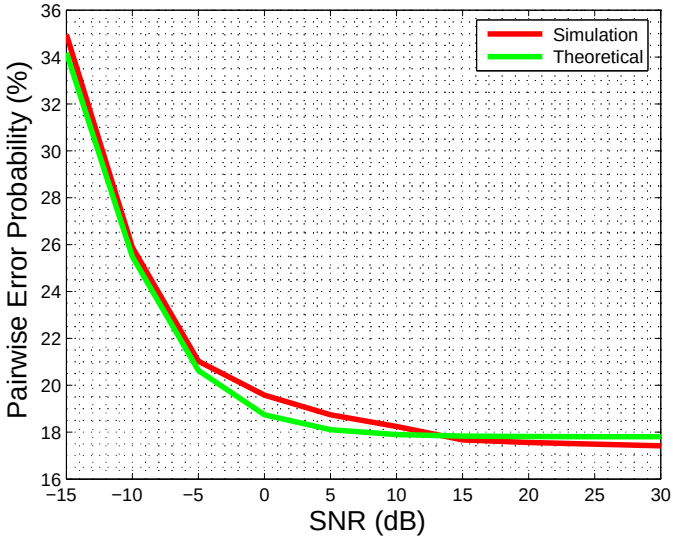

Figure C.19: PEP pour $\beta = 1.1$.

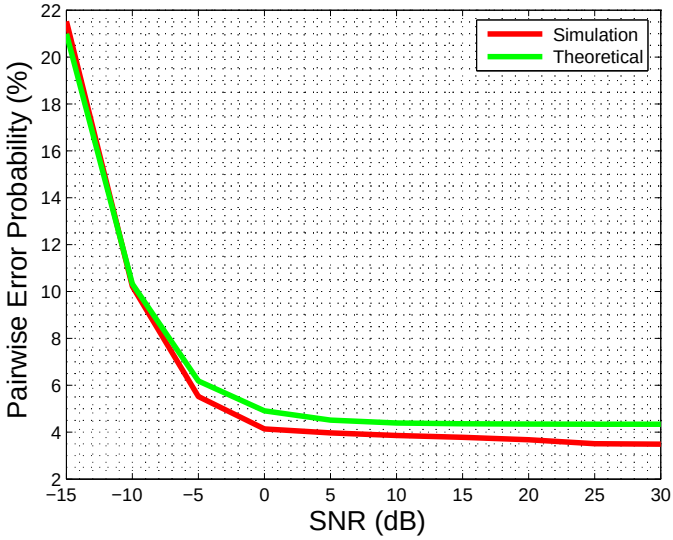

Figure C.20: PEP pour $\beta = 1.2$.

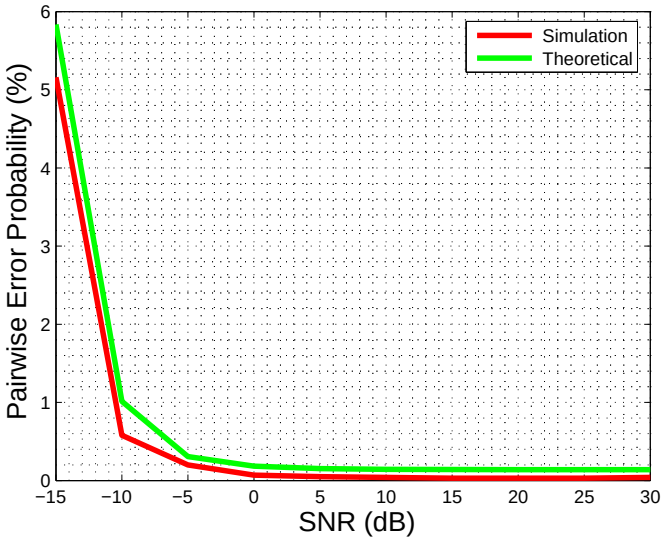

Figure C.21: PEP pour $\beta = 1.4$.

cas de chemin amplitudes déterministe en utilisant LS comme la fonction objectif, nous avons observé que la PEP diminue avec l'augmentation de la différence des vecteurs du PDP. Il s'agissait d'un résultat attendu, en fait. Pour le cas GML Rayleigh fading, nous avons vu la saturation de la PEP autour d'une valeur non nulle dans le régime de la haute SNR. Cela est dû à l'aléatoire dans les coefficients du canal et améliore seulement avec le nombre d'observations de canaux (T). Nous pouvons dire que l'atténuation totale de la saturation n'est possible que pour des valeurs élevées de T.

C.8 Résumé du Chapitre 6

Dans ce chapitre, dans le contexte de projet WHERE, nous avons effectué un suivi des MT basée sur adaptative KF. Nous avons essayé d'adapter les paramètres de mobilité à nos mesures GPS. Différents modèles de mobilité ont été étudiées. Parmi eux, AR (1) modèle de vitesse de Markov a atteint la plus haute performance en termes d'erreur de prédiction de position. Cependant temporellement blanc spatialement couleur bruit modèle de accélération moins bons résultats que le modèle d'accélération spatio-temporel du bruit blanc qui était surprenant pour nous.

Les histogrammes des erreurs de prédiction de position que nous avons obtenus pour chacun des modèles de mobilité peut être vu dans les figures C.22, C.23 and C.24:

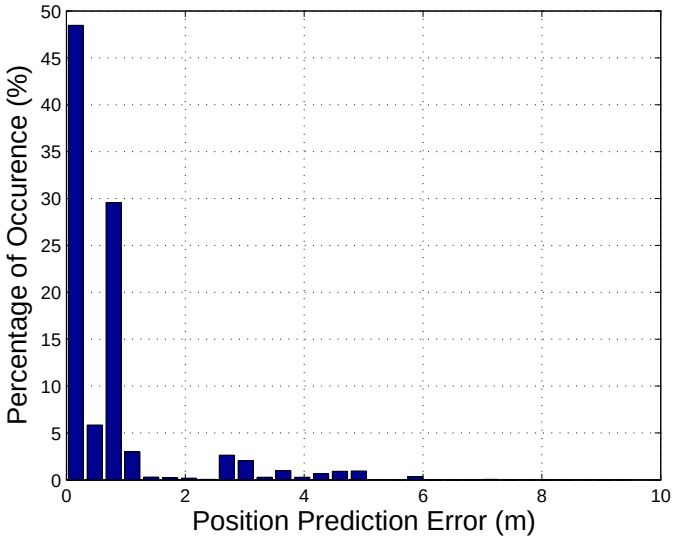

Figure C.22: Histogramme erreur de position de prédiction pour le blanc temporellement, le bruit spatialement coloré modèle d'accélération.

C.9 Futures travaux

Bien que nous avons essayé de répondre à une grande variété de questions relatives à la localisation basée sur les empreintes, il ya certainement de nombreux problèmes ouverts laissés objet d'aucune enquête, et de nombreuses questions sans réponse. Certains d'entre eux ci-dessous peut y avoir des problèmes intéressants pour la recherche supplémentaire ou pour d'autres chercheurs qui s'intéressent à ce domaine:

- Pour le domaine fréquentiel PDDP-F, sélection de fenêtre est également un facteur important à effectuer de la performance de l'algorithme. Par conséquent, l'impact de sélection de une fenêtre appropriée pourrait être étudiée plus avant. Tant pour le domaine temporel et domaine fréquentiel PDDP-F algorithmes, commune cas l'estimation du vecteur vitesse doit être améliorée.

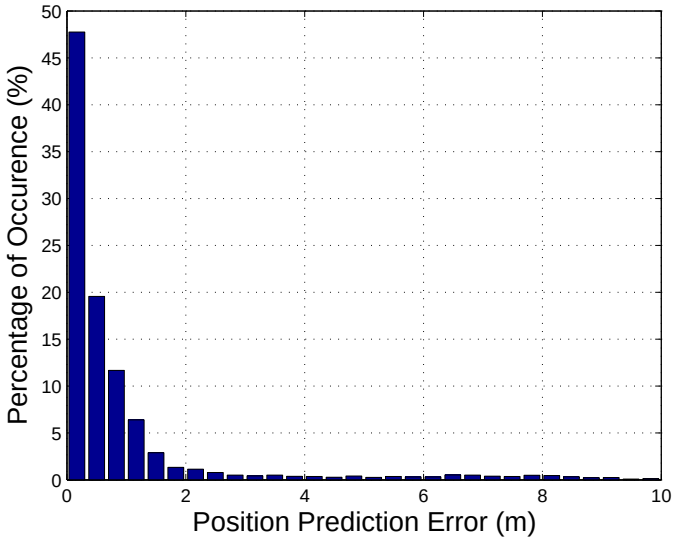

Figure C.23: Histogramme erreur de position de prédiction pour l'accélération spatio-temporellement blanc modèle.

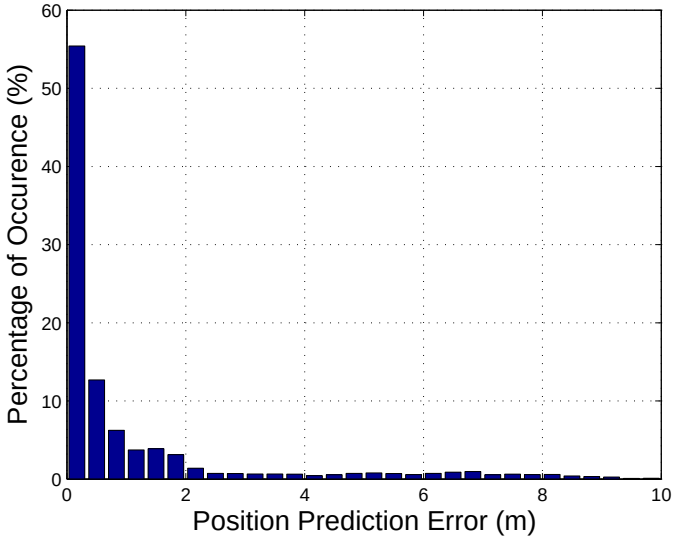

Figure C.24: Histogramme erreur de position de prédiction pour AR(1) (Markov) modèle de vitesse.

- Même si nous avons introduit la PDDSP plus générale, nous venons d'explorer le cas SISO par la suite. Par conséquent, multi-antennes extensions des algorithmes d'empreintes proposées semblent prometteuses solutions définitivement à la localisation NLoS.

- Un autre paramètre important est la forme de l'impulsion pour les algorithmes d'empreintes que nous avons introduites. Par conséquent, ses effets sur la performance de localisation peuvent être étudiées, par exemple, formes d'impulsions différentes, divers facteurs roll-off pour les filtres raised cosine, etc.

- Les études montrent que, les paramètres de canal diffus pourrait avoir une importance considérable dans la modélisation de canal [49, 50]. En dépit du fait que nos algorithmes d'empreintes peut fonctionner à la fois dans des environnements de canaux spéculaires et diffus, nous avons seulement tiré les CRBs pour les canaux spéculaires. Par conséquent, il s'agit d'une question ouverte de savoir comment les bornes pour les LDPs, vecteur de position sera effectuée après l'intégration des composants du canal diffuses.

- Dérivation des CRBs pour PDDSP-F méthodes pourraient être considérés parmi les importants travaux de recherche futurs. Comme on peut s'y attendre, plus de paramètres seront impliqués dans ce processus. Réponses de réseau d'antenne en fonction des AoAs, AoDs jouera un rôle important et en même temps, il faudra de nouveaux paramètres pour être estimées conjointement.

- Nous avons mis la majorité de nos efforts sur l'exploration des limites de performance de PDP-F algorithmes. Par conséquent, nous croyons qu'il ya encore des choses à faire au sujet des CRBs de PDDP-F méthodes. La plupart du temps, les CRBs pour les LDPs ont été tirées. Même si nous savons que, après la transformation des LDPs au vecteur position $\mathbf{r}$, PDDP-F méthodes aboutissent à des performances de localisation strictement meilleure, il serait agréable de voir les résultats closed-form lorsque les deux retards et Doppler shifts sont utilisés de la processus de localisation. Cela exige l'intégration des Doppler shifts (f_i's) dans la matrice $\mathbf{F}$. Nous l'avons fait en partie dans le chemin amplitudes section isotrope écarts. Toutefois, elle exige une enquête plus poussée, une plus grande simplification pour obtenir des résultats intuitifs dans l'expression CRB globale pour l'estimation de $\mathbf{r}$.

- Nous avons ouvert une voie de recherche nouvelle par l'importation de l'analyse PEP dans le PDP-F. Cependant, nous sommes juste au début. Il ya encore des scénarios les plus intéressants à simuler. En outre, pour le cas non-ergodique dans la base de GML PDP-F, nous ne pouvions pas avancer beaucoup. Seule l'expression a été déclaré. Cependant elle doit être évaluée (en moyenne au cours des coefficients du canal). Nous nous attendons à voir des résultats intéressants il en termes de diversité à venir à des décalages de retard entre la position CIRs vrai et le faux du. Naturellement, d'autres travaux futurs serait intéressant d'étendre l'analyse à PDDP-F méthodes qui semble un peu difficile mais pas impossible.

- Suivi des MT basée sur adaptative KF nécessite plus pensé parce que nous nous attendions à un résultat différent (nous nous attendions à ce que blanc temporellement modèle spatialement couleur accélération du bruit serait plus efficace que le modèle spatio-temporel du bruit d'accélération blanc) à la fin. Par conséquent, les raisons de ce résultat surprenant doit être découvert.

De loin l'élément le plus crucial, nous n'avons pas mis en avant dans la liste précédente a été d'appliquer les algorithmes d'empreintes à *multi-BS* scénarios. En fait, l'une des caractéristiques les plus intéressantes de prise d'empreintes digitales des méthodes de localisation basés sur est leur capacité à localiser le MT avec juste un BS unique. C'est sans doute un très grand avantage par rapport aux méthodes classiques de localisation qui nécessitent des signaux provenant de plusieurs stations de base. Parce que la coordination entre les stations de base, la fusion des informations obtenues dans chaque lien BS-MT exigent un effort énorme. Cependant, si il ya une infrastructure déjà établie, alors pourquoi ne pas l'exploiter? Plus intéressant encore, les approches hybrides pourrait être suggéré (existe déjà entre les méthodes classiques géométriques tels que ToA / RSS, etc) entre les empreintes digitales et d'autres classiques ou des méthodes de localisation avancée géométriques pour améliorer la précision de la localisation encore plus loin. En conséquence, on peut dire que " il ya beaucoup de place à l'amélioration ".

Bibliography

[1] H. Koshima and J. Hoshen, "Personal locator services emerge," *Spectrum, IEEE*, vol. 37, no. 2, pp. 41 –48, feb 2000.

[2] J. Caffery and G. Stuber, "Overview of radiolocation in cdma cellular systems," *Communications Magazine, IEEE*, vol. 36, no. 4, pp. 38 –45, apr 1998.

[3] G. Sun and W. Guo, "Robust mobile geo-location algorithm based on ls-svm," *Vehicular Technology, IEEE Transactions on*, vol. 54, no. 3, pp. 1037 – 1041, may 2005.

[4] S. Swales, J. Maloney, and J. Stevenson, "Locating mobile phones and the us wireless e-911 mandate," in *Novel Methods of Location and Tracking of Cellular Mobiles and Their System Applications (Ref. No. 1999/046), IEE Colloquium on*, 1999, pp. 2/1 –2/6.

[5] I. Jami, M. Ali, and R. Ormondroyd, "Comparison of methods of locating and tracking cellular mobiles," in *Novel Methods of Location and Tracking of Cellular Mobiles and Their System Applications (Ref. No. 1999/046), IEE Colloquium on*, 1999, pp. 1/1 –1/6.

[6] Z. Sahinoglu, S. Gezici, and I. Guvenc, *Ultra-wideband Positioning Systems.* Cambridge: Cambridge University Press, 2008.

[7] G. Turin, "An introduction to matched filters," *Information Theory, IRE Transactions on*, vol. 6, no. 3, pp. 311 –329, june 1960.

[8] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 3, pp. 672 – 681, march 2006.

[9] J. Caffery, *Wireless Location in CDMA Cellular Radio Systems.* Boston: Kluwer Academic Publishers, 2000.

[10] A. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *Signal Processing Magazine, IEEE*, vol. 22, no. 4, pp. 24 – 40, july 2005.

[11] A. Urruela, J. Sala, and J. Riba, "Average performance analysis of circular and hyperbolic geolocation," *Vehicular Technology, IEEE Transactions on*, vol. 55, no. 1, pp. 52 – 66, jan. 2006.

[12] W. Figel, N. Shepherd, and W. Trammell, "Vehicle location by a signal attenuation method," *Vehicular Technology, IEEE Transactions on*, vol. 18, no. 3, pp. 105 – 109, nov 1969.

[13] Y. Qi, "Wireless geolocation in a non-line-of-sight environment," Ph.D. dissertation, Princeton University, Dec. 2004.

[14] M. Laaraiedh, S. Avrillon, and B. Uguen, "Hybrid data fusion techniques for localization in uwb networks," in *Positioning, Navigation and Communication, 2009. WPNC 2009. 6th Workshop on*, march 2009, pp. 51 –57.

[15] S. Gezici, "A Survey on Wireless Position Estimation," *Wireless personal communications*, vol. 44, no. 3, pp. 263 – 282, 2008.

[16] M. Najar, J. Huerta, J. Vidal, and J. Castro, "Mobile location with bias tracking in non-line-of-sight," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 3, may 2004, pp. iii – 956–9 vol.3.

[17] F. Althaus, F. Troesch, and A. Wittneben, "Uwb geo-regioning in rich multipath environment," in *Vehicular Technology Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd*, vol. 2, sept., 2005, pp. 1001 – 1005.

[18] C. Steiner, F. Althaus, and A. Wittneben, "On the performance of uwb geo-regioning," in *Signal Processing Advances in Wireless Communications, 2006. SPAWC '06. IEEE 7th Workshop on*, july 2006, pp. 1 –5.

[19] C. Nerguizian, C. Despins, and S. Affes, "Geolocation in mines with an impulse response fingerprinting technique and neural networks," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 3, pp. 603 – 611, march 2006.

[20] M. Triki, D. Slock, V. Rigal, and P. Francois, "Mobile terminal positioning via power delay profile fingerprinting: Reproducible validation simulations," in *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th*, sept. 2006, pp. 1 –5.

[21] H. Laitinen, J. Lahteenmaki, and T. Nordstrom, "Database correlation method for gsm location," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 4, 2001, pp. 2504 –2508 vol.4.

[22] O. Hilsenrath and M. Wax, "Radio Transmitter Location Finding for Wireless Communication Network Service and Management," 2000.

[23] M. Wax, Y. Meng, and O. Hilsenrath, "Subspace signature matching for location ambiguity resolution in wireless communication systems," 2000.

[24] M. Wax, O. Hilsenrath, and A. Bar, "Radio Transmitter Location Finding in CDMA Wireless Communication Systems," 2001.

[25] S. Ahonen and P. Eskelinen, "Mobile terminal location for umts," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 18, no. 2, pp. 23 – 27, feb 2003.

[26] M. Triki and D. Slock, "Mobile localization for nlos propagation," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, sept. 2007, pp. 1 –4.

[27] H. V. Poor, *An Introduction to Signal Detection and Estimation, 2nd. Ed.* Springer, 1994.

[28] C. E. Cook and M. Bernfeld, *Radar signals- An introduction to theory and application(Book).* Academic Press.

[29] I. Telatar and D. Tse, "Capacity and mutual information of wideband multipath fading channels," *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1384 –1400, jul 2000.

[30] T. J. Rothenberg, "Identification in Parametric Models," *Econometrica*, no. 3, pp. 577–591.

[31] G. Durgin, N. Patwari, and T. Rappaport, "Improved 3d ray launching method for wireless propagation prediction," *Electronics Letters*, vol. 33, no. 16, pp. 1412 –1413, jul 1997.

[32] J.-M. Molina-Garcia-Pardo, J.-V. Rodriguez, and L. Juan-Llacer, "Parametric mimo model for ray tracing/launching simulations," in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 1, sept. 2004, pp. 6 – 8 Vol. 1.

[33] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, ser. signal processing series, A. V. Oppenheim, Ed. Prentice Hall, no. 2.

[34] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory.* Prentice Hall PTR, no. 4.

[35] Y. Qi and H. Kobayashi, "A Unified Analysis for Cramer-Rao Lower Bound for Geolocation," in *Proc. 36th Annual Conference on Information Sciences and Systems (CISS 2002)*, 2002.

[36] N. Levanon, "Lowest gdop in 2-d scenarios," *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 147, no. 3, pp. 149 –155, jun 2000.

[37] T. Oktem and D. Slock, "Power Delay Doppler Profile Fingerprinting for mobile localization in NLOS," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, 2010, pp. 876–881.

[38] J. Proakis, *Digital Communications 4th edition.* Englewood Cliffs, NJ: McGraw-Hill, 2000.

[39] A. Goldsmith, *Wireless communications.* Cambridge University Press, 2005.

[40] S. M. Kay, *Intuitive probability and random processes using MATLAB.* Springer, 2006.

[41] D. Maiwald and D. Kraus, "Calculation of moments of complex wishart and complex inverse wishart distributed matrices," *Radar, Sonar and Navigation, IEE Proceedings* -, vol. 147, no. 4, pp. 162 –168, aug 2000.

[42] B. Anderson and J. Moore, *Optimal Filtering.* Prentice Hall (or more recently: Dover, 2005), 1979.

[43] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.* Cambridge, MA: MIT Press, 1983.

[44] L. Ljung, *System Identification: Theory for the User.* Upper Saddle River, NJ: Prentice-Hall, 2002, 2nd edition.

[45] J. Wiklander, "Performance comparison of the Extended Kalman Filter and the Recursive Prediction Error Method," Master's thesis, Linköping Univ., 2003, reg nr: LiTH-ISY-EX-3351, http://www.essays.se/essay/5eb48d45e1/.

[46] Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via EM algorithm," *J. Royal Statisti. Soci.*, vol. 39, 1977.

[47] E. Weinstein, A. Oppenheim, M. Feder, and J. Buck, "Iterative and Sequential Algorithms for Multisensor Signal Enhancement," *IEEE Trans. Signal Proc.*, vol. 42, no. 4, 1994.

[48] W. Gao, S. Tsai, and J. Lehnert, "Diversity Combining for DS/SS Systems With Time-Varying, Correlated Fading Branches," *IEEE Trans. Communications*, vol. 51, no. 2, 2003.

[49] N. Czink, A. Richter, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "Including diffuse multipath parameters in mimo channel models," in *Vehicular Technology Conference, 2007. VTC-2007 Fall. 2007 IEEE 66th*, 30 2007-oct. 3 2007, pp. 874 –878.

[50] O. Norklit and J. Andersen, "Diffuse channel model and experimental results for array antennas in mobile environments," *Antennas and Propagation, IEEE Transactions on*, vol. 46, no. 6, pp. 834 –840, jun 1998.