ETH zürich

ETH Library

Location fingerprinting for ultra-wideband systems

The key to efficient and robust localization

Doctoral Thesis

Author(s):
Steiner, Christoph

Publication date:
2010

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-a-006193910

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Series in Wireless Communications 11

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

Diss. ETH No. 19147

Location Fingerprinting for Ultra-Wideband Systems

The Key to Efficient and Robust Localization

A dissertation submitted to the
ETH Zurich

for the degree of
Doctor of Sciences (Dr. sc. ETH Zurich)

presented by
CHRISTOPH STEINER
Dipl.-Ing., Graz University of Technology
born March 11, 1980
citizen of Austria

accepted on the recommendation of
Prof. Dr. Armin Wittneben, examiner
Prof. Dr. Moe Z. Win, co-examiner

2010

Danke

An dieser Stelle möchte ich mich bei all denen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Lieber Armin, ich möchte mich ganz besonders bei dir bedanken. Ich erinnere mich an viele interessante Gespräche mit dir und - was mich nach wie vor fasziniert - nach jedem Gespräch habe ich die Dinge strukturierter gesehen und darüber hinaus eine Handvoll neuer Ideen mitbekommen. Ich bewundere deine freundschaftliche und fröhliche Art, deine Motivationskünste und deine Energie. Vielen Dank für deine ausgezeichnete Betreuung!

Dear Professor Moe Z. Win, I would like to thank you for being my co-examiner and reviewing my PhD thesis. I am honored.

Dear Georgios and dear Celal, I am very happy that we three shared an office for almost five years. I will remember our numerous interesting discussions on the white board, our activities outside the office, and - of course - seemingly endless phone conversations in Greek and Turkish.

Lieber Oliver, es hat mich auch sehr gefreut, mit dir das Büro für fast zwei Jahre zu teilen. Schade, dass wir nicht mehr regelmäßig Klettern gehen können.

Lieber Heinrich, lieber Florian, lieber Thomas und lieber Frank, ich bin sehr froh mit euch in der UWB Gruppe zusammengearbeitet zu haben. Gemeinsam haben wir viel erreicht und dabei jede Menge Spaß gehabt.

Special thanks go to Claudia, Barbara, Azadeh, Marc, Jörg, Gabriel, Ingmar, Boris, Etienne, Stefan, Jian, Raphael, Zemene, Eric, Carlo and Christoph. You have been and are great colleagues!

Abstract

In this thesis, a novel position location concept is proposed and studied, which provides accurate position estimates in dense multipath and non-line-of-sight propagation environments. The main idea is to apply the location fingerprinting paradigm of position location to channel impulse responses with ultra-wide bandwidth. The large bandwidth enables a fine temporal resolution of the multipath propagation channel, which in turn acts as a unique location fingerprint of the positions of transmitter and receiver.

In the first part of this thesis, a location fingerprinting framework is developed from a communication theoretic perspective. The position location problem is formulated as hypothesis testing problem, such that fundamental methods from statistical detection theory can be applied. Location fingerprints are modeled by parameterized probability density functions. Different hypotheses are distinguished by these parameters, which have to be estimated during a training phase. This framework generalizes a wide class of location fingerprinting approaches and enables the systematic derivation of optimal classification rules and theoretical performance analysis.

In the second part, location fingerprinting with two specific ultra-wideband receiver structures is studied in detail. The first receiver is able to perform channel estimation. The corresponding location fingerprints are chosen as Nyquist sampled versions of the estimated channel impulse responses. The second receiver is a low complexity generalized energy detection receiver, where the energy samples at the output of the analog front-end serve as location fingerprints. In order to derive optimal classification rules, it is necessary to establish a stochastic description of the location fingerprints. This stochastic modeling is performed on the basis of measured data and a model selection criterion. The position location performance of both receiver structures is analyzed theoretically and experimentally with measured data. It is shown that decimeter accuracy is achievable with both receiver structures in dense multipath and non-line-of-sight propagation environments.

However, the performance analysis reveals also a major shortcoming of the proposed method: In order to achieve high position location accuracy, a large amount of training

data is required. This issue is addressed in the third part of this thesis, where two promising techniques are proposed, which increase the efficiency of the training phase. At first, the position location problem is reformulated, such that the training phase can be combined with the localization phase in an iterative manner. Results from the localization phase are used as additional training data. Based on experimental performance results it is shown that the required amount of training data can be significantly reduced. The second technique is even more promising. Only very few measured channel impulse responses - theoretically only three per hypothesis for two-dimensional localization - are required during the training phase for parameter estimation. This efficient training phase is based on a geometrical channel model and exploits a priori knowledge about the geometry of the propagation environment. An experimental performance evaluation shows the high potential of this approach to achieve minimal training phase complexity.

The thesis concludes with a summary of the major findings and with a list of interesting future research topics in the field of location fingerprinting for ultra-wideband systems.

Kurzfassung

In dieser Arbeit wird eine neuartige Methode zur Ortung von drahtlosen Kommunikationsgeräten präsentiert und analysiert. Das Verfahren verspricht eine genaue und zuverlässige Positionsbestimmung, insbesondere für einen Übertragungskanal mit dichter Mehrwegeausbreitung und für den Fall, dass keine direkte Sichtverbindung zwischen Sender und Empfänger besteht. Die grundsätzliche Idee besteht darin, die Impulsantwort des Übertragungskanals als *Fingerabdruck* für die Positionen von Sender und Empfänger zu verwenden. Eine feine zeitliche Auflösung des Übertragungskanals ist dabei ausschlaggebend. Dies kann durch eine hinreichend grosse Signalbandbreite erreicht werden. Dieser örtliche Fingerabdruck wird im Folgenden als *Location Fingerprint* bezeichnet. Lokalisierung mittels Location Fingerprints wird *Location Fingerprinting* genannt.

Diese Dissertation beginnt mit der Entwicklung eines theoretischen Grundgerüstes zur systematischen Beschreibung und Analyse von Location Fingerprinting. Das Ortungsproblem wird als Hypothesentest formuliert, wodurch die Anwendung von Methoden aus der statistischen Entscheidungstheorie ermöglicht wird. Die stochastische Modellierung der Location Fingerprints erfolgt über eine parametrisierte Wahrscheinlichkeitsdichtefunktion, wobei die Parameter einzelne Hypothesen voneinander unterscheiden und während einer Trainingsphase empirisch geschätzt werden müssen. Mit diesem theoretischen Grundgerüst können viele andere, in der Literatur vorgeschlagene, Location Fingerprinting Ansätze beschrieben und analysiert werden. Ausserdem ermöglicht dieses Grundgerüst die systematische Herleitung von optimalen Entscheidungsregeln und die theoretische Analyse von Fehlerwahrscheinlichkeiten.

Im zweiten Teil wird die Lokalisierung mit zwei unterschiedlichen Empfängerstrukturen behandelt. Zunächst wird ein kohärenter Empfänger betrachtet, der Kanalimpulsantworten mit grosser Bandbreite schätzen kann. Die zeitlich gefensterte Kanalimpulsantwort wird mit Nyquist-Rate abgetastet, um den entsprechenden Location Fingerprint zu erhalten. Als zweiter Empfänger wird ein verallgemeinerter Energiedetektor untersucht. Die Abtastwerte am Ausgang des analogen Frontends bilden den entsprechenden Location Fingerprint. Damit

statistische Methoden angewendet und optimale Entscheidungsregeln abgeleitet werden können, wird eine genaue stochastische Beschreibung der Location Fingerprints benötigt. Die Auswahl des stochastischen Modells basiert auf der statistischen Analyse von gemessenen Daten.

Die Genauigkeit und Zuverlässigkeit der Positionsbestimmung wird theoretisch und experimentell untersucht. Mit beiden Empfängerstrukturen kann in sehr vielen Fällen eine Genauigkeit im Bereich von wenigen Dezimetern erreicht werden, obwohl alle gemessenen Übertragungskanäle eine dichte Mehrwegeausbreitung und einige Übertragungskanäle keine direkte Sichtverbindung zwischen Sender und Empfänger aufweisen. Diese experimentelle Analyse offenbart aber auch eine grosse Schwäche: Es werden sehr viele Trainingsdaten zur empirischen Parameterschätzung benötigt, um eine hohe Genauigkeit zu erreichen.

Der dritte Teil dieser Arbeit widmet sich dieser Schwäche. Zwei vielversprechende Methoden zur Verbesserung der Effizienz der Trainingsphase werden vorgestellt und analysiert. Zunächst wird das Ortungsproblem so umformuliert, dass die Trainingsphase mit der Lokalisierungsphase kombiniert werden kann. Iterativ werden Resultate der Lokalisierungsphase dazu verwendet, um die Trainingsergebnisse zu verbessern. Bessere Trainingsergebnisse bedeuten wiederum genauere Ortungsresultate. Die experimentelle Analyse zeigt, dass die benötigte Anzahl an Trainingsdaten dadurch erheblich reduziert werden kann. Die zweite Methode ist noch vielversprechender. Theoretisch werden für eine zweidimensionale Lokalisierung nur drei gemessene Kanalimpulsantworten pro Hypothese zur Parameterschätzung während der Trainingsphase benötigt. Diese effiziente Trainingsphase basiert auf einem geometrischen Kanalmodell und nutzt a priori Wissen über die Geometrie des Raumes. Eine experimentelle Analyse dieser Parameterschätzmethode zeigt deren grosses Potential, um die Komplexität der Trainingsphase auf ein Minimum zu reduzieren.

Als Abschluss dieser Abhandlung werden alle Ergebnisse resümiert, Schlussfolgerungen gezogen, und interessante weiterführende Forschungsfragen diskutiert.

Contents

Danke i

Abstract iii

Kurzfassung v

1 Introduction 1

1.1 Motivation . . . 1
1.2 Ultra-Wideband Wireless Communication . . . 2
1.3 Ultra-Wideband Positioning Systems . . . 2
1.3.1 Geometric Position Location . . . 3
1.3.2 Location Fingerprinting . . . 4
1.4 Contributions . . . 7

2 Location Fingerprinting: A Communication Theoretic Perspective 9

2.1 Problem Formulation . . . 9
2.2 Deterministic Location Fingerprints . . . 11
2.3 Novel Position Location Concept - "Geo-Regioning" . . . 12
2.4 Location Fingerprinting as Hypothesis Testing Problem . . . 14
2.4.1 Average Positioning Error . . . 15
2.4.2 Total Probability of Error . . . 15
2.4.3 Location Fingerprinting with Multiple Observations per Agent . . . 17
2.4.4 Pairwise Error Probabilities . . . 17

3 Channel Measurement Campaign 19

3.1 Measurement Setup . . . 19
3.2 Measurement Scenario . . . 20
3.3 Impact of Antennas . . . 20
3.4 Measurement Signal-to-Noise Ratio . . . 22

3.5 Energy Normalization 23
3.6 Temporal Alignment 23
3.6.1 Problem Formulation 23
3.6.2 Maximum Absolute Value Alignment 24

4 Location Fingerprinting with a Coherent Receiver 25
4.1 Choice of the Location Fingerprints 26
4.2 Regional Channel Model 27
4.2.1 Marginal Distribution of Channel Taps 28
4.2.2 Second Order Statistics 35
4.2.3 Summary 39
4.3 Distortions of the Location Fingerprints 39
4.3.1 Antenna Patterns 39
4.3.2 Thermal Noise and Time Variations 40
4.4 Position Location and Clustering Systems 41
4.4.1 Training Phase 41
4.4.2 Localization Phase 42
4.5 Analytical Pairwise Error Probabilities 43
4.6 Theoretical and Experimental Performance Analysis 47
4.6.1 Randomized Cross-Validation Method 48
4.6.2 Default System Parameters 48
4.6.3 Number of Training Signals 49
4.6.3.1 Analytical Results 49
4.6.3.2 Experimental Results 52
4.6.4 Conditional and Pairwise Error Probabilities 53
4.6.5 Distortion Analysis 54
4.6.6 Signal Bandwidth 55
4.6.7 Observation Window Size 57
4.7 Summary and Conclusions 58

5 Location Fingerprinting with a Generalized Energy Detection Receiver 59
5.1 Generalized Energy Detection Receiver 60
5.1.1 Choice of the Location Fingerprints 61
5.1.2 System Parameters 62
5.2 Exact Distribution of Energy Samples for a Gaussian Channel Model 62

5.3 Stochastic Modeling of Energy Samples 65
5.3.1 Gaussian Channel Model 66
5.3.2 Measured Channel Impulse Responses 69
5.3.3 Summary 71
5.4 Correlation among Energy Samples 72
5.4.1 Empirical Correlation Matrices 73
5.4.2 Analytical Expression for Correlation Coefficients 73
5.5 Position Location and Clustering Systems 75
5.5.1 Training Phase 75
5.5.2 Localization Phase 76
5.6 Accuracy of Closed Form Approximations 77
5.7 Experimental Performance Analysis 79
5.7.1 Default System Parameters 79
5.7.2 Number of Training Vectors 80
5.7.3 Distortion Analysis 81
5.7.4 Sampling Frequency of Energy Detector 81
5.7.5 Number of Regions 82
5.8 Summary and Conclusions 84

6 Combination of Training Phase and Localization Phase 85
6.1 Location Fingerprinting via Parameter Estimation of Mixture Densities . . 86
6.2 Expectation Maximization Algorithm 87
6.2.1 Maximum Likelihood Parameter Estimation for Mixture Densities . 87
6.2.2 Gaussian Mixture Densities 89
6.3 Position Location and Clustering Systems 90
6.3.1 Training Phase 90
6.3.2 Localization Phase 91
6.4 Experimental Performance Analysis 93
6.4.1 Singular Initial Covariance Matrices 94
6.4.2 Number of Training Signals L and Observations L_b 96
6.4.3 Multiple Observations per Agent 97
6.5 Summary and Conclusions 99

7 Efficient Training Phase 101
7.1 Electromagnetic Wave Propagation 101
7.2 Input Output Relations in Passband and Equivalent Baseband 104

7.3 Analysis of Path Delays . . . 105
7.3.1 Geometrical Description of Interacting Objects . . . 105
7.3.2 Nonlinear Mapping of Transmitter Position on Path Delays . . . 105
7.3.3 Upper Bound on Path Delay Variation . . . 108
7.3.4 Linearization of Path Delays . . . 108
7.3.5 Temporal Alignment . . . 113
7.4 Linearization of Path Gains . . . 114
7.5 Simplified Input Output Relation . . . 115
7.6 Estimation of Linear Model Parameters . . . 116
7.6.1 Joint Estimation of Path Gains and Path Delays . . . 116
7.6.2 Least Squares Estimation of Υ_{LM} . . . 118
7.6.3 The Path Pairing Problem . . . 120
7.7 Experimental Analysis of Prediction Accuracy . . . 122
7.7.1 Measurement Campaign in Anechoic Chamber . . . 122
7.7.2 Figures of Merit . . . 124
7.7.3 Performance Results . . . 126
7.8 Parameter Estimation based on Predicted Channel Responses . . . 131
7.9 Summary and Conclusions . . . 132

8 Conclusions and Outlook 135
8.1 Conclusions . . . 135
8.2 Outlook on Future Research . . . 137

A Distance Matrix for UWB Channel Measurement Campaign 139

B Akaike's Information Criterion 141

C Derivation of Approximation (5.5) 145

D Wishart Distribution 149

E Expectation of the Product of Four Gaussian Random Variables 151

Acronyms 153

Notation 155

Bibliography 157

Chapter 1

Introduction

1.1 Motivation

Position location information is essential for the emergence of applications that will revolutionize the daily life of many people: Starting from monitoring of fire fighters and police men, over precise and ubiquitous equipment tracking, up to the surveillance of swarm behavior [1–4]. It is of particular interest for communication engineers that position location information opens a new dimension - or rather three new dimensions - which can be exploited for optimization of various figures of merit such as bit error probability of a communication link [5], interference mitigation in wireless networks, or delay and throughput of routing algorithms.

Nowadays, high localization accuracy is still an ambitious goal in dense multipath and non-line-of-sight (LoS) propagation environments for state-of-the-art position location approaches [3, 6]. However, exactly such environments will prevail in most of the aforementioned application scenarios, which makes it essential to invent and explore novel position location methods featuring good performance in such harsh propagation environments.

Furthermore, low complexity is a key requirement for a ubiquitous implementation of position location systems. In the context of localization the term complexity subsumes the required position location infrastructure, necessary a priori knowledge, additional hardware requirements for receivers (RXs) and transmitters (TXs), and processing power of localization algorithms.

The unreliability of state-of-the-art localization methods in dense multipath and non-LoS propagation environments and the need for low complexity position location systems motivate the research presented in this thesis. The ultimate goal is a robust localization or

clustering[1] method, which provides accurate position location information or cluster information with affordable complexity merely as a side product of data transmission between wireless nodes.

1.2 Ultra-Wideband Wireless Communication

Ultra-wideband (UWB) communication systems are thoroughly discussed in [7–9]. We recall important advantages over conventional narrowband communication systems, which motivate the consideration of UWB signals for position location. UWB systems are allowed to transmit *license-free* in the spectrum from 3.1 GHz to 10.6 GHz with regulated transmit power [10]. Since the spectrum can be used for free, a lot of new applications requiring cheap wireless communication are supported. It is expected that most of these applications benefit additionally from position location information. Furthermore, the large bandwidth of several GHz enables very high data rates and, with that, the possibility to design ultra low power transceivers by exploiting low duty-cycle communication [11]. Ultra low power consumption supports the emergence of applications requiring mobile, battery powered wireless devices. Altogether it is envisioned that a lot of cheap UWB transceivers will coexist and form an ad-hoc wireless network. It is of essential importance, from the application perspective and from the network management perspective, that the positions of these transceivers can be estimated accurately and with low complexity.

1.3 Ultra-Wideband Positioning Systems

The research presented in this thesis focuses on indoor position location and clustering for short range wireless communication systems using radio frequency (RF) signals with ultra-wide bandwidths. However, many theoretical concepts are introduced in a generic way, which makes them immediately applicable to other research areas such as ultrasound signals or underwater communication. The objective of position location systems is the estimation of the unknown positions of wireless nodes denoted as *agents* based on a set of observations at or from reference nodes with known positions denoted as *anchors*[2]. In general, the following two position location scenarios can be distinguished [12]:

[1] Clustering of a wireless network is defined as the process of grouping nodes into clusters based on a certain distance measure between the nodes. In this thesis, we consider Euclidian distance as distance measure, such that physically close nodes are grouped into the same cluster.

[2] Throughout this thesis, we refer to the agents as TXs and to the anchors as RXs without loss of generality, because their roles are interchangeable due to channel reciprocity.

1. Self-positioning: An agent estimates its own position based on observed RF signals, which have been transmitted by the anchors.
2. Remote positioning: The position of an agent is estimated by a central unit (CU), which uses the observations of the transmitted RF signal of the agent at the anchors.

In the following two sections, two conceptually different position location techniques are discussed. The main purpose is to review state-of-the-art position location techniques with emphasis on their shortcomings in harsh propagation environments and on their complexity. The interested reader is referred to the book entitled "Ultra-wideband Positioning Systems" by Sahinoglu et al. [13], which provides a comprehensive introduction to UWB positioning systems in Chapter 1 and to UWB signals in Chapter 2.

1.3.1 Geometric Position Location

Geometric position location techniques consist of two steps. First, signal metrics[3] depending on the relative positions of agent and anchor are estimated from the observed RF signals at each anchor. In the second step these metrics are used for multi-lateration or multi-angulation, respectively [2, 14]. There exists a lower bound on the required number of anchors such that an unambiguous position location estimate is obtained. This lower bound is based on geometrical considerations. For example, at least three anchors are required for unambiguous two-dimensional position location estimation, if time of arrival (ToA) estimates are used.

The algorithmic complexity of multi-lateration or multi-angulation algorithms (e.g. non-linear least squares optimization [15–17]) is well understood and for current computers not obstructive, since the processing can be performed at a CU which is connected to a power supply. However, the estimation of signal metrics from RF signals at the anchors is performed by a wireless node, which might be just battery powered. Therefore, there exists a large amount of research concerned with low complexity estimation of ToA, time difference of arrival (TDoA), received signal strength (RSS) or angle of arrival (AoA). For example, low complexity ToA estimation rules are proposed in [18–23].

The accuracy of geometric position location techniques is determined by the signal-to-noise ratio (SNR) of the RF signals, the hardware abilities and processing power of the anchors, the propagation environment, and the number of anchors. Fundamental performance limits based on the Cramér-Rao lower bound and the Ziv-Zakai lower bound for

[3] Examples for signal metrics are time of arrival, time difference of arrival, received signal strength, and angle of arrival.

various RX architectures and RF signals are derived in [6]. If the propagation channel has a strong LoS component, then geometric position location techniques show a remarkable good performance. The experimental demonstration of position location and tracking described in [24] is just one example for the potential of position location based on ranging[4] with UWB signals. However, a rich multipath propagation environment and especially non-LoS channel conditions degrade the quality of ToA, TDoA, RSS and AoA measurements drastically. Increasing the RF signal bandwidth increases the robustness against multipath errors, which implies that UWB signals are preferable in indoor environments [14, 25]. Despite the large bandwidth, non-LoS situations cause positively biased ToA and TDoA estimates [26] and inaccurate RSS and AoA estimates. Therefore, it is essential to deploy the anchors carefully to ensure LoS conditions for all possible agent positions. If this is not feasible, the number of anchors must be increased at the expense of hardware and infrastructure complexity.

Recently, cooperative localization has gained a lot of attention [27, 28]. The idea is to allow for cooperation among agents to improve the overall localization performance and the coverage area. Signal metrics between adjacent agents are used additionally by the position location algorithms, which implies that the lower bound on the required number of anchors does not exist anymore. Furthermore, the robustness against non-LoS situations can be increased, since the number of LoS communication links is increased with high probability. The cooperative paradigm requires additional data transfer among agents and anchors, which might limit the applicability of this approach.

1.3.2 Location Fingerprinting

Location fingerprinting [3, 29, 30] is a conceptually different position location technique, which also consists of two phases. In phase one (training phase), the position location system gathers coordinates of training points and related signal metrics (location fingerprints) extracted from RF signals at a number of anchors and stores them into a central database. In phase two (localization phase), the anchors observe location fingerprints of agents and find the best matching entries in the database. There exists a large number of algorithms for this pattern matching such as support vector machines, neural networks, weighted k nearest neighbors, or Bayesian approaches [31, 32]. Note that in contrast to geometric position location techniques, there does not exist a lower bound on the required number of anchors. Thus, theoretically it is possible to locate agents unambiguously based on the observed RF

[4] The range information is obtained from ToA, TDoA, or RSS measurements.

signal at a single anchor. Further note that location fingerprinting is a classification problem, whereas geometric position location is a regression problem [32].

The complexity of these pattern matching (classification) algorithms is of minor importance, since the processing can be done at a CU which is connected to a power supply. The major part of the complexity is caused by the training phase, since location fingerprints from the whole surveillance area[5] must be gathered. An important parameter is the grid spacing, which determines the total number of training points and the resolution of the spatial quantization. The complexity of the training phase is further discussed in Chapter 2.

The performance of location fingerprinting algorithms is commonly measured with *accuracy* and *precision* [30]. Accuracy measures the distance error between the estimated and the true location and precision states the percentage of position location estimates with a certain accuracy. The performance of location fingerprinting algorithms is mainly determined by the grid spacing, the choice and modeling of the location fingerprints, the pattern matching algorithm, the SNR of the RF signals, the variability of the propagation environment, and the number of anchors. Assume that the training points lie on a two-dimensional grid with a spacing of 0.2 m in each dimension and that the surveillance area is an office room with an area of 5 m times 5 m. In this example the total number of training points would be $(5/0.2)^2 = 625$. The true position of an agent is assumed to be uniformly distributed in the uncertainty area of 0.2 m times 0.2 m around a grid point. The figure of merit for the accuracy is the root mean square (RMS) error. Therefore, the lower bound on the achievable localization accuracy would be $0.2\sqrt{1/6} \approx 8.2$ cm.

The choice and modeling of location fingerprints is a very important step, which impacts performance, complexity, and robustness of the system. For UWB systems there exists very few research on this topic, which motivates the thorough treatment of this modeling problem in this thesis. Nowadays, the most common choice for location fingerprints is RSS of Wi-Fi signals [30, 33]. The corresponding localization systems are implemented as software add-on to existing Wi-Fi infrastructure. There exist virtually no additional costs except for the training phase, since every Wi-Fi card is able to measure the RSS and the coordinates of the Wi-Fi access points acting as anchors are known a priori. Each access point provides only one degree of freedom (one RSS value), which implies that many access points are required, in order to achieve a satisfactorily performance. However, in contrast to geometric position location techniques, a strong LoS component is not required. Experimental performance analyses of recent location fingerprinting systems based on RSS can be found in [33, 34],

[5] The surveillance area subsumes all possible agent positions. Note that the surveillance area does not have to be a connected set. Depending on the application isolated regions might be sufficient.

where accuracies from 2 m to 3 m for roughly two-third of all test positions are reported.

In [35], it is proposed to use seven parameters (mean excess delay, RMS delay spread, maximum excess delay, total received power, number of multipath components, power of the first path, and the arrival time of the first path), which are estimated from channel impulse responses (CIRs) with 200 MHz bandwidth, as location fingerprint. The pattern matching algorithm is implemented with an artificial neural network using 367 training signals to train the neural network. The authors report an accuracy of 2 m for 80 percent of the test cases using a single anchor. The grid spacing of the training points is 0.5 m in one direction and 1 m in the other direction. In [36], the authors propose to apply a wavelet compression to extract features of CIRs with 200 MHz bandwidth and use them as location fingerprints. Pattern matching is done again with an artificial neural network and similar accuracy as in [35] with a precision of 67 percent is achieved. Note that both approaches in [35, 36] require time synchronization between TX and RX.

The application of UWB signals to location fingerprinting has been first proposed in [37, 38], where sampled CIRs are used as location fingerprints. A Bayesian approach[6] is pursued to derive a maximum likelihood (ML) pattern matching algorithm. In these works, it is shown based on measured CIRs with 3 GHz bandwidth that excellent accuracy at high precision can be achieved with a single anchor. These promising results have initiated the research presented in this thesis. The authors of [39] apply UWB CIRs for location fingerprinting in underground mines in the same way as proposed in [37, 38] and report similar performance results. In [40], the authors propose to use UWB CIRs in the frequency band from 3.1 GHz to 10.6 GHz as location fingerprints. Their database consists of a high-resolution map of CIRs, which requires a time consuming measurement process and a large storage capacity. A very accurate positioning device is used to measure CIRs at a grid spacing of 0.01 m in each dimension. The surveillance area is 1 m times 1 m resulting in 10000 grid points. The position of an agent is found by the maximization of the CIR cross-correlation coefficient. The performance results for LoS and non-LoS situations show a remarkable high accuracy of around 2 cm. Note that time synchronization between TX and RX is assumed for the approach proposed in [40].

In general, location fingerprinting is a promising alternative to geometric position location techniques especially in harsh indoor environments, where the propagation channel does not have a strong LoS component. Provided the RF signal bandwidth is large enough, the multipath propagation structure, which is related to the positions of the agent and anchor, can be exploited for position location. In case of UWB signals, the RF signals at each

[6] Further details can be found in Chapter 2.

anchor have already a large number of degrees of freedom, which implies that good performance with a single anchor is achievable. This is essential for applications, which require an operation without pre-installed infrastructure in an ad-hoc manner. An example would be self-localization of nodes in a wireless network [41].

1.4 Contributions

In this thesis, a novel position location concept based on UWB RF signals is proposed and studied. The distinct advantage of this method over state-of-the-art position location solutions is that accurate position estimates can be obtained in dense multipath and non-LoS propagation environments. The main idea is to apply the location fingerprinting approach to CIRs with an ultra-wide bandwidth, which provides an accurate temporal resolution of the multipath propagation channel. This multipath structure is a unique fingerprint of the relative positions of TX and RX, i.e. agent and anchor. In a constant propagation environment, the CIR depends only on the positions of agent and anchor. In contrast to geometric position location techniques, the proposed location fingerprinting system benefits from dense multipath propagation, because the location fingerprints at different agent positions become more distinct. Since signal metrics related to the direct path are not required for localization, the performance of the proposed location fingerprinting scheme is robust to non-LoS conditions. Accurate time synchronization between anchors and agents is not required, since only the shape of the CIR is chosen as signal metric and the absolute timing information is neglected. First investigations of the proposed location fingerprinting approach, originally called UWB geo-regioning, are published in [37, 38]. The work presented in this thesis continues this research and provides the following contributions:

In Chapter 2, a location fingerprinting framework is developed from a communication theoretic perspective. The position location problem is formulated as hypothesis testing problem, such that methods from statistical detection theory can be applied. It is proposed to relax the requirement of estimating the exact location of an agent to rather deciding, whether this agent is located in a region of a certain size. Stochastic location fingerprints are assumed and parameterized probability density functions (PDFs) are used to model their behavior. An immediate consequence of this stochastic modeling approach is that the training phase reduces to a statistical parameter estimation problem. The theoretical framework generalizes a wide class of location fingerprinting approaches and enables the systematic derivation of optimal classification rules and a theoretical performance analysis.

In the second part, location fingerprinting with two specific UWB RX structures is studied in detail. The first RX (cf. Chapter 4) is able to perform channel estimation. The corresponding location fingerprints are chosen as Nyquist sampled versions of the estimated CIRs. The second RX is a low complexity generalized energy detection (ED) RX (cf. Chapter 5) and the energy samples at the output of the analog front-end serve as location fingerprints. In order to derive optimal classification rules it is necessary to establish a stochastic description of the location fingerprints. This stochastic modeling is performed on the basis of measured CIRs and a model selection criterion. These modeling results are also important for other fields in wireless communications, such as channel modeling or RX design (e.g. [42]). Especially the stochastic modeling of energy samples based on measured CIRs has not been treated so far.

The performance of both RXs is analyzed theoretically and experimentally using measured CIRs. The corresponding channel measurement campaign is outlined in Chapter 3. The impact of important system parameters such as RF signal bandwidth on the performance is investigated. It is shown that decimeter position location accuracy is achievable for both RX structures in a dense multipath and non-LoS propagation environment. However, these performance investigations reveal also a major shortcoming of the proposed method: A large amount of training data is required, in order to achieve high position location accuracy and precision. This issue is addressed in the third part of this thesis, where two promising techniques are developed, in order to increase the efficiency of the training phase. In Chapter 6, it is proposed to combine the parameter estimation step during the training phase with the localization phase, in order to iteratively improve the quality of the parameter estimates and the position location accuracy. It is shown with experimental performance results that the required amount of training data can be significantly reduced.

The second technique proposed in Chapter 7 is even more promising. Only very few measured channel responses (CRs)[7] - theoretically only three for a two-dimensional localization problem - are required to obtain initial parameter estimates of the stochastic location fingerprint model. This estimation method is based on a geometrical UWB channel model and exploits a priori knowledge about the geometry of the propagation environment. The performance is evaluated with dedicated measurements in a controlled propagation environment. It is expected that these parameter estimates provide only a partial description of a realistic propagation channel, but can serve as initial parameter estimates for the aforementioned iterative procedure.

[7] A CR is defined as convolution of a CIR with a transmit pulse. If the transmit pulse is a sinc-pulse then CR is identical to CIR.

Chapter 2

Location Fingerprinting: A Communication Theoretic Perspective

In this chapter[1], a location fingerprinting framework is developed from a communication theoretic perspective applying fundamental results from statistical detection theory. This framework generalizes all location fingerprinting methods applying classification algorithms based on a Bayesian formulation. For example neural networks or support vector machines are not covered by this framework. The following investigations are done for the two-dimensional Euclidian space for the sake of visualization. Extensions to three dimensions are straightforward.

2.1 Problem Formulation

We consider an indoor scenario with a given surveillance area as depicted in Fig. 2.1. The surveillance area is sampled on a grid with M points. During a training phase, location fingerprints denoted by $\mathbf{s}_m$ and the corresponding coordinates denoted by $(x, y)_m$ are gathered for all $m = 1, 2, \ldots, M$ grid points. The set of all location fingerprints and coordinates constitutes the location fingerprint database[2]. In case of remote positioning, this database is stored at the CU, which also executes the pattern matching algorithm and might additionally act as anchor. If there exist multiple distributed anchors (cf. Fig. 2.1), for example distributed access points in RSS based location fingerprinting systems, then all individual location fingerprints observed at each anchor are transmitted to the CU and are stacked into $\mathbf{s}_m$. In case of self-positioning, each agent act as CU and, thus, must have a copy of the database and

[1] Parts of this chapter have been published in [43].
[2] Such a database is also known as radio map.

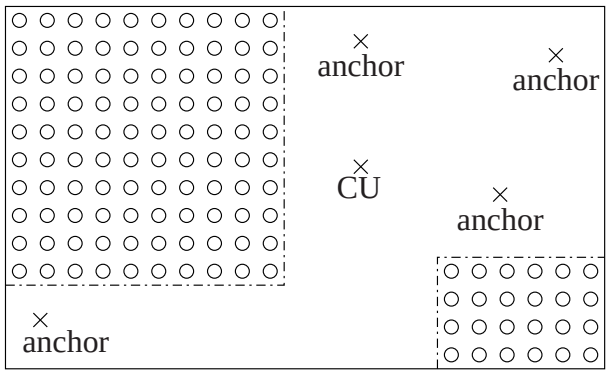

Fig. 2.1: Grid points within a given surveillance area, anchors and central unit.

has to execute the pattern matching algorithm. The number of components in the vector $\mathbf{s}_m$ is denoted by N. Throughout this thesis, it is assumed that the agents transmit (i.e. act as TXs) and the anchors receive (i.e. act as RXs), which corresponds to the remote positioning setup. Due to channel reciprocity the roles of agents and anchors are interchangeable, which implies that this assumption does not reduce generality.

The location fingerprint $\mathbf{s}_m$ is related to the coordinates $(x,y)_m$ and the positions of the anchors via the corresponding propagation channels. The actual shape of $\mathbf{s}_m$ depends on the signal processing at the agent and anchors, but does not matter for the here proposed theoretical framework. For example, in RSS based location fingerprinting systems each element in $\mathbf{s}_m$ corresponds to the average RSS at a specific access point (anchor). Note that the RSS depends also on the transmit power of the agent, which implies that the transmit powers of all agents either have to be known or must be set to the same value. For location fingerprinting systems employing wideband or UWB signals, $\mathbf{s}_m$ could be a sampled CIR or channel parameters like path delays and gains.

During the localization phase, the CU observes the vector $\mathbf{y}$ from an agent with unknown position (x,y), which corresponds to remote positioning. The inference of (x,y) based on $\mathbf{y}$ is formulated as an M-ary hypothesis testing problem [44] according to

Hypothesis H_m : The agent exciting $\mathbf{y}$ is located at position $(x,y)_m$.

Note that the CU could be considered as agent, which would correspond to self-positioning. Then the hypotheses are reformulated according to

Hypothesis H_m : The agent observing $\mathbf{y}$ is located at position $(x,y)_m$.

Depending on the modeling of $\mathbf{s}_m$, different detection problems can be distinguished.

2.2 Deterministic Location Fingerprints

The simplest model assumes that $\mathbf{s}_m$ is deterministic. Deterministic location fingerprints are commonly used in RSS based systems due to the rather stable behavior of the average RSS, which is mainly influenced by large scale propagation effects such as path loss (PL) or shadowing. The advantages of the deterministic formulation are low complexity pattern matching algorithms, such as Euclidian distance calculation between the observation and each location fingerprint [30, 45] or the weighted k nearest neighbors algorithm [31].

In general, the observation $\mathbf{y}$ does not correspond exactly to one of the M location fingerprints due to random distortions like thermal noise, interference, and variations of the propagation channel over time. Therefore, the observations are commonly modeled by

$$\mathbf{y} = \mathbf{s}_m + \mathbf{n},$$

where the random vector $\mathbf{n}$ accounts for the distortions. This model assumes that the $\mathbf{s}_m$ are the true location fingerprints, which implies that the distortions during the training phase have to be mitigated. This is usually achieved through averaging several location fingerprints from the same position over time.

Each $\mathbf{s}_m$ can be interpreted as a constellation point in an N-dimensional signal space. Thus, there are totally M constellation points. The PDF of $\mathbf{y}$ given H_m is determined by $\mathbf{s}_m$ and the PDF of $\mathbf{n}$. These PDFs for all M hypotheses suffice to derive optimal pattern matching (classification) algorithms and calculate the corresponding error probability. These topics are further discussed in Section 2.4.

In most cases the unknown position of the agent denoted by (x, y) is not equal to one of the grid points $(x, y)_m$, which poses an additional source of error. The impact of this deviation on the performance of the classification algorithm depends on its robustness to spatial variations of the location fingerprints. In order to combat this problem, the grid spacing can be reduced and the number of training points increased. This, however, increases the complexity of the training phase and localization phase. Moreover, the probability of making decision errors increases as well.

As already mentioned, RSS based location fingerprints can be made robust to temporal and spatial variations of the propagation channel by averaging the RSS over small scale propagation effects. On the contrary, wideband CIRs are very sensitive to these small scale effects, which implies that the deterministic formulation is not applicable to location fingerprints related to CIRs.

2.3 Novel Position Location Concept - "Geo-Regioning"

In this section, a novel position location concept based on a stochastic description of the location fingerprints $\mathbf{s}_m$ is proposed. The main idea is to relax the requirement of exact grid coordinates. In the above formulation assuming deterministic location fingerprints, the objective is to estimate the exact coordinates of an agent. In the here proposed formulation, the goal is to estimate the region, in which an agent is located in. This is illustrated in Fig. 2.2, where the squares define the regions. Note that there are four grid points in each region.

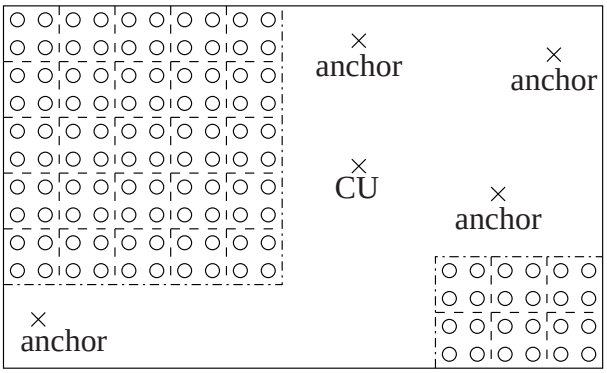

Fig. 2.2: Regions of an indoor position location system.

This relaxed problem formulation has some distinct advantages over the deterministic formulation. First, the training coordinates need not be known exactly. It is sufficient to know that the received training signal is generated by a TX located within the considered region. Furthermore, it is possible to trade off accuracy against complexity. The larger the regions are chosen, the fewer of them are needed in order to cover the surveillance area. Fewer and larger regions imply less complexity of training and localization. The drawback of larger regions is a larger uncertainty area, which determines the fundamental position location inaccuracy due to the spatial quantization into regions. Assuming a uniform distribution of the agent positions within these regions, the corresponding RMS errors can be calculated.

The hypothesis testing problem for remote positioning is reformulated according to

Hypothesis H_m : The agent exciting $\mathbf{y}$ is located in region m.

In case of self-positioning the hypotheses are stated as

Hypothesis H_m : The agent observing $\mathbf{y}$ is located in region m.

Since the location fingerprints represent now an area[3] instead of a grid point, it is essential to model the spatial small scale behavior of the location fingerprints. Consequently, we propose stochastic location fingerprints and assume that there exists a parameterized PDF, which models the behavior of the location fingerprints over space. This PDF depends on the shape of the location fingerprints, i.e. on the signal processing at TX and RX, the dimensions of the regions, and the assumptions about the propagation channel. Different regions/hypotheses are distinguished by the parameters of this PDF. These parameters need to be estimated during the training phase for each region. The database consists of M parameter sets denoted by Θ_m with the corresponding coordinates $(x, y)_m$, which represent the centers of the regions. The PDF of **y** is determined by the PDF of $\mathbf{s}_m$ and the PDF of **n**.

The following fundamental questions arise:

1. How should the location fingerprints $\mathbf{s}_m$ be chosen?
2. What parameterized PDF should be used to model these location fingerprints?
3. Which parameters of this PDF are relevant to distinguish the regions?
4. What performance can be achieved in a dense multipath and non-LoS propagation environment?
5. What is the impact of system parameters on performance and complexity of the location fingerprinting system?
6. How much training data is needed to obtain accurate parameter estimates such that a desired performance is achievable?
7. How can the efficiency of the training phase be improved?

In the course of this thesis, these questions will be answered.

In the next section, we complete the location fingerprinting framework by defining two figures of merit for the position location and clustering problem, and by deriving the corresponding optimal classification algorithms.

[3] In the three-dimensional case they represent a volume.

2.4 Location Fingerprinting as Hypothesis Testing Problem

We start the discussion by repeating the hypotheses:

$$\text{Hypothesis } H_m : \text{The agent exciting } \mathbf{y} \text{ is located at } (x,y)_m, \tag{2.1}$$

where $(x,y)_m$ can be interpreted as exact position (cf. Section 2.2) or as the center of a region (cf. Section 2.3).

If we are not interested in the physical position, but rather in a clustering of a wireless network, then it is sufficient to formulate the hypotheses as

$$\text{Hypothesis } H_m : \text{The node exciting } \mathbf{y} \text{ is located in cluster } m. \tag{2.2}$$

We refer to (2.1) as position location problem and to (2.2) as clustering problem.

In order to solve the hypothesis testing problem, the Bayesian paradigm [46] is followed. It is assumed that a priori probabilities denoted by $\pi_1, \pi_2, \ldots, \pi_M$ can be assigned to the hypotheses. These a priori probabilities can be considered as relative number of agents located in each region. If these a priori probabilities are unknown, then they are set to $1/M$. The average cost or Bayes risk ($\mathcal{R}$) is defined as [44, 46]

$$\mathcal{R} = \sum_{i=1}^{M}\sum_{j=1}^{M} \pi_j C_{i,j} \mathrm{P}\left(H_i|H_j\right) = \sum_{i=1}^{M}\sum_{j=1}^{M} \pi_j C_{i,j} \int_{Z_i} f_j\left(\mathbf{y}|H_j\right) \mathrm{d}\mathbf{y}, \tag{2.3}$$

where $\mathrm{P}\left(H_i|H_j\right)$ is the probability of deciding for H_i, when H_j is true, $f_j\left(\mathbf{y}|H_j\right)$[4] is the conditional PDF of $\mathbf{y}$ given H_j, and Z_i defines the part of the observation space in which H_i is chosen. The parameters $C_{i,j}$ denote the costs for deciding for H_i when H_j is true.

Following the derivations in [46], the decision rule minimizing (2.3) is found as

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmin}} \left(\sum_{j=1}^{M} C_{m,j} \mathrm{P}\left(H_j|\mathbf{y}\right) \right),$$

where $\hat{m}$ denotes the estimate for the true hypothesis and $\mathrm{P}\left(H_j|\mathbf{y}\right)$ is the a posteriori probability of H_j given observation $\mathbf{y}$. Applying Bayes' rule and noticing that $f\left(\mathbf{y}\right) = \sum_{m=1}^{M} \pi_m f_m\left(\mathbf{y}|H_m\right)$ does not depend on m, an equivalent decision rule is given

[4] The conditioning on H_j determines the parameter set Θ_j.

by

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmin}} \left(\sum_{j=1}^{M} \pi_j C_{m,j} f_j\left(\mathbf{y}|H_j\right) \right). \tag{2.4}$$

In the following two sections, we define two performance measures, which are used for the position location problem and for the clustering problem, respectively.

2.4.1 Average Positioning Error

The figure of merit for the position location problem is the average positioning error, which is defined as

$$\begin{aligned} \mathcal{D}_e &\triangleq \sum_{j=1}^{M} \pi_j \sum_{i=1}^{M} d_{i,j} \mathrm{P}\left(H_i|H_j\right) = \sum_{j=1}^{M} \pi_j \left(d_{j,j} \mathrm{P}\left(H_j|H_j\right) + \sum_{i=1, i\neq j}^{M} d_{i,j} \mathrm{P}\left(H_i|H_j\right) \right) \\ &= \sum_{j=1}^{M} \pi_j d_{j,j} + \sum_{j=1}^{M} \pi_j \sum_{i=1, i\neq j}^{M} \left(d_{i,j} - d_{j,j}\right) \mathrm{P}\left(H_i|H_j\right), \end{aligned} \tag{2.5}$$

where $d_{i,j} = d_{j,i}$ is the Euclidian distance between the center points $(x,y)_i$ and $(x,y)_j$ of regions i and j. For this derivation we have used the fact that $\sum_{i=1}^{M} \mathrm{P}\left(H_i|H_j\right) = 1$. We notice that the average positioning error results from Bayes risk by setting the costs $C_{i,j}$ equal to $d_{i,j}$. Thus, if all $d_{i,j}$ are known a priori between all regions, then the decision rule

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmin}} \left(\sum_{j=1}^{M} \pi_j d_{m,j} f_j\left(\mathbf{y}|H_j\right) \right) \tag{2.6}$$

minimizes $\mathcal{D}_e$. Note that we can account for the fundamental position location inaccuracies within the regions by assigning non-zero values to $d_{j,j}$. This is especially important, if different regions have different dimensions. The values for $d_{j,j}$ and π_j determine the minimal average positioning error, which is achieved if all conditional error probabilities $\mathrm{P}\left(H_i|H_j\right)$ for $i \neq j$ are zero.

2.4.2 Total Probability of Error

When considering the clustering problem, the $d_{i,j}$ are not relevant. Therefore, the costs of wrong decisions are set to one ($C_{i,j} = 1$ for $i \neq j$) and the costs of correct decisions are set to zero ($C_{i,i} = 0$). This assignment leads to a minimization of the total probability of error [44],

which is given by

$$\mathcal{P}_e = \sum_{j=1}^{M} \pi_j \sum_{i=1, i\neq j}^{M} \mathrm{P}\left(H_i | H_j\right). \tag{2.7}$$

The decision rule, which minimizes (2.7) is found by inserting the assumptions on the costs into (2.4) and by simplifying the expression, which results in the M-ary maximum a posteriori probability (MAP) decision rule

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \mathrm{P}\left(H_m | \mathbf{y}\right) = \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \pi_m \, f_m\left(\mathbf{y} | H_m\right). \tag{2.8}$$

For equally likely hypotheses, i.e. $\pi_m = 1/M$ for all m, this rule simplifies further to the M-ary ML decision rule.

By defining the likelihood ratios $\Lambda_i\left(\mathbf{y}\right)$ for $i = 1, 2, \ldots, M$ as

$$\Lambda_i\left(\mathbf{y}\right) = \frac{f_i\left(\mathbf{y} | H_i\right)}{f_1\left(\mathbf{y} | H_1\right)}, \tag{2.9}$$

the M-ary MAP decision rule can be reformulated as

$$\pi_i \Lambda_i\left(\mathbf{y}\right) \underset{\hat{m}\neq i}{\overset{\hat{m}\neq j}{\gtrless}} \pi_j \Lambda_j\left(\mathbf{y}\right) \text{ for } i, j = 1, 2, \ldots, M \text{ and } i > j. \tag{2.10}$$

The $M-1$ likelihood ratios $\Lambda_i\left(\mathbf{y}\right)$ for $i = 2, 3, \ldots, M$ define the coordinate system of the $(M-1)$-dimensional decision space and the $M\left(M-1\right)/2$ decision rules in (2.10) define the $(M-2)$-dimensional hyper planes, which are the boundaries of the decision regions. For $M = 2$ the decision space is one-dimensional and there exists one decision threshold. For $M = 3$ the decision space is two-dimensional and the three decision boundaries are straight lines, which define decision areas. For $M = 4$ the decision space is three dimensional and the six hyper planes are planes.

Applying all $M\left(M-1\right)/2$ decision rules in (2.10) to an observation vector $\mathbf{y}$ produces the MAP estimate $\hat{m}$ for the true hypothesis. Depending on the particular expressions for $f_m\left(\mathbf{y} | H_m\right)$ the likelihood ratios in (2.9) can be considerably simplified, which reduces the complexity of the MAP decision rule and mitigates numerical problems due to a potentially high dimensionality of $\mathbf{y}$. Therefore, the decision rule stated in (2.10) might be easier to implement than the original decision rule stated in (2.8).

2.4.3 Location Fingerprinting with Multiple Observations per Agent

In a realistic position location scenario, it is conceivable that the CU has multiple observations from each agent available for the localization (region detection) process. The essential assumption is that these multiple observations are caused by an agent, which is located in *one* region.

For example the agent could be equipped with multiple antennas and could use them either sequentially to transmit its signals to the anchors or use orthogonal signals to transmit them simultaneously. Furthermore, if the agent is mobile it can transmit signals periodically always from a slightly different position. These are just two examples how such multiple observations could be generated. Note that the same ideas can be also applied to self-positioning.

In the following, we derive the optimal decision rule taking multiple observations into account. The CU has recorded K observations $\{\mathbf{y}_1, \ldots, \mathbf{y}_K\}$. It is assumed that these observations are independent and are caused by an agent located in the same region. The individual observations are stacked into a larger vector $\bar{\mathbf{y}} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_K^T\right]^T$. The conditional PDF of $\bar{\mathbf{y}}$ is given by the product of the individual PDFs of $\mathbf{y}_k$ according to

$$f_m\left(\bar{\mathbf{y}}|H_m\right) = \prod_{k=1}^{K} f_m\left(\mathbf{y}_k|H_m\right). \tag{2.11}$$

This follows from the assumption that the individual observations are independent. By inserting the PDFs for $\bar{\mathbf{y}}$ given H_m into (2.6) or (2.8) we can modify the decision rules such that multiple observations are exploited.

2.4.4 Pairwise Error Probabilities

For most decision problems with large M, it is mathematically intractable to obtain analytic expressions for $\mathcal{D}_e$ or $\mathcal{P}_e$ due to complex decision regions in an $(M-1)$-dimensional space. In order to circumvent this issue, it is useful to assume that only two hypotheses, say H_i and H_j, are possible. The corresponding binary decision rule following the Bayesian paradigm is given by

$$l \triangleq \text{function}\left(\frac{f_i\left(\mathbf{y}|H_i\right)}{f_j\left(\mathbf{y}|H_j\right)}\right) \underset{\hat{m}=j}{\overset{\hat{m}=i}{\gtrless}} \delta_{i,j},$$

which is a threshold test on the sufficient statistic l. A sufficient statistic is a function of the observation $\mathbf{y}$ which has the property that the likelihood ratio $\frac{f_i(\mathbf{y}|H_i)}{f_j(\mathbf{y}|H_j)}$ can be written as a function of the sufficient statistic [44]. The threshold $\delta_{i,j}$ depends on the considered performance measure. For $\mathcal{P}_e$ it is given by $\delta_{i,j} = \frac{\pi_j}{\pi_i}$ and for $\mathcal{D}_e$ it is $\delta_{i,j} = \frac{\pi_j\left(d_{i,j}-d_{j,j}\right)}{\pi_i\left(d_{j,i}-d_{i,i}\right)}$. The sufficient statistic l is itself a random variable and has a PDF denoted by $f_{l|H_j}(l)$, if it is assumed that H_j is the true hypothesis.

Considering the binary hypothesis testing problem, the pairwise error probabilities are defined by

$$\mathrm{P}\left(H_j \to H_i\right) \triangleq \int_{\delta_{i,j}}^{\infty} f_{l|H_j}(l)\,\mathrm{d}l \text{ and } \mathrm{P}\left(H_i \to H_j\right) \triangleq \int_{-\infty}^{\delta_{i,j}} f_{l|H_i}(l)\,\mathrm{d}l. \tag{2.12}$$

The quantity $\mathrm{P}\left(H_j \to H_i\right)$ is the probability of deciding for H_i, when H_j is true and no other hypotheses are possible. Note that the pairwise error probabilities are in general not symmetric, i.e. $\mathrm{P}\left(H_j \to H_i\right) \neq \mathrm{P}\left(H_i \to H_j\right)$. The pairwise error probabilities can be used to upper bound both performance measures by applying the union bound [47], which gives

$$\begin{aligned}
\mathcal{D}_e &\leq \sum_{j=1}^{M} \pi_j d_{j,j} + \sum_{j=1}^{M} \pi_j \sum_{i=1, i\neq j}^{M} \left(d_{i,j} - d_{j,j}\right) \mathrm{P}\left(H_j \to H_i\right), \\
\mathcal{P}_e &\leq \sum_{j=1}^{M} \pi_j \sum_{i=1, i\neq j}^{M} \mathrm{P}\left(H_j \to H_i\right).
\end{aligned} \tag{2.13}$$

If we can derive analytical expressions for $\mathrm{P}\left(H_j \to H_i\right)$, we can provide analytical upper bounds for the performance of the position location or clustering problem without performing time consuming Monte Carlo simulations. Furthermore, analytical error expressions usually provide insights into fundamental effects determining the performance and are viable tools to optimize system parameters.

Chapter 3

Channel Measurement Campaign

This chapter describes a UWB channel measurement campaign, which is designed to demonstrate the feasibility of the proposed location fingerprinting method and evaluate its performance in a dense multipath propagation environment with non-LoS conditions. A more detailed description of this measurement campaign can be found in [48]. Note that we do not specify the type of the location fingerprints $\mathbf{s}_m$ yet, but they will be extracted from the measured CIRs obtained by this campaign.

3.1 Measurement Setup

The measurements were performed in a cellar room (cf. Fig. 3.1) with a size of about 7.4 m times 15 m and a height of 6 m. There are many metallic objects in the room such as metallic shelves, heating pipes, cabinets and metal cores, implying a dense multipath environment with some non-LoS situations. The propagation environment was kept mostly static during the measurements, although people, who performed the measurement, were always present and sometimes moving.

A time-domain correlation method was used to estimate the CIRs. The principle is to perform a cross-correlation between the received signal and the known transmit signal at the RX. In practice, the transmit signal is often generated using pseudo noise (PN) bit streams or m-sequences. The transmit signal is fed to a power amplifier (PAMP) and finally to the transmit antenna. This signal propagates through the channel, is received by the receive antenna, amplified by an low noise amplifier (LNA), and sampled by an oscilloscope with a sampling frequency of 20 GHz. The measurement frequency range was roughly limited from 3 GHz to 6 GHz by the transfer functions of the UWB antenna, PAMP, LNA, and input filter of the oscilloscope. The reference signal for the cross-correlation, i.e. the transmit

signal, was pre-stored in the oscilloscope, such that no wired connection between TX and RX was required. This implies, however, that the absolute time references are unknown. Thus, each measured CIR has an unknown time shift. Furthermore, the impulse responses of the transmit and receive antenna are comprised in the measured CIRs, which is further discussed in Section 3.3.

3.2 Measurement Scenario

A large number of CIRs between one static RX equipped with four antennas and a mobile TX also equipped with four antennas has been measured. The RX acts as CU. The height of TX and RX antennas was chosen as 1.8 m. The receive antennas were mounted in the corners of a square with a side length of 38 cm and the transmit antennas were mounted in the corners of a square with a side length of 20 cm on a two-dimensional positioning device. This device allows to move the TX array within an area of 27 cm times 56 cm. The positioning device was moved to 22 different locations. These locations define the center points $(x, y)_m$ of the $M = 22$ regions, which are depicted in Fig. 3.1. The maximum distance between two regions is approximately 17 m, whereas the minimum separation of two transmit antenna positions in two different regions is approximately 10 cm. In Appendix A the complete distance matrix collecting the distances $d_{i,j}$ between all center points is listed.

The TX array was moved with an almost constant speed of 1 cm/s within each region. The trigger at the oscilloscope was not synchronized with the movement of the TX array, which means that the exact positions of the TX antennas are unknown. However, since triggering was done periodically every 1.7 s, the spacing of subsequent measured CIRs is approximately 1.7 cm. Roughly 155 trigger events were performed implying 2480 measured CIRs per region.

Fig. 3.2 depicts the received signal at receive antenna 3 after cross-correlation with the reference PN bit stream. The four visible CIRs origin from four time-shifted transmit signals at the four transmit antennas.

3.3 Impact of Antennas

The measured CIRs comprise the patterns and the transfer functions of the transmit and receive antennas. Skycross SM3TO10MA antennas, as depicted in Fig. 3.3, were used for the measurements.

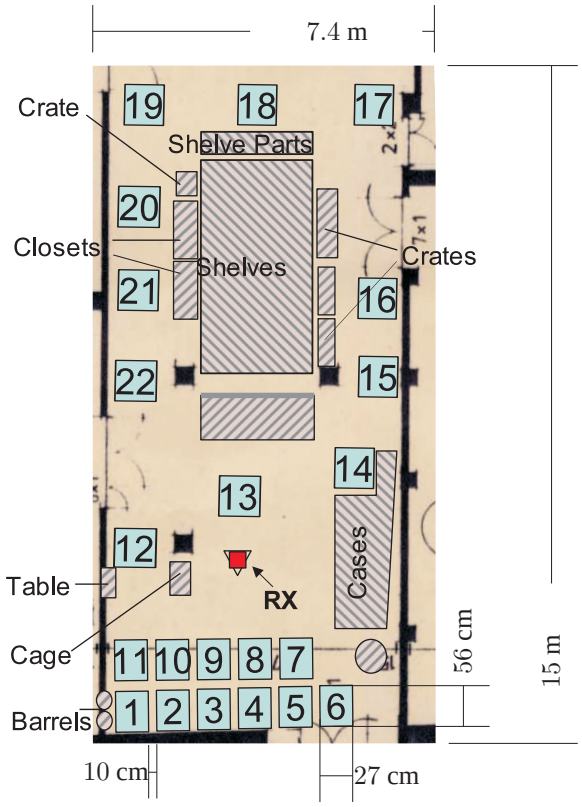

Fig. 3.1: Measurement scenario: Regions, propagation environment, and RX.

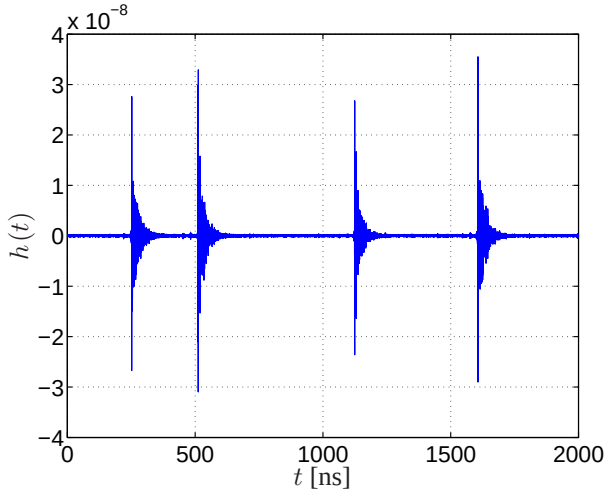

Fig. 3.2: Output of the cross-correlation at receive antenna 3.

These antennas are specified to work in the frequency range from 3.1 to 10 GHz. Furthermore, the data sheet[1] claims that these antennas are omnidirectional in the azimuth plane,

[1]SM3TO10MA data sheet is online available at www.skycross.com/Products/PDFs/SMT-3TO10M-A.pdf.

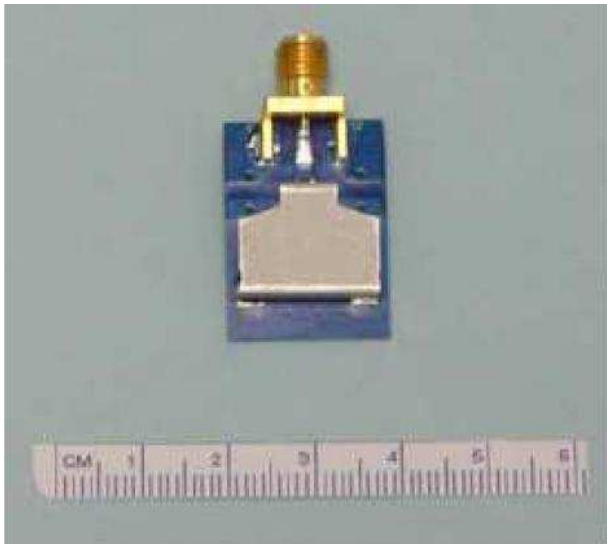

Fig. 3.3: Skycross SM3TO10MA UWB antenna.

which is defined as the plane perpendicular to the axis of the antenna. The measured antenna patterns in the azimuth plane for different frequencies are depicted in Fig. 6.1 on page 59 in [49]. It can be concluded from these measurement results that the largest gain difference is approximately 5 dB, which occurs between the azimuth angles 90 degrees and 270 degrees. Further, it can be concluded that the antenna gain stays roughly constant for small changes in the azimuth angle.

The orientations of the receive antennas were fixed during the whole measurement campaign. Further, the transmit antennas were mounted with fixed orientations on the positioning device. Due to the movement of this device within each region the orientations of the transmit antennas were only slightly changing. Therefore, it is assumed that the influence of the antenna patterns on the CIRs measured in each region is negligible.

3.4 Measurement Signal-to-Noise Ratio

The thermal noise due to the LNA and the electronics of the oscilloscope is assumed to be an additive zero mean white Gaussian noise process. The corresponding noise samples are therefore Gaussian distributed with zero mean and variance σ^2_{mea}, which is given by the room temperature, the noise figures of the LNA and oscilloscope, and the measurement bandwidth. The measurement SNR is denoted by SNR_{mea} and is defined as the average energy of the CIRs within one region divided by $2\sigma^2_{\text{mea}}$. The measurement SNR ranges from 45 dB to 55 dB depending on the distances between the center points of the regions and the

RX. Due to this high measurement SNR, it is assumed that the measured CIRs are essentially distortion-free.

3.5 Energy Normalization

In order to remove the dependence of the location fingerprints on the transmit power of the agents, the average energy of CIRs within one region is normalized to the same value for all regions. This is necessary, whenever it cannot be guaranteed that all agents use the same transmit power. For example this can happen, when the data rate is adapted to the current work load or channel conditions. It is important to notice that energy normalization implies that the PL information is not used for position location.

3.6 Temporal Alignment

The measurement setup does not provide absolute timing information, which means that there exists an unknown time offset for each measured CIR. We have already pointed out that signal metrics related to the direct path are not required for location fingerprinting. However, we require some sort of temporal alignment of the measured CIRs, in order to obtain meaningful location fingerprints and estimate meaningful statistical parameters.

In every wireless communication system, the RX needs to synchronize its symbol timing to the symbol timing of the TX. Such a synchronization algorithm would be responsible for the temporal alignment task in a realistic RX implementation. In the following, simple temporal alignment strategies are discussed.

3.6.1 Problem Formulation

We consider a sampled received signal in an observation window $[0, K-1]$ denoted by $\mathbf{r}$ and assume that the CIR denoted by $\mathbf{h}$ of length $N < K$ is located somewhere within the observation window. The objective is to identify a reference sample, estimate it for each measured CIR and shift the time axis accordingly such that all measured CIRs have a common time reference. Such a reference sample could be, for example, the sample with the maximum absolute value.

3.6.2 Maximum Absolute Value Alignment

The simplest strategy is to identify the sample with the maximum absolute value as the reference sample. This can be written as

$$n_{\text{maxabs}} = \underset{n\in\{0,\ldots,K-1\}}{\text{argmax}} |r[n]|.$$

Depending on the SNR, the noise samples present in $\mathbf{r}$ may produce large variations in n_{maxabs}. Thus, this alignment strategy requires a high SNR.

In LoS situations, the CIRs are most likely aligned to the direct path, which implies that samples before the reference sample are only noise samples. However, in case of a non-LoS situation, the CIRs are aligned to the strongest path, which is not necessarily the direct path. This means that also samples before the reference sample can carry significant CIR energy and contain position location information. There is a trade-off between accounting for possible non-LoS situations and wasting samples by accounting for noise-only samples.

In general, it can be assumed that the position location system can afford more RX complexity, for example a higher SNR, during the training phase than during the localization phase. Therefore, we restrict the SNR analysis to the localization phase and assume that we can provide SNR_{mea} during the training phase. Thus, the temporal alignment of CIRs utilized during the training phase is done at SNR_{mea}.

Time-of-Arrival Alignment A more sophisticated temporal alignment strategy is to identify the reference sample as first sample in $\mathbf{h}$, i.e. $h[0]$. The first sample $h[0]$ is usually interpreted as ToA of the CIR. The interested reader is referred to [6, 22, 23] for a detailed discussion of various methods with different complexities for ToA estimation. The simulation results provided in this thesis consider exclusively maximum absolute value alignment as discussed above.

Chapter 4

Location Fingerprinting with a Coherent Receiver

In this chapter[1], location fingerprinting with a *single* anchor using a *coherent* RX is studied. The RX is called coherent, because it can perform CIR estimation. Note that we do not require anything from the agents except from being able to send a training sequence with a certain bandwidth for CIR estimation. Especially time synchronization between agents and RX is not required. Thus, the whole complexity of the position location system is shifted to the anchor, which also acts as CU in accordance to the framework developed in Chapter 2.

We adopt the formulation of stochastic location fingerprints proposed in Section 2.3, which means that the location fingerprints $\mathbf{s}_m$ within each region m are modeled by a PDF with parameter set Θ_m. In the course of this chapter, we will answer some of the fundamental questions raised in Chapter 2. The first two questions, which are treated in Sections 4.1 and 4.2, are concerned with the choice of the location fingerprints and with an accurate and mathematically tractable stochastic description of them.

[1] Parts of this chapter have been published in [50, 51].

4.1 Choice of the Location Fingerprints

A physical multipath propagation channel with L paths can be interpreted as tapped delay line [52] with L taps according to

$$h(t,\tau) = \sum_{l=1}^{L} \gamma_l(t)\delta\left(\tau - \tau_l(t)\right),$$

where $\gamma_l(t)$ is the time-varying gain and $\tau_l(t)$ is the time-varying delay of path l. Time variations of gains and delays are caused by movement of TX or RX (or both) and/or movement of objects in the environment.

Since any wireless transmission system is band-limited, the continuous time CIR of interest is as well and can be represented by discrete-time samples denoted by $h[t,n] \triangleq h_{\text{band-limited}}\left(t, \frac{n}{f_s}\right)$ with f_s as the sampling rate. If f_s fulfills the sampling theorem, the continuous time CIR can be recovered by interpolation with time-shifted sinc functions.

We consider an equivalent baseband representation of the continuous time CIR, which implies complex valued samples $h[t,n]$. Since the equivalent baseband samples provide a complete description of the continuous time CIR, it is proposed to use these samples within an observation window as location fingerprint and define $\mathbf{s}_m$ with length N according to

$$\mathbf{s}_m(t) \triangleq \left[h_m[t,0], h_m[t,1], \ldots, h_m[t,N-1]\right]^T,$$

where the subscript m denotes the agent's region. Thus, $h_m[t,n]$ denotes the n-th sample of an equivalent baseband CIR from an agent located in region m observed at time instant t. The time dependency of these samples is still considered, which causes time-varying location fingerprints. Due to the time windowing, the continuous time CIR cannot be reconstructed based on $\mathbf{s}_m(t)$ without error. The observation window should be chosen such that the main part of the CIR energy is captured. Furthermore, it is assumed that $\mathbf{s}_m(t)$ does not change in t during one observation window. Since the observation duration is typically in the order of tens of nanoseconds, this is assumption is well justified in indoor scenarios due to the limited speed of TX, RX, and objects in the environment.

If f_s is larger than Nyquist rate this corresponds to oversampling, which provides a redundant representation of the continuous time CIR. This implies that the length of $\mathbf{s}_m$ increases for the same observation window duration and, consequently that more statistical parameters have to be estimated. This, in turn, increases the complexity of the training phase without increasing the information content. Thus, f_s is chosen as Nyquist rate.

4.2 Regional Channel Model

For design, implementation, and evaluation of a location fingerprinting system, a statistical model of $\mathbf{s}_m(t)$ is required. The statistical description should be robust to time-variations caused by a randomly changing propagation environment, for example, by moving people. More importantly, we have to accurately model the variations of the location fingerprints within a small region of space. This change in position in the order of a few multiples of the carrier wavelength causes small scale fading of the samples $h[t, n]$, which can be modeled with various probability distributions [52, 53]. In the following we use the terms channel taps and samples interchangeably.

A common assumption on the small scale fading behavior of equivalent baseband channel taps is that they are complex Gaussian distributed. The justification of this assumption is given by the central limit theorem. Many reflected and scattered partial waves of similar average power from different directions superimpose at the receive antenna and contribute to one channel tap. If the number of partial waves is large enough, the central limit theorem can be applied and the resulting distribution of the channel tap can be assumed to be Gaussian [54, 55]. However, as the sampling frequency becomes larger due to increasing signal bandwidth (approaching UWB bandwidths) less of these partial waves contribute to one channel tap. This fact questions the applicability of the central limit theorem.

In literature, there exist various studies on the tap statistics for UWB channels. For the tap amplitudes, the Nakagami-m [56], Log-normal [49, 57], and Weibull [58] distributions are proposed. However, also Rayleigh and Rice amplitude distributions arising from a complex Gaussian channel model are supported by some channel measurement campaigns [59, 60]. The phase distribution of the channel taps is commonly assumed to be uniform between $-\pi$ and π.

In the next section, we apply Akaike's information criterion (AIC) [61] to channel taps obtained by the measurement campaign described in Chapter 3, in order to find the best distribution out of a candidate set for the modeling of the small scale fading behavior. A short review of AIC can be found in Appendix B. Furthermore, thorough discussions about UWB channel modeling can be found in [52, 53, 60, 62, 63]. The AIC is applied to channel taps after measurement post-processing steps in the following order: Passband filtering from 3 to 6 GHz, equivalent baseband transformation, threefold interpolation (upsampling to 60 GHz sampling rate), temporal alignment, downsampling to Nyquist rate (downsampling to 3 GHz sampling rate), and energy normalization.

4.2.1 Marginal Distribution of Channel Taps

The considered ensemble consists of all measured CIRs from one region (all four transmit antennas) to receive antenna 3. The number of CIRs in this ensemble is 620. Region 2 with LoS CIRs and region 18 with non-LoS CIRs are chosen for presentation in this section. Further investigations show that the obtained results and drawn conclusions hold for all other regions as well. Although the measurement process for one region lasted for roughly 250 seconds, it is assumed that the ensemble is measured at a single time instant $t = t_0$. Thus, the time dependency of the location fingerprints is omitted in the following considerations. The propagation environment was kept mostly static during the measurements, which justifies this assumption.

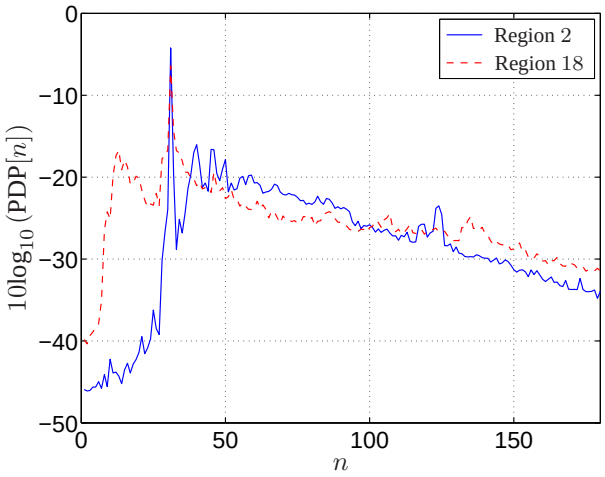

Fig. 4.1: Normalized PDPs in dB for regions 2 and 18.

Fig. 4.1 depicts the normalized power delay profiles (PDPs) for regions 2 and 18 in dB. The reference sample is set to $n_{\text{maxabs}} = 31$. It can be seen that samples before the reference sample carry significant CIR energy for region 18, which is not the case for region 2. The observation window has a duration of 60 ns, which corresponds to $N = 180$ samples obtained with 3 GHz Nyquist sampling.

Amplitude Distribution The candidate set consists of the following six distributions: Nakagami-m, Log-normal, Weibull, Gamma, Rayleigh, and Rice. All distributions have two parameters except for the Rayleigh distribution, which has only one parameter. This choice is based on channel modeling literature as discussed above. Fig. 4.2 and Fig. 4.3 depict the respective Akaike Weights for all fitted candidate distributions for region 2 and region 18. The Akaike Weights (cf. Appendix B) are estimates for the probability that the

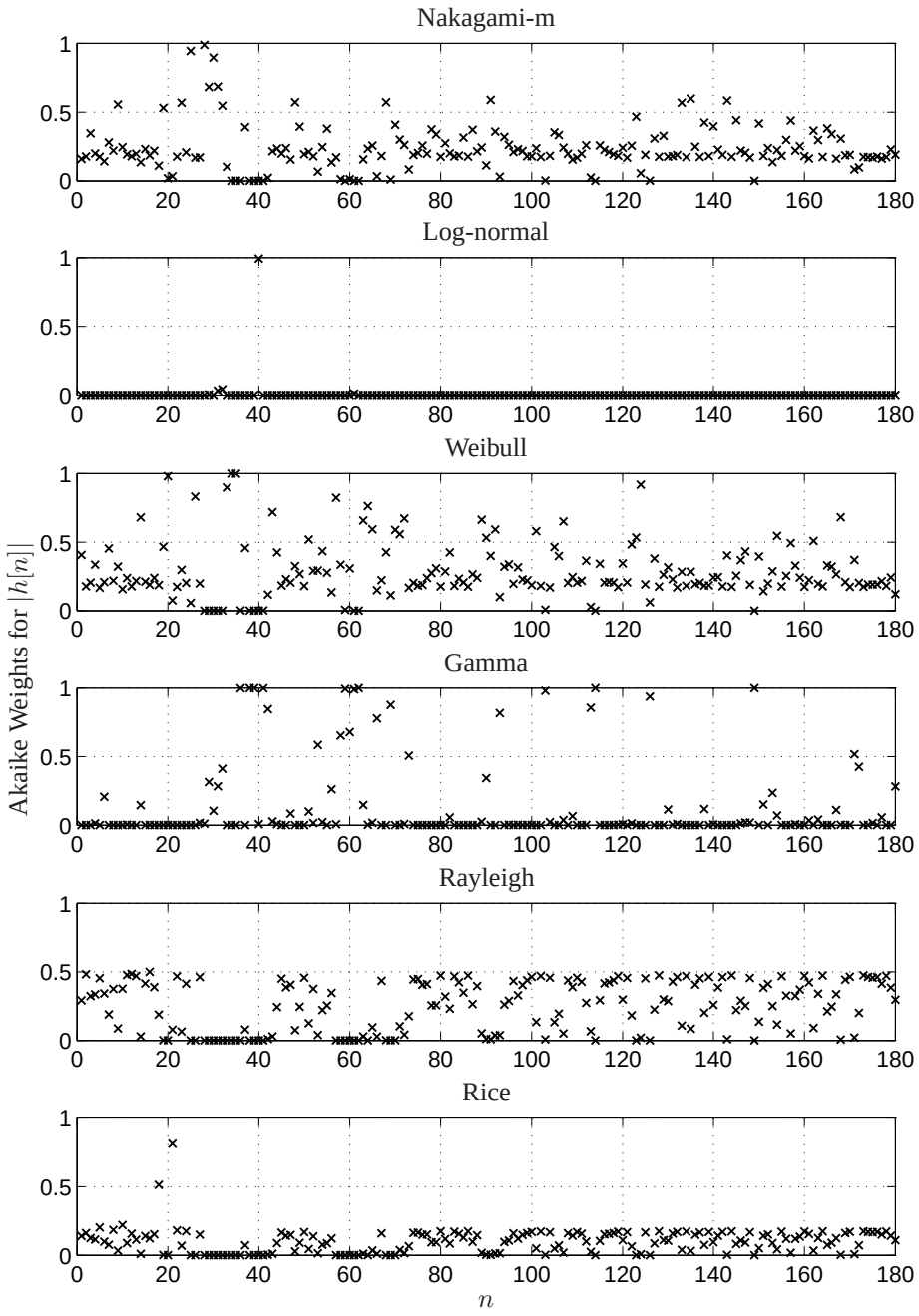

Fig. 4.2: Akaike Weights for tap amplitudes $|h[1]|, |h[2]|, \ldots, |h[180]|$ for CIRs from region 2.

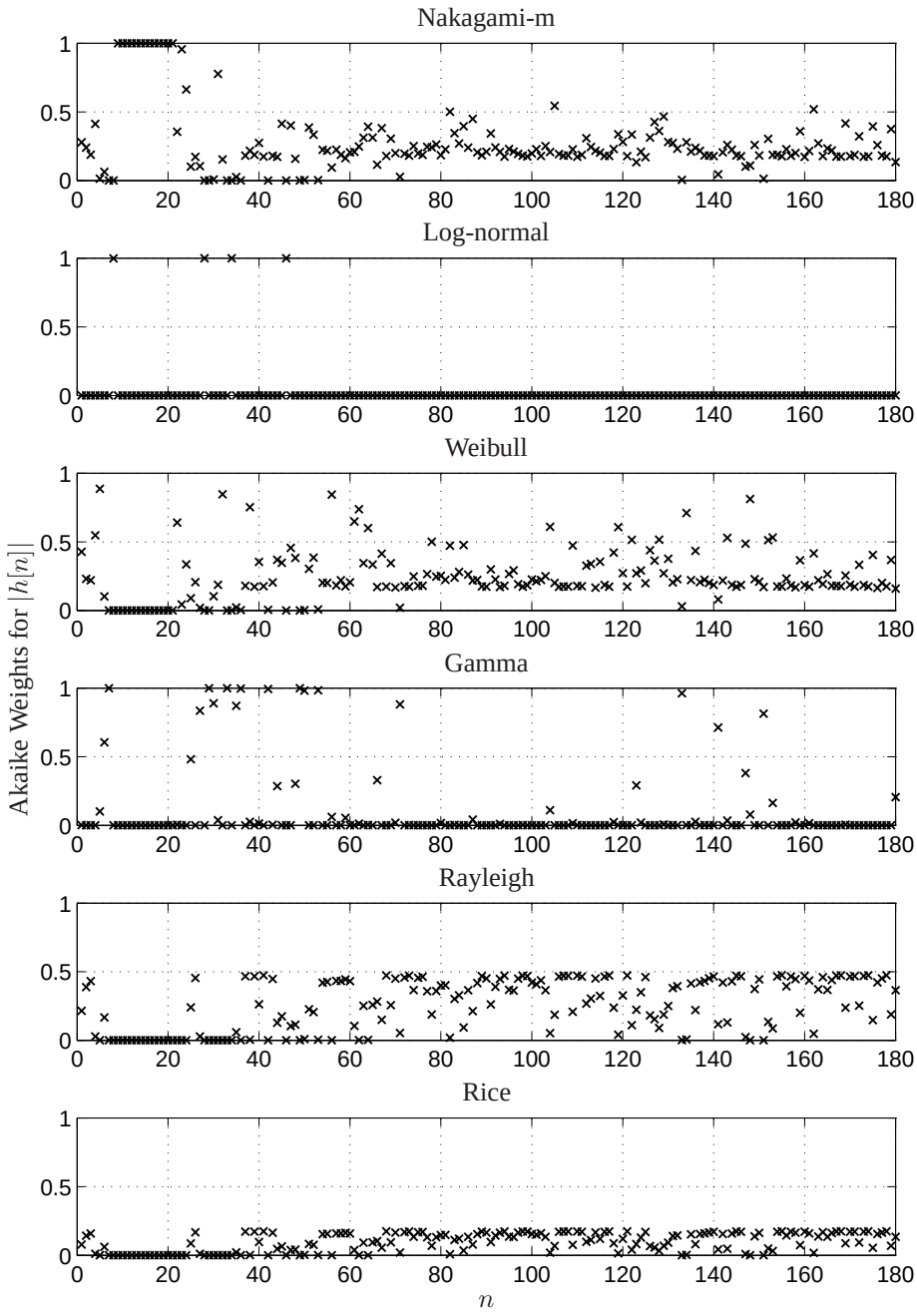

Fig. 4.3: Akaike Weights for tap amplitudes $|h[1]|, |h[2]|, \ldots, |h[180]|$ for CIRs from region 18.

corresponding distribution shows the best modeling quality within the candidate set. Thus, a high Akaike Weight indicates that this distribution is a good candidate to model amplitude variations of the corresponding channel tap. The Akaike Weights sum up to one for each n.

Following observation can be made:

1. There does not exist a single distribution, which shows the best fit for all n.
2. Log-normal and Gamma distributions are not suited for the majority of taps.
3. The channel taps can be separated into two categories with different behavior:
   - The taps surrounding n_{maxabs} are not well modeled with a Rayleigh or Rice distribution due to the temporal alignment, which enforces large tap amplitudes for all CIRs.

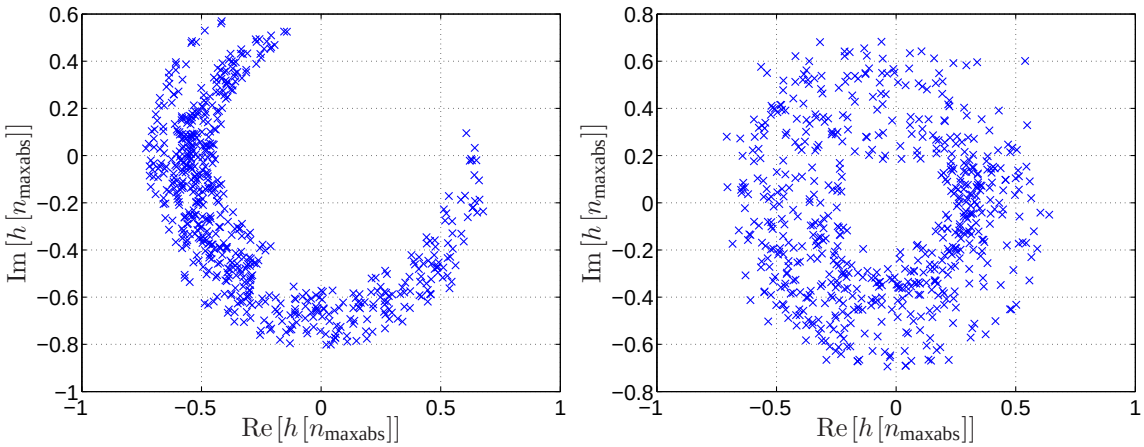

Fig. 4.4: Scatter plot for $h\,[n_{\text{maxabs}}]$ for region 2 (left) and region 18 (right).

Fig. 4.4 shows a scatter plot for $h\,[n_{\text{maxabs}}]$ for region 2 and region 18. It can be seen that the reference samples do not follow a complex Gaussian distribution. This can be also observed for surrounding channel taps. These taps, however, do not carry significant information for location fingerprinting, since the temporal alignment procedure is done in the same way for all regions.

   - The majority of all other taps is best modeled with the Rayleigh distribution. However, the Nakagami-m or the Weibull distribution has sometimes the largest Akaike Weight. When we look at the parameter values for Nakagami-m and Weibull in these cases, it can be observed that they are very close to the values that reduce the respective distribution to a Rayleigh distribution.

4. There exists no significant difference between LoS region 2 and non-LoS region 18 except for the first arriving cluster (samples 9 to 21). Here, Nakagami-m distribution

fits the tap statistics from region 18 best. However, the values for the parameter m are again close to 1, which would reduce the Nakagami-m distribution to the Rayleigh distribution.

5. Fig. 4.5 and Fig. 4.6 depict the fitted PDFs and the histograms of the reference sample and a multipath channel tap for regions 2 and 18. It can be seen that Rayleigh and Rice show a bad fit for the reference sample. For the multipath channel tap all PDFs except Gamma and Log-normal almost overlap. Note that the Akaike Weights for Rayleigh and Rice for the chosen multipath channel taps are smaller than for Weibull and Nakagami-m.

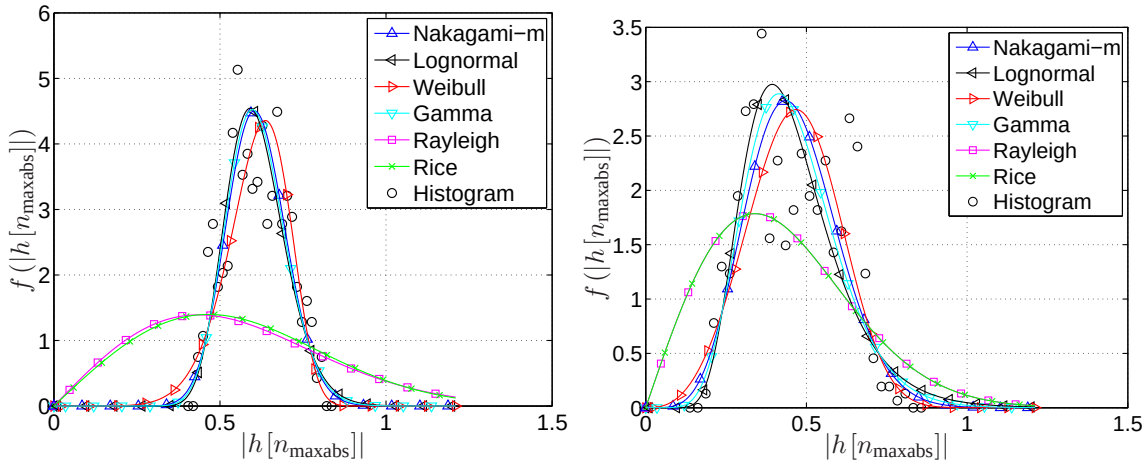

Fig. 4.5: PDFs of $|h[n_{\text{maxabs}}]|$ from region 2 (left) and region 18 (right).

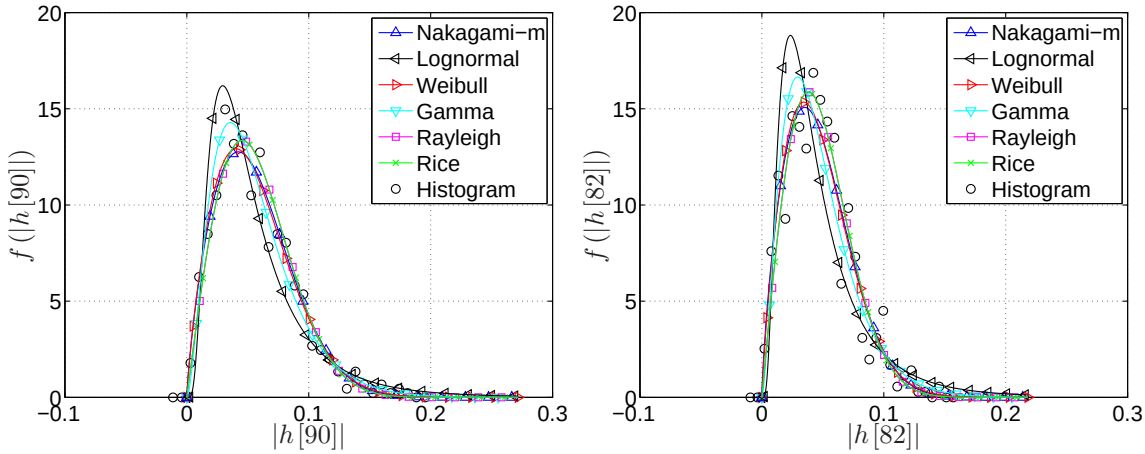

Fig. 4.6: PDFs of multipath channel taps from region 2 (left) and region 18 (right).

It can be concluded from these empirical results that the Rayleigh distribution is a meaningful candidate to model the small scale fading behavior of most of the channel tap amplitudes. This is also in line with the results from measurement campaign I in [59], where a measurement scenario similar to the one described in Chapter 3 is considered. The ensemble in [59] consists of CIRs from a TX array to one receive antenna in a static environment.

However, there is an obvious modeling mismatch of the Rayleigh distribution for taps surrounding n_{maxabs}. These taps would be better modeled by any other distribution than Rayleigh or Rice, which can be seen from Fig. 4.5. According to the theory of AIC, we would have to use multiple candidate distributions to accurately model all channel tap amplitudes. Unfortunately, this would prevent us from defining a suitable joint distribution of the channel taps, which is discussed later in Section 4.2.2. Therefore, it is proposed to stick to the Rayleigh distribution for all channel taps. With the application of location fingerprinting in mind, this is a sensible modeling decision, because the channel taps carrying the least information about the agent's location are the samples surrounding n_{maxabs}. The reason for this is the temporal alignment process, which is the same for all CIRs from all regions.

The distributions of the tap phases remain to be investigated in the following. The working hypothesis is that the phases are uniformly distributed between $-\pi$ and π. This is motivated by propagation effects, because already small changes in the path delays cause large changes of the tap phases.

Phase Distribution As stated above, the phase distribution is generally assumed to be uniform in literature. Thus, we seek to verify this assumption based on visual inspection of the histograms obtained from measured channel taps.

Fig. 4.7 depicts the histograms of the tap phases $\phi[n] \triangleq \arg[h[n]]$ for the two considered regions. The gray scale encodes relative occurrence, i.e., darker means more frequently. Visual inspection generally confirms the uniform phase distribution hypothesis, which is also supported by the experimental findings in [60]. An exception are the samples for $n \leq 32$ from region 2, for which some phases have almost zero relative occurrence. This implies that, in case of LoS propagation and sufficiently small multipath interference from succeeding paths, the temporal alignment works accurately enough to partially recover phase information of the LoS path. This is not the case for CIRs from region 18 due to the multipath interference affecting also preceding samples of n_{maxabs}.

This effect of the temporal alignment process is emphasized, if alignment is done with highly oversampled passband CIRs. For the here presented results, 9-fold interpolation is

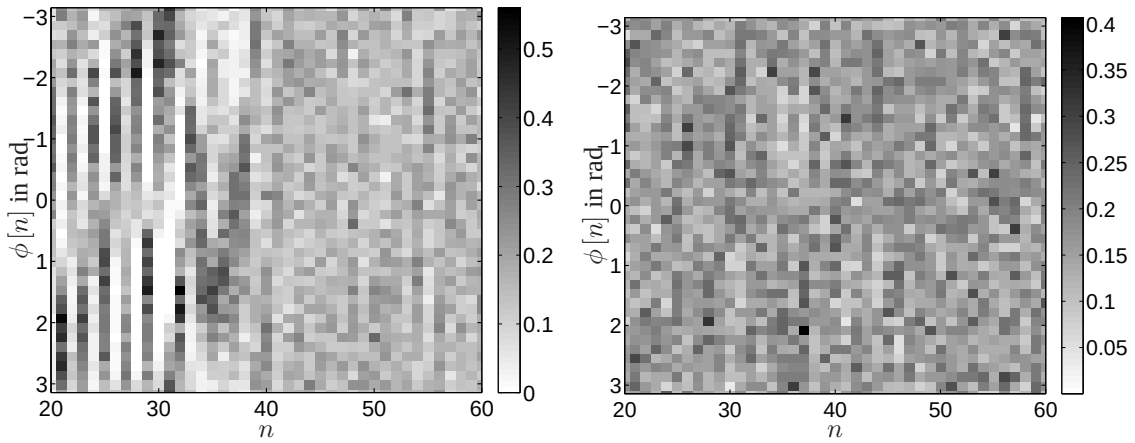

Fig. 4.7: Histograms of the tap phases for CIRs from regions 2 (left) and 18 (right). The gray scale encodes relative occurrence.

applied implying a sampling rate of 180 GHz. The new measurement post-processing order is: 9-fold interpolation, temporal alignment, passband filtering from 3 to 6 GHz, equivalent baseband transformation, downsampling to Nyquist rate, and energy normalization.

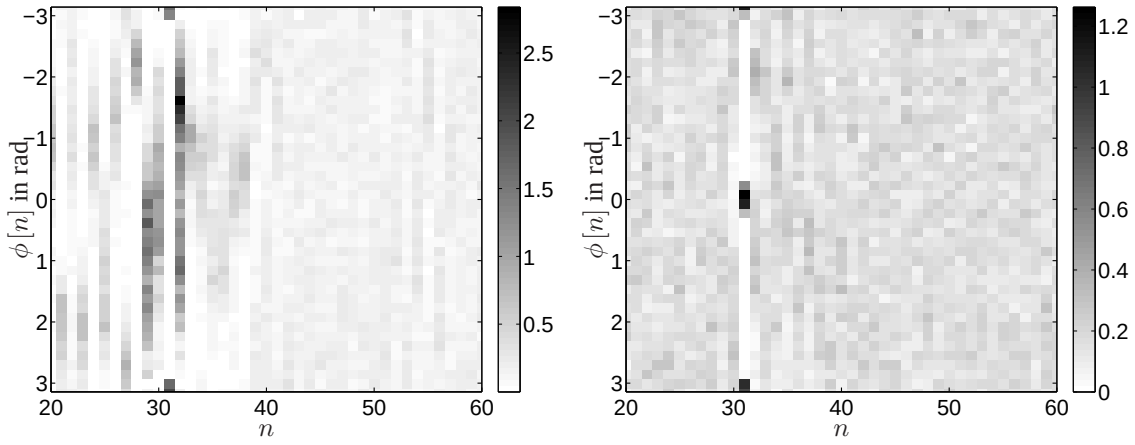

Fig. 4.8: Histograms of the tap phases for CIRs from region 2 (left) and 18 (right). The gray scale encodes relative occurrence. Temporal alignment of passband CIRs with 180 GHz sampling rate.

Fig. 4.8 depicts the corresponding histograms of the tap phases for the two considered regions. For region 2 it can be observed that the phase of the reference sample $n_{\text{maxabs}} = 31$ is almost deterministic, i.e. $\phi[n_{\text{maxabs}}] = \pi$. Furthermore, the surrounding samples show similar phase distributions with zero relative occurrence for many values. For region 18 two very likely values for $\phi[n_{\text{maxabs}}] \in \{0, \pi\}$ can be observed, but the phases of the remaining channel taps are well modeled by a uniform distribution.

Most of the channel tap phases are very well modeled by a uniform distribution, if the temporal alignment process is done with baseband signals. For the remaining simulation results in this thesis temporal alignment is exclusively performed with baseband signals.

Conclusions Based on this empirical analysis of measured channel taps, we come to the conclusion that the most promising stochastic model of the channel taps is circular symmetric Gaussian [64, 65], which leads to Rayleigh distributed amplitudes and uniformly distributed phases. Interestingly, this model has the least complexity (only one free real parameter per channel tap) compared to all other candidates. Furthermore, the Gaussian distribution has many advantages in terms of mathematical tractability. It is, for example, straightforward to extend the Gaussian distribution to a multivariate (MV) Gaussian distribution.

4.2.2 Second Order Statistics

So far, we have investigated the marginal channel tap statistics of equivalent baseband CIRs over rectangular shaped regions of dimension 27 cm times 56 cm. However, a full stochastic description of the location fingerprints requires a joint distribution of $\mathbf{s}_m(t)$. In this section, we consider the same ensembles of measured CIRs from region 2 and region 18 as for the analysis of the marginal channel tap statistics. It is again assumed that the ensembles are observed at a single time instant $t = t_0$. This implies that the time dependency of the location fingerprints is neglected. Therefore, the joint distribution of $\mathbf{s}_m \triangleq \mathbf{s}_m(t_0)$ is sought.

In order to apply model selection techniques like AIC as used for the marginal tap statistics a huge amount of independent CIR realizations needs to be measured. Since this is not feasible, the following analysis focuses only on the second order statistics, i.e. correlations among channel taps. In general, correlations among channel taps of an equivalent baseband CIR exists due to the following reasons (cf. [52, 53, 60]):

- Band-limited filters with non-rectangular frequency responses. This includes also the antenna transfer functions.
- Spatial variations over multiple wavelengths cause multipath components to extend over several channel taps.
- Frequency dependent propagation mechanisms, which are further discussed in Section 7.1.
- Correlated scatterers or correlated multipath components.

The last three effects are related to the propagation environment and are, therefore, location dependent. It is expected that these correlations can be exploited to improve the location fingerprinting performance by providing a more distinct stochastic description of the location fingerprints originating from different regions.

We propose to model the location fingerprint vectors $\mathbf{s}_m$ as complex-valued jointly Gaussian random vectors. Hence, the distribution of $\mathbf{s}_m$ is fully determined by the first and second order statistics, i.e. by the mean and the covariance. This modeling assumption is in accordance to the results of Section 4.2.1, because the marginal distributions of a complex-valued jointly Gaussian random vector are complex-valued Gaussian.

Complex-Valued Jointly Gaussian Random Vectors In this paragraph, we review the definitions of complex-valued jointly, proper, and circular symmetric Gaussian random vectors.

A complex-valued random vector $\mathbf{h}$ of length N is said to be jointly Gaussian distributed, if the real-valued random vector $\tilde{\mathbf{h}}$ of length $2N$ defined as

$$\tilde{\mathbf{h}} \triangleq \begin{bmatrix} \mathrm{Re}\,[\mathbf{h}] \\ \mathrm{Im}\,[\mathbf{h}] \end{bmatrix}$$

is jointly Gaussian distributed. Thus, in order to specify the distribution of $\mathbf{h}$, we require the mean vector $\tilde{\boldsymbol{\mu}} = \mathrm{E}\left[\tilde{\mathbf{h}}\right]$ and the covariance matrix $\tilde{\boldsymbol{\Sigma}} = \mathrm{E}\left[\left(\tilde{\mathbf{h}} - \tilde{\boldsymbol{\mu}}\right)\left(\tilde{\mathbf{h}} - \tilde{\boldsymbol{\mu}}\right)^T\right]$.

The distribution of $\mathbf{h}$ is said to be proper [64, 65] if and only if $\tilde{\boldsymbol{\Sigma}}$ can be written as

$$\tilde{\boldsymbol{\Sigma}} = \begin{bmatrix} \mathrm{Re}\,[\mathbf{A}] & \mathrm{Im}\,[\mathbf{A}] \\ -\mathrm{Im}\,[\mathbf{A}] & \mathrm{Re}\,[\mathbf{A}] \end{bmatrix},$$

for some non-negative definite Hermitian matrix $\mathbf{A}$. In this case we have for $\boldsymbol{\mu} = \mathrm{E}\,[\mathbf{h}]$ that

$$\begin{aligned} \boldsymbol{\Sigma} &= \mathrm{E}\left[(\mathbf{h} - \boldsymbol{\mu})(\mathbf{h} - \boldsymbol{\mu})^H\right] = \mathbf{A} \\ \mathbf{Z} &= \mathrm{E}\left[(\mathbf{h} - \boldsymbol{\mu})(\mathbf{h} - \boldsymbol{\mu})^T\right] = \mathbf{0}. \end{aligned}$$

Note that the *pseudo-covariance matrix* $\mathbf{Z}$ becomes zero. Thus, the proper Gaussian PDF is fully determined by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and is given by

$$f(\mathbf{h}) = \frac{1}{\pi^N |\boldsymbol{\Sigma}|} \exp\left(-(\mathbf{h} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{h} - \boldsymbol{\mu})\right).$$

In order to specify a conventional jointly Gaussian random vector of length N, $3N+2N^2$ free real parameters are required. For a proper Gaussian random vector of length N, only $2N+N^2$ free real parameters are required.

If $\boldsymbol{\mu}$ is the zero vector, then the distribution of $\mathbf{h}$ is called circular symmetric Gaussian. In order to specify a circular symmetric Gaussian random vector, $\boldsymbol{\Sigma}$ with N^2 free real parameters is required.

Empirical Correlation Matrices In this paragraph, we present empirical correlation matrices of location fingerprints from region 2 and region 18. In accordance to the results from Section 4.2.1, the channel taps are assumed to have zero mean, which implies that $\mathrm{E}[\mathbf{s}_m] = \mathbf{0}$ is assumed.

The empirical covariance and correlation matrices are estimated based on a set of L location fingerprints from region m: $\left\{\mathbf{s}_m^{(1)}, \mathbf{s}_m^{(2)}, \ldots, \mathbf{s}_m^{(L)}\right\}$. The ML estimates $\hat{\boldsymbol{\Sigma}}_m$ of the covariance matrix $\boldsymbol{\Sigma}_m$ and $\hat{\mathbf{Z}}_m$ of the pseudo-covariance matrix $\mathbf{Z}_m$ are given by

$$\hat{\boldsymbol{\Sigma}}_m = \frac{1}{L}\sum_{l=1}^{L} \mathbf{s}_m^{(l)} \left(\mathbf{s}_m^{(l)}\right)^H \text{ and } \hat{\mathbf{Z}}_m = \frac{1}{L}\sum_{l=1}^{L} \mathbf{s}_m^{(l)} \left(\mathbf{s}_m^{(l)}\right)^T. \tag{4.1}$$

By normalizing the entries in these covariance matrices according to

$$\left[\hat{\mathbf{R}}_m\right]_{i,j} = \frac{\left[\hat{\boldsymbol{\Sigma}}_m\right]_{i,j}}{\sqrt{\left[\hat{\boldsymbol{\Sigma}}_m\right]_{i,i}\left[\hat{\boldsymbol{\Sigma}}_m\right]_{j,j}}} \text{ and } \left[\hat{\mathbf{K}}_m\right]_{i,j} = \frac{\left[\hat{\mathbf{Z}}_m\right]_{i,j}}{\sqrt{\left[\hat{\boldsymbol{\Sigma}}_m\right]_{i,i}\left[\hat{\boldsymbol{\Sigma}}_m\right]_{j,j}}},$$

the corresponding empirical correlation matrix $\hat{\mathbf{R}}_m$ and pseudo-correlation matrix $\hat{\mathbf{K}}_m$ are obtained.

Fig. 4.9 and Fig. 4.10 depict the absolute values of the empirical correlation matrices and pseudo-correlation matrices for region 2 and region 18 using ensembles of $L = 620$ measured CIRs per region. For both regions, several correlation coefficients with an absolute value larger than 0.5 - especially in the vicinity of the main diagonal - can be observed.

The correlation values for channel taps with $n \leq n_{\text{maxabs}} = 31$ in the left plot of Fig. 4.9 are larger than 0.5, which is not the case in the left plot of Fig. 4.10. The reason is again the temporal alignment process. Since region 2 is a LoS region and the multipath interference from succeeding paths on channel taps $n \leq n_{\text{maxabs}}$ is weak, these channel taps have similar amplitudes and phases for all temporally aligned CIRs. This results in high channel tap correlation values. In region 18 there exists also multipath interference for $n \leq n_{\text{maxabs}}$,

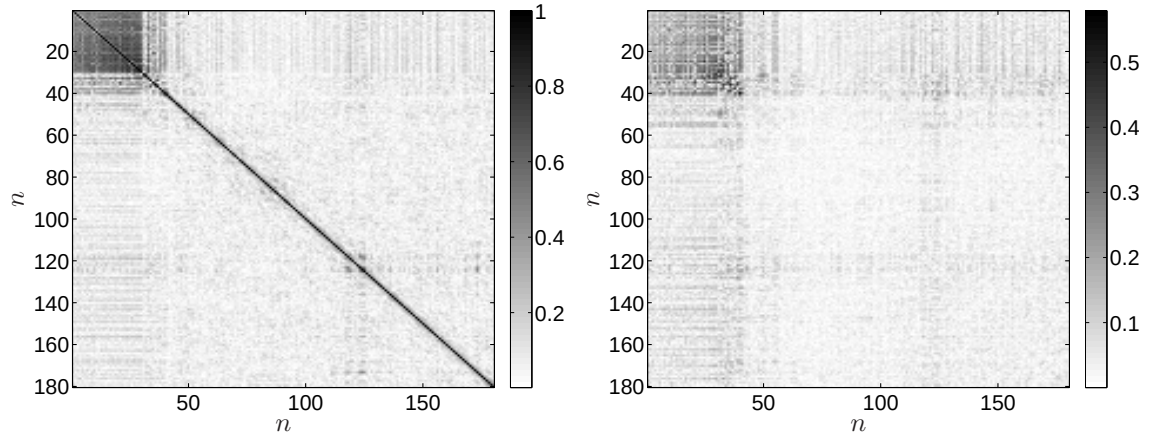

Fig. 4.9: Absolute values of correlation matrix $\hat{\mathbf{R}}_2$ (left) and pseudo-correlation matrix $\hat{\mathbf{K}}_2$ (right).

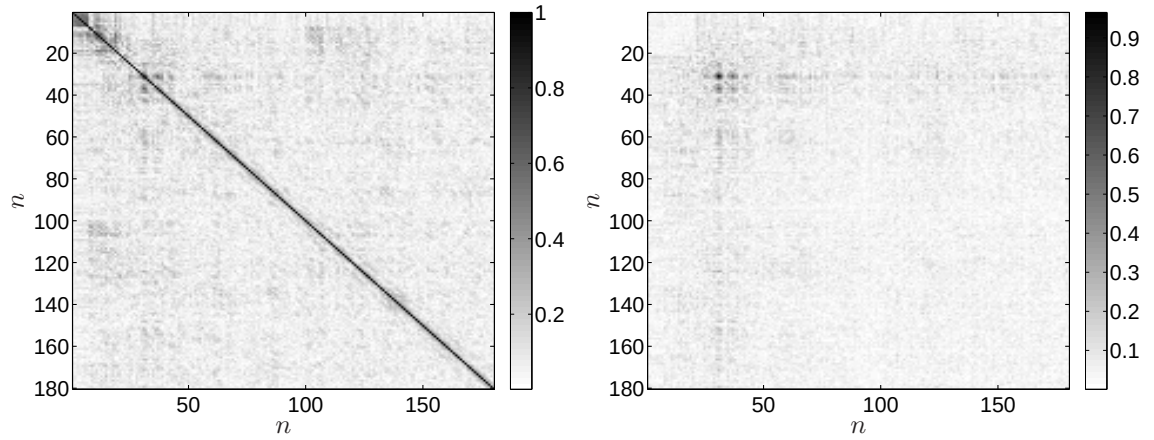

Fig. 4.10: Absolute values of correlation matrix $\hat{\mathbf{R}}_{18}$ (left) and pseudo-correlation matrix $\hat{\mathbf{K}}_{18}$ (right).

which explains why the corresponding correlation values are significantly smaller. Although these high correlation values cannot improve the localization performance for LoS regions - because they are equal for all LoS regions - they can help to distinguish LoS and non-LoS regions.

In general, it cannot be assumed that the pseudo-correlation matrices and, consequently, the pseudo-covariance matrices vanish. Visual inspection of the right plots in Fig. 4.9 and Fig. 4.10 indicates that there exist some non-negligible pseudo-correlation values. Therefore, it will be again a modeling decision, whether the location fingerprints are assumed to be circular symmetric or not. In order to reduce the complexity of the training phase as much as possible, only as few free modeling parameters as possible should be used. This reasoning

favors the assumption of circular symmetric location fingerprints. On the contrary, if a lot of training data is available, a better performance could be achieved assuming non-zero pseudo-covariance matrices.

4.2.3 Summary

It is proposed to model the location fingerprints $\mathbf{s}_m$ as circular symmetric Gaussian random vectors. The zero mean assumption stems from the empirical analysis of the marginal channel tap statistics performed in Section 4.2.1. Hence, the PDF of location fingerprints from agents located in region m is given by

$$f_m\left(\mathbf{s}|H_m\right) = \frac{1}{\pi^N \left|\mathbf{\Sigma}_m\right|} \exp\left(-\mathbf{s}^H \mathbf{\Sigma}_m^{-1} \mathbf{s}\right).$$

In order to specify this PDF for each region the estimation of $\mathbf{\Sigma}_m$ with N^2 free real parameters is required. Thus, the parameter set $\Theta_m = \{\mathbf{\Sigma}_m\}$ fully describes the location fingerprints from agents located in region m, which are obtained with a coherent RX.

For the sake of performance comparison we present in the coarse of this chapter analytical and empirical performance results for the following location fingerprint models:

- COV model:

  Circular symmetric Gaussian PDF with $\Theta_m = \left\{\mathbf{\Sigma}_m = \mathrm{E}\left[\mathbf{s}_m \mathbf{s}_m^H\right]\right\}$ and N^2 free real parameters.

- PCV model:

  Zero mean complex-valued jointly Gaussian PDF with $\Theta_m = \left\{\mathbf{\Sigma}_m, \mathbf{Z}_m = \mathrm{E}\left[\mathbf{s}_m \mathbf{s}_m^T\right]\right\}$ and $2N^2 + N$ free real parameters.

- PDP model:

  Circular symmetric Gaussian PDF assuming independent channel taps with $\Theta_m = \left\{\mathrm{PDP}_m \triangleq \mathrm{diag}\left[\mathbf{\Sigma}_m\right]\right\}$ and N free real parameters.

4.3 Distortions of the Location Fingerprints

4.3.1 Antenna Patterns

If the agents are mobile, there exists inevitable variations (distortions) of the location fingerprints due to antenna patterns of TX and RX antenna. If the antenna patterns deviate

significantly from omnidirectional antennas[2], the location fingerprints depend strongly on the antenna orientations. This effect has to be accounted for during the training phase by either separate training (parameter estimation) for all relevant orientations or by averaging the location fingerprints over all orientations. If the impact of the antenna patterns is ignored, the varying antenna orientations due to movements of the agent will distort the location fingerprints in comparison to the training phase and act as error source.

Throughout this thesis, we assume approximately omnidirectional antennas in the azimuth plane and consider only application scenarios which require two-dimensional localization. Note that the UWB antennas described in Section 3.3, which were used for the channel measurements, fulfill this assumption. Therefore, it can be assumed that distortions of the location fingerprints caused by antenna patterns are negligible.

4.3.2 Thermal Noise and Time Variations

So far, we have proposed a stochastic model for location fingerprints in a quasi-static environment and for large SNR. We assume that these conditions can be fulfilled during the training phase, such that the parameter sets Θ_m can be estimated for all M regions. However, during the localization phase we might face lower SNR and, moreover, time variations of the propagation environment caused by moving objects like people, machines, or other agents. Thus, we discuss and model the impact of these distortions on the location fingerprints in the following.

The measurement campaign II in [59] is concerned with time variations. The measurement scenario is composed of a TX and an RX at fixed positions and a time varying propagation environment caused by moving people. The authors find based on AIC that the channel taps are best modeled by a non-zero mean proper Gaussian distribution. This implies that there exists a deterministic mean component caused by the static part of the environment like walls, floor and ceiling, and a stochastic zero mean component caused by the movements. This stochastic component is according to the results in [59] best modeled by a circular symmetric Gaussian distribution.

The distribution of thermal noise is commonly modeled as zero mean Gaussian. Thus, we can conclude that a sensible model for the distortions is a circular symmetric Gaussian random vector $\mathbf{n}$. Without any prior information about the statistics of the distortions, the only reasonable covariance matrix of $\mathbf{n}$ is a weighted identity matrix.

[2] This is especially relevant for three-dimensional position location scenarios, because it is very hard to design an antenna which is approximately omnidirectional in all three dimensions, i.e. isotropic.

Consequently, the observation vector $\mathbf{y}$ of a coherent UWB RX during the localization phase is modeled as

$$\mathbf{y} = \mathbf{s}_m + \mathbf{n},$$

with independent random vectors $\mathbf{s}_m$ and $\mathbf{n}$. The conditional PDF of $\mathbf{y}$ assuming the agent is located in region m is given by

$$f_m(\mathbf{y}|H_m) = \frac{1}{\pi^N |\mathbf{\Sigma}_m + \sigma^2 \mathbf{I}|} \exp\left(-\mathbf{y}^H \left(\mathbf{\Sigma}_m + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{y}\right).$$

We define the SNR in dB accounting for all existing distortions as

$$\mathrm{SNR}_m \triangleq 10 \log_{10}\left(\frac{\mathrm{E}\left[\mathbf{s}_m^H \mathbf{s}_m\right]}{2\sigma^2}\right) = 10 \log_{10}\left(\frac{\mathrm{Tr}\left[\mathbf{\Sigma}_m\right]}{2\sigma^2}\right). \tag{4.2}$$

Note that the SNR will be generally different for each region m due to the varying distances $(x,y)_m$ to the RX. However, after the energy normalization step in the post-processing stage, we have assured that $E = \mathrm{Tr}\left[\mathbf{\Sigma}_m\right]$ for all m and, thus, $\mathrm{SNR} = \mathrm{SNR}_m$ for all m. Further, note that this SNR definition is consistent to the widely used E_b/N_0 expression, where E_b denotes the energy per bit of the received signal and $N_0/2$ is the power spectral density of an additive white Gaussian noise process.

4.4 Position Location and Clustering Systems

4.4.1 Training Phase

The parameters of the conditional PDF $f_m(\mathbf{y}|H_m)$ must be estimated during a training phase based on L location fingerprints caused by agents located in region m. This has to be done for all M regions. The ML estimates for the covariance and the pseudo-covariance matrix are given in (4.1). Due to the zero mean assumption, these two estimates are unbiased, i.e. $\mathrm{E}\left[\hat{\mathbf{\Sigma}}_m\right] = \mathbf{\Sigma}_m$ and $\mathrm{E}\left[\hat{\mathbf{Z}}_m\right] = \mathbf{Z}_m$. Due to a finite number of training signals, the estimated covariance matrices are random quantities. It can be shown that $\hat{\mathbf{\Sigma}}_m$ and $\hat{\mathbf{Z}}_m$ are distributed according to a Wishart distribution [66, 67]. Further information on the Wishart distribution is given in Appendix D.

The quality of the covariance matrix estimate is analyzed in the following. The region index m is dropped for notational convenience. The expected value of the squared absolute

estimation error of the (i,j)-th component of the covariance matrix is given by

$$\begin{aligned}
&\mathrm{E}\left[\left|\left[\hat{\boldsymbol{\Sigma}}\right]_{i,j}-\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\right]=\mathrm{E}\left[\left|\left[\hat{\boldsymbol{\Sigma}}\right]_{i,j}\right|^2\right]-\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\\
&=\frac{1}{L^2}\sum_{l=1}^{L}\sum_{k=1}^{L}\mathrm{E}\left[\mathbf{s}[i]^{(l)}\left(\mathbf{s}[j]^{(l)}\right)^*\left(\mathbf{s}[i]^{(k)}\left(\mathbf{s}[j]^{(k)}\right)^*\right)^*\right]-\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\\
&=\frac{1}{L^2}\left(\sum_{l=1}^{L}\sum_{k=1,k\neq l}^{L}\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2+\sum_{l=1}^{L}\mathrm{E}\left[\mathbf{s}[i]^{(l)}\left(\mathbf{s}[j]^{(l)}\right)^*\left(\mathbf{s}[i]^{(l)}\right)^*\mathbf{s}[j]^{(l)}\right]\right)-\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\\
&=\frac{1}{L^2}\left(L\left(L-1\right)\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2+L\left(\left[\boldsymbol{\Sigma}\right]_{i,i}\left[\boldsymbol{\Sigma}\right]_{j,j}+2\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\right)\right)-\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\\
&=\frac{1}{L}\left(\left[\boldsymbol{\Sigma}\right]_{i,i}\left[\boldsymbol{\Sigma}\right]_{j,j}+\left|\left[\boldsymbol{\Sigma}\right]_{i,j}\right|^2\right).
\end{aligned}$$

The expected value of the product of four Gaussian random variables, i.e. $\mathrm{E}\left[\mathbf{s}[i]^{(l)}\left(\mathbf{s}[j]^{(l)}\right)^*\left(\mathbf{s}[i]^{(l)}\right)^*\mathbf{s}[j]^{(l)}\right]$, is derived in Appendix E. We notice that the estimation quality increases linearly with L. Further, we note that N does not influence the estimation quality of a single component of the covariance matrix. However, as N increases for a fixed L the overall structure of the covariance matrix is harder to estimate as the estimation errors of individual components accumulate. As a rule of thumb, L must be approximately as large as the number of free real parameters, such that the stochastic description becomes accurate [32]. In order to guarantee that the estimated covariance matrix is not rank deficient, $L \geq N$ is required. The impact of L on the position location performance is investigated in Section 4.6.3.

The proposed location fingerprinting framework in Chapter 2 enables the combination of training and localization phase via iterative algorithms, which improve the quality of the parameter estimates during the localization phase. This idea is further discussed in Chapter 6.

4.4.2 Localization Phase

The localization phase consists of evaluating (2.6) or (2.8), depending on the chosen figure of merit, with an observation vector $\mathbf{y}$. However, it is computationally expensive to evaluate M MV Gaussian PDFs for (2.8) and even more expensive to evaluate M^2 MV Gaussian PDFs for (2.6). In the latter case we cannot simplify the expression by computing the natural logarithm due to the summation of MV Gaussian PDFs. A possible solution would be a stored lookup table for all MV Gaussian PDFs which in turn requires large storage capacity.

However, for (2.8) taking the natural logarithm gives

$$\begin{aligned}\hat{m} &= \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \ln(\pi_m) + \ln(f_m(\mathbf{y}|H_m)) \\ &= \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \ln(\pi_m) + \ln\left(\frac{1}{\pi^N \left|\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right|} \exp\left(-\mathbf{y}^H \left(\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right)^{-1}\mathbf{y}\right)\right) \\ &= \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \ln(\pi_m) - \ln\left(\left|\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right|\right) - \mathbf{y}^H \left(\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right)^{-1}\mathbf{y}.\end{aligned}$$

From this equation it can be seen that there are many terms which can be computed once and stored. The only operations which have to be performed for each observation are the evaluations of M quadratic forms $\mathbf{y}^H \left(\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right)^{-1}\mathbf{y}$ with pre-computed matrices $\left(\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right)^{-1}$.

Multiple Observations If there are K independent observations $\bar{\mathbf{y}} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_K^T\right]^T$ available (cf. Section 2.4.3), we apply (2.11) and obtain a slightly modified decision rule:

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \ln(\pi_m) - K\ln\left(\left|\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right|\right) - \sum_{k=1}^{K} \mathbf{y}_k^H \left(\hat{\boldsymbol{\Sigma}}_m + \sigma^2\mathbf{I}\right)^{-1}\mathbf{y}_k.$$

PCV model The PCV model assumes a non-zero pseudo-covariance matrix. Therefore, $\mathbf{y}$ is split into real and imaginary part resulting in $\tilde{\mathbf{y}}$, which is a real-valued zero mean MV Gaussian random vector with length $2N$. The corresponding covariance matrix can be computed from $\hat{\boldsymbol{\Sigma}}_m$ and $\hat{\mathbf{Z}}_m$ according to

$$\hat{\tilde{\boldsymbol{\Sigma}}}_m = \frac{1}{2}\begin{bmatrix} \operatorname{Re}\left[\hat{\boldsymbol{\Sigma}}_m + \hat{\mathbf{Z}}_m\right] & \operatorname{Im}\left[-\hat{\boldsymbol{\Sigma}}_m + \hat{\mathbf{Z}}_m\right] \\ \operatorname{Im}\left[\hat{\boldsymbol{\Sigma}}_m + \hat{\mathbf{Z}}_m\right] & \operatorname{Re}\left[\hat{\boldsymbol{\Sigma}}_m - \hat{\mathbf{Z}}_m\right] \end{bmatrix}.$$

4.5 Analytical Pairwise Error Probabilities

For the theoretical analysis and optimization of a position location (or communication) system it is essential to be able to calculate or at least bound the figure of merits analytically. If this is mathematically not feasible one has to resort to computationally expensive Monte Carlo simulations for performance evaluation. Moreover, optimization of a large number of system parameters is infeasible with Monte Carlo simulations. In this section, analytical expressions for the pairwise error probabilities are derived. Applying the bounds in (2.13) results in upper bounds for the average positioning error $\mathcal{D}_e$ and the total probability of error $\mathcal{P}_e$.

In order to calculate the pairwise error probabilities $\mathrm{P}(H_j \to H_i)$ and $\mathrm{P}(H_i \to H_j)$, we consider the binary hypothesis testing problem given by

$$\frac{f_i(\mathbf{y}|H_i)}{f_j(\mathbf{y}|H_j)} \underset{\hat{m}=j}{\overset{\hat{m}=i}{\gtrless}} \delta_{i,j}.$$

Taking the natural logarithm and collecting terms, the likelihood ratio can be simplified to

$$\mathbf{y}^H \left(\left(\mathbf{\Sigma}_j + \sigma^2 \mathbf{I} \right)^{-1} - \left(\mathbf{\Sigma}_i + \sigma^2 \mathbf{I} \right)^{-1} \right) \mathbf{y} \underset{\hat{m}=j}{\overset{\hat{m}=i}{\gtrless}} \bar{\delta}_{i,j}, \tag{4.3}$$

where the threshold $\bar{\delta}_{i,j}$ is given by

$$\bar{\delta}_{i,j} = \ln(\delta_{i,j}) + \ln\left(\frac{\left| \mathbf{\Sigma}_i + \sigma^2 \mathbf{I} \right|}{\left| \mathbf{\Sigma}_j + \sigma^2 \mathbf{I} \right|} \right). \tag{4.4}$$

We note that for this problem the sufficient statistic l is given by

$$l = \mathbf{y}^H \left(\left(\mathbf{\Sigma}_j + \sigma^2 \mathbf{I} \right)^{-1} - \left(\mathbf{\Sigma}_i + \sigma^2 \mathbf{I} \right)^{-1} \right) \mathbf{y} = \mathbf{y}^H \mathbf{A} \mathbf{y}. \tag{4.5}$$

Note that the matrix $\mathbf{A}$ is Hermitian but neither positive semidefinite nor negative semidefinite. Thus, some eigenvalues of $\mathbf{A}$ are positive and some are negative. For the calculation of the pairwise error probabilities we require the PDF of the sufficient statistic given either H_i or H_j. We will derive the PDF of l given H_i which implies that $\mathbf{y}$ is distributed according to a zero mean MV Gaussian PDF with covariance matrix $\mathbf{\Sigma}_i + \sigma^2 \mathbf{I}$. The derivation of the PDF given H_j follows analogously.

The first step is to *whiten* $\mathbf{y}$ by a linear transformation according to $\mathbf{w} = \mathbf{\Lambda}^{-1/2} \mathbf{U}^H \mathbf{y}$, where $\mathbf{\Lambda}$ and $\mathbf{U}$ are the diagonal eigenvalue matrix and the unitary eigenvector matrix of $\mathbf{\Sigma}_i + \sigma^2 \mathbf{I}$, such that we have the decomposition $\mathbf{\Sigma}_i + \sigma^2 \mathbf{I} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$. The covariance matrix of $\mathbf{w}$ is the identity matrix which is shown by

$$\mathrm{E}\left[\mathbf{w}\mathbf{w}^H \right] = \mathbf{\Lambda}^{-1/2} \mathbf{U}^H \mathrm{E}\left[\mathbf{y}\mathbf{y}^H \right] \mathbf{U} \mathbf{\Lambda}^{-1/2} = \mathbf{I}.$$

Consequently, the sufficient statistic l can be written in terms of $\mathbf{w}$ according to

$$l = \mathbf{w}^H \left(\mathbf{\Lambda}^{1/2} \mathbf{U}^H \mathbf{A} \mathbf{U} \mathbf{\Lambda}^{1/2} \right) \mathbf{w} = \mathbf{w}^H \mathbf{B} \mathbf{w}. \tag{4.6}$$

The matrix $\mathbf{B}$ is again Hermitian and can be decomposed into eigenvalue and eigenvector matrices according to $\mathbf{B} = \mathbf{V} \mathbf{\Delta} \mathbf{V}^H$. The unitary transformation $\mathbf{v} = \mathbf{V}^H \mathbf{w}$ renders the covari-

ance matrix of $\mathbf{v}$ as identity and diagonalizes the quadratic form in (4.5). Thus, the sufficient statistic in terms of $\mathbf{v}$ is given by

$$l = \mathbf{v}^H \mathbf{\Delta} \mathbf{v} = \sum_{n=1}^{N} [\mathbf{\Delta}]_{n,n} |v[n]|^2 . \tag{4.7}$$

We notice that all N random variables $|v[n]|^2$ are i.i.d. exponentially distributed with expected value 1. Thus, the PDF of the sufficient statistic l is given by a weighted sum of i.i.d. exponential random variables. Note that the weights are equal to the real valued eigenvalues of $\mathbf{B}$ and depend on the assumed true hypothesis. Further, some eigenvalues of $\mathbf{B}$ can be positive and some can be negative. In order to proceed, we split the sum in (4.7) into two parts collecting the positive and negative weights, respectively. Thus, we have that $l = l^{(+)} - l^{(-)}$ and

$$l^{(+)} = \sum_{n=1}^{N^{(+)}} \phi^{(+)}[n] \left| v^{(+)}[n] \right|^2 \text{ and } l^{(-)} = \sum_{n=1}^{N^{(-)}} \phi^{(-)}[n] \left| v^{(-)}[n] \right|^2 ,$$

where the sets of the corresponding weights are given by $\left\{\phi^{(+)}\right\} = \left\{ [\mathbf{\Delta}]_{n,n} \,|\, [\mathbf{\Delta}]_{n,n} \geq 0 \right\}$ and $\left\{\phi^{(-)}\right\} = \left\{ -[\mathbf{\Delta}]_{n,n} \,|\, [\mathbf{\Delta}]_{n,n} < 0 \right\}$. Note that the two random variables $l^{(+)}$ and $l^{(-)}$ are independent. According to [47], the PDFs of $l^{(+)}$ and $l^{(-)}$ assuming mutually distinct weights are given in closed form by

$$f_{l^{(+)}|H_i}(x) = \begin{cases} \sum_{n=1}^{N^{(+)}} C^{(+)}[n] \exp\left(-\frac{x}{\phi^{(+)}[n]} \right) & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} \text{,where}$$

$$C^{(+)}[n] = \prod_{i=1, i \neq n}^{N^{(+)}} \frac{1}{\phi^{(+)}[n] - \phi^{(+)}[i]},$$

and

$$f_{l^{(-)}|H_i}(y) = \begin{cases} \sum_{n=1}^{N^{(-)}} C^{(-)}[n] \exp\left(-\frac{y}{\phi^{(-)}[n]} \right) & \text{for } y \geq 0 \\ 0 & \text{for } y < 0 \end{cases} \text{,where}$$

$$C^{(-)}[n] = \prod_{i=1, i \neq n}^{N^{(-)}} \frac{1}{\phi^{(-)}[n] - \phi^{(-)}[i]}.$$

The definition of the pairwise error probability is $\mathrm{P}(H_i \to H_j) \triangleq \int_{-\infty}^{\bar{\delta}_{i,j}} f_{l|H_i}(l)\, \mathrm{d}l$. Following the presentation in [65] on page 185 we can readily use $f_{l^{(+)}|H_i}(x)$ and $f_{l^{(-)}|H_i}(y)$ to

calculate $\mathrm{P}(H_i \to H_j)$ for $\bar{\delta}_{i,j} \geq 0$ according to

$$\begin{aligned}\mathrm{P}(H_i \to H_j) &= \int_{y=0}^{\infty} \int_{x=0}^{\bar{\delta}_{i,j}+y} f_{l^{(+)}|H_i}(x)\, f_{l^{(-)}|H_i}(y)\,\mathrm{d}x\mathrm{d}y \\ &= \int_{y=0}^{\infty} F_{l^{(+)}|H_i}\left(\bar{\delta}_{i,j}+y\right) f_{l^{(-)}|H_i}(y)\,\mathrm{d}y,\end{aligned}$$

where $F_{l^{(+)}|H_i}(x)$ is the cummulative distribution function (CDF) of $l^{(+)}$ given H_i. Plugging in the expressions for $f_{l^{(+)}|H_i}(x)$ and $f_{l^{(-)}|H_i}(y)$ and carrying out the integration we get the following closed form expression for the pairwise error probabilities:

$$\mathrm{P}(H_i \to H_j) = \sum_{n=1}^{N^{(+)}} \sum_{k=1}^{N^{(-)}} C^{(+)}[n]\phi^{(+)}[n]C^{(-)}[k]\phi^{(-)}[k] \left(1 - \frac{\phi^{(+)}[n] \exp\left(-\frac{\bar{\delta}_{i,j}}{\phi^{(+)}[n]}\right)}{\phi^{(+)}[n]+\phi^{(-)}[k]}\right).$$

The situation changes slightly for $\bar{\delta}_{i,j} < 0$. In this case we have

$$\begin{aligned}\mathrm{P}(H_i \to H_j) &= \int_{y=-\bar{\delta}_{i,j}}^{\infty} \int_{x=0}^{\bar{\delta}_{i,j}+y} f_{l^{(+)}|H_i}(x)\, f_{l^{(-)}|H_i}(y)\,\mathrm{d}x\mathrm{d}y \\ &= \int_{y=-\bar{\delta}_{i,j}}^{\infty} F_{l^{(+)}|H_i}\left(\bar{\delta}_{i,j}+y\right) f_{l^{(-)}|H_i}(y)\,\mathrm{d}y \\ &= \sum_{n=1}^{N^{(+)}} \sum_{k=1}^{N^{(-)}} C^{(+)}[n]\phi^{(+)}[n]C^{(-)}[k]\left(\phi^{(-)}[k]\right)^2 \frac{\exp\left(\frac{\bar{\delta}_{i,j}}{\phi^{(-)}[k]}\right)}{\phi^{(+)}[n]+\phi^{(-)}[k]}.\end{aligned}$$

The pairwise error probability $\mathrm{P}(H_j \to H_i)$ can be calculated analogously by assuming that H_j is the true hypothesis and by using

$$\mathrm{P}(H_j \to H_i) \triangleq \int_{\bar{\delta}_{i,j}}^{\infty} f_{l|H_j}(l)\,\mathrm{d}l = 1 - \int_{-\infty}^{\bar{\delta}_{i,j}} f_{l|H_j}(l)\,\mathrm{d}l.$$

In order to verify these derivations, the analytical results are compared to numerical pairwise error probabilities obtained from a Monte Carlo simulation with 60000 realizations per hypothesis. The matrices $\mathbf{\Sigma}_i$ and $\mathbf{\Sigma}_j$ are estimated using $L = 620$ measured CIRs from region $i = 4$ and $L = 620$ measured CIRs from region $j = 5$ (cf. Chapter 3). It is assumed

that $\mathbf{\Sigma}_4$ and $\mathbf{\Sigma}_5$ are the true covariance matrices. Moreover, it is assumed that the empirical estimation of the covariance matrices is error-free. Theoretically this would require $L = \infty$ training signals. The size of the observation window is set to $T_s = 20$ ns and the signal bandwidth is $B = 3$ GHz. The length of the location fingerprints results as $N = 60$ samples.

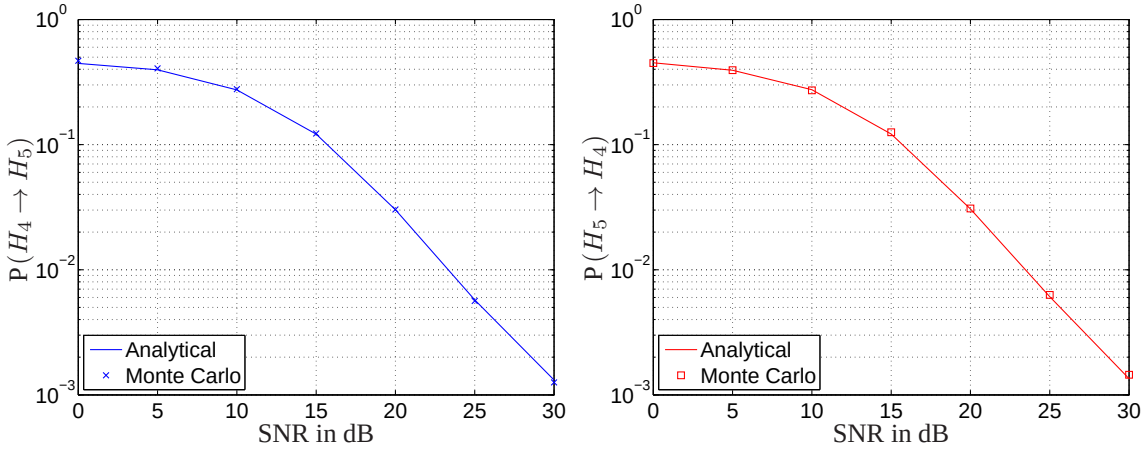

Fig. 4.11: Analytical pairwise error probabilities and numerical Monte Carlo results.

Fig. 4.11 depicts the analytical pairwise error probabilities and the results of the Monte Carlo simulations. It can be seen that the numerical results coincide with the analytical results. In order to obtain analytical upper bounds on $\mathcal{P}_e$ and $\mathcal{D}_e$, these analytical expressions for the pairwise error probabilities have to be substituted to the bounds given in (2.13).

4.6 Theoretical and Experimental Performance Analysis

In this section, the performance of the proposed location fingerprinting system with a coherent RX is analyzed. The figures of merit are $\mathcal{P}_e$ for the clustering problem and $\mathcal{D}_e$ for the position location problem. We consider the ensemble of measured CIRs from all four transmit antennas to receive antenna 3, which resembles a remote positioning scenario. The UWB channel measurement campaign is described in Chapter 3. The total number of measured CIRs per region is 620. Further, the CIRs are normalized, such that the average energy is the same for all regions. This implies that the PL information is not used for position location as already discussed in Section 3.5.

It is assumed that the cluster application requires a performance of $\mathcal{P}_e \approx 10^{-2}$. This would imply that 99 percent of the observations are classified to the correct clusters. Further, it is assumed that the position location application requires a performance of $\mathcal{D}_e \approx 25$ cm. Note that the fundamental positioning uncertainty is $d_{i,i} \approx 18$ cm for all regions.

The following two approaches are pursued for performance analysis:

1. We use measured CIRs and perform Monte Carlo simulations. In order to obtain realistic estimates of the figures of merit we resort to a randomized cross-validation method [68], which is described in Section 4.6.1.
2. We apply the analytical results on the pairwise error probabilities obtained in Section 4.5. Note that these results assume that the channel taps are indeed jointly Gaussian distributed.

There is an inevitable mismatch between experimental and theoretical results due to the stochastic modeling of location fingerprints. A reasonable small mismatch indicates that the stochastic model is accurate and applicable.

4.6.1 Randomized Cross-Validation Method

The first step is to partition the set of measured CIRs per region into L CIRs for parameter estimation (training phase) and the remaining CIRs for performance evaluation (localization phase). The choice of the training signals is random. If the impact of distortions (SNR) is analyzed, then 1000 realizations of $\mathbf{n}$ are generated for each partition. Using Monte Carlo simulations the figures of merit for each partition are calculated. The last step is to repeat the partitioning a number of times (here 100 times) and to average the obtained results.

4.6.2 Default System Parameters

Important system parameters are the number of training signals L per region, the SNR, the one-sided bandwidth of the passband signals denoted by B, the observation window size or symbol period denoted by T_s, and the number of regions M. Note that T_s and B determine the length of the location fingerprints according to $N = T_s B$ due to sampling at Nyquist rate with $f_s = B$. The default system parameter values for the following simulation results are summarized as follows.

- Default model is the COV model with $\Theta_m = \{\mathbf{\Sigma}_m\}$.
- Default number of observations per agent is $K = 1$, i.e. $\bar{\mathbf{y}} = \mathbf{y}$.
- Default number of training signals is $L = 400$. This leaves 220 test signals per region.
- Default SNR is the measurement SNR denoted by SNR_{mea} in Chapter 3. This corresponds to adding no distortion signals $\mathbf{n}$, or equivalently, setting σ^2 to zero.

- Default bandwidth is $B = 3$ GHz. This is the largest possible signal bandwidth determined by the hardware constraints of the measurement equipment.
- Default observation window size is $T_s = 30$ ns. Together with the default bandwidth this gives the default length of the location fingerprints with $N = 90$ samples. Note that 30 ns corresponds roughly to the largest RMS delay spread among all CIRs from all regions. The RMS delay spreads of the measured CIRs are reported in [69]. However, the maximum excess delays of the CIRs are in general larger than 30 ns and can become as large as 80 ns.
- Default number of regions is $M = 22$, which corresponds to the whole measurement set.
- Default a priori probabilities are $\pi_m = 1/M$.
- Default reference sample is $n_{\text{maxabs}} = 8$.

4.6.3 Number of Training Signals

In this section, the impact of L on the performance is investigated. We seek to answer the question, how many training signals are required such that the target performance can be achieved. Furthermore, interactions with other system parameters are investigated.

4.6.3.1 Analytical Results

An immediate application of the analytical results from the Section 4.5 is the analysis of the impact of L on the performance. The advantage of the analytical approach is that L is not limited to the total number of measured CIRs, which implies that the asymptotic behavior can be studied.

First, we analyze the binary hypothesis testing problem with H_4 and H_5, for which we know that the pairwise error probabilities are equal to the conditional error probabilities. We calculate $\mathbf{\Sigma}_4$ and $\mathbf{\Sigma}_5$ using all 620 measured CIRs from each region and assume that these are the true covariance matrices. In contrast to the results in Fig. 4.11, finite values for L are chosen. This implies that the matrix $\mathbf{A}$ in (4.5) and the threshold $\bar{\delta}_{4,5}$ in (4.4) are given by

$$\mathbf{A} = \left(\hat{\mathbf{\Sigma}}_5 + \sigma^2\mathbf{I}\right)^{-1} - \left(\hat{\mathbf{\Sigma}}_4 + \sigma^2\mathbf{I}\right)^{-1} \text{ and } \bar{\delta}_{4,5} = \ln\left(\delta_{4,5}\right) + \ln\left(\frac{\left|\hat{\mathbf{\Sigma}}_4 + \sigma^2\mathbf{I}\right|}{\left|\hat{\mathbf{\Sigma}}_5 + \sigma^2\mathbf{I}\right|}\right),$$

where $\hat{\mathbf{\Sigma}}_4$ and $\hat{\mathbf{\Sigma}}_5$ are estimates of $\mathbf{\Sigma}_4$ and $\mathbf{\Sigma}_5$ in accordance to (4.1) based on L training signals. We utilize the fact that $\hat{\mathbf{\Sigma}}_4$ and $\hat{\mathbf{\Sigma}}_5$ are distributed according to a Wishart distribution and generate 1000 realizations of them for each value of L. Note that L is a parameter of the Wishart distribution (cf. Appendix D). For increasing L the realizations are more and more similar to the true covariance matrices $\mathbf{\Sigma}_4$ and $\mathbf{\Sigma}_5$. For each realization of $\hat{\mathbf{\Sigma}}_4$ and $\hat{\mathbf{\Sigma}}_5$ with a given L the corresponding analytical pairwise error probabilities are calculated and these results are averaged for final presentation in Fig. 4.12.

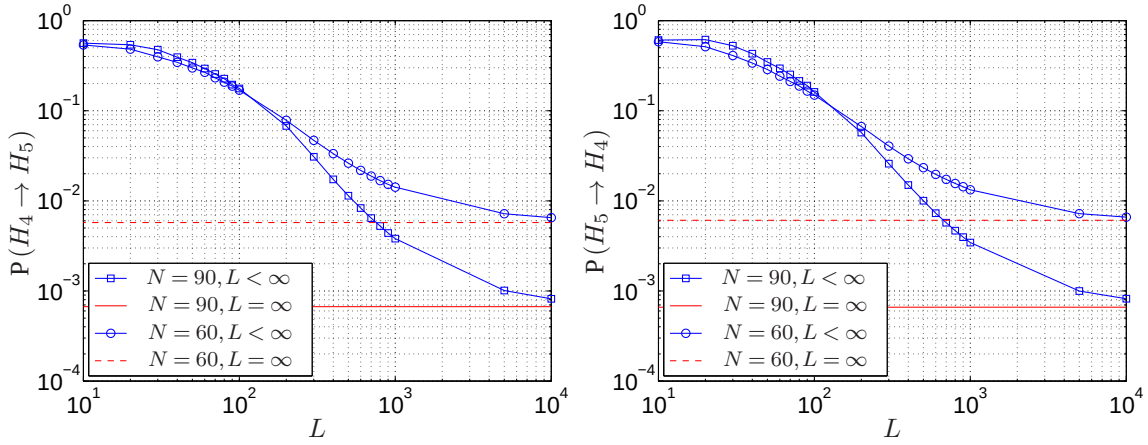

Fig. 4.12: Analytical pairwise error probabilities as function of L at SNR $= 25$ dB for the COV model.

Fig. 4.12 depicts the analytical pairwise error probabilities as function of $L < \infty$ and for perfect estimates, i.e. $L = \infty$. The curves for $N = 90$ samples are obtained with the default system parameters except for SNR $= 25$ dB. The curves for $N = 60$ are obtained by reducing T_s to 20 ns. These results allow us to get some insight into the performance loss due to a finite number of training signals. At first, it can be seen that the error probability for $L < 100$ is larger than 0.1, which is not acceptable for many applications. However, 100 training signals per region are already a lot. It can be concluded that the proposed location fingerprinting approach with a coherent RX requires a clever procedure to perform parameter estimation with few training signals, while still achieving small error probabilities. Chapters 6 and 7 propose means to achieve this goal.

Further, we can compare the error probabilities in Fig. 4.12 for two different N. Decreasing N implies that the number of degrees of freedom of the location fingerprints is decreased. Consequently, the location fingerprints lose in differentiability and the performance for perfect estimates ($L = \infty$) decreases. However, the situation changes for $L < \infty$. For $L \approx 100$ the two performance curves for $N = 60$ and $N = 90$ intersect. If less than 100 training signals are available, then decreasing N will reduce the error probabilities. It can be concluded

from these results that decreasing N improves the performance for small L. The drawback is, however, that small N implies already bad performance. In Fig. 4.12, N is decreased by reducing T_s from 30 ns to 20 ns at fixed B. Similar results are obtained when B is reduced at fixed T_s.

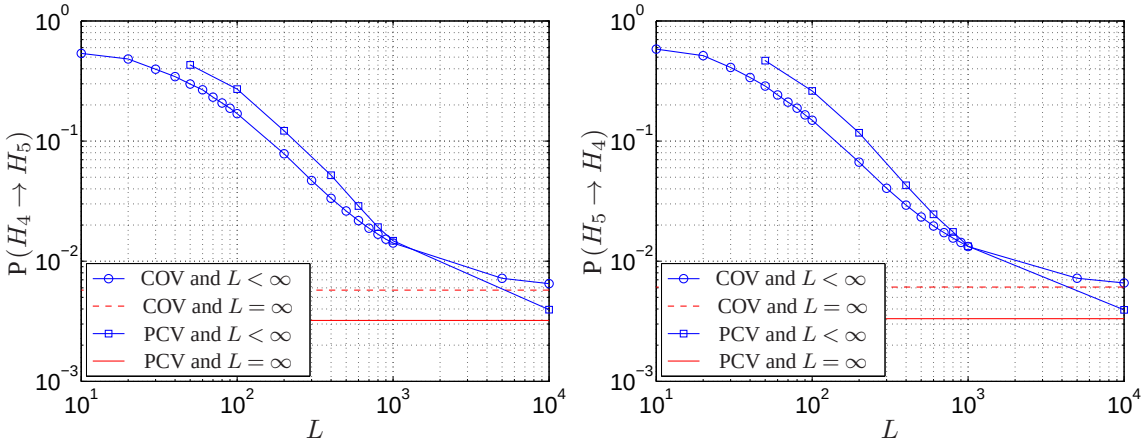

Fig. 4.13: Analytical pairwise error probabilities as function of L at SNR $= 25$ dB with $T_s = 20$ ns for the COV and the PCV model.

A further application of the analytical results is the analysis of the performance depending on L for different modeling assumptions. In the following, we compare the performance results of the COV model and the more complex PCV model accounting also for the pseudo-covariance matrix. Note that the number of free real parameters increases from N^2 to $2N^2 + N$ in this case. Further note that the analytical results derived in Section 4.5 cannot be applied to the PCV model, because the decision problem in (4.3) changes from a complex Gaussian problem with circular symmetric Gaussian random vectors to a general Gaussian problem [44] with real valued random vectors. In [44], it is shown that there does not exist a closed form expression for the pairwise error probabilities of the general Gaussian problem. For performance analysis one can resort to upper bounds on the pairwise error probabilities like the Chernoff bound or approximations, which are given in closed form. The exact pairwise error probabilities can be calculated only numerically.

In Section 5.2 we derive a numerical algorithm to compute the PDFs for $l^{(+)}$ and $l^{(-)}$ for real Gaussian random variables $v^{(+)}[n]$ and $v^{(-)}[n]$. Having these PDFs we can compute the corresponding pairwise error probabilities for the general Gaussian problem numerically. Fig. 4.13 depicts the pairwise error probabilities as function of L for different modeling assumptions. Note that the results for the COV model are the same as in Fig. 4.12 for $T_s = 20$ ns. The conclusion from these curves is that $L > 1000$ is required in order to be

able to benefit from the higher complexity of the PCV model. Based on this result, it can be concluded that the PCV model is only of minor applicability and importance, because a very large number of training signals is required.

4.6.3.2 Experimental Results

In this section, we use measured CIRs to study the impact of L on $\mathcal{P}_e$ and $\mathcal{D}_e$. The following performance results are obtained for the default system parameters considering all $M = 22$ regions.

For $L < N$ the empirical covariance matrices $\hat{\mathbf{\Sigma}}_m$ are rank deficient, i.e. singular[3]. For singular $\hat{\mathbf{\Sigma}}_m$ and no artificial distortion signal, i.e. $\sigma^2 = 0$, the decision rules in (2.6) and (2.8) cannot be evaluated. In order to cope with this problem, we propose to apply a perturbation method: A weighted identity matrix $\sigma_p^2\mathbf{I}$ is added to the singular $\hat{\mathbf{\Sigma}}_m$. This ensures full rank of the overall covariance matrix. For the performance results in Fig. 4.14, σ_p^2 is chosen such that an SNR of 30 dB is emulated. However, note that no artificial distortion signals are added to the test signals during the localization phase.

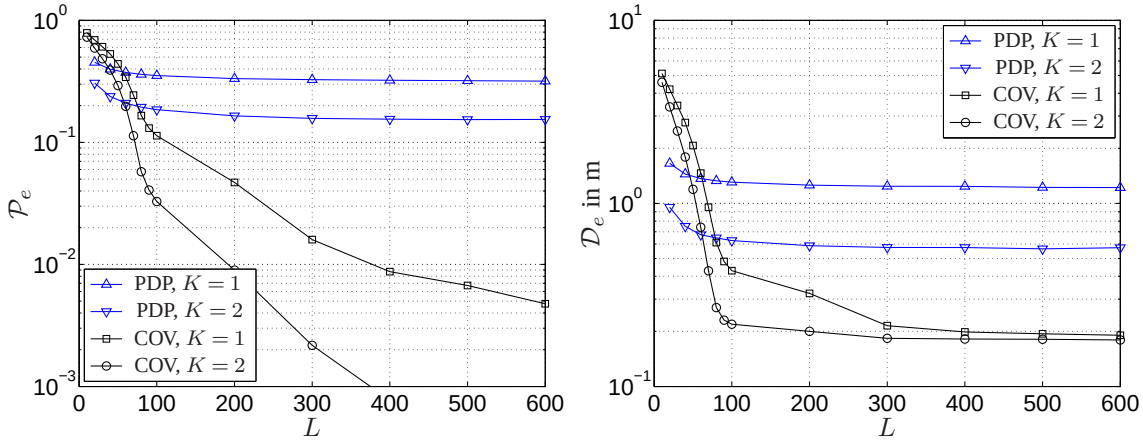

Fig. 4.14: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L for default system parameters with $K = 1$ and $K = 2$ observations per agent.

Fig. 4.14 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of L for measured CIRs. The PDP model and the COV model are considered for $K = 1$ and $K = 2$ available observations per agent. The perturbation method is applied for $L \leq 100$ and it can be observed that reasonable performance can be achieved with the COV model although the empirical covariance matrices are singular. However, further decreasing L degrades the performance of the COV model drastically.

[3]Note that $L \geq N$ is a necessary but not sufficient condition for regular covariance matrices.

For $L < 60$, it is better to ignore all off-diagonal elements of the covariance matrices and apply the PDP model instead of the COV model.

A significant performance gain for both models can be observed, when $K = 2$ observations per agent are exploited. Especially for the COV model, almost error-free performance is achieved for $L > 300$. For the performance results in Fig. 4.14 the observation pairs per agent are chosen randomly from the set of test signals. Further simulations show a continuous performance improvement as K is increased.

4.6.4 Conditional and Pairwise Error Probabilities

In this section, all numerically obtained conditional error probabilities $\mathrm{P}(H_i|H_j)$ using measured CIRs and all analytical pairwise error probabilities $\mathrm{P}(H_j \to H_i)$ are investigated. Note that $\mathrm{P}(H_j \to H_i)$ are calculated for finite L and that the default system parameters are applied.

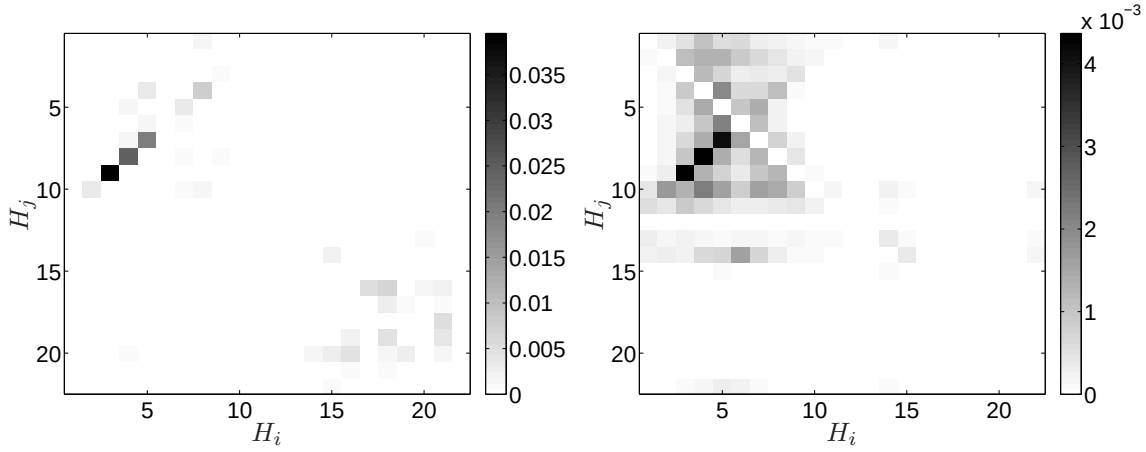

Fig. 4.15: $\mathrm{P}(H_i|H_j)$ for measured CIRs (left) and analytical $\mathrm{P}(H_j \to H_i)$ (right) for the COV model.

Fig. 4.15 depicts all $\mathrm{P}(H_i|H_j)$ for measured CIRs and $\mathrm{P}(H_j \to H_i)$ for the COV model. It can be observed that the values for $\mathrm{P}(H_j \to H_i)$ are generally smaller than the values for $\mathrm{P}(H_i|H_j)$, because the measured CIRs are not exactly Gaussian distributed and there exists an inevitable modeling mismatch. Nevertheless, it can be seen that in both cases the largest error probabilities are $\{\mathrm{P}(H_3|H_9), \mathrm{P}(H_9 \to H_3)\}$ followed by $\{\mathrm{P}(H_4|H_8), \mathrm{P}(H_8 \to H_4)\}$, and $\{\mathrm{P}(H_5|H_7), \mathrm{P}(H_7 \to H_5)\}$. These results emphasize the importance of the analytical pairwise error probabilities. If a large pairwise error probability is predicted then also the conditional error probability for measured data will be high. With this knowledge one has the

opportunity to optimize system parameters, region dimensions and region positions. If the pairwise error probabilities are very large, it is conceivable to merge the two clusters/regions into a single cluster/region.

Another interesting observation from Fig. 4.15 is that most classification errors happen among nearby regions. This holds for experimental as well as analytical results. For example there are almost no classification errors between the region set $\{1, 2, \ldots, 11\}$ and the region set $\{12, 13, \ldots, 22\}$. According to these results, it can be expected that the proposed location fingerprinting approach is scalable which means that small $\mathcal{P}_e$ and $\mathcal{D}_e$ can be sustained, while M and the surveillance area are increased.

In order to calculate $\mathcal{P}_e$ analytically one has to resort to bounds (2.13) as discussed in Section 2.4.4. Note that this provides only an approximation (not an upper bound) to the empirical $\mathcal{P}_e$ for measured CIRs due to the modeling mismatch.

4.6.5 Distortion Analysis

Let us recall the SNR definition (4.2) after CIR energy normalization as

$$\text{SNR} = 10 \log_{10} \left(\frac{E}{2\sigma^2} \right),$$

where E denotes the average energy of CIRs and σ^2 the variance of each sample of the distortion vector $\mathbf{n}$. By changing σ^2 we can emulate different SNR operating points. Note that $\sigma^2 = 0$ corresponds to the default SNR case.

We assume a high SNR for the training phase. Thus, the temporal alignment and parameter estimation during the training phase is done at SNR_{mea}. However, the temporal alignment of location fingerprints observed during the localization phase is generally influenced by the distortion signal. Therefore, we present two results:

1. The temporal alignment is done with distorted location fingerprints. The corresponding curve in Fig. 4.16 is denoted by *with align noise*.
2. The temporal alignment is done with distortion-free location fingerprints. The corresponding curves in Fig. 4.16 are denoted by *w/o align noise*.

The second result shows the best achievable performance. This could be obtained with more sophisticated alignment algorithms, which perform better than the maximum absolute value alignment described in Section 3.6.2. An extensive analysis of such alignment algorithms is, however, beyond the scope of this thesis and recommended for future work.

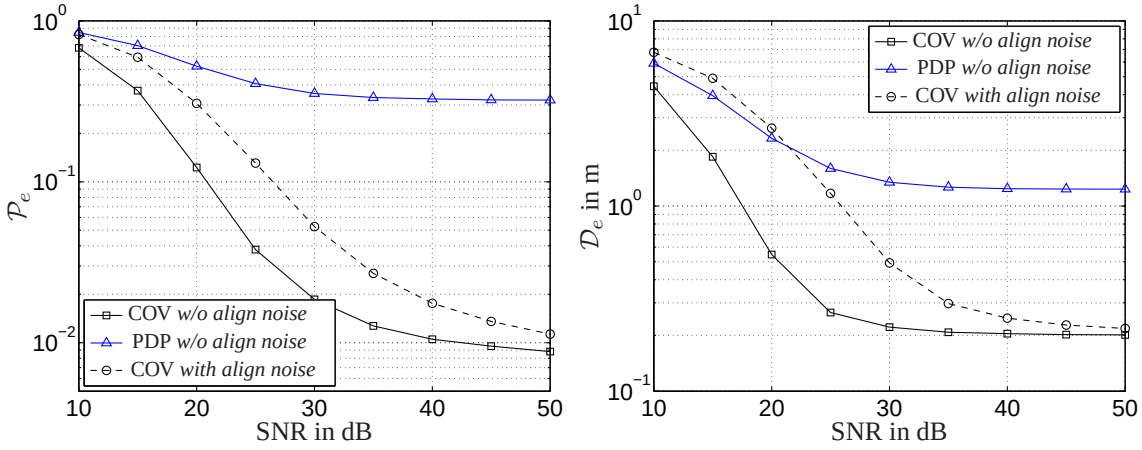

Fig. 4.16: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of SNR for default system parameters and measured CIRs.

Fig. 4.16 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of the SNR for measured CIRs and for the COV and PDP model. In order to reach the target performance for distortion-free alignment and the COV model, an SNR of around 30 dB is required. In case of noisy alignment the required SNR increases to around 40 dB.

Multiple Observations per Agent Multiple observations per agent can be exploited in order to improve the performance at low SNR. The mechanics are exactly the same as discussed in Section 4.6.3.2. There is, however, a major difference: The observations need not be collected from different transmitter positions, because the main error sources at low SNR are random distortions, which change independently from observation to observations. Thus, also a static agent with only a single antenna can produce exploitable multiple observations.

4.6.6 Signal Bandwidth

In this section, the impact of the signal bandwidth on the performance is investigated. The largest bandwidth supported by the measurement equipment is 3 GHz in the frequency band from 3 GHz to 6 GHz. In order to reduce the bandwidth an ideal bandpass filter centered around 4.5 GHz with the desired bandwidth B is applied to the measured CIRs. The temporal alignment is performed after bandwidth reduction. Note that the reference sample is adapted to the bandwidth (sampling rate) such that the reference time instant is kept roughly the same for all bandwidths. For example, consider the default bandwidth of $B = 3$ GHz with the default reference sample of $n_{\text{maxabs}} = 8$, which implies a reference time instant of $8/3$ ns.

Now, let B be 1 GHz. In order to achieve roughly the same reference time instant, we have to set n_{maxabs} to 3.

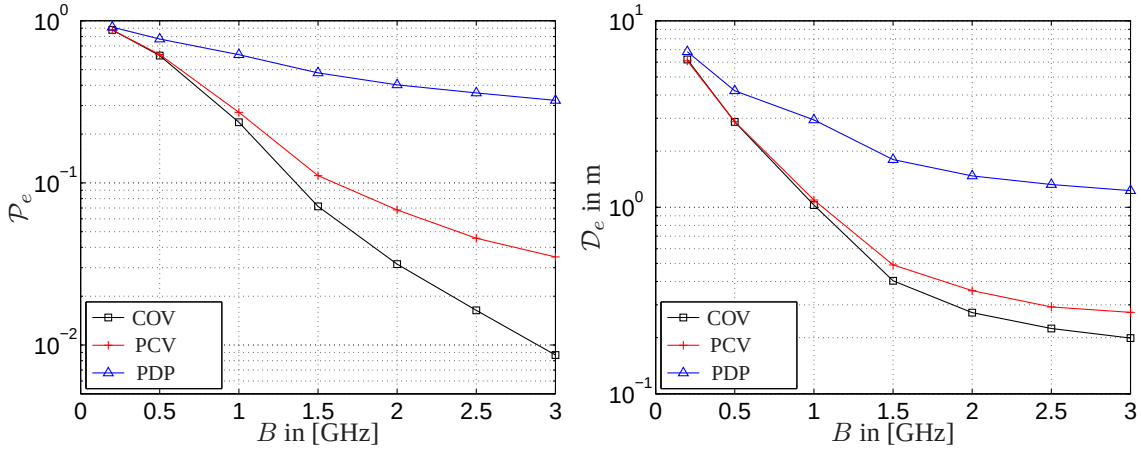

Fig. 4.17: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of B for default system parameters and measured channels.

Fig. 4.17 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of B for all three considered models and measured CIRs. The remaining system parameters are set to their default values. A constant performance improvement with increasing B can be observed for all models. The decreasing performance gap between the COV and the PCV model for decreasing B is explained by the different number of free parameters of these models. For small N even $L = 400$ training signals are sufficient to obtain accurate parameter estimates for the PCV model. The performance of the PDP model is plotted for the sake of completeness.

Reducing B has two effects:

1. The temporal resolution of the multipath propagation channel is reduced and an increasing number of multipath components contribute to the individual channel taps. As a consequence, the statistics of the location fingerprints from different regions become more similar, which is one reason for the performance degradation as B decreases.

2. The length of the location fingerprints is reduced for decreasing B. This implies (i) a reduced number of degrees of freedom and (ii) fewer free model parameters. Less degrees of freedom cause a performance degradation, whereas less free parameters imply that less training signals are required for accurate estimation of covariance matrices. For example, let $B = 2$ GHz and $T_s = 30$ ns, which implies $N = 60$. Simulation results show that L can be reduced to 300, while still achieving a performance of $\mathcal{P}_e \approx 0.03$. In order to achieve the same $\mathcal{P}_e$ in Fig. 4.17 with $N = 90$, $L = 400$ training signals are required.

It can be concluded that a larger signal bandwidth yields better performance of the proposed location fingerprinting system provided that enough training signals are available. However, it is important to keep in mind that increasing B increases the number of free model parameters, which in turn implies that more training signals for parameter estimation are required.

4.6.7 Observation Window Size

In this section, the impact of the observation window size on the performance is investigated. We vary T_s from 5 ns up to 50 ns, which implies an N from 15 to 150 with the default bandwidth of $B = 3$ GHz. Similar to increasing B, increasing T_s has two effects:

1. Larger T_s implies larger N, with the same implications as above.
2. Increasing T_s implies that also multipath components arriving later at the receive antenna are observed and can be exploited for localization. However, the impact of later multipath components on the position location performance vanishes due to the increased propagation distance and the corresponding larger PL. This effect is emphasized in case of low SNR, because all multipath components with an amplitude below the noise floor have only marginal impact on the performance.

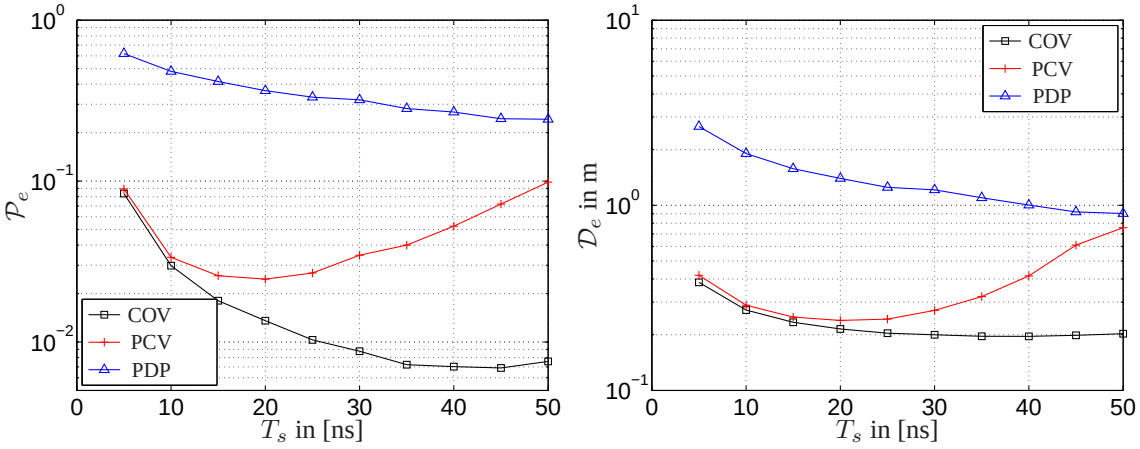

Fig. 4.18: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of T_s for default system parameters and measured CIRs.

Fig. 4.18 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of T_s for all three considered models and measured CIRs. The remaining system parameters are set to their default values. For the three different models different optimal values for T_s can be observed. Since the PDP model has only N

free real parameters, the largest T_s of 50 ns is optimal. On the contrary the PCV model has $2N^2 + N$ free real parameters. Thus, with $L = 400$ training signals the optimal value for T_s is 20 ns. For the default COV model we find that $T_s = 40$ ns achieves best performance, although the performance difference to $T_s = 35$ ns and $T_s = 45$ ns is only marginal. Note that these results are obtained for the measurement SNR. In general, the optimal T_s also depends on the SNR. We further observe that the default value of T_s, which is based on the largest RMS delay spread, is a sensible choice providing nearly optimal performance.

4.7 Summary and Conclusions

This chapter presents a thorough study of a location fingerprinting system employing a *single* anchor, which is able to estimate CIRs with large bandwidths. Performance results based on measured CIRs in a dense multipath environment with non-LoS situations show the high potential of the proposed location fingerprinting method. It is feasible to achieve a total probability of error of less than 10^{-2} and an average positioning error of less than 20 cm for the considered surveillance area, which is quantized into $M = 22$ regions. As for all location fingerprinting systems, the complexity of the training phase dominates the overall complexity. Chapters 6 and 7 propose means to reduce the complexity of the training phase. However, before techniques to improve the efficiency of the training phase are investigated, we turn our attention towards reduced RX (anchor) complexity in the next chapter.

Chapter 5

Location Fingerprinting with a Generalized Energy Detection Receiver

In this chapter[1], we aim at location fingerprinting with a low complexity UWB RX structure. The fundamental difference to a coherent RX considered in Chapter 4 is that it cannot estimate the CIR but rather performs nonlinear analog signal processing to obtain energy samples. The corresponding system model of the generalized ED RX is described in Section 5.1.

The main motivation for these investigations is anchor complexity. This is especially important for the self-positioning scenario. The coherent RX considered in Chapter 4 has to perform the transformation into equivalent baseband and to sample this signal with Nyquist rate. For large signal bandwidths as considered in UWB systems these operations require very powerful and power-hungry hardware components. This high RX complexity would be feasible if the anchor is static and connected to the power supply. However, for an ad-hoc network it is beneficial, when all nodes regardless of complexity are able to perform location fingerprinting. Furthermore, it is envisioned that in such an ad-hoc network there will be mainly low complexity devices, which are battery powered.

We first describe the system model of the considered low complexity UWB RX and present the corresponding choice of the location fingerprints. Then we tackle the problem of their stochastic description. Similar to Chapter 4, we conclude this chapter with a performance analysis.

[1] Parts of this chapter have been published in [43].

5.1 Generalized Energy Detection Receiver

The system model of a generalized ED RX is depicted in Fig. 5.1.

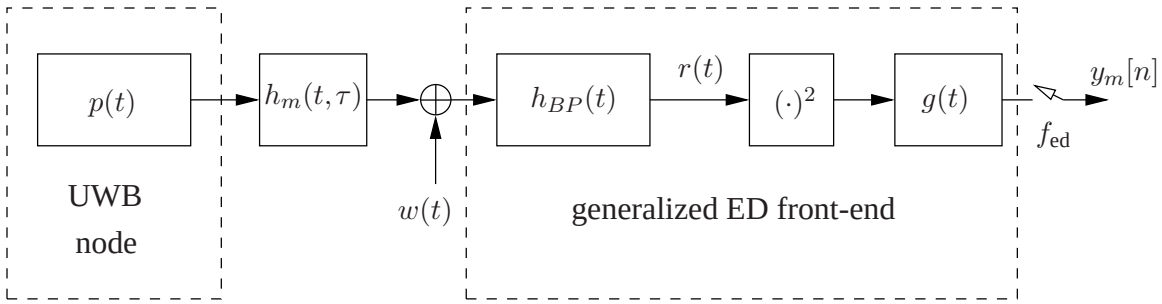

Fig. 5.1: Communication system with a generalized ED RX.

The TX employs either pulse position modulation or on-off keying and the RX uses an analog ED front-end with a generalized integration filter $g(t)$. The conventional ED front-end mainly investigated in literature considers only rectangular integration filters. However, it is shown in [5, 70] that $g(t)$ can be optimized based on a priori knowledge of channel and interference statistics such that significant gains in bit error probability are achievable. Therefore, we try to be as general as possible in this system model and allow for arbitrary integration filters $g(t)$.

The unmodulated input signal to the squaring device $r(t)$ within one symbol period $[0, T_s]$ is given by

$$r(t) = \int_{\tau} h_m(t,\tau) \underbrace{q(t-\tau)}_{p(t)*h_{BP}(t)} \mathrm{d}\tau + \int_{\tau} w(t-\tau) h_{BP}(\tau) \mathrm{d}\tau, \tag{5.1}$$

where $q(t)$ is the equivalent transmit pulse after convolution of $p(t)$ with the front-end band-pass filter $h_{BP}(t)$ and $h_m(t,\tau)$ is the generally time-varying CIR from an agent located in region m. It is assumed that $h_m(t,\tau)$ does not change in t during one symbol period (observation window). Since T_s is in the order of tens of nanoseconds, this assumption is well justified. This assumptions simplifies (5.1) to the standard input-output relation of linear time-invariant systems. The stochastic process $w(t)$ accounts for possible distortions as described in Section 4.3. We stick to the SNR definition in (4.2) in order to be able to compare the performance of both approaches. The SNR has been defined according to

$$\text{SNR} = 10 \log_{10} \left(\frac{E_b}{N_0} \right),$$

where E_b is the average energy of continuous time CIRs and $N_0/2$ is the power spectral density of the white Gaussian process $w(t)$.

It is assumed, that the generalized ED RX is synchronized to the symbol timing of the TX. For example, a low complexity solution for symbol synchronization with generalized ED RXs is proposed in [71]. For the simulation results in this chapter we apply the same maximum absolute value alignment strategy as in Chapter 4, which is described in Section 3.6. The temporal alignment is performed before the squaring device using the signal $r(t)$ at SNR_{mea}.

The signal after the generalized integration filter $g(t)$ is sampled at a rate of f_{ed} producing the energy samples $y_m[n]$. The sampling rate f_{ed} and T_s specify the number of observed energy samples $N = T_s f_{\text{ed}}$. For $N = 1$ the only available information for location fingerprinting is the RSS. This information is far too coarse for reliable localization with only one anchor. However, by increasing f_{ed} for a fixed T_s a higher spatio-temporal resolution of the UWB channel is obtained, and it is expected that reasonable localization accuracy can be achieved with a single anchor. The influence of f_{ed} on the performance is investigated in Section 5.7.

5.1.1 Choice of the Location Fingerprints

The location fingerprint $\mathbf{y}$ of length N for the generalized ED RX is obtained by stacking all energy samples $y[n]$ for $n = 1, 2, \ldots, N$ into one vector. As discussed in Chapter 2 we require the conditional PDFs $f_m(\mathbf{y}|H_m)$ for $m = 1, 2, \ldots, M$ for evaluation of the decision rules in (2.6) or (2.8). The stochastic modeling of the energy samples $y[n]$ for UWB propagation channels is not treated in literature to the best of the author's knowledge. An accurate stochastic description is, however, essential for any kind of ML or MAP operation on the energy samples. There exist two possible approaches to this modeling problem:

- $\mathcal{A}$ Take existing and accepted stochastic models for the UWB propagation channel as discussed in Section 4.2 and try to come up with a mathematical derivation of the distribution of the energy samples $y[n]$.
- $\mathcal{B}$ Apply a model selection criterion like AIC directly to measured energy samples $y[n]$.

Section 5.2 is concerned with approach $\mathcal{A}$ and provides a numerical algorithm to calculate the theoretic distribution of the energy samples based on the regional channel model proposed in Section 4.2. Section 5.3 follows approach $\mathcal{B}$ and applies AIC to measured energy samples, which are obtained by processing measured CIRs according to the generalized ED RX front-end in Fig. 5.1.

5.1.2 System Parameters

If not mentioned otherwise, the system parameters are fixed as follows for the calculation of computer simulation results. The transmit pulse $p(t)$ is assumed to be flat in the desired frequency band. The front-end bandpass filter $h_{BP}(t)$ is rectangular in frequency domain and has 3 GHz bandwidth from 3 to 6 GHz. The sampling frequency of the generalized ED RX is equal to $f_{\text{ed}} = 1$ GHz. A symbol period of $T_s = 30$ ns is considered, which leads to $N = 30$. Continuous-time signals are represented in the computer by samples obtained with $f_s = 40$ GHz sampling.

5.2 Exact Distribution of Energy Samples for a Gaussian Channel Model

In the following, the exact PDFs of the energy samples $y_m[n]$ are calculated under the assumption that the propagation channel $h_m(t, \tau)$ and the distortion signal $w(t)$ are realizations of real Gaussian processes. The specific parameters of these Gaussian processes are not relevant for these theoretical derivations. There are three reasons to consider a Gaussian channel model:

1. There exist analyses of measurement data (cf. Section 4.2), which support the Gaussian assumption also for UWB propagation channels. Furthermore, it is shown in Chapter 4 that the regional channel model applying jointly Gaussian channel taps provides an accurate stochastic description of the equivalent baseband CIRs and excellent location fingerprinting performance.
2. This derivation enables performance prediction based on first and second order channel statistics (mean and covariance functions) without requiring extensive channel measurements. Mean and covariance functions fully determine the Gaussian process and can be obtained via ray tracing tools or deterministic channel models.
3. For all other stochastic models the mathematical derivation will most likely fail due to intractability.

The interval of the real and band-limited signal $r(t)$ causing sample $y_m[n]$ is represented by the vector $\mathbf{r}_{n,m}$ of length $J = \frac{f_s}{f_{\text{ed}}}$, which is obtained by sampling $r(t)$ with f_s. Note that $J = 40$ for the default system parameters. Following the Gaussian assumption it is proposed that $\mathbf{r}_{n,m}$ is a realization of a MV Gaussian distribution with mean vector $\mathbf{v}_{n,m}$ and

covariance matrix $\mathbf{S}_{n,m}$. In the following derivations the sample index n and the region index m are dropped for notational convenience.

The generalized integration filter $g(t)$ is assumed to be time-limited from 0 to $1/f_{\text{ed}}$. Additionally assuming that $g(t)$ is a real and positive function, the energy sample y can be written as positive definite Gaussian quadratic form according to

$$\begin{aligned} y_m[n] &= \mathbf{r}_{n,m}^T \mathbf{G} \mathbf{r}_{n,m} \\ &\downarrow \ldots \text{ dropping sample index } n \text{ and region index } m. \\ y &= \mathbf{r}^T \mathbf{G} \mathbf{r}. \end{aligned} \tag{5.2}$$

The PDF of the random energy sample y is denoted by $f(y)$. The entries in the diagonal matrix $\mathbf{G}$ are obtained by sampling $g(-t)$ for $-1/f_{\text{ed}} \leq t \leq 0$ with f_s. As derived in [72], the quadratic form in (5.2) can be diagonalized, such that $f(y)$ remains unchanged, according to

$$\mathbf{r}^T \mathbf{G} \mathbf{r} \overset{\text{PDF}}{\Longleftrightarrow} \sum_{j=1}^{J} \lambda_j z_j^2,$$

where λ_j denotes the j-th eigenvalue of $\mathbf{S}^{1/2}\mathbf{G}\mathbf{S}^{1/2}$. The random variables z_j are independent and Gaussian distributed with variance one and their means γ_j are given by

$$\gamma_j = \left(\mathbf{S}^{-1/2}\mathbf{u}_j\right)^T \mathbf{v},$$

where $\mathbf{u}_j$ denotes the j-th eigenvector of $\mathbf{S}^{1/2}\mathbf{G}\mathbf{S}^{1/2}$.

If all λ_j are equal and there exists at least one $\gamma_j \neq 0$, then y is distributed according to a Noncentral Chi-square distribution[2]. Such a PDF is obtained, when the conditional PDF of y given a channel realization is sought (e.g. [73]). Note that this is only valid for rectangular integration filters, i.e. $g(t) =$ constant for $0 \leq t \leq 1/f_{\text{ed}}$ and zero otherwise.

For arbitrary λ_j, a closed form expression for $f(y)$ does not exist. Grenander, Pollak, and Slepian present an efficient and numerically stable approach to calculate $f(y)$ for zero mean random variables z_j ($\gamma_j = 0\ \forall j$) in [74]. This method is extended to the general case of nonzero means in the following. The algorithm is based on Fourier's inversion of the characteristic function of y, which is for nonzero means given by [75]

$$\Psi(t) = \prod_{j=1}^{J} (1 - 2t\lambda_j)^{-1/2} \exp\left(\frac{t\gamma_j^2\lambda_j}{1 - 2t\lambda_j}\right) \text{ for } t = \imath\omega. \tag{5.3}$$

[2] If all $\gamma_j = 0$, then y is distributed according to a Chi-square distribution.

The logarithmic derivative of (5.3) is given by

$$\frac{\partial \ln\left(\Psi\left(t\right)\right)}{\partial t} = \frac{\partial \Psi\left(t\right)}{\partial t}\frac{1}{\Psi\left(t\right)} = \imath \sum_{j=1}^{J} \frac{\lambda_j}{1-2t\lambda_j} + \frac{\gamma_j^2 \lambda_j}{\left(1-2t\lambda_j\right)^2}. \tag{5.4}$$

Applying the inverse Fourier transform to (5.4) results in an integral equation for $f\left(y\right)$ given by

$$y f\left(y\right) = \int_0^y f\left(y-\tau\right) a\left(\tau\right) \mathrm{d}\tau,$$

where

$$a\left(\tau\right) = \frac{1}{2}\left(\sum_{j=1}^{J} \exp\left(-\frac{\tau}{2\lambda_j}\right)\left(1+\frac{\gamma_j^2 \tau}{2\lambda_j}\right)\right).$$

This integral equation can be numerically solved with the trapezoidal integration rule according to [74], which gives

$$f\left(k\Delta\right) = \frac{1}{k-\frac{1}{2}a(0)} \sum_{l=1}^{k-1} f(l\Delta) a((k-l)\Delta) \text{ for } k = 2,3,\ldots,$$

where Δ is the mesh size. This recursive equation requires nonzero initial values, which are obtained by using an analytic approximation of $f\left(y\right)$ around $y = 0$ based on Taylor series expansion of (5.3) and inverse Fourier transform. The approximation is given by

$$f\left(y\right) = \frac{y^{(J/2)-1} \exp\left(\frac{1}{2}\sum_{j=1}^{J}\gamma_j^2\right)}{2^{J/2}\Gamma\left(J/2\right)\prod_{j=1}^{J}\lambda_j^{1/2}}\left(1+\frac{y}{2J}\sum_{j=1}^{J}\frac{\gamma_j^2-1}{\lambda_j}+\mathcal{E}\left(y^2\right)\right). \tag{5.5}$$

All exponents of y in the error term $\mathcal{E}\left(y^2\right)$ are larger than or equal to 2. Thus, this error term is negligible for sufficiently small values of y. The derivation of this approximation is outlined in Appendix C. Note that the term $\exp\left(\frac{1}{2}\sum_{j=1}^{J}\gamma_j^2\right)$ might lead to numerical instabilities for large values of γ_j.

Fig. 5.2 depicts the numerically obtained PDF $f(y)$ for $\Delta = 0.005$, $J = 10$, $\lambda_j = 1\ \forall j$, and $\sum_{j=1}^{J}\gamma_j^2 = 6.9$, and the analytical PDF of y, which is known to be Noncentral Chi-square. The numerical values show a perfect match to the analytic curve. For illustration purposes not all numerical values are plotted.

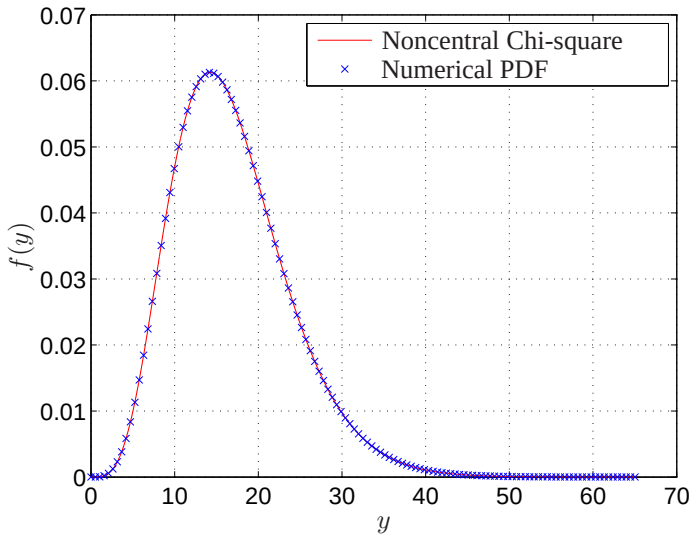

Fig. 5.2: Analytic PDF (Noncentral Chi-square) and numerically obtained PDF for $\Delta = 0.005$, $J = 10$, $\lambda_j = 1\ \forall j$, and $\sum_{j=1}^{J} \gamma_j^2 = 6.9$.

With this numerical algorithm the exact PDFs of energy samples for arbitrary Gaussian processes at the input of the ED and arbitrary positive integration filters can be computed. This enables the development of ML and MAP synchronization algorithms [42] and data detection rules for generalized ED RXs. In this thesis the theoretically derived exact PDFs are used for benchmarking purposes. We propose closed form PDFs in Section 5.3 to model the energy samples and would like to know how well they match to the exact PDF. Furthermore, the exact PDFs can be used to predict the performance of the location fingerprinting system, if the parameters of the Gaussian channel model are known.

5.3 Stochastic Modeling of Energy Samples

In this section AIC (cf. Appendix B) is applied to measured energy samples, which are obtained by processing measured CIRs from one region according to the generalized ED RX front-end in Fig. 5.1. The goal is to find closed form PDFs that provide an accurate stochastic description of the energy samples. Note that we require simple closed form PDFs $f_m\left(\mathbf{y}|H_m\right)$ for a computationally efficient calculation of (2.6) or (2.8). The candidate set considered here consists of commonly used PDFs for continuous nonnegative random variables in closed form with one or two parameters: Chi-square, Gamma, Weibull, Log-normal, Noncentral Chi-square, Nakagami-m, and Rayleigh.

We investigate two different impulse responses $g(t)$ depicted in Fig. 5.3: (i) A rectangular integration window and (ii) the impulse response of a first-order low-pass filter with a 3 dB bandwidth of 1 GHz.

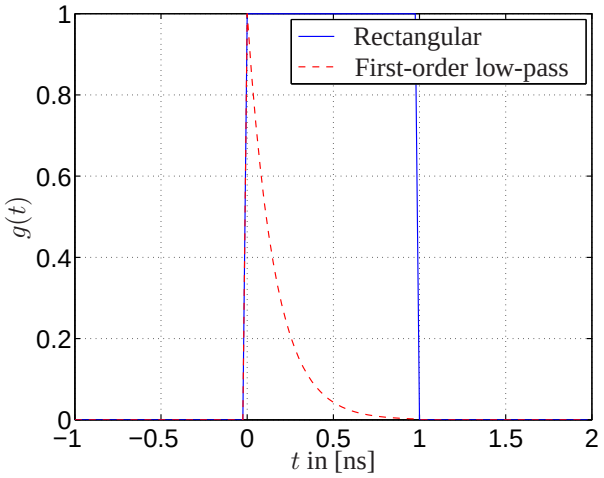

Fig. 5.3: Investigated impulse responses $g(t)$.

In the next section artificially generated energy samples based on a Gaussian channel model are considered. In this case the exact PDFs of the energy samples are known from Section 5.2. Thus, we can compare the exact PDF to the closed form PDF and measure their *approximation* quality, since the operating model is known. In Section 5.3.2 we apply AIC to measured energy samples, for which the exact PDF (operating model) is unknown.

5.3.1 Gaussian Channel Model

It is assumed that $w(t)$ and $h_m(t,\tau)$ are realizations of real Gaussian processes. Therefore $r(t)$ is itself a realization of a real Gaussian process. Note that continuous-time signals are assumed to be sampled with f_s, such that $r(t)$ in the interval $[0, T_s]$ can be described by the vector $\mathbf{r}$, which is a realization of a MV Gaussian distribution. The corresponding statistical parameters of $\mathbf{r}$ (mean vector and covariance matrix) are estimated based on 620 measured CIRs from region 4, which is chosen for presentation. The obtained results and drawn conclusions in this section carry over to all other 21 regions. Given these parameters, 10000 artificial realizations of $\mathbf{r}$ are generated. Processing them according to the generalized ED RX front-end depicted in Fig. 5.1, artificial realizations of the energy samples $y_4[n]$ are obtained.

Rectangular Integration Window Fig. 5.4 shows the Akaike Weights for all candidate PDFs and a rectangular integration window at low and high SNR. At low SNR (10 dB), all

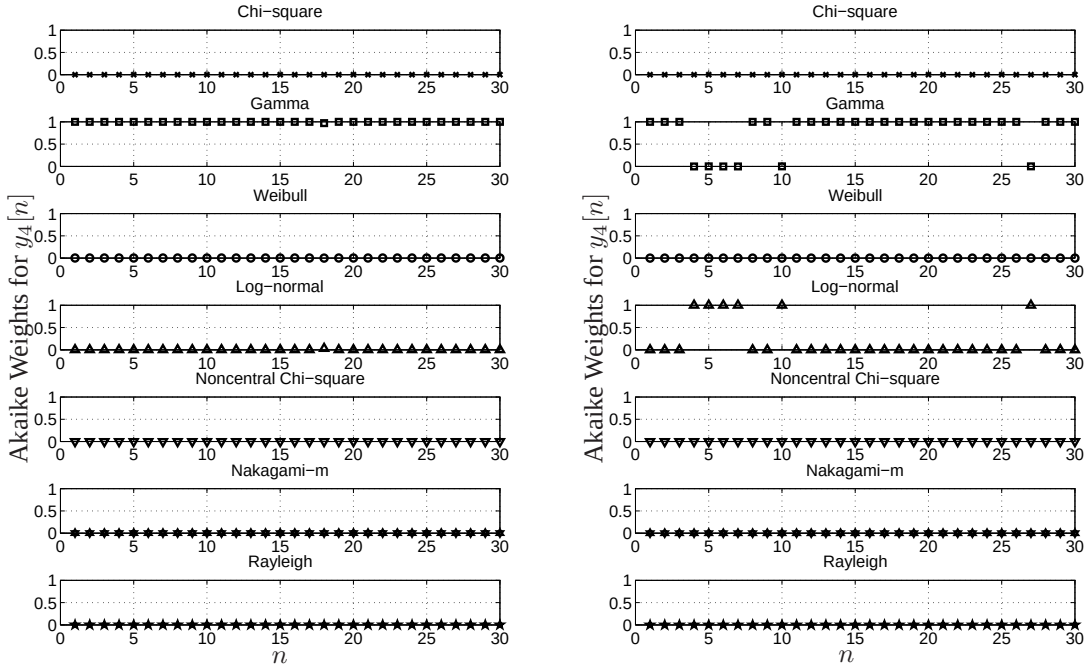

Fig. 5.4: Akaike Weights at SNR $= 10$ dB (left) and SNR $= 40$ dB (right) for Gaussian channel model using a rectangular integration window.

energy samples are best modeled by the Gamma PDF. This homogeneity is caused by the dominating effect of the distortion signal $w(t)$ with identical statistics for all n. The energy samples cannot be modeled with a Chi-square PDF because the zero mean and white random process $w(t)$ is no longer white after bandpass filtering. Note that only the integration of a squared white Gaussian process produces a Chi-square distribution.

At high SNR (40 dB), the non-stationary statistics of the propagation channel are dominating the distribution of $y_4[n]$. Since the statistics of the channel change with n due to multipath propagation (non-stationary channel statistics), the best modeling PDFs for $y_4[n]$ might change with n as well. It can be seen from Fig. 5.4 that still mostly the Gamma PDF achieves highest weights, but for example $y_4[5]$ is better modeled with a Log-normal PDF. Thus, depending on the channel statistics and the SNR, either the Gamma PDF or the Log-normal PDF are best suited among the PDFs in the candidate set to describe the statistics of $y_4[n]$.

Fig. 5.5 shows the fitted closed form PDFs and the exact PDFs obtained numerically (cf. Section 5.2) of $y_4[5]$ and $y_4[9]$ for SNR $= 40$ dB. It can be seen that the Log-normal PDF approximates the PDF of $y_4[5]$ better and the Gamma PDF approximates the PDF of

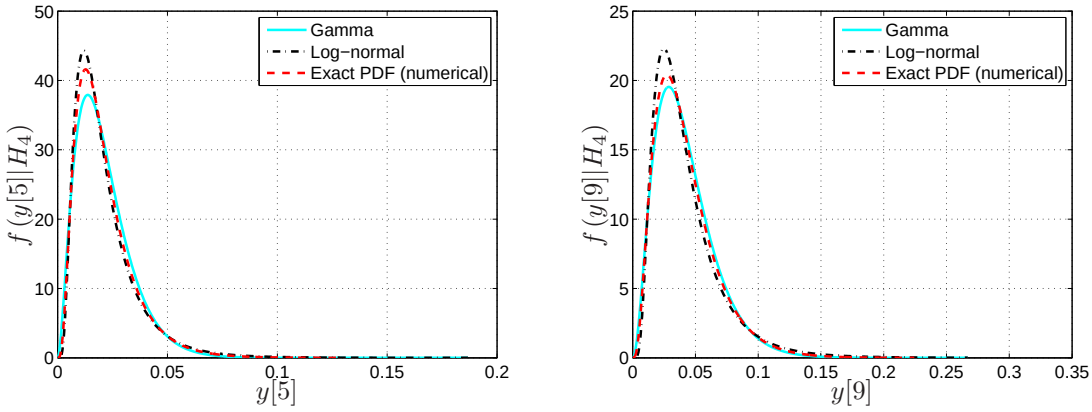

Fig. 5.5: PDFs for $y_4[5]$ (left) and $y_4[9]$ (right) for Gaussian channel model at SNR $= 40$ dB using a rectangular integration window.

$y_4[9]$ better. However, both approximations show a very good match to the exact PDF, which supports the applicability of these closed form PDF approximations.

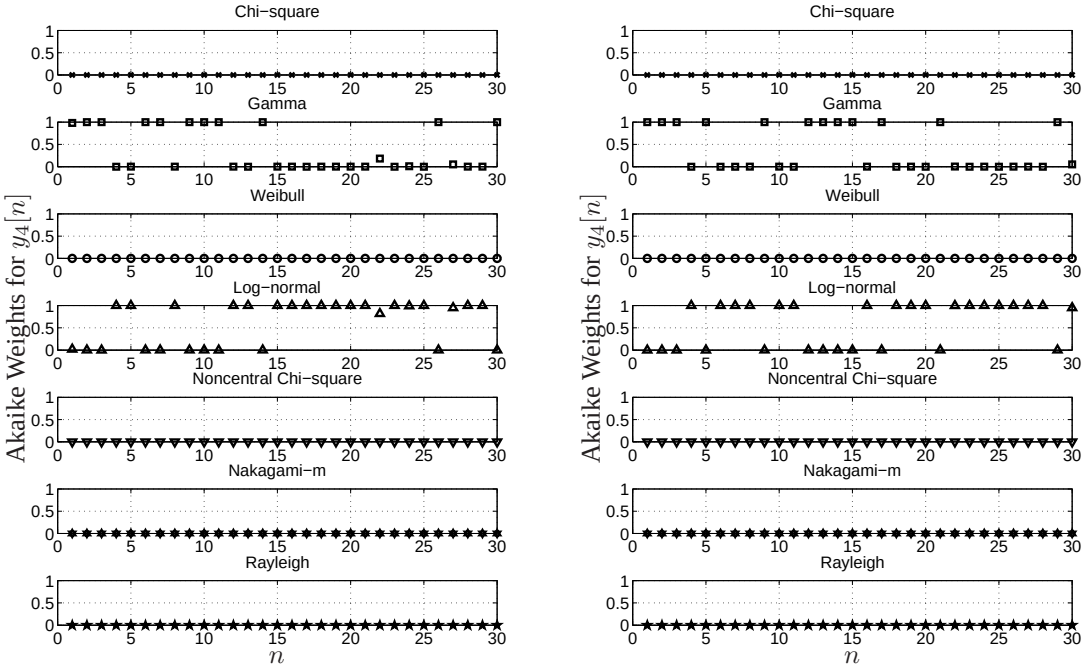

Fig. 5.6: Akaike Weights at SNR $= 10$ dB (left) and SNR $= 40$ dB (right) for Gaussian channel model using a first-order low-pass filter.

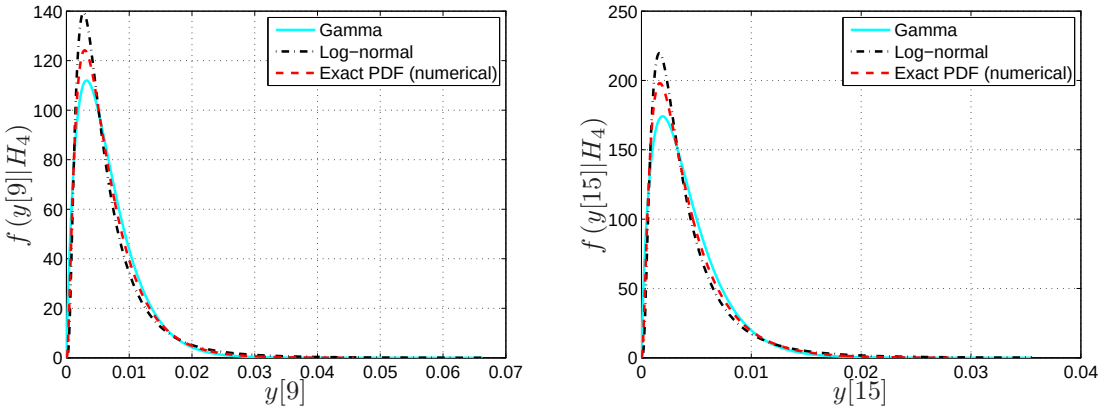

Fig. 5.7: PDFs for $y_4[9]$ (left) and $y_4[15]$ (right) for Gaussian channel model at SNR $= 40$ dB using a first-order low-pass filter.

First-order Low-pass Filter Fig. 5.6 depicts the Akaike Weights for all candidate PDFs and a first-order low-pass filter at low and high SNR. Similar conclusions as for the rectangular integration window can be drawn. However, we can observe that the homogeneity at low SNR is lost. Log-normal and Gamma PDF alternate although the input statistics are essentially the same for each n. This effect can also be observed for distortion signal $w(t)$ only by neglecting the channel. Thus, it depends on the particular realizations of $\mathbf{r}$, whether Log-normal or Gamma PDF has the higher Akaike Weight. At high SNR mostly the Log-normal PDF is winning. Fig. 5.7 shows the fitted closed form PDF and the exact PDF of energy samples $y_4[9]$ and $y_4[15]$ at high SNR using a first-order low-pass filter. It can be observed that Log-normal and Gamma PDF show a very good match to the exact PDF.

5.3.2 Measured Channel Impulse Responses

For the calculation of the Akaike Weights, an ensemble of 620 measured CIRs from all four transmit antennas located in region 4 to receive antenna 3 is used. We only show results for a rectangular integration window, but the results and conclusions carry over to a first-order low-pass filter and, moreover, to all other 21 regions.

The Akaike Weights in Fig. 5.8 give essentially the same picture as the Akaike Weights for the Gaussian channel model. We conclude that also for measured CIRs the Gamma and Log-normal PDF are best suited to model the stochastic behavior of the energy samples. It can be seen from Fig. 5.9 that the exact PDF can differ from all other PDFs, since it is computed under the assumption of a Gaussian channel model. Especially in the left plot of Fig. 5.9 the exact PDF does not match to Gamma PDF, Log-normal PDF, and the histogram. The reason

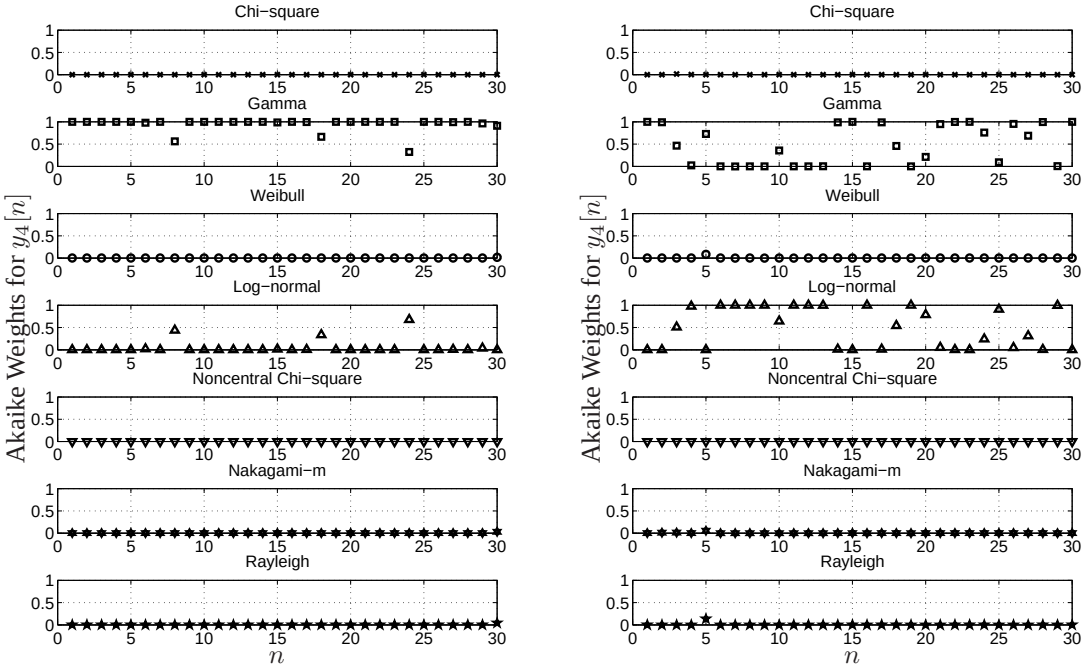

Fig. 5.8: Akaike Weights at SNR $= 10$ dB (left) and SNR $= 40$ dB (right) for measured CIRs from region 4 using a rectangular integration window.

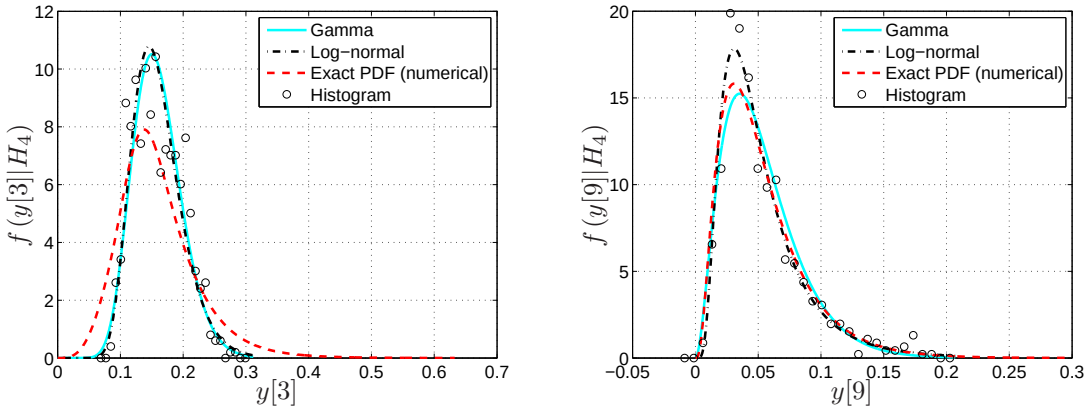

Fig. 5.9: PDFs for $y_4[3]$ (left) and $y_4[9]$ (right) for measured CIRs from region 4 at SNR $= 40$ dB using a rectangular integration window.

for this mismatch is the temporal alignment and the fact that the channel taps causing $y_4[3]$ deviate significantly from a Gaussian distribution. This observation indicates that approach $\mathcal{B}$ may be more promising than approach $\mathcal{A}$ for the stochastic modeling of energy samples.

5.3.3 Summary

In summary we consider the following PDFs for the stochastic modeling of energy samples $y_m[n]$:

- The exact PDF for a Gaussian channel model is derived in Section 5.2 and is denoted by $f_{\mathrm{E}}\left(y[n]|H_m\right)$. The parameters of the Gaussian channel model can be estimated from measured CIRs from region m or can be obtained by a geometrical channel model and a ray tracing approach.
- The Gamma PDF with shape parameter $\alpha_{n,m}$ and scale parameter $\beta_{n,m}$:

$$f_{\mathrm{G}}\left(y[n]|H_m\right) = \begin{cases} \frac{(y[n])^{\alpha_{n,m}-1}\exp(-y[n]/\beta_{n,m})}{\beta_{n,m}^{\alpha_{n,m}}\,\Gamma\left(\alpha_{n,m}\right)} & \text{for } y[n] \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

- The Log-normal PDF with mean $\mu_{n,m}$ and variance $\sigma_{n,m}^2$:

$$f_{\mathrm{L}}\left(y[n]|H_m\right) = \begin{cases} \frac{1}{y[n]\sqrt{2\pi\sigma_{n,m}^2}}\exp\left(-\frac{(\ln(y[n])-\mu_{n,m})^2}{2\sigma_{n,m}^2}\right) & \text{for } y[n] \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

If statistically independent energy samples are assumed, then the joint PDF of the location fingerprint vector $\mathbf{y}$ is obtained as product of the marginal PDFs according to

$$\text{1.) Exact: } f_{m,\mathrm{E}}\left(\mathbf{y}|H_m\right) = \prod_{n=1}^{N} f_{\mathrm{E}}\left(y[n]|H_m\right), \tag{5.6}$$

$$\text{2.) Gamma: } f_{m,\mathrm{G}}\left(\mathbf{y}|H_m\right) = \prod_{n=1}^{N} f_{\mathrm{G}}\left(y[n]|H_m\right), \tag{5.7}$$

$$\text{3.) Log-normal: } f_{m,\mathrm{L}}\left(\mathbf{y}|H_m\right) = \prod_{n=1}^{N} f_{\mathrm{L}}\left(y[n]|H_m\right). \tag{5.8}$$

In general, the joint PDF given in (5.6) is an approximation to the exact joint PDF of $\mathbf{y}$ assuming a Gaussian channel model, because correlations in $\mathbf{r}$ cause dependencies in $\mathbf{y}$. However, it is mathematically and computationally intractable to compute the exact joint PDF,

which would require the numerical inversion of an N-dimensional characteristic function. Correlations among energy samples are further discussed in the next section.

5.4 Correlation among Energy Samples

So far, we have investigated the marginal distributions of the energy samples in $\mathbf{y}$. However, a full stochastic description of $\mathbf{y}$ requires a joint PDF as already discussed in Section 4.2.2. It has been shown in Section 4.6 that the channel tap correlations, which are caused by the propagation channel, differ from region to region and can be exploited to improve the location fingerprinting performance. Based on this observation for a coherent RX, it is conjectured that also correlations among energy samples exist, which are caused by propagation effects and a region specific. Thus, accounting for correlations among energy samples is a means to improve the position location performance for a generalized ED RX.

We are looking for a mathematically tractable MV PDF for $\mathbf{y}$, which is able to model correlations among energy samples. There exists a comprehensive discussion of continuous MV distributions in [76] covering MV Gaussian, MV Exponential, MV Extreme Value, MV Beta, MV Gamma, MV Logistic, MV Liouville, and MV Pareto distributions. However, the mathematical expressions for the respective PDFs, except for the MV Gaussian PDF, are involved and mostly not given in closed form. This essentially prevents the derivation of ML parameter estimation and ML or MAP decision algorithms. Therefore, there is a lack of alternatives. One possible solution would be the MV Gaussian distribution. However, since the components in $\mathbf{y}$ are energy samples and therefore non-negative, this model is obviously not accurate.

Based on the analysis of the marginal distributions the best stochastic description of $\mathbf{y}$ would be either the MV Gamma or the MV Log-normal distribution. The MV Gamma PDF cannot be given in closed form [76], which leaves as only alternative the MV Log-normal distribution. The corresponding PDF is given in closed form by

$$f_{m,\mathrm{MVL}}\left(\mathbf{y}|H_m\right) = \begin{cases} \frac{\exp\left(-\frac{1}{2}\left((\ln(\mathbf{y})-\boldsymbol{\mu}_m)^T\boldsymbol{\Sigma}_m^{-1}(\ln(\mathbf{y})-\boldsymbol{\mu_m})\right)\right)}{(2\pi)^{N/2}|\boldsymbol{\Sigma}_m|^{1/2}\prod_{n=1}^{N}y[n]} & \text{for } \mathbf{y} \geq \mathbf{0}, \\ 0 & \text{otherwise}, \end{cases} \tag{5.9}$$

where the natural logarithm in $\ln(\mathbf{y})$ is taken element-wise. The random variable transformation $\mathbf{x} = \ln(\mathbf{y})$ renders $\mathbf{x}$ as jointly Gaussian distributed with mean vector $\boldsymbol{\mu}_m$ and covariance matrix $\boldsymbol{\Sigma}_m$. This can be shown by the MV random variable transformation theo-

rem $f_X(\mathbf{x}) = f_{\mathrm{MVL}}(\mathbf{y})\,|\mathbf{J}|$ and by noticing that the determinant of the Jacobian $\mathbf{J}$ is given by $|\mathbf{J}| = \prod_{n=1}^{N} y[n]$.

5.4.1 Empirical Correlation Matrices

Fig. 5.10 depicts the empirical correlation matrices of $\mathbf{y}$ for 620 artificially generated CIRs based on the Gaussian channel model in the left plot and for 620 measured CIRs from region 2 in the right plot. The parameters for the Gaussian channel model are estimated based on 620 measured CIRs from region 2. Fig. 5.11 depicts the corresponding empirical correlation matrices of $\mathbf{x} = \ln(\mathbf{y})$. It can be observed that in both cases the correlation coefficients are considerably smaller for the Gaussian channel model than for measured CIRs. The reasons for the large correlation values for $n \leq 4$ are again the temporal alignment and the weak multipath interference on the LoS path.

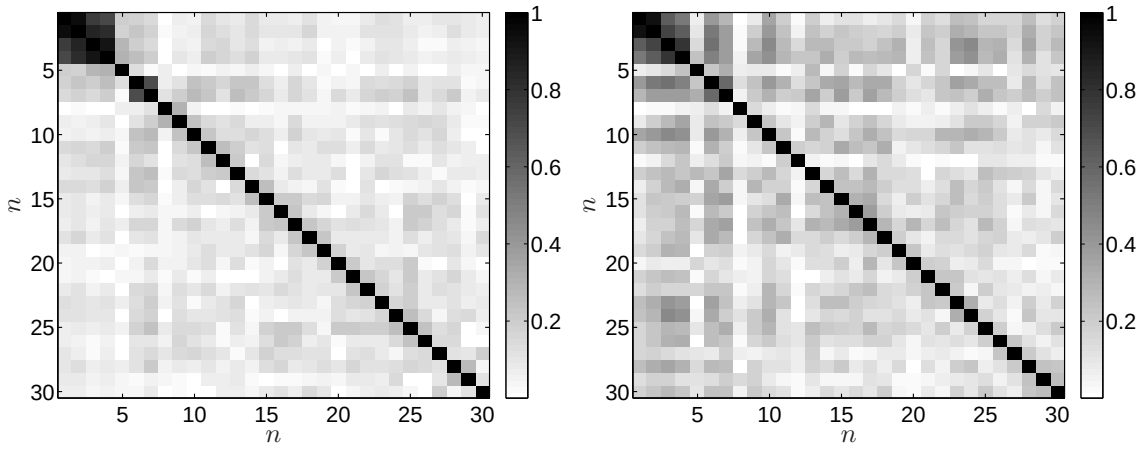

Fig. 5.10: Absolute values of empirical correlation matrices of $\mathbf{y}$ for Gaussian channel model (left) and measured CIRs (right) from region 2.

5.4.2 Analytical Expression for Correlation Coefficients

The analytical expression for the correlation coefficient between two energy samples $y[i]$ and $y[j]$ based on the statistical parameters of $\mathbf{r}$ is derived in the following. The equations for the energy samples $y[i]$ and $y[j]$ are given by

$$y[i] = \mathbf{r}_i^T \mathbf{G} \mathbf{r}_i \text{ and } y[j] = \mathbf{r}_j^T \mathbf{G} \mathbf{r}_j.$$

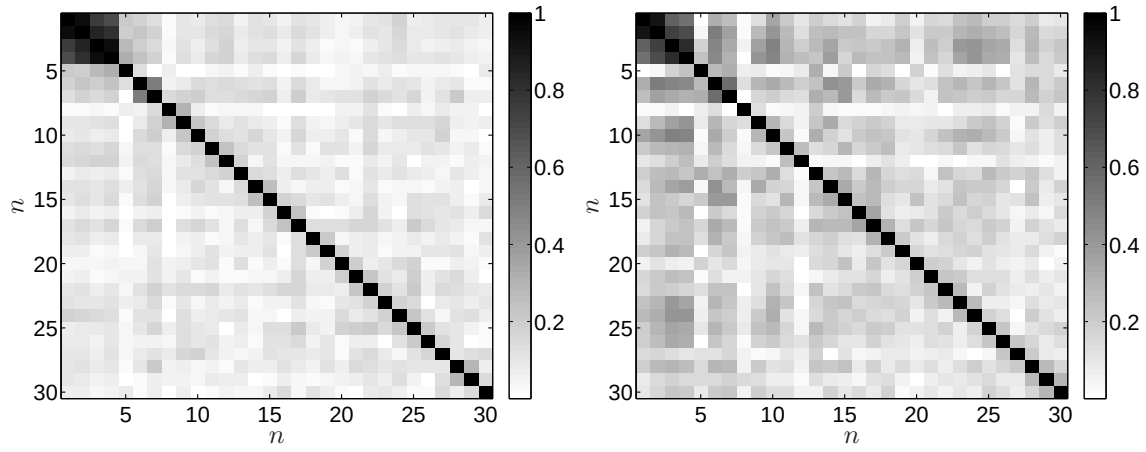

Fig. 5.11: Absolute values of empirical correlation matrices of **x** for Gaussian channel model (left) and measured CIRs (right) from region 2.

The mean vectors of $\mathbf{r}_i$ and $\mathbf{r}_j$ are denoted by $\mathbf{v}_i$ and $\mathbf{v}_j$, respectively. The covariance matrices of $\mathbf{r}_i$ and $\mathbf{r}_j$ are denoted by $\mathbf{S}_i$ and $\mathbf{S}_j$, respectively. The covariance matrix between $\mathbf{r}_i$ and $\mathbf{r}_j$ is denoted by $\mathbf{S}_{(i,j)}$. Note that the region index m is neglected in this derivation for notational convenience. The integration filter $g(t)$ is assumed to be restricted to the interval $[0, 1/f_{\text{ed}}]$ such that it influences only a single energy sample.

The correlation coefficient between $y[i]$ and $y[j]$ is defined as

$$\rho_{i,j} = \frac{\mathrm{E}\left[\left(y[i] - \mathrm{E}\left[y[i]\right]\right)\left(y[j] - \mathrm{E}\left[y[j]\right]\right)\right]}{\sqrt{\mathrm{E}\left[\left(y[i] - \mathrm{E}\left[y[i]\right]\right)^2\right]\mathrm{E}\left[\left(y[j] - \mathrm{E}\left[y[j]\right]\right)^2\right]}}. \tag{5.10}$$

The expected value of an energy sample $y[n]$ is given by

$$\mathrm{E}\left[y[n]\right] = \mathrm{E}\left[\mathbf{r}_n^T \mathbf{G} \mathbf{r}_n\right] = \mathrm{Tr}\left[\mathbf{S}_n \mathbf{G}\right] + \mathbf{v}_n^T \mathbf{G} \mathbf{v}_n = \sum_{k=1}^{J} g\left[k\right]\left(\left[\mathbf{S}_n\right]_{k,k} + \left(v_n\left[k\right]\right)^2\right).$$

The second equality follows from the fact that $\mathbf{G}$ is a diagonal matrix with $\mathbf{g} = \mathrm{diag}\left[\mathbf{G}\right]$. Further, the expected value of the product of two energy samples is given by

$$\mathrm{E}\left[y[i]y[j]\right] = \mathrm{E}\left[\mathbf{r}_i^T \mathbf{G} \mathbf{r}_i \mathbf{r}_j^T \mathbf{G} \mathbf{r}_j\right] = \sum_{k=1}^{J}\sum_{l=1}^{J} g\left[k\right] g\left[l\right] \mathrm{E}\left[\left(r_i\left[k\right]\right)^2 \left(r_j\left[l\right]\right)^2\right].$$

This expression holds for arbitrary distributions of $\mathbf{r}_i$ and $\mathbf{r}_j$ and a diagonal matrix $\mathbf{G}$. Unfortunately, the expectation is hard to calculate for arbitrary distributions, since moments up to the fourth order are involved. However, we obtain for jointly Gaussian random vectors $\mathbf{r}_i$

and $\mathbf{r}_j$

$$\mathrm{E}\left[y[i]y[j]\right] = \sum_{k=1}^{J}\sum_{l=1}^{J} g\left[k\right] g\left[l\right] \Bigg(\left[\mathbf{S}_i\right]_{k,k} \left[\mathbf{S}_j\right]_{l,l} + 2\left(\left[\mathbf{S}_{(i,j)}\right]_{k,l}\right)^2 + 4v_i\left[k\right] v_j\left[l\right] \left[\mathbf{S}_{(i,j)}\right]_{k,l} + \left(v_i\left[k\right]\right)^2 \left[\mathbf{S}_j\right]_{l,l} + \left(v_j\left[l\right]\right)^2 \left[\mathbf{S}_i\right]_{k,k} + \left(v_i\left[k\right]\right)^2 \left(v_j\left[l\right]\right)^2 \Bigg).$$

The expectation of the product of four correlated jointly Gaussian random variables ($\mathrm{E}\left[X_1X_2X_3X_4\right]$) is derived in Appendix E. The above result follows directly by choosing $X_1 = X_2 = r_i\left[k\right]$ and $X_3 = X_4 = r_j\left[l\right]$.

Consequently, the covariance coefficient between $y[i]$ and $y[j]$ is given by

$$\begin{aligned}&\mathrm{E}\left[\left(y[i] - \mathrm{E}\left[y[i]\right]\right)\left(y[j] - \mathrm{E}\left[y[j]\right]\right)\right] = \mathrm{E}\left[y[i]y[j]\right] - \mathrm{E}\left[y[i]\right]\mathrm{E}\left[y[j]\right] \\ &= \sum_{k=1}^{J}\sum_{l=1}^{J} g\left[k\right] g\left[l\right] \left(2\left(\left[\mathbf{S}_{(i,j)}\right]_{k,l}\right)^2 + 4v_i\left[k\right] v_j\left[l\right] \left[\mathbf{S}_{(i,j)}\right]_{k,l}\right).\end{aligned}$$

The correlation coefficient $\rho_{i,j}$ follows by using this expression in (5.10) with the corresponding indices i and j. Assuming that the vectors $\mathbf{r}_i$ and $\mathbf{r}_j$ are zero mean then the covariance coefficient depends only on $\mathbf{S}_{(i,j)}$ and $\mathbf{G}$. It can be concluded that the energy samples $y[i]$ and $y[j]$ are strongly correlated if the two vectors $\mathbf{r}_i$ and $\mathbf{r}_j$ are strongly correlated.

5.5 Position Location and Clustering Systems

5.5.1 Training Phase

For the evaluation of (2.6) and (2.8), the parameters of the conditional PDFs of $\mathbf{y}$ given H_m for all m must be estimated during a training phase. In contrast to location fingerprinting with a coherent RX, the distortion signal $w(t)$ is comprised in the location fingerprints of a generalized ED RX. This implies that the training phase must be repeated for each SNR operating point, since the SNR influences the stochastic description of $\mathbf{y}$. Consequently, a high SNR during the training phase is no longer required.

The ML parameter estimates for $f_{m,\mathrm{G}}\left(y[n]|H_m\right)$ based on L training samples

$y_m^{(1)}[n], y_m^{(2)}[n], \ldots, y_m^{(L)}[n]$ from nodes located in region m are given by [77]

$$\ln\left(\hat{\alpha}_{n,m}\right) - \frac{\Gamma^{'}\left(\hat{\alpha}_{n,m}\right)}{\Gamma\left(\hat{\alpha}_{n,m}\right)} = \ln\left(\frac{1}{L}\sum_{l=1}^{L} y_m^{(l)}[n]\right) - \frac{1}{L}\sum_{l=1}^{L}\ln\left(y_m^{(l)}[n]\right), \tag{5.11}$$

$$\hat{\beta}_{n,m} = \frac{1}{\hat{\alpha}_{n,m} L}\sum_{l=1}^{L} y_m^{(l)}[n].$$

Although there exists no closed form solution for $\hat{\alpha}_{n,m}$ in (5.11), numerical solvers can be applied.

The ML parameter estimates for $f_{m,\text{MVL}}\left(\mathbf{y}|H_m\right)$ are given by [76]

$$\hat{\boldsymbol{\mu}}_m = \frac{1}{L}\sum_{l=1}^{L}\ln\left(\mathbf{y}_m^{(l)}\right) = \frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_m^{(l)}$$

$$\hat{\boldsymbol{\Sigma}}_m = \frac{1}{L}\sum_{l=1}^{L}\left(\ln\left(\mathbf{y}_m^{(l)}\right) - \hat{\boldsymbol{\mu}}_m\right)\left(\ln\left(\mathbf{y}_m^{(l)}\right) - \hat{\boldsymbol{\mu}}_m\right)^T = \frac{1}{L}\sum_{l=1}^{L}\left(\mathbf{x}_m^{(l)} - \hat{\boldsymbol{\mu}}_m\right)\left(\mathbf{x}_m^{(l)} - \hat{\boldsymbol{\mu}}_m\right)^T.$$

The ML parameter estimates for $f_{m,\text{L}}\left(\mathbf{y}|H_m\right)$ follow directly from the ML parameter estimates for $f_{m,\text{MVL}}\left(\mathbf{y}|H_m\right)$.

We notice that there are $2N$ free real parameters to estimate if it is assumed that the energy samples are independent. If we model also the correlation among energy samples there are $N + N(N+1)/2$ free real parameters to estimate. The accuracy of the parameter estimates is determined by L as already discussed in Section 4.4.1. The impact of L on the position location performance is investigated in Section 5.7.2.

The proposed location fingerprinting framework in Chapter 2 enables the combination of training and localization phase via iterative algorithms, which improve the quality of the parameter estimates during the localization phase. This idea is further discussed in Chapter 6.

5.5.2 Localization Phase

The localization phase consists of evaluating (2.6) or (2.8), depending on the figure of merit, with an observation vector $\mathbf{y}$. The actual metrics depend on the PDF of $\mathbf{y}$. As an example we consider decision rule (2.8) for the MV Log-normal model, which can be formulated as

$$\hat{m} = \underset{m=1,2,\ldots,M}{\text{argmax}} \frac{\pi_m}{|\hat{\boldsymbol{\Sigma}}_m|^{1/2}} \exp\left(-\frac{\left(\ln\left(\mathbf{y}\right) - \hat{\boldsymbol{\mu}}_m\right)^T \hat{\boldsymbol{\Sigma}}_m^{-1}\left(\ln\left(\mathbf{y}\right) - \hat{\boldsymbol{\mu}}_m\right)}{2}\right),$$

where we have neglected all terms that are independent of m. Taking the natural logarithm gives

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \ln\left(\frac{\pi_m}{|\hat{\boldsymbol{\Sigma}}_m|^{1/2}}\right) - \frac{\left(\ln(\mathbf{y}) - \hat{\boldsymbol{\mu}}_m\right)^T \hat{\boldsymbol{\Sigma}}_m^{-1} \left(\ln(\mathbf{y}) - \hat{\boldsymbol{\mu}}_m\right)}{2}.$$

This decision rule is similar to the decision rule, which has been obtained in Section 4.4.2 for a coherent RX. Prior to the calculation of the quadratic form with the pre-calculated and stored matrix $\hat{\boldsymbol{\Sigma}}_m^{-1}$, the natural logarithm of the observation vector $\mathbf{y}$ must be calculated and the mean $\hat{\boldsymbol{\mu}}_m$ must be subtracted. Furthermore, it can be observed that the impact of the distortion signal $w(t)$ is no longer visible in the decision rule as compared to the rule in Section 4.4.2. This is due to the nonlinear RX front-end and the stochastic modeling approach, which accounts for the impact of $w(t)$ in $\hat{\boldsymbol{\mu}}_m$ and $\hat{\boldsymbol{\Sigma}}_m$.

Multiple Observations per Agent If there are K observations $\bar{\mathbf{y}} = \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_K^T\right]^T$ available (cf. Section 2.4.3), we apply (2.11) and obtain a slightly modified decision rule:

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmax}} \ln\left(\frac{\pi_m}{|\hat{\boldsymbol{\Sigma}}_m|^{K/2}}\right) - \frac{1}{2}\sum_{k=1}^{K} \left(\ln(\mathbf{y}_k) - \hat{\boldsymbol{\mu}}_m\right)^T \hat{\boldsymbol{\Sigma}}_m^{-1} \left(\ln(\mathbf{y}_k) - \hat{\boldsymbol{\mu}}_m\right).$$

5.6 Accuracy of Closed Form Approximations

In Section 5.3, visual inspection has shown that the proposed closed form PDFs are reasonable approximations to the exact PDFs of $y_m[n]$ for a Gaussian channel model. In this section, we support this observation by analyzing $\mathcal{P}_e$. For this analysis, it is sufficient to restrict the number of hypotheses to $M = 2$. Let us consider H_5 and H_{10} with $\pi_5 = \pi_{10} = 0.5$. We present $\mathcal{P}_e$ results assuming a Gaussian channel model as well as for measured CIRs. The parameters of the Gaussian channel model are estimated using 620 measured CIRs from each region. Based on these parameters the exact marginal PDFs are calculated and consequently $f_{5,\mathrm{E}}(\mathbf{y}|H_5)$ and $f_{10,\mathrm{E}}(\mathbf{y}|H_{10})$ are obtained according to (5.6). The parameters of the closed form PDFs are estimated based on $L = 400$ training signals per region according to Section 5.5.1.

The left plot of Fig. 5.12 depicts $\mathcal{P}_e$ for artificially generated CIRs based on a Gaussian channel model. The energy samples in $\mathbf{y}$ are generally dependent due to correlations in $\mathbf{r}$. In order to obtain independent energy samples the Gaussian channel model is modified, such that the covariance matrices $\mathbf{S}_{(i,j)}$ between the vectors $\mathbf{r}_i$ and $\mathbf{r}_j$ for $i \neq j$ are zero

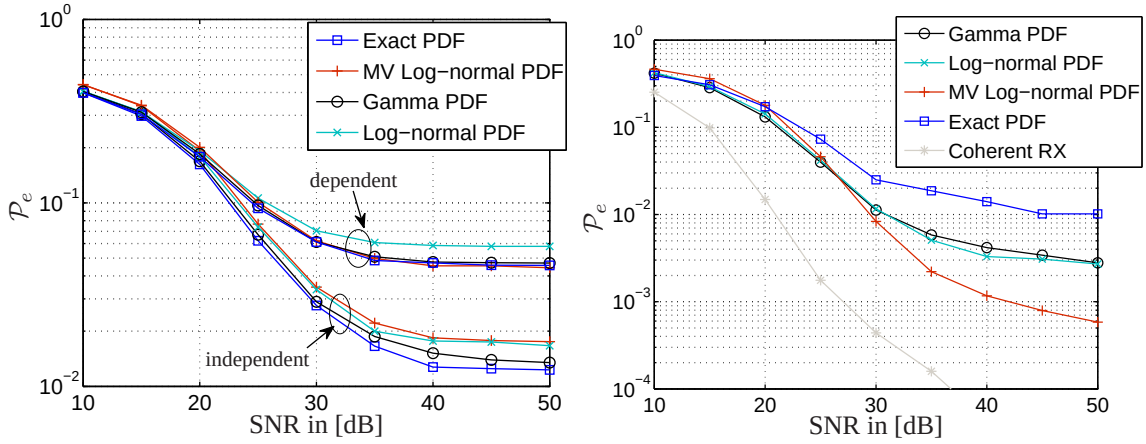

Fig. 5.12: $\mathcal{P}_e$ for H_5 and H_{10} with $\pi_5 = \pi_{10} = 0.5$ assuming a Gaussian channel model (left) and for measured CIRs (right).

matrices. This is done to ensure that decision rule (2.8) using the exact PDF (5.6) is optimal. The corresponding performance curves can be seen in the left plot of Fig. 5.12 marked with *independent*. It can be seen that all closed form PDFs achieve $\mathcal{P}_e$ close to the optimum, at which the Gamma PDF comes closest. These performance results support the accuracy and applicability of the proposed closed form PDFs.

The unmodified Gaussian channel model causes dependencies in $\mathbf{y}$. The corresponding curves are marked with *dependent* in the left plot of Fig. 5.12. There is a performance loss due to the wrong independence assumption. Note that the exact PDF given in (5.6) does not provide minimal $\mathcal{P}_e$ anymore. Only the MV Log-normal PDF given in (5.9) accounts for the correlations among energy samples. However, it can be seen that for the Gaussian channel model the achievable performance improvement compared to the PDFs assuming independent energy samples is only marginal.

The right plot of Fig. 5.12 depicts $\mathcal{P}_e$ for measured CIRs. It can be observed that all closed form PDFs outperform the exact PDF. This supports the conclusion from Section 5.3 that the direct modeling of energy samples (approach $\mathcal{B}$) is more promising than the mathematical derivation of the exact PDF based on a Gaussian channel model (approach $\mathcal{A}$). Nevertheless, the exact PDF derived from the Gaussian channel model is useful for performance prediction providing conservative values for $\mathcal{P}_e$.

A significant performance improvement of the MV Log-normal PDF compared to the other PDFs can be observed from the right plot of Fig. 5.12. Such an improvement cannot be observed in the left plot of Fig. 5.12 assuming a Gaussian channel model. These results are conform with the empirical correlation matrices shown in Fig. 5.11, where it can be seen

that the correlation coefficients assuming a Gaussian channel model are significantly smaller than for measured CIRs. It can be concluded from these results that the MV Log-normal PDF is the preferable PDF for modeling the location fingerprint $\mathbf{y}$ generated by a generalized ED RX.

For comparison the performance of a coherent RX with default parameters (cf. Chapter 4) is plotted as reference. In order to reach $\mathcal{P}_e = 10^{-2}$, an SNR of around 21 dB is required for the coherent RX, whereas an SNR of around 30 dB is required for the ED RX. Furthermore, the error floor at high SNR can be significantly reduced by applying a coherent RX.

5.7 Experimental Performance Analysis

In this section, the performance of location fingerprinting with a generalized ED RX is investigated. The figures of merit are $\mathcal{P}_e$ for the clustering problem and $\mathcal{D}_e$ for the position location problem. We consider the ensemble of measured CIRs from all four transmit antennas to receive antenna 3, which resembles a remote positioning scenario. The UWB channel measurement campaign is described in Chapter 3. The total number of measured CIRs per region is 620. The CIRs are normalized such that the average energy is the same for all regions implying that the PL information is not used for localization. In order to obtain realistic estimates for the figures of merit we resort to the randomized cross-validation method described in Section 4.6.1. The performance metrics $\mathcal{P}_e$ and $\mathcal{D}_e$ are calculated assuming Gamma PDF (5.7), Log-normal PDF (5.8), and MV Log-normal PDF (5.9).

5.7.1 Default System Parameters

Important system parameters are the number of training signals L per region, the SNR, the sampling frequency of the energy detector f_{ed}, the observation window size or symbol period denoted by T_s, and the number of regions M. Note that T_s and f_{ed} determine the length of the location fingerprints according to $N = T_s f_{\text{ed}}$. The default values of these parameters are summarized as follows.

- Default number of training signals is $L = 400$. This leaves 220 test signals per region.
- Default SNR is the measurement SNR denoted by SNR_{mea} in Chapter 3. This corresponds to adding no artificial distortion signals.
- Default sampling rate is $f_{\text{ed}} = 1$ GHz.

- Default observation window size is $T_s = 30$ ns. Together with the default sampling rate this gives a default length of the location fingerprints of $N = 30$ energy samples.
- Default number of regions is $M = 22$, which corresponds to the whole measurement set.
- Default a priori probabilities are $\pi_m = 1/M$.
- Default integration filter $g(t)$ is rectangular.
- Default reference sample is $n_{\text{maxabs}} = 3$.

Two performance curves for each closed form PDF are calculated and plotted:

(1) The ED RX has $K = 1$ observation $\mathbf{y}$ available for region detection.

(2) The ED RX has $K = 2$ observations $\bar{\mathbf{y}} = \left[\mathbf{y}_1^T, \mathbf{y}_2^T\right]^T$ from the same region available. The observation pairs are picked randomly from the test data set.

5.7.2 Number of Training Vectors

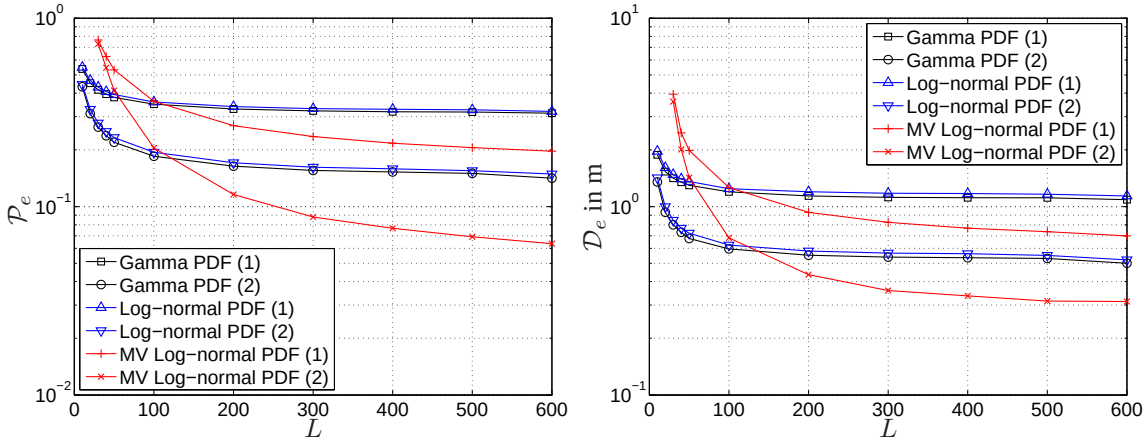

Fig. 5.13: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L.

Fig. 5.13 depicts the position location performance as a function of L. The remaining parameters are set to their default values. The length of $\mathbf{y}$ is $N = 30$, which means that $N + N(N+1)/2 = 495$ free real parameters must be estimated for the MV Log-normal PDF. The other two PDFs assuming independent taps require the estimation of $2N = 60$ free real parameters. Therefore, the performance of the MV Log-normal PDF drops drastically for small L, since the stochastic description becomes inaccurate. However, for $L > 100$ the higher model complexity of MV Log-normal pays off. Further, it can be observed that two available observations[3] give a significant performance improvement. For example, the aver-

[3]The corresponding performance curves in Fig. 5.13 are marked with (2).

age positioning error can be decreased from 93 cm to 43 cm using the MV Log-normal model with $L = 200$ training signals. Simulation results show that more than two observations give additional performance gains.

5.7.3 Distortion Analysis

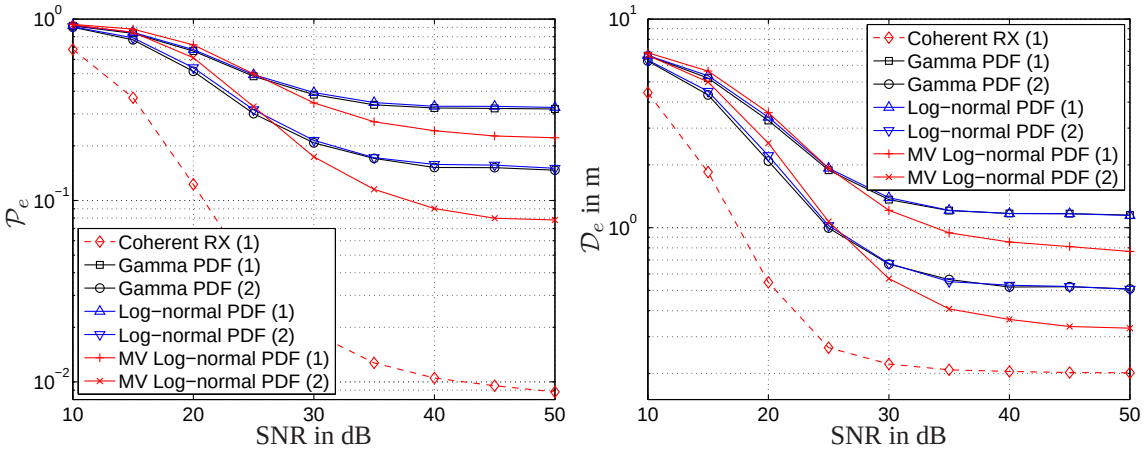

Fig. 5.14: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of SNR.

Fig. 5.14 depicts the performance as function of the SNR. The remaining parameters are set to their default values. As reference, the performance of a coherent RX with default system parameters is shown. These curves are the same as the solid black curves with square markers depicted in Fig. 4.16. It can be seen that there is a substantial performance loss of the generalized ED RX in comparison to the coherent RX. The error floor at high SNR increases from $\mathcal{P}_e \approx 0.009$ to $\mathcal{P}_e \approx 0.22$ and from $\mathcal{D}_e \approx 0.18$ m to $\mathcal{D}_e \approx 0.76$ m, when the MV Log-normal model is assumed. This performance loss can be recovered by exploiting multiple observations per agent, which can be seen by the curves marked with (2) in Fig. 5.14. Since the distortion signal $w(t)$ changes from observation to observation, the agents need not be mobile and do not require multiple antennas. However, in order to obtain a reasonable performance for the generalized ED RX an SNR of at least 25 dB is required. In case of a coherent RX the required SNR reduces to approximately 15 dB.

5.7.4 Sampling Frequency of Energy Detector

Fig. 5.15 shows the performance as function of f_{ed}. The remaining parameters are set to their default values. Increasing f_{ed} has two effects:

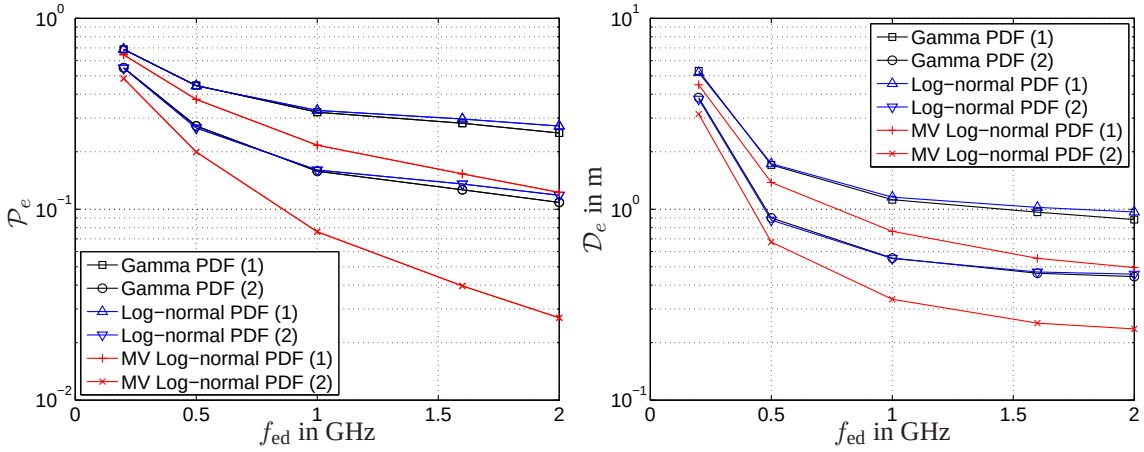

Fig. 5.15: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of f_{ed}.

1. The bandwidth of the filter $g(t)$ is increased and the multipath components of the propagation channel are better resolved in time. This leads to better performance due to more distinguished location fingerprints with more degrees of freedom.
2. The number of free parameters is increased, which requires larger L for an accurate stochastic description.

It can be observed from Fig. 5.15 that $L = 400$ is large enough such that performance gains are achievable when f_{ed} is increased to 2 GHz. It can be also seen that performance drops fast for decreasing f_{ed} implying that the temporal resolution of the propagation channel becomes too coarse. It can be concluded that a sampling frequency of $f_{\text{ed}} \geq 500$ MHz is required for reasonable performance.

5.7.5 Number of Regions

The number of regions is an important system parameter. As M increases, it can be expected that $\mathcal{P}_e$ increases as well. The communication theoretic analogy would be a signal space with a fixed number of dimensions and an increasing number of constellation points for a fixed SNR. How fast $\mathcal{P}_e$ increases, depends on the dimensionality of this signal space or in the position location setup on the degrees of freedom of the location fingerprints. The larger the degrees of freedom (signal space dimensions) are, the slower will be the performance degradation for increasing M. Note that RSS location fingerprints have only one degree of freedom per access point (anchor), which implies a rapid performance degradation. Since we are considering location fingerprints based on the UWB propagation channel we expect a graceful performance degradation for increasing M due to many degrees of freedom. The

following results support our conjecture of graceful performance degradation for increasing M.

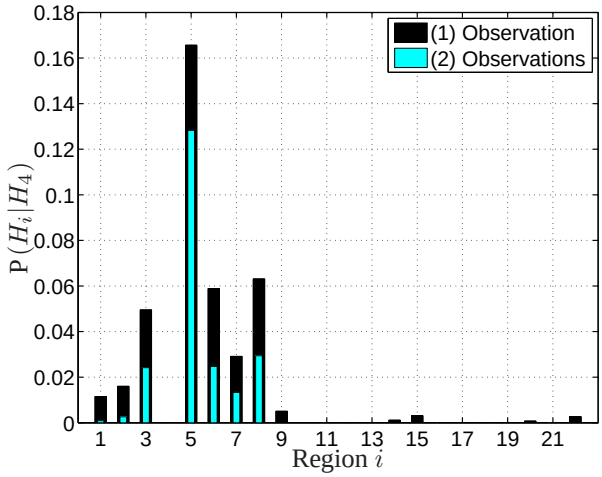

Fig. 5.16: Conditional error probabilities $\mathrm{P}(H_i|H_4)$ for observations from region 4 for $L = 200$ and MV Log-normal PDF.

Fig. 5.16 depicts the conditional error probabilities $\mathrm{P}(H_i|H_4)$ for observations from region 4 assuming one or two available observations. The stochastic model for the location fingerprints is the MV Log-normal PDF. The default system parameters except for $L = 200$ are applied. It can be seen that most of the classification errors occur among physically close regions, which implies that classification errors are mainly local phenomenons. Thus, it is expected that increasing M will not lead to a significant increase in $\mathcal{P}_e$ as long as the additional regions are not very close to already considered regions. Furthermore, $\mathcal{D}_e$ will stay essentially constant, since classification errors among nearby regions have only marginal impact on $\mathcal{D}_e$.

If the position location application requires larger or full coverage of the room, it is conceivable that the region dimensions are increased at the cost of a larger fundamental positioning uncertainty while maintaining reasonably small values for M and $\mathcal{P}_e$. For example the total probability of error considering the eight adjacent regions $\{1, 2, 3, 4, 8, 9, 10, 11\}$ is calculated as $\mathcal{P}_e = 0.22$. If regions $\{1, 2, 3, 4\}$ and regions $\{8, 9, 10, 11\}$ are merged into two super-regions the total probability of error reduces to $\mathcal{P}_e = 0.12$. For this example the default parameter settings except for $L = 200$ are applied, the MV Log-normal PDF is assumed, and one observation per agent is considered.

5.8 Summary and Conclusions

This chapter presents a detailed analysis of a location fingerprinting system employing a *single* anchor with low complexity. The location fingerprints correspond to vectors of energy samples, which can be modeled with a Gamma, a Log-normal, or a MV Log-normal PDF. An experimental performance analysis based on measured CIRs reveals that the MV Log-normal PDF provides the best performance as long as the number of training signals is sufficiently large.

An average positioning error of $\mathcal{D}_e \approx 93$ cm at high SNR, $f_{\text{ed}} = 1$ GHz, $L = 200$, and $M = 22$ can be achieved with a single observation per agent. If two observations are available, the error decreases to $\mathcal{D}_e \approx 43$ cm. If the distortions are non-negligible (low SNR) the corresponding performance degradation can be compensated by using more observations per agent for the region detection process.The agents need not be mobile and do not require multiple antennas in this case, because the distortions change from observation to observation.

In this chapter, it is demonstrated how location fingerprinting can be performed with a low complexity generalized ED RX. With the MV Log-normal PDF we propose an accurate and mathematically tractable stochastic model for a vector of energy samples, which is able to account for correlations among the energy samples.

In the upcoming Chapters 6 and 7, two promising methods to increase the efficiency of the training phase are proposed. The objective is to reduce the required number of training signals per region, while maintaining a certain position location performance.

Chapter 6

Combination of Training Phase and Localization Phase

In this chapter[1], we propose to combine the parameter estimation phase with the localization phase in order to jointly improve the quality of the parameter estimates and the performance metrics. To the best of the author's knowledge, this idea to jointly train and localize has not been known in the location fingerprinting community so far. It is, however, well known that the complexity of the training phase is the main limitation, which prevents a practical implementation of location fingerprinting systems. This fact emphasizes the importance of this idea even more.

First, we discuss in Section 6.1, how the hypothesis testing framework introduced in Section 2.4 must be adapted such that joint parameter estimation and localization can be implemented with the expectation maximization (EM) algorithm. Then, we review the required theory to understand and apply the EM algorithm in Section 6.2, discuss the position location and clustering system in Section 6.3, and finally provide an experimental performance analysis in Section 6.4.

[1] Parts of this chapter have been published in [78].

6.1 Location Fingerprinting via Parameter Estimation of Mixture Densities

We consider the same scenario as discussed in Section 2.1: The surveillance area is quantized into M regions, in which the agents can be located. This implies that each agent has to be located in one of those M regions. Further, there exists a parameterized PDF, which models the stochastic behavior of the location fingerprints and whose parameters distinguish the regions. The conditional PDF of a location fingerprint $\mathbf{y}$ of an agent located in region m is denoted by $f_m(\mathbf{y}|\Theta_m)$ with parameter set Θ_m.

Let $\mathbf{y}$ be the location fingerprint of an agent with unknown position. Since the agent could be located in any region, the parameter set of the PDF describing $\mathbf{y}$ could be any of the M parameter sets Θ_m. Consequently, it is proposed to model the unconditional PDF of $\mathbf{y}$ as a mixture density [79] according to

$$f(\mathbf{y}|\Theta) = \sum_{m=1}^{M} \pi_m f_m(\mathbf{y}|\Theta_m), \tag{6.1}$$

where the weights π_m of the component PDFs sum up to one. These weights are in analogy to the a priori probabilities defined in Section 2.4. The set of parameters describing $f(\mathbf{y}|\Theta)$ is given by

$$\Theta = \{\pi_1, \pi_2, \ldots, \pi_M, \Theta_1, \Theta_2, \ldots, \Theta_M\}.$$

If the individual parameter sets Θ_m and the weights π_m are known perfectly[2], then the optimal position location procedure is to apply the M-ary hypothesis testing framework developed in Section 2.4. In practice, however, π_m may be unknown and Θ_m must be estimated from a finite number of training signals. It has been shown in Section 4.6.3 and Section 5.7.2 that a large number of training signals is required, in order to obtain accurate parameter estimates and good position location performance.

The modeling of $f(\mathbf{y}|\Theta)$ as mixture density enables the combination of parameter estimation and region detection. Each new observation $\mathbf{y}$ is first assigned to the most likely component PDF (region detection step) and then used to improve the parameter estimates. An iterative algorithm accomplishing this task is the EM algorithm, which is reviewed in the next section.

[2] Recall that this would require infinitely many training signals.

6.2 Expectation Maximization Algorithm

The original problem, which led to the design of the EM algorithm, was the search for ML parameter estimates of a PDF based on an empirical data set with incomplete or missing values [80].

Later on, it was found that the EM approach is also very useful for optimizing likelihood functions, which cannot be optimized analytically. For example such involved likelihood functions arise from ML parameter estimation of mixture densities [81]. The idea is to simplify the likelihood function by assuming the existence of missing parameters, which is outlined in the next section. A detailed tutorial about the EM algorithm can be found in [82] and in Chapter 9 of [32].

6.2.1 Maximum Likelihood Parameter Estimation for Mixture Densities

We consider the mixture density from (6.1) and search for the ML parameter estimates $\hat{\Theta}$ of Θ based on L_b independent observations $\mathcal{Y} = \left\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{L_b}\right\}$. The according log-likelihood function is given by

$$\Lambda\left(\Theta|\mathcal{Y}\right) \triangleq \ln\left(f\left(\mathcal{Y}|\Theta\right)\right) = \ln\left(\prod_{l=1}^{L_b} f\left(\mathbf{y}_l|\Theta\right)\right) = \sum_{l=1}^{L_b} \ln\left(\sum_{m=1}^{M} \pi_m f_m\left(\mathbf{y}_l|\Theta_m\right)\right). \tag{6.2}$$

The ML parameter estimates $\hat{\Theta}$ are obtained by maximizing $\Lambda\left(\Theta|\mathcal{Y}\right)$ with respect to Θ, which is mathematically complex due to the logarithm of a sum of component PDFs.

We apply the trick of assuming the existence of an unobserved (missing) data set denoted by $\mathbf{u} = \left[u_1, u_2, \ldots, u_{L_b}\right]$, where the value of each element $u_l \in \{1, 2, \ldots, M\}$ indicates the component PDF of the corresponding observation $\mathbf{y}_l$. For example, if $u_l = m$ then $\mathbf{y}_l$ belongs to the m-th component PDF $f_m\left(\mathbf{y}_l|\Theta_m\right)$.

We denote $\mathcal{Y}$ as incomplete data set and $\{\mathcal{Y}, \mathbf{u}\}$ as complete data set. Assuming knowledge of $\mathcal{Y}$ and $\mathbf{u}$, the complete log-likelihood function is given by

$$\begin{aligned}\Lambda\left(\Theta|\mathcal{Y}, \mathbf{u}\right) &\triangleq \ln\left(f\left(\mathcal{Y}, \mathbf{u}|\Theta\right)\right) = \ln\left(\prod_{l=1}^{L_b} f\left(\mathbf{y}_l, u_l|\Theta\right)\right) \\ &= \sum_{l=1}^{L_b} \ln\left(f_{u_l}\left(\mathbf{y}_l|\Theta_{u_l}\right)\right) + \ln\left(\pi_{u_l}\right).\end{aligned} \tag{6.3}$$

Depending on the particular expressions for the component PDFs the complete log-likelihood function can be maximized with respect to Θ with several methods. For some PDFs like MV Gaussian there exist even analytical expressions for $\hat{\Theta}$. The main problem is, however, that $\mathbf{u}$ is usually unknown.

We proceed by assuming that $\mathbf{u}$ is a random vector, which makes the complete log-likelihood function $\Lambda(\Theta|\mathcal{Y},\mathbf{u})$ itself a random variable. The *expectation* step of the EM algorithm calculates the expected value of $\Lambda(\Theta|\mathcal{Y},\mathbf{u})$ with respect to $\mathbf{u}$ given $\mathcal{Y}$ and the parameter estimates from the last iteration $i-1$ denoted by $\Theta[i-1]$. Mathematically, this is written as

$$\begin{aligned} Q(\Theta,\Theta[i-1]) &= \mathrm{E}_{\mathbf{u}}\left[\Lambda(\Theta|\mathcal{Y},\mathbf{u})\,|\mathcal{Y},\Theta[i-1]\right] \\ &= \int_{\mathbf{u}} \Lambda(\Theta|\mathcal{Y},\mathbf{u})\, p(\mathbf{u}|\mathcal{Y},\Theta[i-1])\,\mathrm{d}\mathbf{u}, \end{aligned} \tag{6.4}$$

where $p(\mathbf{u}|\mathcal{Y},\Theta[i-1])$ denotes the conditional PDF of the unobserved data given the observation set $\mathcal{Y}$ and $\Theta[i-1]$. The cost function $Q(\Theta,\Theta[i-1])$ is maximized in the *maximization* step according to

$$\Theta[i] = \underset{\Theta}{\operatorname{argmax}}\left(Q(\Theta,\Theta[i-1])\right), \tag{6.5}$$

which gives the new parameter estimates $\Theta[i]$. Equations (6.4) and (6.5) are general and must be further derived until the EM algorithm can be implemented. In order to proceed, we require an expression for $p(\mathbf{u}|\mathcal{Y},\Theta[i-1])$.

For the special case of ML parameter estimation of mixture densities, the conditional probability mass function (PMF)[3] of $\mathbf{u}$ is given by

$$\begin{aligned} p(\mathbf{u}|\mathcal{Y},\Theta[i-1]) &= \prod_{l=1}^{L_b} p(u_l|\mathbf{y}_l,\Theta[i-1]) = \prod_{l=1}^{L_b} \frac{f(\mathbf{y}_l,u_l|\Theta[i-1])}{f(\mathbf{y}_l|\Theta[i-1])} \\ &= \prod_{l=1}^{L_b} \frac{\pi_{u_l}[i-1]\, f_{u_l}(\mathbf{y}_l|\Theta_{u_l}[i-1])}{\sum_{m=1}^{M}\pi_m[i-1]\, f_m(\mathbf{y}_l|\Theta_m[i-1])}, \end{aligned}$$

where the first equality is due to the definition of $\mathbf{u}$, which states that each u_l is only linked to the corresponding $\mathbf{y}_l$. This implies independent components in $\mathbf{u}$. The second equality resembles Bayes' rule and the third equality follows from plugging in known expressions from (6.2) and (6.3).

[3] Note that $\mathbf{u}$ is in this case a discrete random vector.

This expression for $p(\mathbf{u}|\mathcal{Y}, \Theta[i-1])$ is inserted into (6.4) and after some calculations [4] we obtain

$$\begin{aligned} Q(\Theta, \Theta[i-1]) &= \sum_{m=1}^{M} \sum_{l=1}^{L_b} \ln(\pi_m f_m(\mathbf{y}_l|\Theta_m)) \, p(m|\mathbf{y}_l, \Theta[i-1]) \\ &= \sum_{m=1}^{M} \sum_{l=1}^{L_b} \ln(\pi_m) \, p(m|\mathbf{y}_l, \Theta[i-1]) + \sum_{m=1}^{M} \sum_{l=1}^{L_b} \ln(f_m(\mathbf{y}_l|\Theta_m)) \, p(m|\mathbf{y}_l, \Theta[i-1]), \end{aligned} \tag{6.6}$$

where $p(m|\mathbf{y}_l, \Theta[i-1])$ denotes the conditional a posteriori PMF of the m-th component PDF and is given by

$$p(m|\mathbf{y}_l, \Theta[i-1]) = \frac{\pi_m[i-1] \, f_m(\mathbf{y}_l|\Theta_m[i-1])}{\sum_{k=1}^{M} \pi_k[i-1] \, f_k(\mathbf{y}_l|\Theta_k[i-1])}. \tag{6.7}$$

Note that $p(m|\mathbf{y}_l, \Theta[i-1])$ is the probability mass of the event that the observation $\mathbf{y}_l$ belongs to the m-th component PDF given $\Theta[i-1]$.

The remaining step is to maximize (6.6) with respect to Θ in order to obtain new parameter estimates. Depending on the component PDFs, this can be still mathematically complex. However, for some component PDFs like MV Gaussian or MV Log-normal[5] it is possible to derive analytical expressions for the parameters, which maximize (6.6). In the next section, we provide update equations for Gaussian mixture densities.

6.2.2 Gaussian Mixture Densities

Gaussian mixture densities are at hand, when all component PDFs are MV Gaussian. If all component PDFs are MV Log-normal, then a Gaussian mixture density in $\mathbf{x}$ is obtained by the random variable transformation $\mathbf{x} = \ln(\mathbf{y})$. The parameter set Θ of Gaussian mixture densities consists of the weights π_m, a set of mean vectors $\boldsymbol{\mu}_m$, and a set of covariance matrices $\boldsymbol{\Sigma}_m$.

The update equations for the new parameter estimates $\Theta[i]$ in terms of the old parameter

[4] This derivation can be found in detail in [82].

[5] Note that the MV Log-normal PDF can be easily transformed into an MV Gaussian PDF (cf. Section 5.4).

estimates $\Theta[i-1]$ are given by[6]

$$\begin{aligned}
\pi_m[i] &= \frac{1}{L_b}\sum_{l=1}^{L_b} p\left(m|\mathbf{y}_l, \Theta\left[i-1\right]\right), \\
\boldsymbol{\mu}_m[i] &= \frac{\sum_{l=1}^{L_b} \mathbf{y}_l p\left(m|\mathbf{y}_l, \Theta\left[i-1\right]\right)}{\sum_{l=1}^{L_b} p\left(m|\mathbf{y}_l, \Theta\left[i-1\right]\right)}, \\
\boldsymbol{\Sigma}_m[i] &= \frac{\sum_{l=1}^{L_b} p\left(m|\mathbf{y}_l, \Theta\left[i-1\right]\right)\left(\mathbf{y}_l - \boldsymbol{\mu}_m[i]\right)\left(\mathbf{y}_l - \boldsymbol{\mu}_m[i]\right)^H}{\sum_{l=1}^{L_b} p\left(m|\mathbf{y}_l, \Theta\left[i-1\right]\right)}.
\end{aligned} \tag{6.8}$$

These update equations perform the *expectation* and *maximization* step simultaneously. The EM algorithm starts with initial parameters estimates denoted by $\Theta[0]$ and iterates through the update equations in (6.8) until convergence. Convergence is reached, when the change of the cost function $Q\left(\Theta[i], \Theta\left[i-1\right]\right) - Q\left(\Theta[i-1], \Theta\left[i-2\right]\right)$ is less than a certain value. It is shown in [80, 81] that each iteration increases the cost function and that the EM algorithm is guaranteed to converge. However, the EM algorithm converges in general only to a local optimum, which implies that the obtained parameter estimates are not necessarily the ML parameter estimates. Which local optimum is reached, depends on the initial parameter estimates $\Theta[0]$.

6.3 Position Location and Clustering Systems

In this section, we outline the required steps of position location and clustering systems using the EM algorithm for joint region/cluster detection and parameter estimation.

6.3.1 Training Phase

As discussed above, the initial parameter estimates $\Theta[0]$ determine the convergence behavior and the outcome of the EM algorithm. It is assumed that the system gathers L training signals per region/cluster during a supervised training phase. These training signals are used to calculate $\Theta[0]$. This step is in analogy to the training phases discussed in Section 4.4.1 and Section 5.5.1. The objective is to reduce L as much as possible, while maintaining a certain $\mathcal{P}_e$ or $\mathcal{D}_e$.

In order to exploit the training signals not only for initialization of the EM algorithm but also for parameter estimation in each iteration step, it is proposed to modify the update

[6] A detailed derivation is provided in [82].

equations in (6.8). The idea is to include the initial parameter estimates $\Theta[0]$ in the update equations by exploiting the a priori knowledge of $p(m|\mathbf{y}_l, \Theta[i-1])$ for the training signals, which takes on the value 1, if the training signal $\mathbf{y}_l$ is caused by a TX located in region m, or the value 0, if the TX causing $\mathbf{y}_l$ is located in any other region. The modified update equations follow as

$$\begin{aligned}
\pi_m[i] &= \frac{1}{L \cdot M + L_b} \left(L + \sum_{l=1}^{L_b} p(m|\mathbf{y}_l, \Theta[i-1]) \right), \\
\boldsymbol{\mu}_m[i] &= \frac{L\boldsymbol{\mu}_m[0] + \sum_{l=1}^{L_b} \mathbf{y}_l p(m|\mathbf{y}_l, \Theta[i-1])}{L + \sum_{l=1}^{L_b} p(m|\mathbf{y}_l, \Theta[i-1])}, \\
\boldsymbol{\Sigma}_m[i] &= \frac{L\boldsymbol{\Sigma}_m[0] + \sum_{l=1}^{L_b} p(m|\mathbf{y}_l, \Theta[i-1]) (\mathbf{y}_l - \boldsymbol{\mu}_m[i]) (\mathbf{y}_l - \boldsymbol{\mu}_m[i])^H}{L + \sum_{l=1}^{L_b} p(m|\mathbf{y}_l, \Theta[i-1])}.
\end{aligned} \tag{6.9}$$

These update equations are obtained, when all $L \cdot M$ training signals are included in the set of observations $\mathcal{Y}$ giving a new set with $L_b + L \cdot M$ observations.

6.3.2 Localization Phase

During the localization phase, the system gathers and stores L_b location fingerprints, i.e. $\mathcal{Y}$, for agents with unknown positions. In case of remote positioning, these observations can be obtained from multiple agents at different positions or also from moving agents at different time instants. In case of self-positioning, these observations are obtained at one moving agent at different time instants. Using $\mathcal{Y}$ and $\Theta[0]$, the update equations in (6.9) are applied until convergence is reached at iteration $i^\star$. The resulting parameter estimates can be extracted from $\Theta[i^\star]$.

Clustering of the individual location fingerprints $\mathbf{y}_l$ is performed according to

$$\begin{aligned}
\hat{m} &= \underset{m \in 1,2,\ldots,M}{\text{argmax}} \; p(m|\mathbf{y}_l, \Theta[i^\star]) = \underset{m \in 1,2,\ldots,M}{\text{argmax}} \left(\frac{\pi_m[i^\star] f_m(\mathbf{y}_l|\Theta_m[i^\star])}{f(\mathbf{y}_l|\Theta[i^\star])} \right) \\
&= \underset{m \in 1,2,\ldots,M}{\text{argmax}} \; \pi_m[i^\star] f_m(\mathbf{y}_l|\Theta_m[i^\star]),
\end{aligned} \tag{6.10}$$

which is in analogy to (2.8). The PMFs $p(m|\mathbf{y}_l, \Theta[i^\star])$ for all M clusters and all L_b observations have been already computed by the EM algorithm.

Region detection of the individual location fingerprints $\mathbf{y}_l$ is performed in accordance

to (2.6):

$$\hat{m} = \underset{m=1,2,\ldots,M}{\operatorname{argmin}} \left(\sum_{j=1}^{M} \pi_j \left[i^\star \right] d_{m,j} f_j \left(\mathbf{y}_l | H_j \right) \right)$$
$$= \underset{m=1,2,\ldots,M}{\operatorname{argmin}} \left(\sum_{j=1}^{M} \pi_j \left[i^\star \right] d_{m,j} f_j \left(\mathbf{y}_l | \Theta_j \left[i^\star \right] \right) \right). \tag{6.11}$$

The number of observed location fingerprints L_b influences the performance of the EM algorithm. The larger L_b the better will be the parameter estimation quality after convergence. More accurate parameter estimates imply improved $\mathcal{P}_e$ and $\mathcal{D}_e$. Therefore, it would be beneficial to collect and store as many location fingerprints as possible before the EM algorithm is applied. There are, however, two drawbacks. First, the region/cluster information is delayed due to the required time for collecting observations. Second, collecting many location fingerprints requires large storage capacity. The *adaptive* EM algorithm described below mitigates these drawbacks.

Adaptive EM Algorithm Let us assume that L_b is limited due to the aforementioned reasons. After recording $\mathcal{Y}$ with the largest allowed L_b, the update equations in (6.9) are applied, which produces parameter estimates $\Theta\left[i^\star \right]$. Clustering or region detection is performed according to (6.10) or (6.11), respectively. If new location fingerprints are recorded the update equations in (6.9) can be applied once again using $\Theta\left[i^\star \right]$ as the new initial parameter estimates. Repeating this step results in an *adaptive* EM algorithm, which has the ability to track significant changes in the propagation environment like the relocation of large objects. The adaptive EM algorithm is outlined as follows:

1. Perform supervised training and measure L location fingerprints per region. Calculate initial parameter estimates $\Theta[0]$.
2. Measure observation set $\mathcal{Y}_1$ and apply (6.9) with initial parameters $\Theta[0]$ until convergence. The resulting parameter estimates are $\Theta_1\left[i_1^\star \right]$.
3. Measure observation set $\mathcal{Y}_2$ by overwriting $\mathcal{Y}_1$ and apply (6.9) with initial parameters $\Theta_1\left[i_1^\star \right]$ until convergence. The resulting parameter estimates are $\Theta_2\left[i_2^\star \right]$.
4. Continue ...

A performance study of the *adaptive* EM algorithm can be found in [78].

Multiple Observations per Agent Multiple location fingerprints per agent can also be used within the mixture density framework, in order to improve the position location performance

as proposed in Section 2.4.3. The corresponding mixture model in (6.1) must be modified in accordance to (2.11), which gives

$$f\left(\bar{\mathbf{y}}|\Theta\right) = \sum_{m=1}^{M} \pi_m f_m\left(\bar{\mathbf{y}}|\Theta_m\right) = \sum_{m=1}^{M} \pi_m \prod_{k=1}^{K} f_m\left(\mathbf{y}_k|\Theta_m\right).$$

This modification requires modified update equations, which can be derived following similar steps as in [82]. First, the cost function $Q\left(\Theta, \Theta\left[i-1\right]\right)$ needs to be modified according to

$$\begin{aligned}\bar{Q}\left(\Theta, \Theta\left[i-1\right]\right) = &\sum_{m=1}^{M}\sum_{l=1}^{\bar{L}_b} \ln\left(\pi_m\right) p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right) \\ &+ \sum_{m=1}^{M}\sum_{l=1}^{\bar{L}_b} \left(\sum_{k=1}^{K} \ln\left(f_m\left(\mathbf{y}_{k,l}|\Theta_m\right)\right)\right) p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right),\end{aligned}$$

where $\bar{\mathbf{y}}_l \triangleq \left[\mathbf{y}_{1,l}^T, \ldots, \mathbf{y}_{K,l}^T\right]^T$, $L_b = K \cdot \bar{L}_b$, and the modified posterior PMF of the m-th component PDF is given by

$$p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right) = \frac{\pi_m\left[i-1\right] \prod_{k=1}^{K} f_m\left(\mathbf{y}_{k,l}|\Theta_m\right)}{\sum_{j=1}^{M} \pi_j\left[i-1\right] \prod_{k=1}^{K} f_j\left(\mathbf{y}_{k,l}|\Theta_j\right)}.$$

The modified update equations follow as

$$\begin{aligned}\pi_m[i] &= \frac{1}{L \cdot M + \bar{L}_b}\left(L + \sum_{l=1}^{\bar{L}_b} p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right)\right), \\ \boldsymbol{\mu}_m[i] &= \frac{L\boldsymbol{\mu}_m[0] + \sum_{l=1}^{\bar{L}_b}\left(\sum_{k=1}^{K} \mathbf{y}_{k,l}\right) p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right)}{L + K\sum_{l=1}^{\bar{L}_b} p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right)}, \\ \boldsymbol{\Sigma}_m[i] &= \frac{L\boldsymbol{\Sigma}_m[0] + \sum_{l=1}^{\bar{L}_b} p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right)\left(\sum_{k=1}^{K}\left(\mathbf{y}_{k,l} - \boldsymbol{\mu}_m[i]\right)\left(\mathbf{y}_{k,l} - \boldsymbol{\mu}_m[i]\right)^H\right)}{L + K\sum_{l=1}^{\bar{L}_b} p\left(m|\bar{\mathbf{y}}_l, \Theta\left[i-1\right]\right)}.\end{aligned} \tag{6.12}$$

6.4 Experimental Performance Analysis

In this section, the performance of a localization or clustering system applying the EM algorithm is investigated. We consider a remote positioning scenario. The anchor node employs either a coherent RX or an ED RX. The figures of merit are $\mathcal{P}_e$ for the clustering prob-

lem and $\mathcal{D}_e$ for the position location problem. $\mathcal{P}_e$ and $\mathcal{D}_e$ are estimated using Monte Carlo simulations and the cross-validation method described in Section 4.6.1.

We consider the same measured CIRs as in Sections 4.6 and 5.7 from all four transmit antennas to receive antenna 3. The CIRs are normalized such that the average energy E_m is the same for all regions implying that the PL information is neglected.

The default system parameters are summarized in the following. If not mentioned otherwise, these values are used to calculate $\mathcal{P}_e$ and $\mathcal{D}_e$.

- Default SNR is the measurement SNR denoted by SNR_{mea} in Chapter 3.
- Default observation window size is $T_s = 30$ ns.
- Default number of regions is $M = 22$.
- Initial a priori probabilities are $\pi_m[0] = 1/M$ for all m.
- Coherent RX:
  - Default bandwidth is $B = 3$ GHz.
  - Default reference sample is $n_{\text{maxabs}} = 8$.
  - Default stochastic model for the location fingerprints is the COV model.
- ED RX:
  - Default sampling rate is $f_{\text{ed}} = 1$ GHz.
  - Default reference sample is $n_{\text{maxabs}} = 3$.
  - Default integration filter $g(t)$ is rectangular.
  - Default stochastic model for the location fingerprints is the MV Log-normal model.

6.4.1 Singular Initial Covariance Matrices

The initial empirical covariance matrices are rank deficient if $L < N$ as already discussed in Section 4.6.3.2. In order to calculate $p(m|\mathbf{y}_l, \Theta[0])$ in (6.7) the likelihoods of the location fingerprints $\mathbf{y}_l$ are required, which in turn requires the inversion of the initial covariance matrices.

This poses a problem especially to coherent RXs, because the location fingerprints have generally more samples than those of the ED RX. For example the default parameter setting of a coherent RX described in Section 4.6.2 implies a length of $N = 90$, which means that

$L > 90$ is a necessary - but in general not sufficient - condition for regular initial covariance matrices.

Two methods are proposed in the following, which cope with the problem of singular[7] initial covariance matrices.

1. Hybrid method: The idea is to use the PDP model instead of the COV model for initialization of the EM algorithm. The off-diagonal elements of the initial covariance matrices are ignored and only the diagonal elements are used. This ensures full rank by assuming independent components in $\mathbf{y}$. Note that this method can be also applied to the MV Log-normal PDF if the anchor employs an ED RX. For initialization the correlations among energy samples are ignored and the Log-normal PDF is applied.

2. Perturbation method: This method has been already discussed and applied in Section 4.6.3.2. The COV model with perturbation is used for initialization of the EM algorithm. A weighted identity matrix $\sigma_p^2\mathbf{I}$ is added to the singular initial covariance matrix. This ensures full rank of the overall covariance matrix. For the performance results in Fig. 6.1 the values of σ_p^2 are chosen such that SNRs of 28, 30, and 32 dB are emulated.

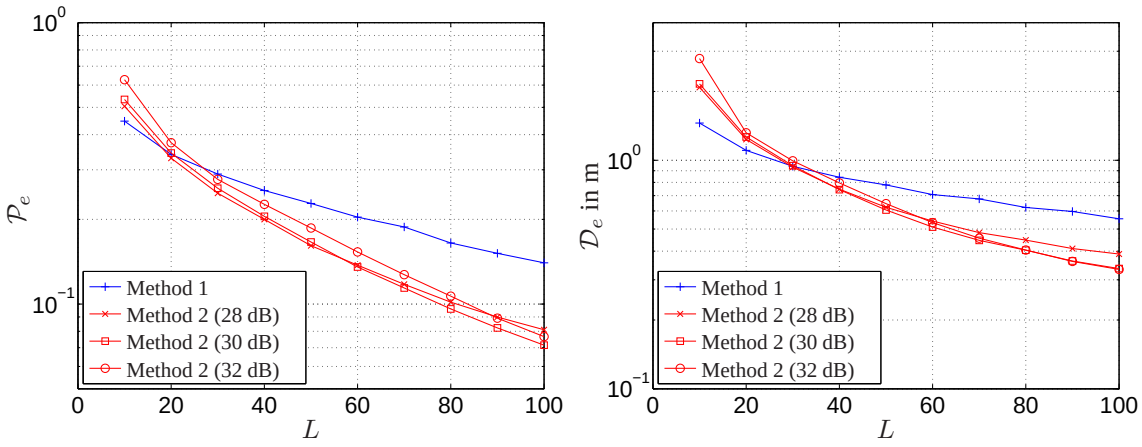

Fig. 6.1: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L for a coherent RX with default system parameters and $L_b = 6600$.

Fig. 6.1 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of L for a coherent RX with default parameters, the two aforementioned initialization methods. From each region, 300 location fingerprints are randomly selected, which results in $L_b = 22 \cdot 300 = 6600$. The hybrid method works better for $L \leq 20$, because the initial estimates of the covariance matrices are very imprecise

[7] This holds also for covariance matrices with very small eigenvalues, i.e. ill-conditioned matrices.

and their structures are mainly dominated by artifacts. Thus, it is better to ignore all off-diagonal elements and restrict the initial statistical region description to the PDP model. However, note that the EM algorithm assumes the COV model, which implies that except for initialization the full covariance matrices are used.

For L larger than 20, the perturbation method outperforms the hybrid method. The perturbation method requires the choice of the weight σ_p^2 of the perturbing identity matrix. A too large weight changes the initial covariance matrices too much, which results in an inaccurate stochastic description. However, a too small weight leads to numerical instabilities during matrix inversion due to very small eigenvalues. As aforementioned σ_p^2 is chosen such that SNRs of 28, 30, and 32 dB are emulated with the SNR definition from (4.2). It can be seen from Fig. 6.1 that the optimal weight depends on L. For larger L, smaller perturbation weights provide better performance, because the initial estimates of the covariance matrices become more accurate. This effect can be observed from Fig. 6.1. The curve for 28 dB cuts the curve for 30 dB at $L = 60$ and the curve for 32 dB at $L = 90$. Altogether, it can be seen that the performance is rather insensitive to the specific choice of σ_p^2. It can be concluded that a perturbation SNR of 30 dB provides a good compromise for all considered L.

6.4.2 Number of Training Signals L and Observations L_b

The performance of the EM approach is compared to the results from Chapters 4 and 5. The results of the COV model and the results of the Gamma model, each with $K = 1$ observation per agent, are used as reference for this comparison. The respective $\mathcal{P}_e$ and $\mathcal{D}_e$ reference curves can be also found in Fig. 4.14 and Fig. 5.13. The results of the PDP model from Fig. 4.14 are not shown here, because they are very similar to the performance of the Gamma model in Fig. 5.13. The Gamma model is chosen as reference for the ED RX, since it provides best performance for $L < 100$ according to Fig. 5.13.

Fig. 6.2 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of L. The curves obtained with a coherent RX are labeled with CO and the curves obtained with an ED RX are labeled with ED. The parameter L_b is set to $L_b = 2200$, respectively $L_b = 4400$, implying 100, respectively 200, observed location fingerprints per region. The perturbation method is applied for the CO results and σ_p^2 is chosen such that a perturbation SNR of 30 dB is emulated. The hybrid method is applied for the ED results due to better performance than the perturbation method. It can be observed from Fig. 6.2 that the EM approach achieves smaller $\mathcal{P}_e$ and $\mathcal{D}_e$ than the reference results. The improvement is more considerable for the coherent RX, because its performance is mainly limited by inaccurate parameter estimates due to few training signals.

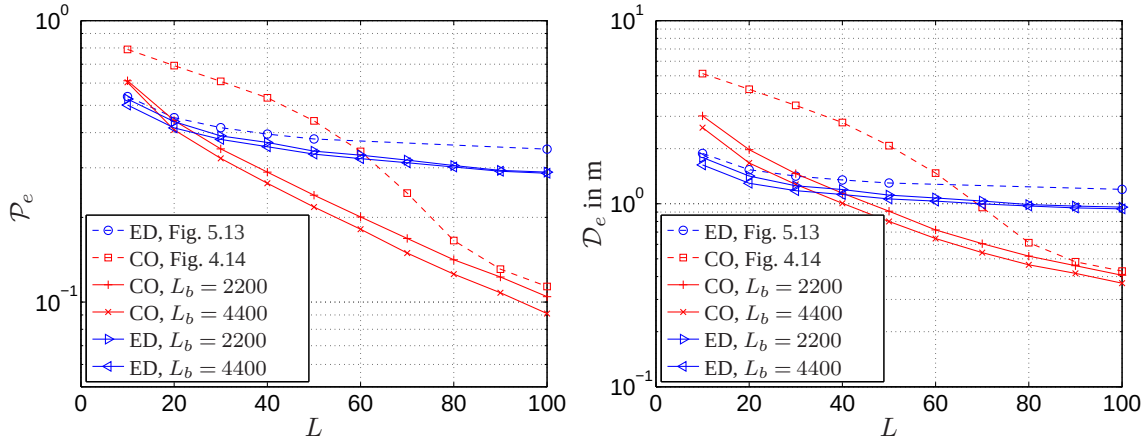

Fig. 6.2: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L.

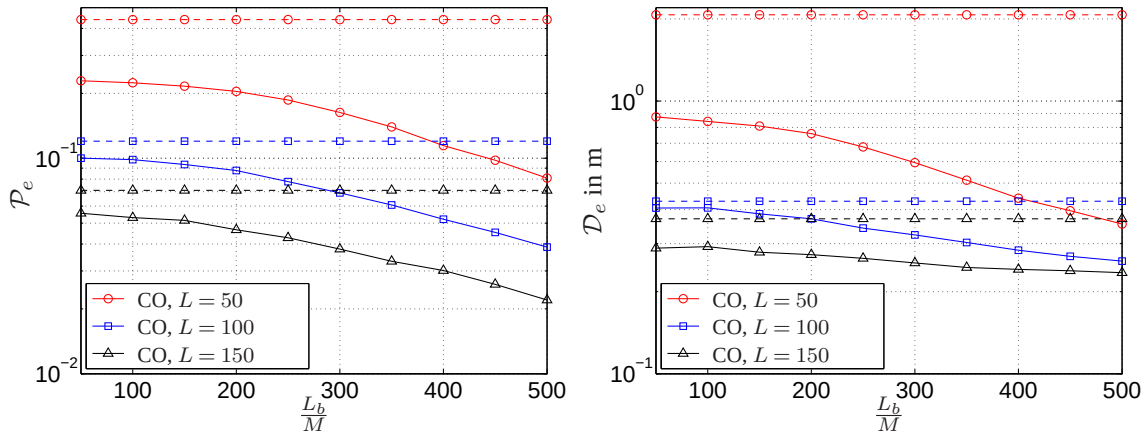

Fig. 6.3: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L_b. The dashed lines indicate the reference values of $\mathcal{P}_e$ and $\mathcal{D}_e$ as depicted in Fig. 4.14.

Fig. 6.3 depicts $\mathcal{P}_e$ and $\mathcal{D}_e$ as function of L_b for a coherent RX and different values of L. The dashed lines indicate the reference values of $\mathcal{P}_e$ and $\mathcal{D}_e$ (cf. Fig. 4.14) for the respective values of L. It can be observed that $\mathcal{P}_e$ and $\mathcal{D}_e$ can be significantly decreased, if the number of observations L_b is increased. This improvement is most significant for the smallest number of training signals, i.e. $L = 50$, because classification errors are mainly caused by the inaccurate parameter estimates in this case. It can be concluded that the EM approach will eventually provide reasonable $\mathcal{P}_e$ and $\mathcal{D}_e$ even for very few training signals as long as L_b is large enough.

6.4.3 Multiple Observations per Agent

In this section, we study the achievable performance improvement of the EM approach, if multiple observations per agent are available.

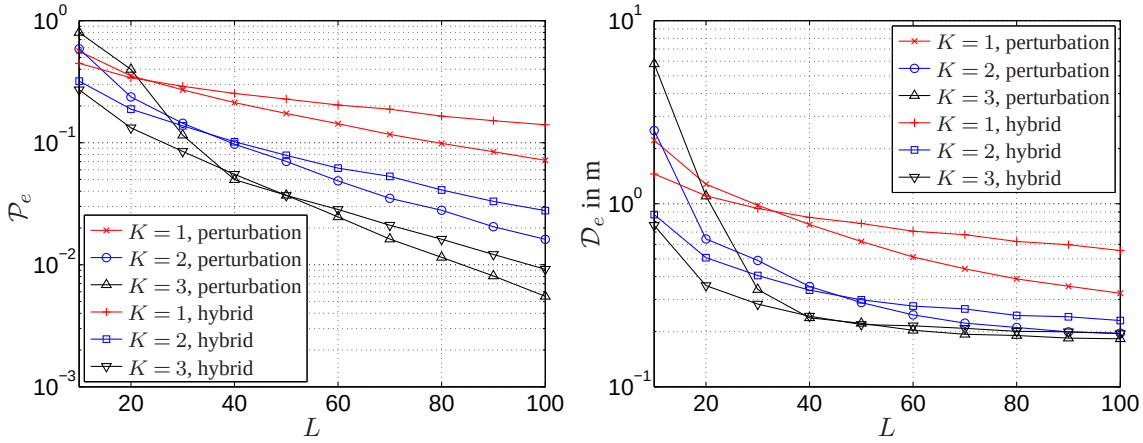

Fig. 6.4: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L for a coherent RX with default system parameters.

Fig. 6.4 shows the performance of a coherent RX with default system parameters as function of L for $K \in \{1,2,3\}$. The total number of observed location fingerprints from agents with unknown positions is kept constant at $L_b = 6600$, which means that $\bar{L}_b \in \{6600, 3300, 2200\}$. The update equations in (6.12) are applied until convergence is reached. Performance curves for both aforementioned initialization methods are plotted in Fig. 6.4. The perturbation SNR is set to 30 dB. For $L \leq 20$ the hybrid method works better than the perturbation method. This effect is emphasized with increasing K. Applying the perturbation method for $K = 3$ and $L \leq 20$ results in the worst performance. For small L only the hybrid method provides a consistent performance improvement with increasing K. As L increases the perturbation method outperforms the hybrid method for all K. Altogether, it can be observed that significant performance improvements are achievable by exploiting multiple observations per agent. Practical ways to obtain these observations are discussed in Section 2.4.3. The average positioning error can be reduced to approximately 25 cm using only $L = 40$ training signals per region and assuming that three location fingerprints per agent are available for position location. The smallest achievable average position location error for $L = 10$ training signals is around 76 cm for $K = 3$ and the hybrid initialization method.

Fig. 6.5 depicts the performance of an ED RX with default system parameters as function of L for $K \in \{1,2,3\}$. The EM algorithm is initialized by applying the hybrid method. A consistent performance improvement with increasing K for all values of L can be observed. A remarkable small average position location error of roughly 30 cm is achievable with $K = 3$ and $L = 20$. The smallest achievable average position location error for $L = 10$ training signals is approximately 40 cm for $K = 3$, which outperforms the coherent RX.

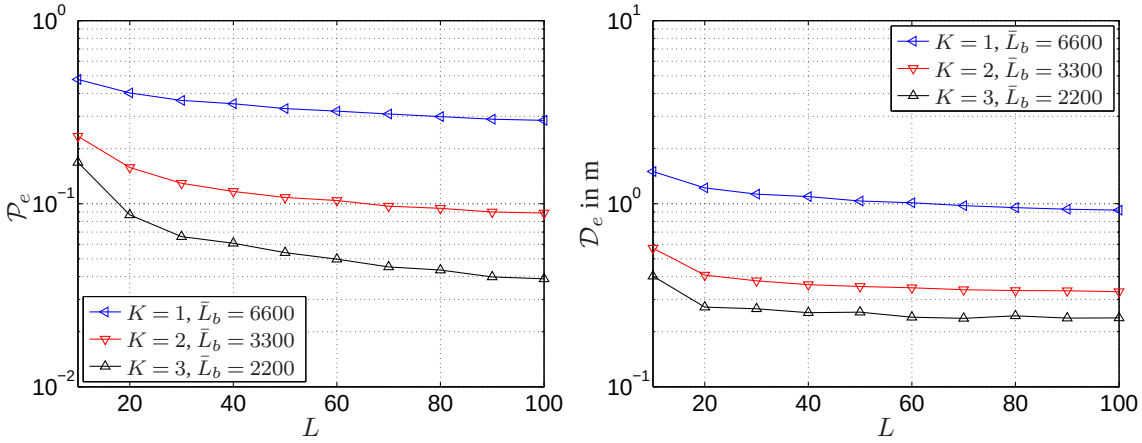

Fig. 6.5: $\mathcal{P}_e$ (left) and $\mathcal{D}_e$ (right) as function of L for an ED RX with default system parameters. Initialization of the EM algorithm is done with the hybrid method.

This is an important results, which states that the low complexity ED RX is the preferable RX structure if only few training signals are available. This behavior is caused by the different lengths of the respective location fingerprints.

6.5 Summary and Conclusions

The theoretical location fingerprinting framework introduced in Chapter 2 is extended by modeling the unconditional PDF of location fingerprints as a mixture of component PDFs. This modeling approach enables the application of the EM algorithm for joint parameter estimation and localization. Already classified location fingerprints are used to improve the quality of the parameter estimates, i.e. act as training signals. This theoretical framework is formulated in a generic way such that it can be applied to all location fingerprinting system, which are based on the Bayesian paradigm.

An empirical performance analysis demonstrates that the required number of training signals can be significantly reduced for both studied RX structures. However, it is also shown that a certain number of a training signals is still required for the initialization of the EM algorithm, in order to achieve acceptable position location and clustering performance. In the next chapter, a novel method for initial parameter estimation is proposed, which has the potential to reduce the complexity of the training phase even further.

Chapter 7

Efficient Training Phase

This chapter proposes an efficient estimation procedure of the parameters Θ_m of the conditional PDFs $f_m(\mathbf{y}|H_m)$. It is based on a simplified UWB channel model, which takes the geometry of the propagation environment into account. The inputs to this efficient training phase are either only very few[1] measured CRs[2] at known TX locations or a floor plan describing the geometry of the propagation environment. The outputs are predicted CRs for arbitrary TX positions within the region of interest, which are used to empirically estimate the parameters Θ_m. Note that virtually infinitely many training signals are available with this approach.

7.1 Electromagnetic Wave Propagation

Theoretically, it would be possible to calculate the electromagnetic field in a given volume of space at a specified time by solving Maxwell's equations. However, such a computation would require the knowledge of the electrical parameters of all materials in the environment and enormous processing power. Fortunately, such high modeling accuracy is not necessary for most engineering problems.

A mathematically convenient simplification is the multipath approximation based on the high-frequency approximation of electromagnetic wave propagation [83]. Its simplest form is called geometrical optics, where it is assumed that the electromagnetic field is composed out of partial waves traveling along geometric propagation paths (rays) and experiencing propagation effects during interaction with objects in the environment. This model provides accurate results as long as the interacting objects are electrically large, i.e. their dimensions

[1] It is shown in this chapter that at least three CRs are required for a two-dimensional localization problem.
[2] A CR is defined as convolution of a CIR with a transmit pulse.

are larger than the considered wavelengths. Furthermore, their electrical properties should remain approximately constant for the wavelengths under study. The wavelengths used by UWB systems range from 3 cm up to 10 cm, which implies that for example walls, floors, ceilings, and furniture can be considered as electrically large interacting objects.

In the following paragraphs the most important propagation effects are summarized. A more detailed discussion can be found in Chapter 4 of [52], Chapter 4 of [84] and in most textbooks on wireless communication.

Free Space Propagation and Antennas The propagation of electromagnetic waves in free space is well understood and is mathematically described in the far field by the Friis transmission equation, which relates the transmitted power P_{TX} to the received power P_{RX} via the free space PL according to

$$P_{\mathrm{RX}} = \mathrm{PL}\,(f,d)\,P_{\mathrm{TX}} = G_{\mathrm{TX}}\,(f)\left(\frac{c_0}{4\pi f d}\right)^2 G_{\mathrm{RX}}\,(f)\,P_{\mathrm{TX}},$$

where the antenna power gains $G_{\mathrm{TX}}\,(f)$ and $G_{\mathrm{RX}}\,(f)$ depend on the frequency f. In general the received power depends also on the orientation of the antennas, which is neglected in the above equation. The received power decreases with the square of the distance d and with the square of the frequency f.

Reflection and Transmission Reflection from and transmission through dielectric or conductive, electrically large objects can be very well modeled with the theory of geometrical optics. For indoor propagation especially walls, floor, ceiling, windows, and large furniture objects can be considered as reflecting objects. How much of the impinging field strength is reflected from the object and how much goes through the object is determined by the reflection and transmission coefficients, respectively. For example, the reflection coefficient of perfect conductors like metallic objects has magnitude one. The overall PL of a reflected partial wave is determined by the sum of the distances between TX, reflector, and RX.

In general, reflection and transmission coefficients are frequency dependent [53, 63, 85]. For example, the reflection coefficient of tempered glass is 0.65 for a frequency of 10.5 GHz and changes to 0.9 for a frequency of 7.5 GHz [86]. This implies that each reflection and transmission should be described by a linear filter with a frequency selective transfer function.

Diffraction Diffraction occurs at object corners, curvatures, wedges, and surface irregularities. This propagation effect cannot be modeled with the theory of geometrical optics. Therefore, the high-frequency approximation is extended with the uniform theory of diffraction [83]. A result of this theory is that the multipath approximation (propagation along rays) can still be applied. The propagation path goes via the interacting object (diffraction point) from TX to RX. Instead of a reflection coefficient, a diffraction coefficient is defined, which decreases with increasing frequency. Diffraction is also frequency dependent with the same implications as for reflection and transmission.

Scattering Scattering describes the interaction of an electromagnetic wave with an electrically small object, whose dimensions are in the order of the wavelength or smaller. Such an interacting object is also called point scatterer and can be considered as the source of a new electromagnetic wave propagating isotropically in all directions according to Huygens' principle. Therefore, the overall PL for point scatterers is determined by the product of the distances between TX, scatterer, and RX.

Diffuse Scattering The typical UWB propagation channel is not only composed out of partial waves caused by electrically large objects, which are well modeled by the multipath approximation, but also out of many components resulting from diffuse scattering or reflections from rough surfaces [52]. These components cannot be accurately described with the multipath approximation and are commonly modeled stochastically [87, 88].

Deterministic and Stochastic Channel Modeling The overall UWB propagation channel can be modeled as a combination of deterministic (concentrated) multipath components mainly caused by specular reflections and point scatterers, and of a stochastic component accounting for diffuse scattering and reflections from rough surfaces.

The UWB location fingerprinting system discussed so far assumes a pure stochastic description of the UWB channel. We will show in the remaining part of this chapter, how the existence of deterministic channel components can be exploited, in order to obtain estimates for the parameter sets Θ_m based on very few measured CRs.

7.2 Input Output Relations in Passband and Equivalent Baseband

We restrict our attention to the two-dimensional Euclidean space in the following considerations. However, all proposed concepts can be extended to three dimensions. The coordinate system is translated such that the position of the RX is in the origin $(0,0)$. The position of the TX is denoted by (x,y). We consider spatial variations of the TX position in the order of a few multiples of the largest considered wavelength of the UWB signal.

Applying the multipath approximation, we obtain an input output relation between the transmitted signal $s_c(t)$ and the received signal $r_c(t,x,y)$. The received signal is also called CR in this chapter. The subscript c stands for *carrier* and indicates that we are dealing with passband signals. Furthermore, the dependency of the received signal on the position of the TX is made explicit. The input output relation for passband signals follows as

$$r_c(t,x,y) = \sum_{l=1}^{L} \gamma_l(x,y) s_c^l(t-\tau_l(x,y)) + w_c(t), \tag{7.1}$$

where L is the total number of deterministic paths, $\gamma_l(x,y)$ and $\tau_l(x,y)$ are the gain and delay of the l-th path, and $w_c(t)$ is a realization of a stochastic process, which accounts for the stochastic channel component and distortions like thermal noise and time variations. The impulse responses of transmitting and receiving antenna, band selection filters, and amplifiers are comprised in $s_c(t)$. The path gain $\gamma_l(x,y)$ accounts for the loss in amplitude of the l-th partial wave due to traveled distance and the l-th interacting object. Due to frequency dependent propagation effects each partial waveform $s_c^l(t)$ will be slightly distorted, which is indicated by the superscript l. It is assumed without loss of generality that the path delays increase with l, i.e. $\tau_i(x,y) < \tau_j(x,y)$ for $i < j$.

For sake of completeness, the corresponding input output relation in equivalent baseband is given by

$$r(t,x,y) = \sum_{l=1}^{L} \gamma_l(x,y) \exp\left(-\imath 2\pi f_c \tau_l(x,y)\right) s^l\left(t-\tau_l(x,y)\right) + w(t),$$

where $w(t)$ and $s^l(t)$ are the equivalent baseband representations of $w_c(t)$ and $s_c^l(t)$, respectively. The carrier frequency or center frequency of $s_c(t)$ is denoted by f_c. It can be observed that the path delays influence the phases of the path gains of the equivalent baseband signal through the term $\exp\left(-\imath 2\pi f_c \tau_l(x,y)\right)$.

7.3 Analysis of Path Delays

In this section, the functional dependence (mapping) of the path delays on the TX position and the geometrical description of interacting objects is derived and analyzed. The knowledge of this mapping enables the prediction of path delays for arbitrary TX positions. Note that the roles of TX and RX are interchangeable due to channel reciprocity. Thus, also path delays for fixed TX position and arbitrary RX positions can be calculated this way.

The estimation of geometrical parameters, which describe the impact of interacting objects on the path delays, can be viewed as extraction of information about the propagation environment from the path delay structure of CIRs. Imaging of the propagation environment as described in [89, 90] is a related technique, which also extracts information of the propagation environment based on a set of CIR measurements obtained from a TX moving along a known trajectory. Time domain imaging algorithms are based on migration techniques [91] such as Kirchhoff migration [89, 90]. However, the goal of such imaging techniques is the generation of a complete image of the environment, which requires the estimation of the shape of the interacting objects. In contrast to this, we only seek to estimate parameters describing the influence of the interacting objects on the path delays, which is a considerably easier task.

7.3.1 Geometrical Description of Interacting Objects

Reflecting objects are approximated by straight lines (or planes) and, hence, can be described in the two-dimensional Euclidian space by their start points $(x_s, y_s)_l$ and end points $(x_e, y_e)_l$, which could be for example the corners of the room. Diffracting and scattering objects are described by the coordinates of the diffraction or scattering point denoted by $(x_d, y_d)_l$. The index l denotes the l-th partial wave, which is connected to the l-th interacting object. In this notation an interacting object could represent a sequence of interactions like for example multiple reflections.

In general the path delay $\tau_l(x, y)$ of the l-th path depends on the TX position (x, y) and the geometrical description of the l-th interacting object. For notational convenience the path index l is dropped in the following considerations.

7.3.2 Nonlinear Mapping of Transmitter Position on Path Delays

For a diffraction or scattering point with coordinates (x_d, y_d), the corresponding path delay is determined by the sum of the distances from TX to (x_d, y_d) and from (x_d, y_d) to the RX

according to

$$\tau(x,y) = \frac{1}{c_0}\left(\sqrt{(x+a)^2+(y+b)^2}+\sqrt{a^2+b^2}\right),$$

with the geometrical parameters $a = -x_d$ and $b = -y_d$.

The situation is slightly more complex for reflecting objects, which are approximated as straight lines in the following. The vector **r** describing a reflecting object and the vector **s** between its start point and the TX position are defined as

$$\mathbf{r} \triangleq \begin{bmatrix} x_e - x_s \\ y_e - y_s \end{bmatrix} \text{ and } \mathbf{s} \triangleq \begin{bmatrix} x - x_s \\ y - y_s \end{bmatrix}.$$

With these vectors the coordinates of the image point[3] (x_i, y_i) and consequently the path delay can be calculated according to

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = 2\frac{\mathbf{r}^T\mathbf{s}}{\|\mathbf{r}\|^2}\mathbf{r} - \mathbf{s} + \begin{bmatrix} x_s \\ y_s \end{bmatrix} \text{ and } \tau(x,y) = \frac{1}{c_0}\sqrt{x_i^2+y_i^2}.$$

Performing algebraic manipulations, the equation for the path delay can be reformulated as

$$\tau(x,y) = \frac{1}{c_0}\sqrt{(x+a)^2+(y+b)^2},$$

with the following geometrical parameters

$$a = \frac{2(y_e-y_s)(y_s x_e - x_s y_e)}{\|\mathbf{r}\|^2} \text{ and } b = -\frac{2(x_e-x_s)(y_s x_e - x_s y_e)}{\|\mathbf{r}\|^2}.$$

We conclude from these derivations that the functional dependence of the path delays on the geometrical description of scattering (diffracting) objects and specular reflecting objects can be stated as

$$\tau(x,y) = \frac{1}{c_0}\left(\sqrt{(x+a)^2+(y+b)^2}+d\right), \tag{7.2}$$

where d is either $\sqrt{a^2+b^2}$ for scattering and diffracting objects or zero for specular reflecting objects. Hence, the knowledge of the geometrical parameters a and b is sufficient to calculate the corresponding path delay for arbitrary TX positions.

The behavior of the first path $(l = 1)$ depends on the LoS condition of the propagation

[3] The image point is obtained by mirroring the source point (x,y) on the reflector.

channel. In case of LoS condition, the first path is equal to the direct path, which implies that $a_1 = b_1 = 0$. In case of non-LoS condition, the direct path might not exist and a_1 and b_1 describe already an interacting object.

Equation (7.2) is derived for a single interaction. It is expected that the functional dependence of τ on (x, y) stays as in (7.2) also for multiple interactions, but the corresponding geometrical parameters a, b, and d become more complicated and include the geometrical descriptions of all involved interacting objects.

Extension to Three Dimensions In this paragraph the functional dependence of the path delay τ on the TX position in the three-dimensional Euclidean space denoted by (x, y, z) is derived. In case of a diffraction or scattering point described by (x_d, y_d, z_d) the relationship is obtained as above according to

$$\tau(x,y,z) = \frac{1}{c_0}\left(\sqrt{(x+a)^2+(y+b)^2+(z+c)^2}+\sqrt{a^2+b^2+c^2}\right),$$

with the geometrical parameters $a = -x_d$, $b = -y_d$, and $c = -z_d$. Reflections in the three-dimensional space happen on surfaces, which are approximated as planes in the following. A reflecting plane is specified in the Hessian normal form, which requires the normal vector $\mathbf{n} \triangleq [n_x, n_y, n_z]^T$ with unit norm and the vector $\mathbf{r}$ from the origin to a point on the plane. Then the equation $\mathbf{n}^T\left([p_x, p_y, p_z]^T - \mathbf{r}\right) = 0$ defines the plane of points (p_x, p_y, p_z). The smallest distance of an arbitrary point (x, y, z) to this plane is given by $\left|\mathbf{n}^T\left([x, y, z]^T - \mathbf{r}\right)\right|$. With this the image point of (x, y, z) at the reflecting plane can be calculated according to

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \left(2\mathbf{n}^T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \mathbf{r}\right)\right)\mathbf{n}.$$

Performing again algebraic manipulations, the path delay equation is obtained as

$$\tau(x,y,z) = \frac{1}{c_0}\sqrt{(x+a)^2+(y+b)^2+(z+c)^2},$$

with the following geometrical parameters

$$\begin{aligned} a &= -2n_x\left(\mathbf{n}^T\mathbf{r}\right), \\ b &= -2n_y\left(\mathbf{n}^T\mathbf{r}\right), \\ c &= -2n_z\left(\mathbf{n}^T\mathbf{r}\right). \end{aligned}$$

7.3.3 Upper Bound on Path Delay Variation

A change of the position of the TX causes a variation of the delays of the multipath components depending on the geometry of the propagation environment. The smallest possible variation is zero and the largest possible variation determines the search interval for an algorithm, which should track path delay variations caused by TX movement.

Assume that the position of the TX is changed in x-direction from (x,y) to $(x+\Delta,y)$. The largest possible delay variation of each path is derived by

$$\begin{aligned}
&\max\ |\tau(x,y)-\tau(x+\Delta,y)| \\
&=\frac{1}{c_0}\max_{a,b}\left|\sqrt{(x+a)^2+(y+b)^2}-\sqrt{(x+a+\Delta)^2+(y+b)^2}\right| \\
&\leq\frac{1}{c_0}\max_{a,b}\left|\sqrt{(x+a-x-a-\Delta)^2+(y+b-y-b)^2}\right|=\frac{|\Delta|}{c_0},
\end{aligned} \tag{7.3}$$

which follows from the reverse triangle inequality. It remains to be shown that this upper bound is achievable, which is done by setting $b=-y$. Thus, the path delays corresponding to observation $r(t,x+\Delta,y)$ must satisfy the bounds

$$\tau(x,y)-\frac{|\Delta|}{c_0}\leq\tau(x+\Delta,y)\leq\tau(x,y)+\frac{|\Delta|}{c_0}.$$

These bounds are only achievable for $b=-y$, which implies that the y-coordinate of the interaction point (e.g. point scatterer) must be equal to the y-coordinate of the TX position. The same chain of consequences holds for the change of the TX position to $(x,y+\Delta)$.

In analogy to the derivation in (7.3), the path delay for the general observation $r(t,x+\Delta_x,y+\Delta_y)$ must satisfy

$$\tau(x,y)-\frac{\sqrt{\Delta_x^2+\Delta_y^2}}{c_0}\leq\tau(x+\Delta_x,y+\Delta_y)\leq\tau(x,y)+\frac{\sqrt{\Delta_x^2+\Delta_y^2}}{c_0}.$$

7.3.4 Linearization of Path Delays

Since we are interested in the behavior of the path delays in the vicinity of a reference point[4] (x_0,y_0), it is proposed to approximate (7.2) with a first order Taylor series expansion

[4]For the location fingerprinting application this reference point acts as a region's center point.

in the point (x_0, y_0) according to

$$\tau(x,y) = \tau(x_0,y_0) + \bar{\alpha}\,(x-x_0) + \bar{\beta}\,(y-y_0) + \mathrm{R}_1\,(x,y)\,, \tag{7.4}$$

with the linear coefficients $\bar{\alpha}$ and $\bar{\beta}$ as

$$\begin{aligned}
\bar{\alpha} &= \left.\frac{\partial\tau(x,y)}{\partial x}\right|_{(x_0,y_0)} = \frac{1}{c_0}\cdot\frac{x_0+a}{\sqrt{(x_0+a)^2+(y_0+b)^2}}\,,\\
\bar{\beta} &= \left.\frac{\partial\tau(x,y)}{\partial y}\right|_{(x_0,y_0)} = \frac{1}{c_0}\cdot\frac{y_0+b}{\sqrt{(x_0+a)^2+(y_0+b)^2}}\,.
\end{aligned} \tag{7.5}$$

The remainder term $\mathrm{R}_1\,(x,y)$ in (7.4) determines the accuracy of the linearization. According to Taylor's Theorem [92] this term can be written as

$$\mathrm{R}_1\,(x,y) = \frac{\left(y_0x - x_0y + b\,(x-x_0) - a\,(y-y_0)\right)^2}{2c_0\left((x_0+s\,(x-x_0)+a)^2+(y_0+s\,(y-y_0)+b)^2\right)^{\frac{3}{2}}}\,,$$

for some $s \in [0,1]$. If the reference point and the geometrical parameters a and b describing a specific interacting object are fixed, the respective linearization error for arbitrary TX positions can be upper bounded by maximizing $\mathrm{R}_1\,(x,y)$ with respect to $s \in [0,1]$. In order to illustrate the quality of this linearization for arbitrary geometries, we present numerically obtained values for the expected and maximum approximation errors as a function of the Euclidian distance $d_e = \sqrt{(x-x_0)^2+(y-y_0)^2}$ between (x,y) and (x_0,y_0). Therefore, the Cartesian coordinates (x,y) are substituted by polar coordinates according to $x = d_e\cos\phi + x_0$ and $y = d_e\sin\phi + y_0$ resulting in the error term $\mathrm{R}_1\,(d_e,\phi)$ as function of d_e and ϕ. The expected and maximum approximation errors are defined as

$$\begin{aligned}
\mathrm{C}_{\mathrm{exp}}\,(d_e) &\triangleq \frac{1}{J}\sum_{j=1}^{J}\mathrm{E}_\phi\left\{\left|\mathrm{R}_1^{(j)}\,(d_e,\phi)\right|\right\},\\
\mathrm{C}_{\mathrm{max}}\,(d_e) &\triangleq \max_{\substack{\phi\in[0,2\pi]\\ j\in\{1,2,\ldots,J\}}}\left|\mathrm{R}_1^{(j)}\,(d_e,\phi)\right|,
\end{aligned}$$

where the expectation and maximum are taken over the angle ϕ and J realizations of (x_0,y_0), a, and b. Each realization of (x_0,y_0), a, and b accounts for a certain reference point and a certain geometry of an interacting object. The ensemble of many of those realizations provide a statistic over all possible propagation geometries. The admissible realizations of a and b are restricted such that the Euclidian distance between the interacting objects and

(x_0, y_0) is larger than 0.3 m and that the Euclidian distance between the interacting objects and the RX is smaller than 5 m.

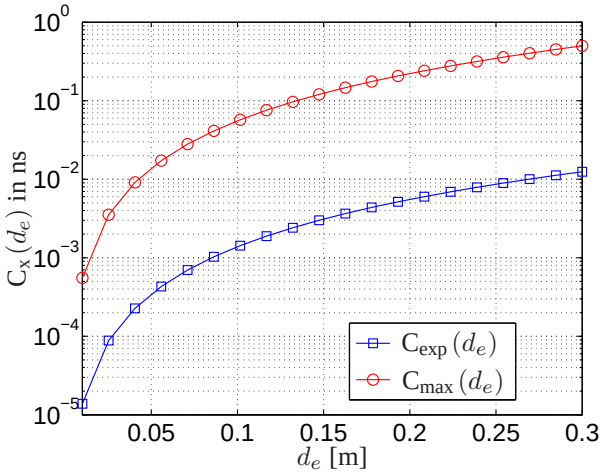

Fig. 7.1: Expected and maximum approximation errors in ns as a function of Euclidian distance to (x_0, y_0).

Fig. 7.1 depicts the expected and maximum approximation errors for $J = 10^6$ geometries. Further numerical results indicate that these curves are stable. Thus, the largest approximation error for position (x, y) with a distance to (x_0, y_0) of 0.3 m is 0.5 ns. However, the expected approximation error is around 0.01 ns at 0.3 m distance, which supports the applicability of the proposed linearization. It can be concluded that for most of the path delays a linear approximation is well suited to describe their behavior in vicinity of the reference point (x_0, y_0).

Analysis of the Linear Coefficients $\bar{\alpha}$ and $\bar{\beta}$ The linear coefficients $\bar{\alpha}$ and $\bar{\beta}$ can be upper and lower bounded according to

$$-\frac{1}{c_0} \leq \bar{\alpha} \leq \frac{1}{c_0} \text{ and } -\frac{1}{c_0} \leq \bar{\beta} \leq \frac{1}{c_0},$$

which follows from the bounds

$$-1 \leq \frac{x_0 + a}{\sqrt{(x_0+a)^2 + (y_0+b)^2}} \leq 1 \text{ and } -1 \leq \frac{y_0 + b}{\sqrt{(x_0+a)^2 + (y_0+b)^2}} \leq 1.$$

Furthermore, the condition $c_0^2\left(\bar{\alpha}^2 + \bar{\beta}^2\right) = 1$ follows from (7.5), which implies that $\bar{\alpha}$ and $\bar{\beta}$ cannot take on arbitrary values between $-\frac{1}{c_0}$ and $\frac{1}{c_0}$. For example if $\bar{\alpha} = \pm\frac{1}{c_0}$ then $\bar{\beta}$ must be

zero. These theoretical bounds and this condition are important for the design of algorithms, which should estimate $\bar{\alpha}$ and $\bar{\beta}$ from empirical data.

Relation of $\bar{\alpha}$ and $\bar{\beta}$ to the Angle of Departure The linear coefficients $\bar{\alpha}$ and $\bar{\beta}$ are related to the angle of departure of the corresponding multipath component. This relationship is analyzed in this paragraph.

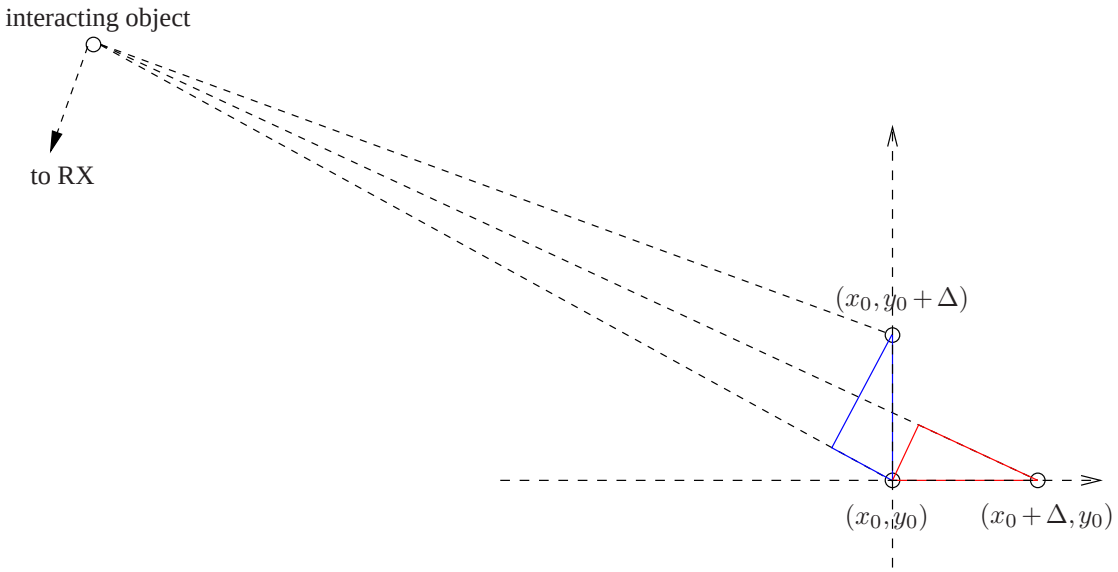

Fig. 7.2: Propagation paths from TX via interacting object to RX.

Fig. 7.2 depicts three propagation paths from three TX positions denoted by (x_0, y_0), $(x_0 + \Delta, y_0)$, and $(x_0, y_0 + \Delta)$ via an interacting object to the RX. The time required for the signal to travel to the RX for TX position (x_0, y_0) is denoted by $\tau(x_0, y_0)$.

Fig. 7.3 depicts a close up of Fig. 7.2. The two right-angled triangles define the angles of departure ϕ_x and ϕ_y with respect to x-axis and y-axis according to

$$\cos\phi_x = \frac{\Delta_a}{\Delta} \text{ and } \sin\phi_y = \frac{\Delta_o}{\Delta},$$

where Δ is the length of the hypotenuses of these triangles. The adjacent leg of ϕ_x has the length Δ_a and the opposite leg of ϕ_y has the length Δ_o.

If the distance from (x_0, y_0) to the interacting object is large compared to Δ, the plane wave assumption can be invoked, which states that the three propagation paths in Fig. 7.2 and Fig. 7.3 can be treated as parallel. In the close up in Fig. 7.3 it can be seen that these paths are indeed approximately parallel. Accordingly, the lengths Δ_a and Δ_o can be approximated

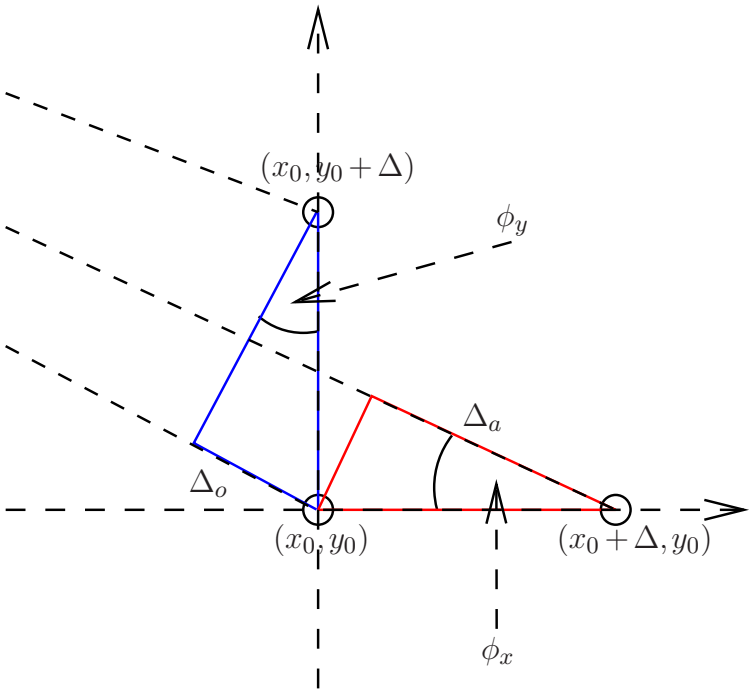

Fig. 7.3: Angles of departure ϕ_x and ϕ_y.

by

$$\Delta_a \approx c_0 \left(\tau(x_0 + \Delta, y_0) - \tau(x_0, y_0)\right) \text{ and } \Delta_o \approx c_0 \left(\tau(x_0, y_0) - \tau(x_0, y_0 + \Delta)\right).$$

Assuming plane waves, the angles ϕ_x and ϕ_y are equal and the approximations for Δ_a and Δ_o are exact, which implies

$$\cos\phi = \cos\phi_x = \frac{c_0 \left(\tau(x_0 + \Delta, y_0) - \tau(x_0, y_0)\right)}{\Delta},$$
$$\sin\phi = \sin\phi_y = \frac{c_0 \left(\tau(x_0, y_0) - \tau(x_0, y_0 + \Delta)\right)}{\Delta}.$$

Another consequence of the plane wave assumption is that the linearization in (7.4) becomes error-free, i.e. $\mathrm{R}_1(x, y) = 0$. Since the plane wave assumption implies parallel paths, the path delay for an arbitrary position (x, y) can be written as

$$\tau(x, y) = \tau(x_0, y_0) + \underbrace{\frac{\cos\phi}{c_0}}_{\bar{\alpha}} (x - x_0) + \underbrace{\left(-\frac{\sin\phi}{c_0}\right)}_{\bar{\beta}} (y - y_0),$$

which can be seen from Fig. 7.3. The relationships between $\bar{\alpha}$ and $\bar{\beta}$ and the angle of depar-

ture ϕ follow as

$$\bar{\alpha} = \frac{\cos\phi}{c_0} = \frac{\tau(x_0+\Delta, y_0) - \tau(x_0, y_0)}{\Delta} \text{ and } \bar{\beta} = -\frac{\sin\phi}{c_0} = \frac{\tau(x_0, y_0+\Delta) - \tau(x_0, y_0)}{\Delta}.$$

It can be concluded from these relationships that the plane wave assumption, applied for angle of departure or AoA estimation with antenna arrays, is in analogy to the linear approximation proposed in (7.4). The advantage of the linearization in (7.4) is the characterization of the error term $\mathrm{R}_1(x,y)$, which enables an accuracy analysis of the linearization. Furthermore, in literature there exist a lot of algorithms, which estimate ϕ with antenna arrays. Those can be applied, in order to obtain approximations for $\bar{\alpha}$ and $\bar{\beta}$. Note that we are interested in the angles of departure (or arrival) of deterministic multipath components.

7.3.5 Temporal Alignment

As already discussed in Section 3.6, the proposed location fingerprinting method does not require time synchronization between TX and RX. This implies that each received signal $r_c(t,x,y)$ has an unknown time reference. In order to remove this uncertainty, temporal alignment strategies have been proposed in Section 3.6. In this chapter, temporal alignment is performed by redefining the path delays according to

$$\begin{aligned}\tilde{\tau}_l(x,y) &\triangleq \tau_l(x,y) - \tau_r(x,y) \\ &= \frac{1}{c_0}\left(\sqrt{(x+a_l)^2+(y+b_l)^2} - \sqrt{(x+a_r)^2+(y+b_r)^2} + d_l - d_r\right),\end{aligned}$$

where $\tau_r(x,y)$ is the path delay of a reference path. Assuming LoS channel conditions, the reference path is defined as the direct path with $r=1$ and $a_1 = b_1 = 0$. For non-LoS channel conditions, the reference path might be defined as the strongest path (cf. Section 3.6.2), i.e. largest $|\gamma_l(x,y)|$. According to the definition of $\tilde{\tau}_l(x,y)$, the reference path of each received signal $r_c(t,x,y)$ arrives at time $t=0$ for all (x,y).

The maximal change of the redefined path delays for a translation of Δ_x in x-direction

and of Δ_y in y-direction can be computed based on (7.3) and follows as

$$\begin{aligned}
&\max\ |\tilde{\tau}_l(x,y) - \tilde{\tau}_l(x+\Delta_x, y+\Delta_y)| \\
&= \max\ |\tau_l(x,y) - \tau_l(x+\Delta_x, y+\Delta_y) + (\tau_r(x+\Delta_x, y+\Delta_y) - \tau_r(x,y))| \\
&\leq \max\ |\tau_l(x,y) - \tau_l(x+\Delta_x, y+\Delta_y)| + \max\ |\tau_r(x,y) - \tau_r(x+\Delta_x, y+\Delta_y)| \\
&= 2\frac{\sqrt{\Delta_x^2 + \Delta_y^2}}{c_0}.
\end{aligned}$$

Thus, the redefined path delays accounting for temporal alignment must satisfy

$$\tilde{\tau}_l(x,y) - 2\frac{\sqrt{\Delta_x^2 + \Delta_y^2}}{c_0} \leq \tilde{\tau}_l(x+\Delta_x, y+\Delta_y) \leq \tilde{\tau}_l(x,y) + 2\frac{\sqrt{\Delta_x^2 + \Delta_y^2}}{c_0}. \tag{7.6}$$

Applying the linear approximation from (7.4) to the redefined path delays results in

$$\tilde{\tau}_l(x,y) \approx \underbrace{\tau_l(x_0,y_0) - \tau_r(x_0,y_0)}_{\tilde{\tau}_l(x_0,y_0)} + \underbrace{(\bar{\alpha}_l - \bar{\alpha}_r)}_{\alpha_l}(x-x_0) + \underbrace{\left(\bar{\beta}_l - \bar{\beta}_r\right)}_{\beta_l}(y-y_0)\,, \tag{7.7}$$

where the redefined linear coefficients α_l and β_l are given by

$$\begin{aligned}
\alpha_l &= \bar{\alpha}_l - \bar{\alpha}_r = \frac{1}{c_0}\left(\frac{x_0 + a_l}{\sqrt{(x_0+a_l)^2 + (y_0+b_l)^2}} - \frac{x_0 + a_r}{\sqrt{(x_0+a_r)^2 + (y_0+b_r)^2}}\right), \\
\beta_l &= \bar{\beta}_l - \bar{\beta}_r = \frac{1}{c_0}\left(\frac{y_0 + b_l}{\sqrt{(x_0+a_l)^2 + (y_0+b_l)^2}} - \frac{y_0 + b_r}{\sqrt{(x_0+a_r)^2 + (y_0+b_r)^2}}\right).
\end{aligned} \tag{7.8}$$

These coefficients measure the linear change of the path delays with (x,y) relative to the linear change of the reference path with (x,y), which implies that $\alpha_r = \beta_r = 0$. Let us assume that $r = 1$ and $\alpha_2 = 0$ implying $\bar{\alpha}_2 = \bar{\alpha}_1$, which means that the linearized delays of the first and the second path change identically as function of x. Upper and lower bounds on α_l and β_l follow from the bounds on $\bar{\alpha}_l$, $\bar{\beta}_l$, $\bar{\alpha}_r$, and $\bar{\beta}_r$ according to

$$-\frac{2}{c_0} \leq \alpha_l \leq \frac{2}{c_0} \text{ and } -\frac{2}{c_0} \leq \beta_l \leq \frac{2}{c_0}. \tag{7.9}$$

7.4 Linearization of Path Gains

The path gains account for large scale effects and vary on a larger spatial scale than the path delays [53, 84]. Therefore, it is commonly assumed that the path gains stay approximately

constant for small changes of the TX position in the order of a few multiples of the carrier wavelength[5]. In this thesis it is proposed to apply a first order Taylor series expansion in order to approximate the path gains in the vicinity of a reference point (x_0, y_0) according to

$$\gamma_l(x,y) \approx \gamma_l(x_0,y_0) + \xi_l(x-x_0) + \zeta_l(y-y_0). \tag{7.10}$$

The model parameters are a bias term $\gamma_l(x_0,y_0)$, and the linear coefficients ξ_l and ζ_l, which account for the linear increase or decrease of the l-th path gain as the TX moves along the x and y direction. In order to be able to predict the path gains for arbitrary TX positions, these three parameters must be known for all L paths.

7.5 Simplified Input Output Relation

Inserting the linear models of the path gains in (7.10) and the path delays accounting for temporal alignment in (7.7) in the input output relation in (7.1), the CR can be approximated according to

$$\begin{aligned} r_c(t,x,y) \approx \sum_{l=1}^{L} &(\gamma_l(x_0,y_0) + \xi_l(x-x_0) + \zeta_l(y-y_0)) \\ &\cdot s_c^l(t - \tilde{\tau}_l(x_0,y_0) - \alpha_l(x-x_0) - \beta_l(y-y_0)) + w_c(t), \end{aligned} \tag{7.11}$$

where the $6L$ linear model parameters are collected in the parameter set

$$\Upsilon_{\mathrm{LM}} = \left\{ \gamma_l(x_0,y_0), \xi_l, \zeta_l, \tilde{\tau}_l(x_0,y_0), \alpha_l, \beta_l \middle| l = 1,2,\ldots,L \right\}.$$

The knowledge of Υ_{LM} and the signals $s_c^l(t)$ is sufficient to predict the deterministic part of the CR caused by a TX located in the vicinity of a reference point (x_0, y_0). The accuracy of this prediction depends on the accuracy of the input output relation (7.1), the linearization accuracy and the accuracy of the estimates of Υ_{LM}. In the following section, the estimation of Υ_{LM} is discussed.

[5] Note that fading due to multipath interference does not influence the individual path gains. This effect is caused after summation over all L multipath components.

7.6 Estimation of Linear Model Parameters

The estimation of Υ_{LM} is partitioned into two steps. First, the path gains and path delays of measured CRs at K known TX positions $\{(x_0, y_0), (x_1, y_1), \ldots, (x_{K-1}, y_{K-1})\}$ are estimated. The joint estimation of path gains and path delays is discussed in Section 7.6.1. Then the linear model parameters are estimated from the obtained path gains and delays via the least squares method. This step is discussed in Section 7.6.2. The flow diagram in Fig. 7.4 illustrates the estimation procedure.

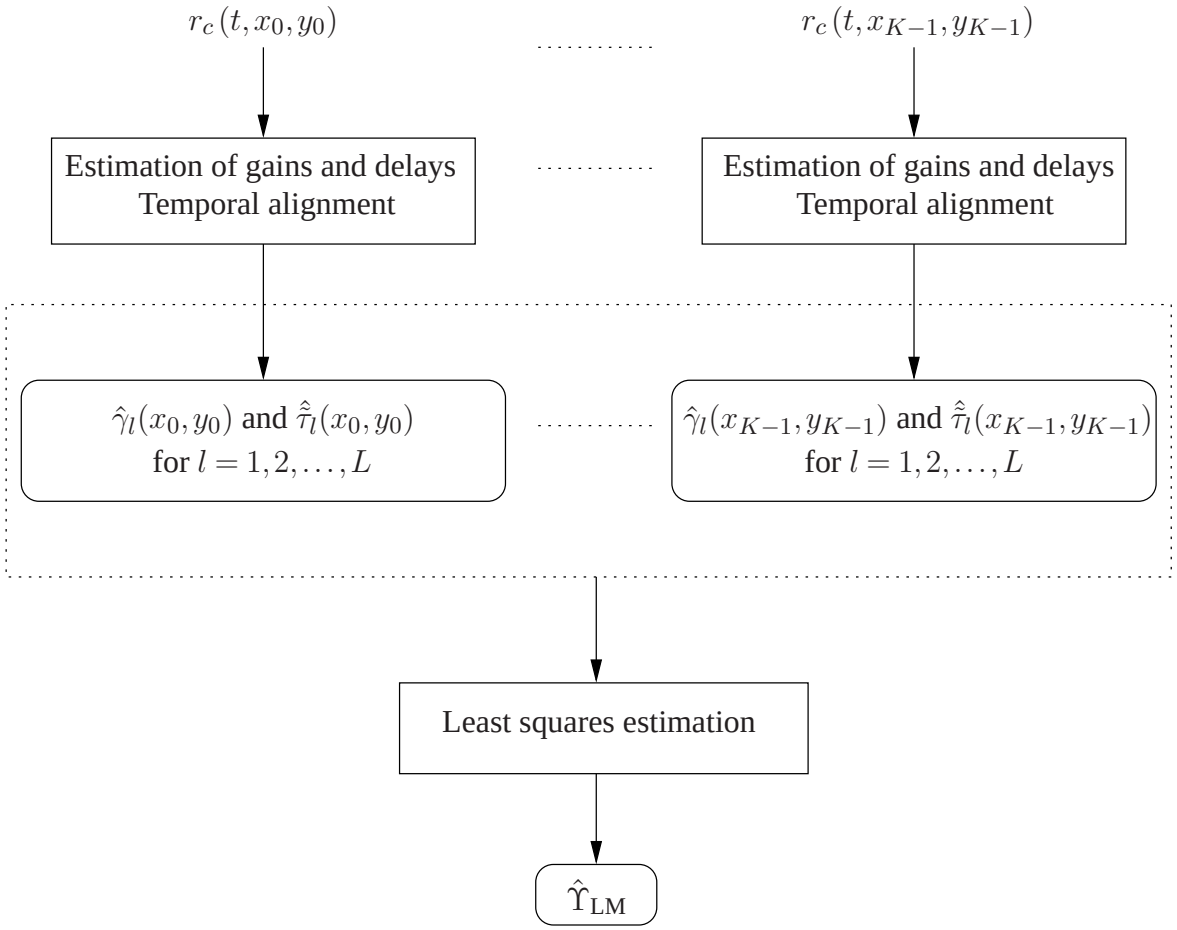

Fig. 7.4: Estimation procedure of Υ_{LM}.

7.6.1 Joint Estimation of Path Gains and Path Delays

This section discusses the joint estimation of path gains and path delays based on a measured CR at a fixed TX position. The dependence of the CR on (x, y) is neglected in the following

for notational convenience. The received signal or CR for fixed TX and RX positions is therefore given by

$$r_c(t) = \sum_{l=1}^{L} \gamma_l s_c(t-\tau_l) + w_c(t), \tag{7.12}$$

where it is additionally assumed that the interactions are frequency independent resulting in undistorted replicas of $s_c(t)$. Most algorithms proposed in literature, which try to estimate the path gains and path delays, are based on the input output relation in (7.12) and assume undistorted replicas of $s_c(t)$. The estimation algorithm used to calculate simulation results in this thesis uses this assumption likewise. The impact of frequency dependent interactions on the estimation quality is further discussed in Section 7.7.

The parameter set $\psi_l \triangleq \{\gamma_l, \tau_l\}$ collects path gain and path delay of the l-th path. The objective is to estimate the parameter set $\Psi \triangleq \{\psi_1, \psi_2, \ldots, \psi_L\}$ based on the observation of $r_c(t)$ in an interval $[0,T]$. For ML parameter estimation, the log-likelihood function of Ψ given $r(t)$ denoted by $\Lambda(\Psi|r_c(t))$ is required. Assuming a zero mean and white Gaussian process $w_c(t)$ with power spectral density $N_0/2$, the log-likelihood function is derived in [93] and is given by

$$\Lambda(\Psi|r_c(t)) = \frac{1}{N_0}\left(2\int_0^T \sum_{l=1}^{L} \gamma_l s_c(t-\tau_l)\, r_c(t)\mathrm{d}t - \int_0^T \left(\sum_{l=1}^{L} \gamma_l s_c(t-\tau_l)\right)^2 \mathrm{d}t\right).$$

The corresponding ML parameter estimation rule follows as

$$\hat{\Psi} = \underset{\Psi}{\operatorname{argmax}} \left(\Lambda(\Psi|r_c(t))\right). \tag{7.13}$$

This optimization problem is computationally difficult due to the large number of parameters, i.e. due to the high dimensionality of Ψ. There exist a number of iterative algorithms like the space-alternating generalized expectation-maximization (SAGE) algorithm [94], or the weighted Fourier transform and relaxation (WRELAX) algorithm [95], which are computationally efficient and provide nearly optimal results. The main idea of these algorithms is to transform the high-dimensional optimization problem to a series of one-dimensional optimization problems, which are then iteratively solved.

For the results presented in this thesis, we apply the WRELAX algorithm, which estimates the parameter set Ψ by minimizing the nonlinear least squares cost function. This corresponds to ML parameter estimation under the assumption of a zero mean and white

Gaussian process $w_c(t)$. The inputs to this algorithm are sampled versions of the transmitted signal $s_c(t)$ denoted by $\mathbf{s}_c$ and the received signal $r_c(t)$ denoted by $\mathbf{r}_c$, the number of expected paths L, and the sampling rate f_s. The output is $\hat{\Psi}$. In practice it is very important to have an accurate measurement of the transmitted signal $s_c(t)$. For details about the implementation of this algorithm and further information the interested reader is referred to [95].

It should be noted that the WRELAX and the SAGE algorithm are based on the input output relation in (7.12) and, therefore, neglect the impact of frequency dependent interactions. The design and application of robust algorithms accounting for frequency dependent interactions is outside the scope of this thesis but recommended for future work in order to improve the estimation performance. For example, in [85] the CLEAN algorithm [96] is modified and used to estimate path gains, path delays, and angle of arrivals under the assumption of frequency dependent interactions. This could be a starting point for future work.

7.6.2 Least Squares Estimation of $\Upsilon_{\mathbf{LM}}$

In this section, least squares estimation of the linear model parameters Υ_{LM} based on estimated path gains and delays at K TX positions is proposed. We discuss only the estimation of the parameters for the linear model of the l-th path gain, since all other estimates follow analogously. The linear model for the l-th path gain is repeated for convenience as

$$\gamma_l(x,y) = \gamma_l(x_0,y_0) + \xi_l(x-x_0) + \zeta_l(y-y_0).$$

In general, the estimates of the path gains and delays will be erroneous, which is accounted for by extending the linear model with an error term according to

$$\hat{\gamma}_l(x,y) = \gamma_l(x,y) + \epsilon = \gamma_l(x_0,y_0) + \xi_l(x-x_0) + \zeta_l(y-y_0) + \epsilon,$$

where $\hat{\gamma}_l(x,y)$ denotes the estimate of $\gamma_l(x,y)$ and ϵ denotes the estimation error. This error is caused by the stochastic component $w_c(t)$ of the measured CR and by limitations of the path gains and delays estimation algorithm. If ϵ can be modeled as a random variable, which would require a model selection procedure like AIC using a large set of empirical estimation errors, the ML estimates for $\gamma_l(x_0,y_0)$, ξ_l, and ζ_l can be derived. Further assuming a prior distribution for the linear model parameters would allow for a Bayesian linear regression as discussed in Chapter 3.3 of [32]. In this section, least squares estimation is proposed, which corresponds to ML parameter estimation assuming a Gaussian PDF for ϵ.

Let us assume that K estimates $\hat{\gamma}_l(x_k, y_k)$ for $k = 0, \ldots, K-1$ at known TX positions (x_k, y_k) in the vicinity of the reference point (x_0, y_0) are available. The joint estimation of path gains and path delays from a measured CR has been discussed in Section 7.6.1. The corresponding sum-of-squares error function [32] is given by

$$R_{\mathrm{SQ}}(\mathbf{w}_l) = \frac{1}{2} \sum_{k=0}^{K-1} \left(\hat{\gamma}(x_k, y_k) - \mathbf{w}_l^T \mathbf{p}_k \right)^2,$$

where the linear model parameters are collected in the vector $\mathbf{w}_l^T \triangleq [\gamma_l(x_0, y_0), \xi_l, \zeta_l]$ and the vector $\mathbf{p}_k$ is defined as $\mathbf{p}_k \triangleq [1, x_k - x_0, y_k - y_0]^T$. Calculating the gradient of $R_{\mathrm{SQ}}(\mathbf{w}_l)$ with respect to $\mathbf{w}_l$, setting it to zero, and solving for $\mathbf{w}_l$ gives the least squares estimate $\mathbf{w}_l^\star$ according to

$$\mathbf{w}_l^\star = \underbrace{\left(\mathbf{P}^T \mathbf{P} \right)^{-1} \mathbf{P}^T}_{\mathbf{P}^\dagger} \mathbf{t}_l, \tag{7.14}$$

where the vector $\mathbf{t}_l$ and the matrix $\mathbf{P}$ are defined by

$$\mathbf{t}_l \triangleq [\hat{\gamma}_l(x_0, y_0), \ldots, \hat{\gamma}_l(x_{K-1}, y_{K-1})]^T \text{ and } \mathbf{P} \triangleq \begin{bmatrix} \mathbf{p}_0^T \\ \vdots \\ \mathbf{p}_{K-1}^T \end{bmatrix}.$$

If the estimation of the linear model parameters of the l-th path delay is considered, then the vector $\mathbf{t}_l$ is defined by $\mathbf{t}_l \triangleq \left[\hat{\tilde{\tau}}_l(x_0, y_0), \ldots, \hat{\tilde{\tau}}_l(x_{K-1}, y_{K-1}) \right]^T$. The matrix $\mathbf{P}^\dagger$ is known as the Moore-Penrose pseudo-inverse of $\mathbf{P}$. The matrix $\mathbf{P}$ has dimensions K times 3, which implies that $K \geq 3$ measurement locations are required in order to ensure that a unique solution $\mathbf{w}_l^\star$ exists. For $K = 3$ the pseudo-inverse $\mathbf{P}^\dagger$ is equal to the inverse of $\mathbf{P}$.

Applying (7.14) to all $2 \cdot L$ vectors $\mathbf{t}_l$ collecting the K estimated path gains and path delays provides the least squares estimates for all linear model parameters Υ_{LM}. The pseudo-inverse $\mathbf{P}^\dagger$ needs to be calculated only once. Note that this formulation assumes that the K estimated path gains and path delays at different TX positions stacked into $\mathbf{t}_l$ correspond to the *same* multipath component, i.e. are caused by the same interacting object. The problem of identifying associated path delays and path gains is called *path pairing* problem and is further discussed in the next section.

7.6.3 The Path Pairing Problem

In Section 7.6.1 the joint estimation of gains and delays of L multipath components from a measured CR is discussed. In Section 7.6.2 least squares estimation of the linear model parameters from a set of estimated path gains and path delays at K different TX positions is proposed. However, before least squares estimation can be applied, path pairing or path association is required. Assume that L path gains and path delays denoted by $\hat{\gamma}_l(x_k, y_k)$ and $\hat{\tau}_l(x_k, y_k)$ have been estimated at K known TX positions, where $l = 1, 2, \ldots, L$ and $k = 1, 2, \ldots, K$. In order to obtain meaningful results from least squares estimation, it has to be ensured that $\hat{\gamma}_l(x_k, y_k)$ and $\hat{\tau}_l(x_k, y_k)$ are associated to the same interacting object with index l for all k.

For simplicity let us assume $L = 2$ multipath components and $K = 2$ TX positions resulting in a set of four path delays and four path gains. The objective of path pairing is to group this set into two sets consisting of two path gains and delays associated to the same interacting object. In practice the following problems can arise:

1. At the second TX position a multipath component is detected from a new interacting object, which has not been detected at the first TX position. This can happen, when one of the detected paths at the first TX position is blocked at the second TX position. This can also happen due to fading effects caused by the influence of other multipath components, which lead to existence (strong gain) and non-existence (weak gain) of paths. This problem causes usually a large difference between the estimated path delays at different TX positions. Thus, by applying the bounds on the path delays derived in (7.3) or (7.6) this estimation error can be identified provided the two TX positions are close to each other.

2. If the two path delays at each TX position are very similar, i.e. the difference is smaller than the bounds in (7.3) and (7.6), then it is not possible to associate the paths properly. In such a case it is likely that the path delays as function of the TX position are intersecting. Thus, for other TX positions the difference between the path delays should increase again. If the difference does not change with the TX position then both identified paths could be associated to a single interacting object.

In a dense multipath environment these two effects make path pairing cumbersome due to the large number of multipath components and diffuse scattering. Increasing the bandwidth of $s_c(t)$, i.e. increasing the temporal resolution, decreases the probability of multipath fading and mitigates the first problem. However, increasing the bandwidth of $s_c(t)$ increases the

impact of frequency dependent interactions, which causes larger estimation errors of the SAGE and WRELAX algorithm.

The following paragraph proposes a technique borrowed from digital image processing, which mitigates the need for path pairing.

The Hough Transform A promising solution to the path pairing problem is the Hough transform, which is a feature extraction technique used in digital image processing [97, 98]. The purpose of this method is to find geometrical shapes in an image by a voting procedure. Originally, the Hough transform was designed to detect straight lines but later on it was extended to detect also more complex shapes like circles or ellipses.

In order to be able to apply the Hough transform, the estimation of the linear model parameters Υ_{LM} from a set of estimated path gains and path delays at different TX positions is reformulated as straight line detection problem. The TX antenna is moved along the x-axis such that CRs are measured at K known grid points denoted by (x_0+x_k, y_0). The WRELAX algorithm[6] is applied to each CR producing $K \cdot L$ path gains and delays, which are given by

$$\gamma_l(x_0+x_k, y_0) = \gamma_l(x_0, y_0) + \xi_l x_k \quad \text{for} \quad l = 1, \ldots, L \quad \text{and} \quad k = 1, \ldots, K,$$
$$\tilde{\tau}_l(x_0+x_k, y_0) = \tilde{\tau}_l(x_0, y_0) + \alpha_l x_k \quad \text{for} \quad l = 1, \ldots, L \quad \text{and} \quad k = 1, \ldots, K.$$

These equations describe $K \cdot L$ points on L lines in the $(\gamma_l(x, y_0), x)$ plane with slope parameters ξ_l and intercept parameters $\gamma_l(x_0, y_0)$, and $K \cdot L$ points on L lines in the $(\tilde{\tau}_l(x, y_0), x)$ plane with slope parameters α_l and intercept parameters $\tilde{\tau}_l(x_0, y_0)$.

Each straight line in the $(\tilde{\tau}_l(x, y_0), x)$ plane can be represented as a single point $(\tilde{\tau}_l(x_0, y_0), \alpha_l)$ in the corresponding parameter space. The case of vertical lines implying infinite slope parameters is ruled out by the bounds on α_l given in (7.9). The Hough transform algorithm uses a two-dimensional accumulator array with quantized values of the parameters α and $\tilde{\tau}(x_0, y_0)$. For each pair $(\tilde{\tau}_l(x_k, y_0), x_k)$ the bins in this array corresponding to possible line parameters are increased by one. After processing all $K \cdot L$ estimated path delays, the bins with the highest values indicate the most likely line parameters. For more details on the Hough transform, the interested reader is referred to [97, 98].

The advantage of the Hough transform over least squares estimation is that path pairing is not required. The drawback is that the Hough transform is less robust against estimation errors than least squares estimation.

[6] Any other algorithm, which jointly estimates the path gains and path delays, can also be applied.

7.7 Experimental Analysis of Prediction Accuracy

In order to assess the quality of CR prediction based on the proposed linear approximations of the path gains and path delays, dedicated CR measurements in an anechoic chamber with two metallic reflectors are performed. Note that path pairing is not an issue in this experiment due to the controlled propagation environment with only three multipath components.

7.7.1 Measurement Campaign in Anechoic Chamber

The time-domain correlation method described in Chapter 3 is used to estimate CRs. Fig. 7.5 shows the block diagram of the measurement setup. Information about the individual hardware blocks can be found in Table 7.1.

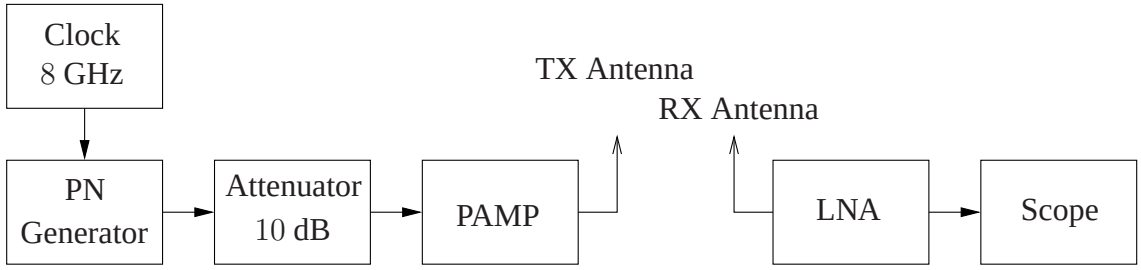

Fig. 7.5: Block diagram of measurement setup.

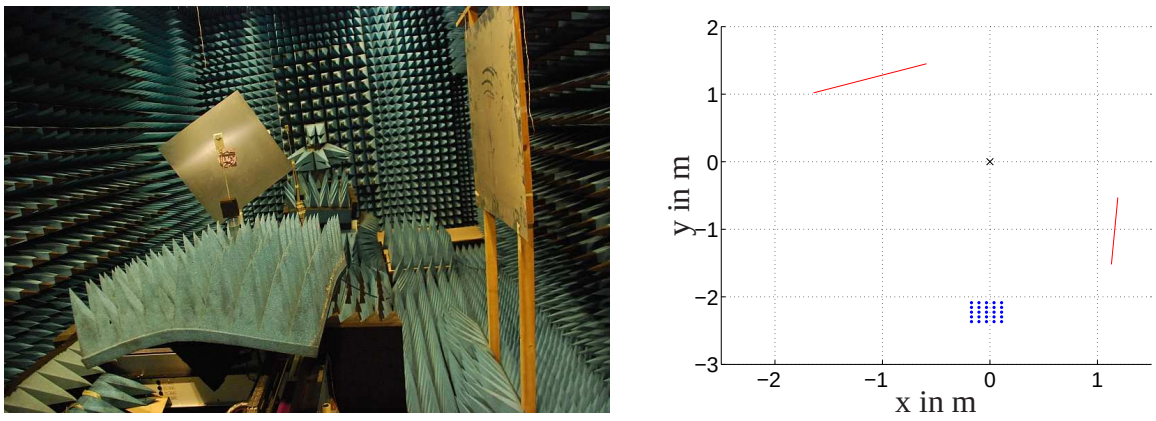

Fig. 7.6: Measurement scenario with two metallic reflectors, fixed receive antenna, and transmit antenna mounted on a positioning device.

The measurement scenario (left) and the corresponding two-dimensional floor plan (right) are depicted in Fig. 7.6. The cross marks the position of the receive antenna at the origin and the dots mark some of the positions of the transmit antenna. The same positioning device

Device	Parameters	Values
Tektronix DPO 72004	sampling rate input filter bandwidth horizontal resolution vertical resolution	50 GS/s 0 - 6 GHz 1 μs/div 200 mV/div
Skycross Antennas SM3TO10MA	height	1.95 m
PN Generator Centellax TG1P1A	bit period order PN sequence duration	1/8 ns 15 ≈ 4096 ns
PAMP Mini Circuits ZVE-8G	gain bandwidth 1 dB compression point	30 dB 2 - 8 GHz 30 dBm
LNA Mini Circuits AFS5	gain bandwidth noise figure 1 dB compression point	42 dB 0.1 - 8 GHz 1.4 dB 10 dBm

Table 7.1: Hardware equipment and parameters.

as for the measurements described in Chapter 3 is used. The transmit antenna is moved on a 1 cm grid over an area of 0.28 m times 0.28 m. The total number of measured CRs is $29 \cdot 29 = 841$. The two metallic reflectors are indicated with lines in the floor plan. The measurement SNR is approximately 70 dB with the SNR definition in (4.2).

In order to get reliable results from the WRELAX algorithm the transmit signal $s_c(t)$ must be known. Note that this signal comprises the impulse responses of cables, 10 dB attenuator, PAMP, transmit antenna, receive antenna, and LNA. The signal $s_c(t)$ is also measured in the anechoic chamber *without* reflectors. In order to mitigate any remaining artifacts caused by the RX antenna mounting and the positioning device, the signal $s_c(t)$ is averaged over all 841 grid positions. Note that this averaging requires temporal alignment.

Fig. 7.7 depicts the averaged transmit signal $s_c(t)$ with the dashed curve and $r_c\,(t,(x_0,y_0))$ with the solid curve. Note that these are passband signals with a frequency range from approximately 2 to 6 GHz. The sampling frequency is 50 GHz. For visualization purpose the signals are scaled such that the maximum absolute value of both is one. The reference point (x_0,y_0) is chosen as the center point of the transmit grid depicted in Fig. 7.6. The exact coordinates of this center point are $(x_0,y_0) = (-0.032,-2.227)$ m. The two reflected paths from the metallic reflectors are clearly visible in Fig. 7.7. The second path arrives approximately 3 ns after the direct path (first path) and is associated to the reflector almost parallel to the y-axis. The third path arrives approximately 10 ns after the first path and is

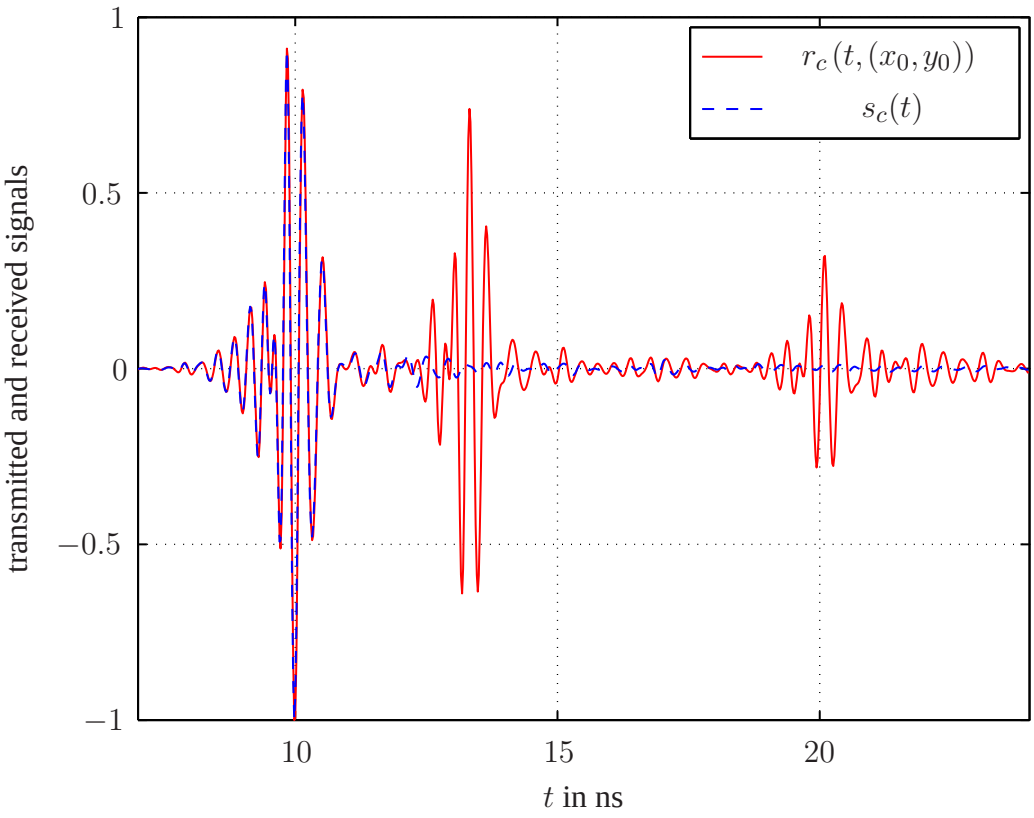

Fig. 7.7: Measured received and transmit signals in anechoic chamber with two reflectors.

associated to the other reflector. Note that the arrival of the direct path, which is defined as the maximum absolute value, is artificially set to $t = 10$ ns by the temporal alignment procedure.

7.7.2 Figures of Merit

For performance analysis, we consider the following three sets of CRs:

1. *Measured* CRs are denoted by $\mathbf{r}_{c,\text{meas}}(x,y)$ and result from the measurement campaign described above. The measured CRs include frequency dependent interactions. Due to the high measurement SNR of 70 dB the stochastic component (noise) is assumed to be negligible and $\mathbf{r}_{c,\text{meas}}(x,y)$ are considered as the true CRs.
2. *Predicted* CRs are denoted by $\mathbf{r}_{c,\text{pred}}((x,y),K)$. First the parameters Υ_{LM} of the linear models of path gains and path delays are estimated based on K measured CRs at known TX positions. Hence $\mathbf{r}_{c,\text{pred}}(x,y,K)$ depends on K. Then the simplified input output relation in (7.11) is used to calculated the CRs, where frequency depen-

dent interactions are neglected, i.e. $s_c^l(t) = s_c(t)$. The geometry of the propagation environment (floor plan information) is not required for this prediction.

3. *Theoretical* CRs are denoted by $\mathbf{r}_{c,\text{theo}}(x,y)$ and are obtained by ray tracing. The geometry of the reflectors and the positions of RX and TX have been measured in the anechoic chamber by using a laser measurement device and a measuring tape. With this information it is possible to construct the two-dimensional floor plan depicted in Fig. 7.6. The geometrical parameters a_l and b_l for all three paths are calculated based on the measured geometry and are listed in Table 7.2. Consequently, the corresponding path delays for all TX grid points can be calculated applying equation (7.2). The reflection coefficients of the two metallic interacting objects are assumed to be -1 for all frequencies. The PL of the l-th path depending on center frequency, bandwidth and TX position is approximated by

$$\text{PL}_l(f_c, B, x, y) \approx \frac{1}{(4\pi)^2 \left(f_c^2 - \frac{B^2}{4}\right) (\tau_l(x,y))^2} = \left(\frac{\gamma_{\text{const}}}{\tau_l(x,y)}\right)^2,$$

which is obtained by averaging $\text{PL}(f,d)$ over the used spectrum. The path gain $\gamma_l(x,y)$ is given by the product of the l-th reflection coefficient and the square root of $\text{PL}_l(f_c, B, x, y)$. The *theoretical* CRs are calculated according to equation (7.1), where frequency dependent interactions are neglected implying $s_c^l(t) = s_c(t)$.

In order to assess the prediction performance, the inner product between normalized versions of these three CRs are used.

1. The inner product $\rho_{\text{meas-pred}}(x,y,K)$ between measured and predicted CRs is defined by

$$\rho_{\text{meas-pred}}(x,y,K) \triangleq \frac{\mathbf{r}_{c,\text{meas}}^T(x,y)\,\mathbf{r}_{c,\text{pred}}(x,y,K)}{\|\mathbf{r}_{c,\text{meas}}(x,y)\| \left\|\mathbf{r}_{c,\text{pred}}(x,y,K)\right\|}$$

and measures the prediction accuracy. Deviations are caused by the linearization of path gains and path delays, by erroneous estimation of Υ_{LM} including WRELAX errors caused by neglecting frequency dependent interactions and antenna patterns, and by the multipath approximation of electromagnetic wave propagation.

2. The inner product $\rho_{\text{meas-theo}}(x,y)$ between measured and theoretical CRs is defined by

$$\rho_{\text{meas-theo}}(x,y) = \frac{\mathbf{r}_{c,\text{meas}}^T(x,y)\,\mathbf{r}_{c,\text{theo}}(x,y)}{\|\mathbf{r}_{c,\text{meas}}(x,y)\| \left\|\mathbf{r}_{c,\text{theo}}(x,y)\right\|}$$

and measures the ray tracing accuracy. Deviations are caused by measurement errors of the geometry of the propagation environment, by neglecting frequency dependent interactions and antenna patterns, and by the multipath approximation of electromagnetic wave propagation.

3. The inner product $\rho_{\text{theo-pred}}(x, y, K)$ between theoretical and predicted CRs is defined by

$$\rho_{\text{theo-pred}}(x, y, K) = \frac{\mathbf{r}_{c,\text{theo}}^{T}(x, y)\, \mathbf{r}_{c,\text{pred}}(x, y, K)}{\left\|\mathbf{r}_{c,\text{theo}}(x, y)\right\| \left\|\mathbf{r}_{c,\text{pred}}(x, y, K)\right\|}$$

and measures the similarity of theoretical CRs obtained by ray tracing and predicted CRs. Deviations are caused by measurement errors of the geometry of the propagation environment, by the linearization of path gains and path delays, and by erroneous estimation of Υ_{LM}.

7.7.3 Performance Results

The WRELAX algorithm is applied to estimate the three path gains and path delays of $r_c(t, (x_0, y_0))$. Fig. 7.8 depicts the result of the WRELAX algorithm together with the measured received signal. From visual inspection it can be seen that the reconstruction is almost perfect. There are, however, visible reconstruction errors around $t = 14$ ns and $t = 22$ ns. These errors are most likely caused by signal distortions due to frequency dependent reflection coefficients.

Path Delay [ns] Path Index	Theoretical (Ray Tracing)	Estimated (WRELAX)	Geometrical Parameters			
			a_l	b_l	$c_0\bar{\alpha}_l$	$c_0\bar{\beta}_l$
$\tilde{\tau}_1(x_0, y_0)$, $l = 1$	0	0	0	0	-0.014	-0.999
$\tilde{\tau}_2(x_0, y_0)$, $l = 2$	3.335	3.332	-2.45	0.15	-0.765	-0.644
$\tilde{\tau}_3(x_0, y_0)$, $l = 3$	10.095	10.100	1.19	-2.90	0.220	-0.976

Table 7.2: Comparison between ray tracing (theoretical path delays) and estimated path delays for the point (x_0, y_0).

Table 7.2 summarizes the temporally aligned path delays for the reference point (x_0, y_0). These path delays are obtained once by ray tracing and once by applying the WRELAX algorithm to the measured CR $r_c(t, (x_0, y_0))$. The parameters a_l and b_l describing the geometry of the interacting objects are also stated in Table 7.2. Furthermore, the theoretical linear coefficients without temporal alignment $c_0\bar{\alpha}_l$ and $c_0\bar{\beta}_l$ are given, where $\bar{\alpha}_l$ and $\bar{\beta}_l$ are

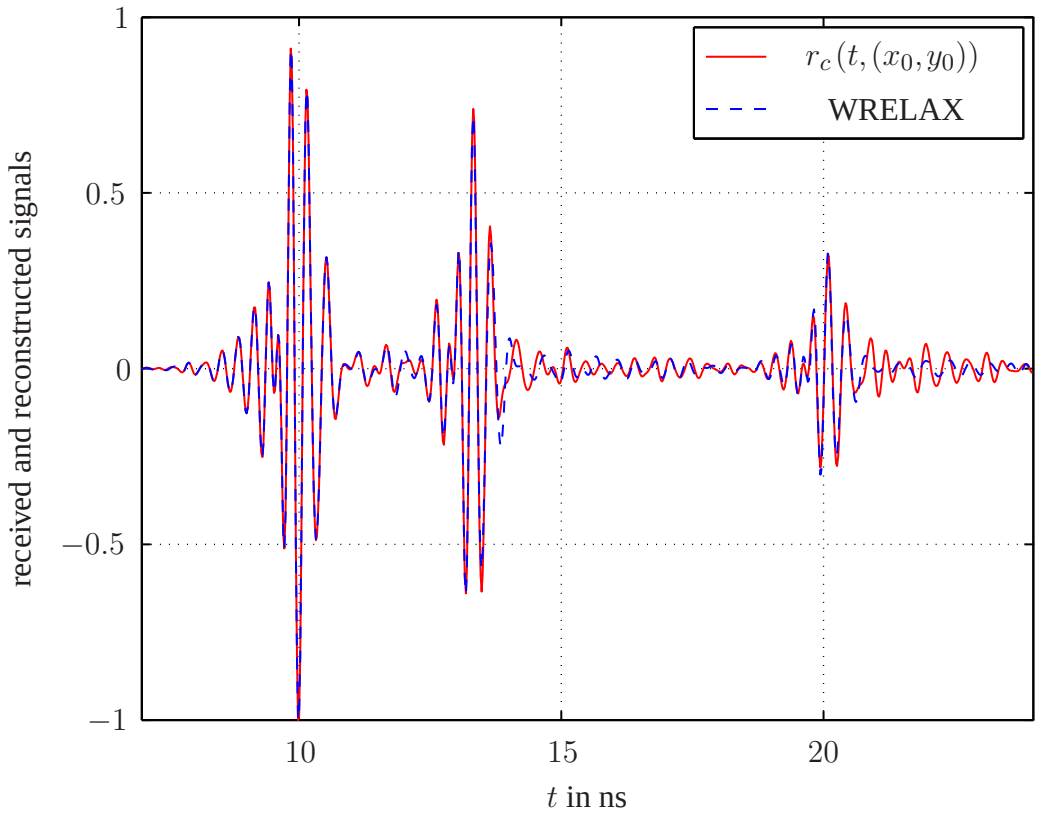

Fig. 7.8: Measured received signal in anechoic chamber with two reflectors and WRELAX reconstruction result.

multiplied with c_0 for a better presentation. It can be observed from Table 7.2 that the WRELAX algorithm is able to accurately estimate the theoretically expected path delays. Note that the time resolution of the WRELAX algorithm is not limited by the sampling frequency, because the estimation is performed in the frequency domain [95].

Table 7.3 summarizes the parameters of the linear model of the path delays. These parameters are once obtained by ray tracing using equation (7.8) with the geometrical parameters in Table 7.2 and once by least squares estimation using K measured CRs in the vicinity of the reference point (x_0, y_0). It can be observed from Table 7.3 that the ray tracing results and the estimation results agree very well for the second path but deviate slightly for the third path. Especially, $c_0\alpha_3$ shows a deviation of around 0.05, which may be due to an inaccurate measurement of the geometry of reflector two. Further, it can be seen that increasing K does not change the linear model parameters significantly. This indicates that already $K = 3$ measurements are sufficient to obtain accurate estimates of the linear model parameters.

Table 7.4 lists the path delays for the lower left corner point $(x_0 - 0.14, y_0 - 0.14)$

Linear Model Parameters	Theoretical (Ray Tracing)	Estimated (Least Squares)		
		$K = 3$	$K = 9$	$K = 16$
$\tilde{\tau}_2(x_0, y_0)$	3.335 ns	3.332 ns	3.332 ns	3.332 ns
$c_0\alpha_2$	-0.751	-0.758	-0.760	-0.755
$c_0\beta_2$	0.355	0.376	0.379	0.380
$\tilde{\tau}_3(x_0, y_0)$	10.095 ns	10.100 ns	10.100 ns	10.100 ns
$c_0\alpha_3$	0.234	0.175	0.183	0.187
$c_0\beta_3$	0.023	0.012	0.014	0.015

Table 7.3: Comparison between linear model parameters obtained from ray tracing and from measured CRs via least squares estimation.

Path Delay [ns]	Theoretical (Ray Tracing)	Estimated (WRELAX)	Predicted	
			$K = 3$	$K = 16$
$\tilde{\tau}_1(x_0 - 0.14, y_0 - 0.14)$	0	0	0	0
$\tilde{\tau}_2(x_0 - 0.14, y_0 - 0.14)$	3.506	3.500	3.510	3.507
$\tilde{\tau}_3(x_0 - 0.14, y_0 - 0.14)$	9.970	9.997	10.013	10.006
$\tilde{\tau}_1(x_0 + 0.14, y_0 + 0.14)$	0	0	0	0
$\tilde{\tau}_2(x_0 + 0.14, y_0 + 0.14)$	3.135	3.130	3.153	3.157
$\tilde{\tau}_3(x_0 + 0.14, y_0 + 0.14)$	10.209	10.188	10.187	10.195

Table 7.4: Comparison between ray tracing (theoretical path delays), estimated path delays via WRELAX, and predicted path delays for upper right and lower left corner point.

and the upper right corner point $(x_0 + 0.14, y_0 + 0.14)$ of the transmit grid. The path delays are once obtained by ray tracing, once by applying the WRELAX algorithm to $r_c(t, (x_0 \pm 0.14, y_0 \pm 0.14))$, and once by using the linear model for prediction. The parameters of the linear model are estimated from K measured CRs in the vicinity of (x_0, y_0). It can be observed from Table 7.4 that the predicted path delays match very accurately to the estimated and to the theoretical ones. These results confirm that the approximation error induced by the linearization of the path delays is indeed negligible for the considered geometry and propagation environment.

Fig. 7.9 depicts the inner product matrix $\rho_{\text{theo-pred}}(x, y, 3)$ for all measured TX position. The predicted CRs are based on linear model parameters estimated from $K = 3$ measured CRs at grid points (x_0, y_0), $(x_0 + 0.01, y_0)$, and $(x_0, y_0 + 0.01)$. It can be seen that the inner products decrease with the Euclidian distance to the reference point (x_0, y_0) due to increasing linearization errors. However, as the smallest inner product is 0.96, it can be concluded that both linear approximations for gains and delays are accurate and applicable for the considered geometry and propagation environment.

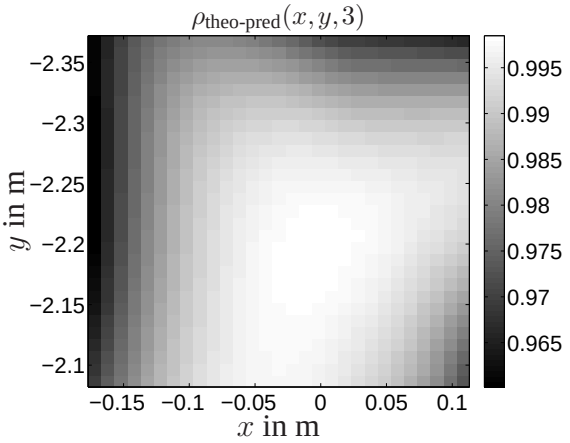

Fig. 7.9: Inner product between theoretical and predicted CRs for $K = 3$.

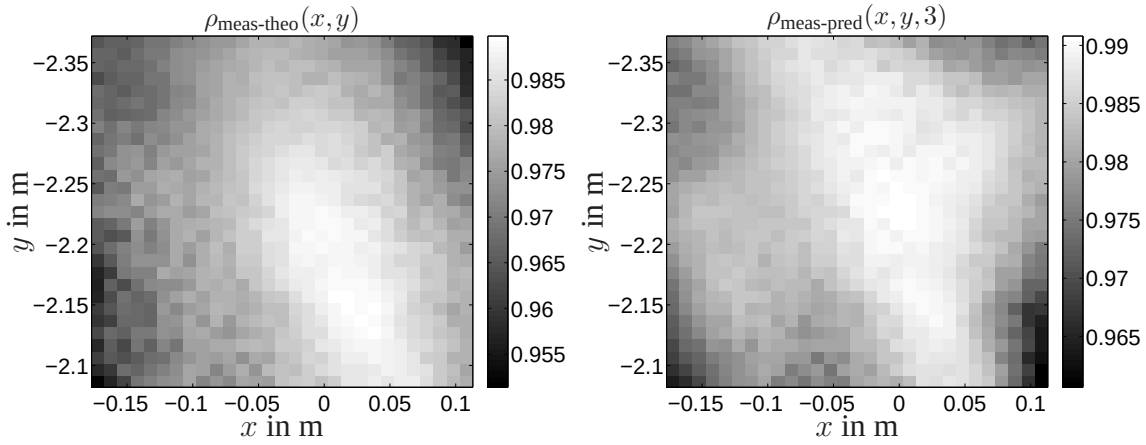

Fig. 7.10: Inner product between measured and theoretical (left) and measured and predicted (right) CRs for $K = 3$.

Fig. 7.10 depicts the inner product matrices $\rho_{\text{meas-theo}}(x, y, 3)$ and $\rho_{\text{meas-pred}}(x, y)$ for all measured TX position. The predicted and theoretical CRs are the same as in Fig. 7.9. The inner products $\rho_{\text{meas-theo}}(x, y, 3)$ range from 0.95 to 0.99 and $\rho_{\text{meas-pred}}(x, y)$ range from 0.96 to 0.99. Due to the marginal decrease of the inner products the following conclusions can be drawn:

1. The multipath approximation of electromagnetic wave propagation is accurate and applicable for large metallic reflectors.
2. The frequency dependent interactions are negligible for the considered frequency band from 2 GHz to 6 GHz and metallic reflectors.
3. The antenna patterns of the applied UWB antennas are negligible.
4. The linear approximation of the path gains and path delays is accurate and applicable

for the considered transmit area.

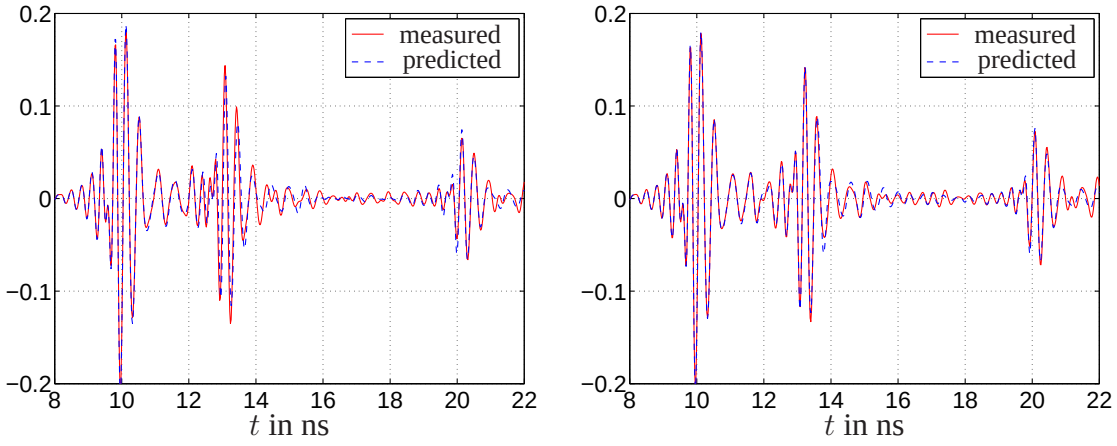

Fig. 7.11: Measured and predicted CRs for smallest inner product (left) and largest inner product (right) in $\rho_{\text{meas-pred}}(x, y, 3)$.

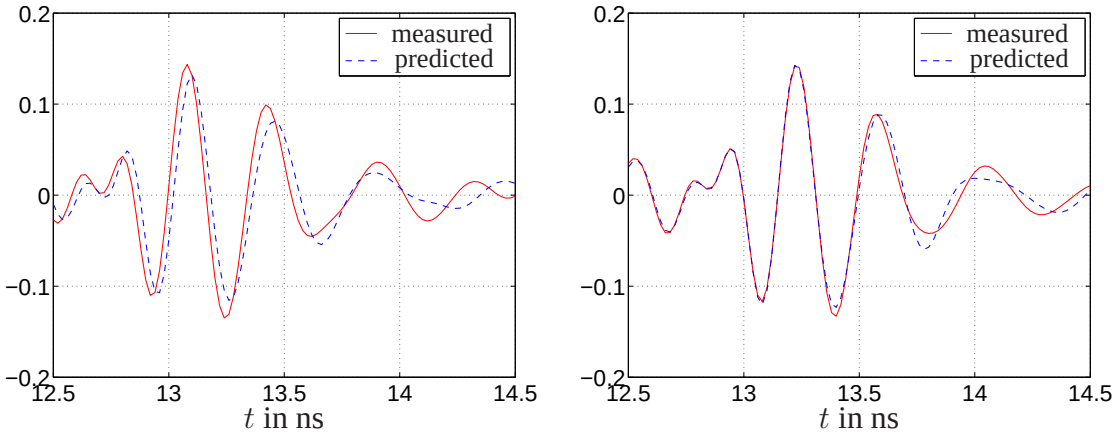

Fig. 7.12: Snapshots of measured and predicted CRs for smallest inner product (left) and largest inner product (right) in $\rho_{\text{meas-pred}}(x, y, 3)$.

Fig. 7.11 depicts the CRs for the smallest and the largest inner products in $\rho_{\text{meas-pred}}(x, y, 3)$. The smallest inner product appears at $(x_0 + 0.14, y_0 + 0.14)$ in the lower right corner[7] of the right plot in Fig. 7.10 and the largest inner product appears at the center point (x_0, y_0). Fig. 7.12 depicts snapshots of these CRs around the direct path. In this figure the worse prediction of the second path in the left plot compared to the right plot is clearly visible.

[7] This corresponds to the upper right corner of the transmit grid depicted in Fig. 7.6.

7.8 Parameter Estimation based on Predicted Channel Responses

So far, a technique for the prediction of CRs at arbitrary TX positions has been developed and the prediction accuracy has been studied. The main purpose of CR prediction in this thesis is the estimation of the parameters Θ_m of the conditional PDFs $f_m(\mathbf{y}|H_m)$. These estimates could be used, for example, to initialize the EM algorithm proposed in Chapter 6. The remaining steps are the prediction of sufficiently many CRs from all M regions, calculation of the respective location fingerprints, and application of the same empirical ML parameter estimation rules as proposed Section 4.4.1 and Section 5.5.1.

As an example, we consider location fingerprints obtained with a coherent RX as proposed in Chapter 4. The corresponding stochastic model for the location fingerprints is complex-valued jointly Gaussian with $\Theta_m = \{\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m, \mathbf{Z}_m\}$. The mean vector does not become zero for the measurements in the anechoic chamber, because the temporal alignment can be performed very accurately due to the negligible impact of multipath fading of succeeding paths on the reference (direct) path. Hence, it can be expected that the reference path produces a strong mean component.

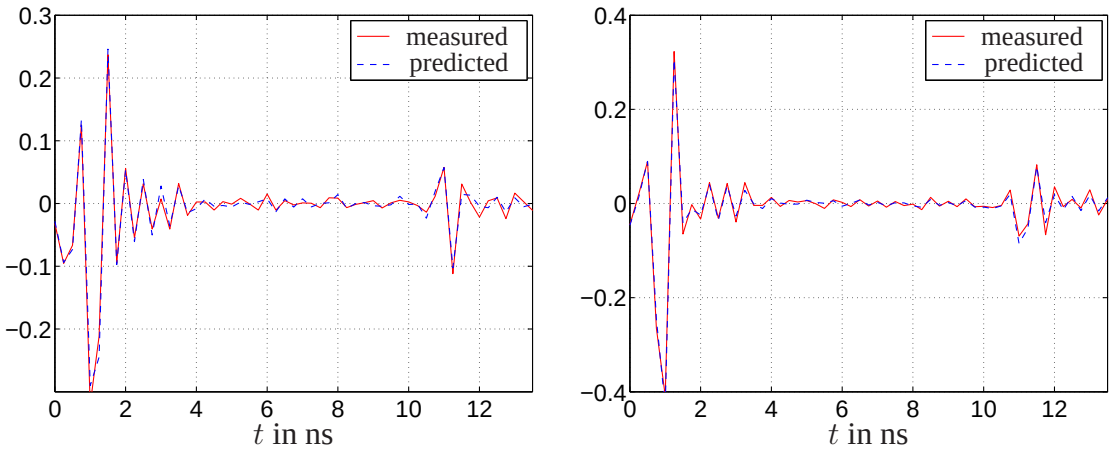

Fig. 7.13: Real parts (left) and imaginary parts (right) of the mean vectors of measured and predicted location fingerprints for a coherent RX.

Fig. 7.13 depicts the real parts and the imaginary parts of the mean vectors, which are obtained from 841 measured location fingerprints and from 841 predicted location fingerprints. The linear model parameters Υ_{LM} are estimated from $K = 3$ measured CRs with the least squares method. It can be seen that the measured and predicted curves show a very good match. As expected, the first path (reference path) has a strong mean component. How-

ever, also the third path exhibits a mean component, whereas the second path does not have a mean component. This can be explained by different delay variations of the second and third relative to the direct path. The maximal delay differences between the lower left and upper right corner points of the transmit grid are given in Table 7.4. The second path has a maximal delay difference of roughly 0.4 ns, whereas the third path has a maximal delay difference of roughly 0.2 ns. Since the carrier frequency is 4 GHz, which implies a carrier period of 0.25 ns, the third path cannot be averaged over a full carrier period. This explains the mean component of the third path.

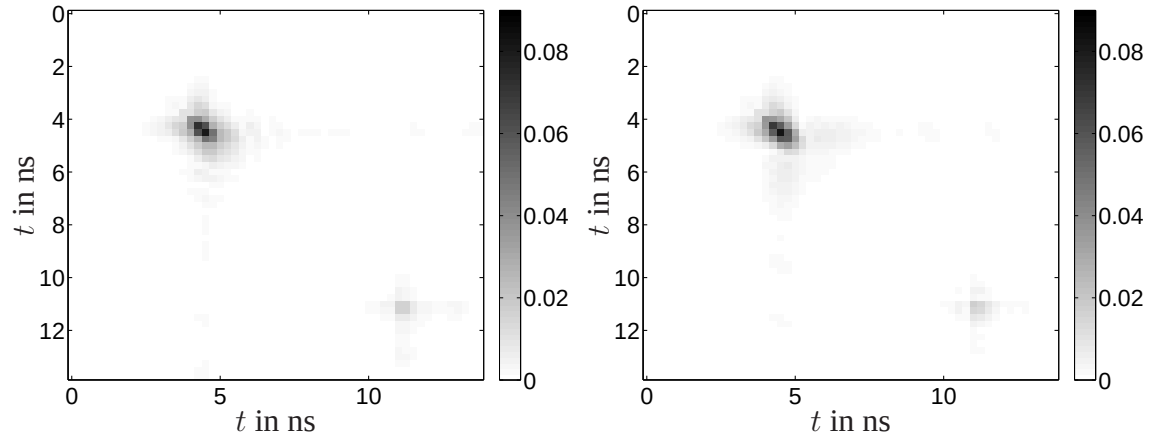

Fig. 7.14: Absolute values of the covariance matrices of measured (left) and predicted (right) location fingerprints for a coherent RX.

Fig. 7.14 shows absolute values of the covariance matrices of measured and predicted location fingerprints for a coherent RX. Only the second path provides a significant contribution to the covariance matrix, since the direct and the third path have a mean component. Visual inspection confirms that the predicted covariance matrix provides a very good match to the measured covariance matrix. The Frobenius norm of the difference of these two matrices results in 0.07, which further indicates the similarity of these two matrices.

7.9 Summary and Conclusions

It can be concluded from this experimental analysis in a controlled propagation environment that the multipath approximation is accurate, the impact of frequency dependent interactions and antenna characteristics is negligible, and that the linearization of path gains and path delays is applicable for the region dimensions of interest. These conclusions hold for the considered frequency band and reflecting objects.

Further, it is shown that the estimation of the linear model parameters Υ_{LM} can be accurately performed by the proposed two step approach with $K = 3$ measured CRs at known TX positions. CRs at arbitrary TX positions can be predicted with the linear model using $\hat{\Upsilon}_{\mathrm{LM}}$. The parameter set Θ_m of the stochastic location fingerprint model are then estimated based on arbitrary many predicted training signals. Hence, the training phase requires only the measurement of three CRs at known TX positions for each region. This holds for all kind of location fingerprints and stochastic models.

Moreover, it is shown that the deterministic components of the CRs can be predicted based on a floor plan. The theoretical CRs, which are calculated via ray tracing, match very well to the measured CRs. Reflections from floor, ceiling, walls, and large objects can be predicted this way irrespective of the remaining propagation environment causing the stochastic channel components.

The crucial part of this efficient parameter estimation method is the identification of multipath components in a propagation environment with more and densely spaced multipaths as discussed in Section 7.6.3. We propose the Hough transform as potential means to solve this problem. The experimental evaluation of the prediction and parameter estimation accuracy in an office or an industrial propagation environment is outside the scope of this thesis but definitely an important open item. The objective of this chapter is merely to introduce the ideas, required theoretical concepts, and estimation algorithms and to evaluate them in a controlled propagation environment.

It would be also conceivable to use the linear model parameters directly as location fingerprints, which implies that a set of path gains and path delays in conjunction with the corresponding linear coefficients fully describe a region. The development and analysis of such a scheme is, however, outside the scope of this thesis.

Chapter 8

Conclusions and Outlook

8.1 Conclusions

Location fingerprinting is a promising alternative to geometrical position location approaches in dense multipath propagation environments with possible non-LoS conditions. However, state-of-the-art location fingerprinting systems suffer from low accuracy and precision due to the use of RF signals with too small bandwidth, which implies location fingerprints with few degrees of freedom per anchor. The only possibility for such systems to achieve more degrees of freedom, is the deployment of additional anchors at the expense of hardware and infrastructure complexity as discussed in Chapter 1.

It is envisioned that location fingerprints extracted from UWB RF signals provide high position location accuracy and precision with a single anchor due to many available degrees of freedom. It is also expected that the multipath structure, which carries position location information, can be resolved in time and, thus, utilized for localization. This thesis features the first comprehensive study of location fingerprinting for UWB systems, confirms these visions and expectations with extensive experimental and theoretical results, and proposes the first location fingerprinting system in literature, which is able to operate with a single anchor. A theoretical location fingerprinting framework is developed that generalizes all location fingerprinting methods, which are based on the Bayesian paradigm, and supports the systematic derivation and analysis of optimal localization (classification) algorithms. A distinct advantage of the proposed Bayesian framework over other classification tools like neural networks or support vector machines is the possibility to optimally combine multiple observations per agent by summation of log-likelihood values, which corresponds to a joint evaluation of soft decisions. Neural networks and support vector machines support only the suboptimal joint evaluation of hard decisions.

Location fingerprinting with two specific UWB RX structures is investigated: a coherent RX and a low complexity generalized ED RX. Experimental performance results based on measured CIRs demonstrate that both UWB RX structures provide excellent position location and clustering performance in a dense multipath propagation environment with non-LoS conditions. It is found that the MV circular symmetric Gaussian PDF is an accurate stochastic location fingerprint model of a coherent RX. Experimental performance results show that the performance can be significantly enhanced by accounting for channel tap correlations. For a generalized ED RX, it is found that the Gamma PDF, the Log-normal PDF, and the MV Log-normal PDF are accurate stochastic descriptions of the location fingerprints. Among these PDFs only the MV Log-normal PDF accounts for correlations among energy samples. The experimental analysis shows that the best performance is achieved with the MV Log-normal PDF.

An experimental performance study reveals a major shortcoming of UWB based location fingerprinting systems, which is the high complexity of the training phase. In order to combat this problem, it is proposed to extend the location fingerprinting framework with mixture models, such that iterative algorithms are applicable. It is noteworthy that such an extension is only supported by the proposed theoretical framework based on a Bayesian formulation. Specifically, the EM algorithm is applied and experimental performance results demonstrate that significant performance gains are achievable, if only few training signals are available.

In the last chapter, we propose to exploit a priori knowledge about the propagation environment to further increase the efficiency of the training phase. It is shown that CRs at arbitrary TX positions can be predicted based on either a floor plan or a geometrical channel model. The parameters of this model can be estimated from three measured CRs at known TX positions. The prediction accuracy is evaluated experimentally in a controlled propagation environment. It is concluded that both techniques are able to accurately predict CRs, which can be further used as training signals.

In order to implement an efficient and robust location fingerprinting system, it is recommended to use predicted CRs in order to obtain initial parameter estimates $\Theta_m[0]$ for all regions, which are required by the EM algorithm. During the localization phase many location fingerprints containing also information about the stochastic and unpredictable channel components are observed. Those location fingerprints are utilized via the EM algorithm to learn the parts of the parameter sets Θ_m, which are caused by stochastic channel components. Finally, such a system would constantly improve over time for a static propagation scenario and, moreover, would be able to cope with small changes of the propagation environment by constantly adapting Θ_m based on new observations and forgetting old location fingerprints.

8.2 Outlook on Future Research

In the following, the most interesting and relevant open topics from the author's perspective are shortly discussed in order to initiate further research in the field of location fingerprinting for UWB systems.

Time Varying Propagation Channels It has to be distinguished between two types of time variations. Time variations caused by moving objects like people cause random distortions of the location fingerprints as discussed in Section 4.3.2. Such distortions can be mitigated, for example, by collecting multiple location fingerprints per agent. If the propagation environment changes deterministically, for example, by relocation of large objects, it might become necessary to redo the training phase. Future research should study the ability of the EM algorithm to adapt to such significant and permanent changes of the propagation environment.

Number and Dimensions of Regions Increasing the number of regions (surveillance area) will most likely lead to a performance degradation, i.e. larger $\mathcal{P}_e$ and $\mathcal{D}_e$. However, the performance results in Section 5.7.5 and Section 4.6.4 indicate a graceful performance degradation, since most classification errors happen among neighboring regions. Therefore, it can be expected that $\mathcal{D}_e$ increases only very slowly with M. Future research should confirm these expectations via further experimental results. Moreover, it would be interesting to develop a theoretical framework, which enables the derivation of optimal region dimensions and placements for a given propagation environment.

Efficient Training Phase In Chapter 7, it is experimentally shown that the deterministic part of CRs can be predicted within a region of 29 cm times 29 cm with high accuracy based on three measured CRs. The experiment has been performed in an anechoic chamber with two metallic reflectors, such that only three multipath components are existent. The crucial part of this prediction method is the association of path gains and path delays, estimated at different TX position, as discussed in Section 7.6.3. This is not an issue for the controlled propagation environment but becomes relevant in an environment with dense multipath propagation like an office or industry hall. An experimental analysis of the prediction performance in such an environment is therefore recommended.

Prototype An important next step would be the implementation of a location fingerprinting prototype system in real time. With such a prototype the impact of realistic error sources like time varying propagation channels and interference can be studied.

Cooperative Location Fingerprinting To the best of the author's knowledge cooperative location fingerprinting has not been investigated or proposed so far. The main reason might be the centralized nature of location fingerprinting due to large databases and centralized processing, which makes it difficult to develop distributed schemes. A big advantage of location fingerprinting with UWB signals, is the need for only one anchor, which stores the database and does the whole processing. Therefore, it might be interesting to develop cooperative location fingerprinting schemes, where some agents have anchor capabilities. This is especially important for low complexity anchors employing a generalized ED RX, where cooperation might be the key to obtain better performance results. Such a cooperative scheme would require the dissemination of the database to these agents only once and afterwards the continuous dissemination of region/cluster information and location fingerprints.

A Distance Matrix for UWB Channel Measurement Campaign

The approximate 22×22 distance matrix **D**, which corresponds to the measurement campaign described in Chapter 3, is given below in meters.

$$\mathbf{D} = \begin{bmatrix} d_{1,1} & d_{1,2} & d_{1,3} & \dots \\ d_{1,2} & d_{2,2} & d_{2,3} & \dots \\ d_{1,3} & d_{2,3} & d_{3,3} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The diagonal entries $d_{i,i}$ are given by

$$d_{i,i} = \sqrt{0.27^2/12 + 0.56^2/12} \approx 0.18 \text{ m},$$

which corresponds to the RMS error of a two-dimensional uniform distribution with lengths 0.27 m times 0.56 m.

$$
\mathbf{D} \approx \begin{bmatrix}
0.18 & 0.37 & 0.74 & 1.11 & 1.48 & 1.85 & 1.62 & 1.29 & 0.99 & 0.76 & 0.66 & 3 & 6 & 9 & 10 & 12 & 17 & 16 & 15 & 13 & 12 & 10 \\
\ddots & 0.18 & 0.37 & 0.74 & 1.11 & 1.48 & 1.29 & 0.99 & 0.76 & 0.66 & 0.76 & 3.02 & 6 & 9 & 10 & 12 & 17 & 16 & 15 & 13 & 12 & 10 \\
 & \ddots & 0.18 & 0.37 & 0.74 & 1.11 & 0.99 & 0.76 & 0.66 & 0.76 & 0.99 & 3.09 & 6 & 9 & 10 & 12 & 17 & 16 & 15 & 13 & 12 & 10 \\
 & & \ddots & 0.18 & 0.37 & 0.74 & 0.76 & 0.66 & 0.76 & 0.99 & 1.29 & 3.20 & 6 & 9 & 10 & 12 & 17 & 16 & 15 & 13 & 12 & 10 \\
 & & & \ddots & 0.18 & 0.37 & 0.66 & 0.76 & 0.99 & 1.29 & 1.62 & 3.35 & 6 & 9 & 10 & 12 & 17 & 16 & 15 & 13 & 12 & 10 \\
 & & & & \ddots & 0.18 & 0.76 & 0.99 & 1.29 & 1.62 & 1.96 & 3.52 & 6 & 9 & 10 & 12 & 17 & 16 & 15 & 13 & 12 & 10 \\
 & & & & & \ddots & 0.18 & 0.37 & 0.74 & 1.11 & 1.48 & 2.77 & 5.34 & 8.34 & 9.34 & 11.34 & 16.34 & 15.34 & 14.34 & 12.34 & 11.34 & 9.34 \\
 & & & & & & \ddots & 0.18 & 0.37 & 0.74 & 1.11 & 2.59 & 5.34 & 8.34 & 9.34 & 11.34 & 16.34 & 15.34 & 14.34 & 12.34 & 11.34 & 9.34 \\
 & & & & & & & \ddots & 0.18 & 0.37 & 0.74 & 2.45 & 5.34 & 8.34 & 9.34 & 11.34 & 16.34 & 15.34 & 14.34 & 12.34 & 11.34 & 9.34 \\
 & & & & & & & & \ddots & 0.18 & 0.37 & 2.37 & 5.34 & 8.34 & 9.34 & 11.34 & 16.34 & 15.34 & 14.34 & 12.34 & 11.34 & 9.34 \\
 & & & & & & & & & \ddots & 0.18 & 2.34 & 5.34 & 8.34 & 9.34 & 11.34 & 16.34 & 15.34 & 14.34 & 12.34 & 11.34 & 9.34 \\
 & & & & & & & & & & \ddots & 0.18 & 3 & 6 & 6.32 & 7.21 & 10 & 8.54 & 10 & 8 & 6 & 4 \\
 & & & & & & & & & & & \ddots & 0.18 & 3 & 3.61 & 5 & 8.54 & 8 & 8.54 & 5.83 & 5 & 3.61 \\
 & & & & & & & & & & & & \ddots & 0.18 & 2 & 4 & 8 & 8.73 & 10.63 & 9.22 & 8.06 & 7.28 \\
 & & & & & & & & & & & & & \ddots & 0.18 & 2 & 6 & 6.95 & 9.22 & 8.06 & 7.28 & 7 \\
 & & & & & & & & & & & & & & \ddots & 0.18 & 4 & 5.32 & 8.06 & 7.28 & 7 & 7.28 \\
 & & & & & & & & & & & & & & & \ddots & 0.18 & 3.5 & 7 & 7.28 & 8.06 & 9.22 \\
 & & & & & & & & & & & & & & & & \ddots & 0.18 & 3.5 & 4.03 & 5.32 & 6.95 \\
 & & & & & & & & & & & & & & & & & \ddots & 0.18 & 2 & 4 & 6 \\
 & & & & & & & & & & & & & & & & & & \ddots & 0.18 & 2 & 4 \\
 & & & & & & & & & & & & & & & & & & & \ddots & 0.18 & 2 \\
 & & & & & & & & & & & & & & & & & & & & \ddots & 0.18
\end{bmatrix}.
$$

B Akaike's Information Criterion

It is assumed that the samples of the equivalent baseband CIR are distributed according to an original PDF f, which is called the operating model. The operating model is usually unknown, since only a finite number of observations (realizations) is available. Therefore, approximating probability models must be specified using the observed data, in order to estimate the operating model. The approximating models are divided into families, where each family is specified by a set of parameters, e.g. mean and variance for the Gaussian family or degrees of freedom for the Chi-square family. An approximating family is denoted as g_{Θ}^{k}, where the subscript Θ stands for the U-dimensional parameter vector, which specifies the PDF. The collection of all a priori defined approximating families $k \in \{1, 2, \ldots, K\}$ forms the candidate set.

It is intuitive that the number of free parameters U determines the flexibility of the models, and therefore also the approximation quality. However, these parameters must be estimated from a finite number of observations. It is again intuitive that the estimates become worse, if more parameters have to be estimated from the same number of observations. Therefore, the approximation of the operating model has two sources of errors or discrepancies:

1. Discrepancy due to approximation: This is the discrepancy between the operating model and the best approximating model within a family. Thus, this discrepancy is the lower bound on the discrepancy between the operating model and the approximating family. Since the operating model is generally unknown, it is impossible to find the best approximating model.
2. Discrepancy due to estimation: This is the discrepancy between the best approximating model and a fitted model with estimated parameters from a finite number of observations. This quantity depends on the actual observations and would change for different observations. Hence, it constitutes a random variable.

The overall discrepancy is a function of these two discrepancies. Therefore, it is necessary to take both discrepancies into account, when approximating families with different complexity (different U) are compared. The discrepancy due to approximation becomes smaller

for more complex models, whereas the discrepancy due to estimation becomes larger. The challenge in selecting the best approximating probability model is to find an appropriate compromise between these two effects.

In practice, none of these discrepancies can be computed. Moreover, the overall discrepancy is a random quantity. It would be of interest to compute the expected value of the overall discrepancy for a given number of observations, which is still impossible, since the operating model is unknown. However, the expected discrepancy can be estimated and the corresponding estimator is called a model selection criterion. The expected discrepancy is generally a complex quantity depending on the operating model, the approximating family, the parameter estimation method, and the number of available observations.

In information theory, the Kullback-Leibler distance $D\left(f||g_\Theta^k\right)$ is a measure for the difference between the two PDFs f and g_Θ^k and is given by

$$\begin{aligned} D\left(f||g_\Theta^k\right) &= \mathrm{E}\left[\log f(X)\right] - \mathrm{E}\left[\log g_\Theta^k(X)\right] \\ &= \int f(x)\log f(x)\mathrm{d}x - \int f(x)\log g_\Theta^k(x)\mathrm{d}x \\ &= -h(X) - \int f(x)\log g_\Theta^k(x)\mathrm{d}x, \end{aligned}$$

where the random variable X is distributed according to the original but unknown PDF f, and $h(\cdot)$ denotes differential entropy. This distance measure is not directly applicable, since the original PDF f is not known. It is known, however, that the Kullback-Leibler distance is nonnegative, i.e. $D\left(f||g_\Theta^k\right) \geq 0$.

The term $-\int f(x)\log g_\Theta^k(x)\mathrm{d}x$ is called Kullback-Leibler discrepancy and approaches the differential entropy of X from above for increasing approximation quality of the model g_Θ^k, which can be seen from

$$-\int f(x)\log g_\Theta^k(x)\mathrm{d}x = h(X) + D\left(f||g_\Theta^k\right). \tag{B.1}$$

The Kullback-Leibler distance $D\left(f||g_\Theta^k\right)$ becomes smaller and smaller and the differential entropy of X is reached if and only if $f = g_\Theta^k$.

Applying the weak law of large numbers, expression (B.1) can be approximated by averaging the log-likelihood values given the model g_Θ^k over N independent observations $x_1, x_2, \ldots, x_N$ according to

$$-\int f(x)\log g_\Theta^k(x)\mathrm{d}x \approx -\frac{1}{N}\sum_{n=1}^{N}\log g_\Theta^k(x_n). \tag{B.2}$$

The Kullback-Leibler discrepancy (B.1) depends also on the estimated parameter vector Θ, which is itself a function of the actual observations $x_1, x_2, \ldots, x_N$. If another set of observations $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N$ is used, a different Kullback-Leibler discrepancy is obtained. Therefore, the expected Kullback-Leibler discrepancy is defined by

$$-\mathrm{E}_\Theta \left[\int f(x) \log g^k_\Theta(x) \mathrm{d}x \right], \tag{B.3}$$

where the expectation is taken with respect to the distribution of the estimated parameter vector Θ. Expression (B.3) has to be minimized over the candidate set, in order to find the best approximating model. As stated, expression (B.3) cannot be computed exactly, but has to be estimated. AIC is an approximately unbiased estimator for (B.3) and is given by [61]

$$\mathrm{AIC}_k = 2U - 2 \sum_{n=1}^{N} \log g^k_{\hat{\Theta}}(x_n). \tag{B.4}$$

The parameter vectors Θ for each family should be estimated using the minimum discrepancy estimator $\hat{\Theta}$, which minimizes the empirical discrepancy. This is the discrepancy between the approximating model and the model obtained by regarding the observations as the whole population. The ML estimator is the minimum discrepancy estimator for the Kullback-Leibler discrepancy.

The Akaike Weights $\{\omega_1, \omega_2, \ldots, \omega_K\}$ are computed following [59] using (B.4), in order to compare the competing candidate families and select the best approximating distribution. The Akaike Weight ω_k is given by

$$\omega_k = \frac{\exp\left(-\frac{\Omega_k}{2}\right)}{\sum_{i=1}^{K} \exp\left(-\frac{\Omega_i}{2}\right)}, \text{ with } \Omega_k = \mathrm{AIC}_k - \min_{i=1,2,\ldots,K} \mathrm{AIC}_i.$$

The Akaike Weight ω_k can be interpreted as estimate for the probability that the k-th candidate family shows the best modeling quality.

C Derivation of Approximation (5.5)

We start by repeating the characteristic function $\Psi(s)$, which is given by

$$\Psi(s) = \prod_{i=1}^{I} (1+2s\lambda_i)^{-1/2} \exp\left(-\frac{s\gamma_i^2\lambda_i}{1+2s\lambda_i}\right) \text{ for } s = -\imath\omega.$$

Note the different definition of s in comparison to t in (5.3). Applying the inverse Fourier transform to this characteristic function the PDF $f(y)$ is obtained as

$$f(y) = \frac{1}{2\pi\imath} \int_{-\imath\infty}^{\imath\infty} \exp(sy) \prod_{i=1}^{I} (1+2s\lambda_i)^{-1/2} \exp\left(-\frac{s\gamma_i^2\lambda_i}{1+2s\lambda_i}\right) \mathrm{d}s. \tag{C.1}$$

Let us define the variable x for a given λ_i according to

$$x \triangleq \frac{1}{2s\lambda_i}.$$

We will use the following Taylor series expansions around $x = 0$:

$$\begin{aligned} \frac{1}{1+x} &= 1 - x + x^2 - x^3 + \ldots = 1 - x + \mathcal{E}\left(x^2\right), \\ (1+x)^{-1/2} &= 1 - \frac{x}{2} + \mathcal{E}\left(x^2\right), \\ \exp\left(ax + \mathcal{E}\left(x^2\right)\right) &= 1 + ax + \mathcal{E}\left(x^2\right) \text{ for } a > 0. \end{aligned} \tag{C.2}$$

The error term $\mathcal{E}\left(x^2\right)$ states that the smallest exponent of x in the error term is 2. The error term in the variable x translates into an error term $\mathcal{E}\left(s^{-2}\right)$ in the variable s by definition of x. Hence, the largest exponent of s in the error term $\mathcal{E}\left(s^{-2}\right)$ is -2. Thus, we have the following approximations:

$$\begin{aligned} (1+2s\lambda_i)^{-1/2} &= (2s\lambda_i)^{-1/2} \left(1 + \frac{1}{2s\lambda_i}\right)^{-1/2} = (2s\lambda_i)^{-1/2} \left(1 - \frac{1}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right), \\ \frac{s\gamma_i^2\lambda_i}{1+2s\lambda_i} &= (2s\lambda_i)^{-1} s\gamma_i^2\lambda_i \left(1 + \frac{1}{2s\lambda_i}\right)^{-1} = \frac{\gamma_i^2}{2} - \frac{\gamma_i^2}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right). \end{aligned}$$

We can now insert these two approximations into (C.1), which gives

$$
\begin{aligned}
f(y) &= \frac{1}{2\pi\imath} \int_{-\imath\infty}^{\imath\infty} \exp(sy) \prod_{i=1}^{I} (2s\lambda_i)^{-1/2} \left(1 - \frac{1}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) \\
&\times \exp\left(-\sum_{i=1}^{I} \frac{\gamma_i^2}{2}\right) \exp\left(\sum_{i=1}^{I} \frac{\gamma_i^2}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) \mathrm{d}s \\
&= \frac{2^{-I/2} \prod_{i=1}^{I} \lambda_i^{-1/2}}{2\pi\imath} \exp\left(-\sum_{i=1}^{I} \frac{\gamma_i^2}{2}\right) \int_{-\imath\infty}^{\imath\infty} \exp(sy)\, s^{-I/2} \prod_{i=1}^{I} \left(1 - \frac{1}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) \\
&\times \prod_{i=1}^{I} \exp\left(\frac{\gamma_i^2}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) \mathrm{d}s.
\end{aligned}
$$

We approximate the product term according to

$$
\prod_{i=1}^{I} \left(1 - \frac{1}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) = 1 - \frac{1}{4s} \sum_{i=1}^{I} \lambda_i^{-1} + \mathcal{E}\left(s^{-2}\right), \tag{C.3}
$$

where the cross terms $\frac{1}{4s\lambda_i} \frac{1}{4s\lambda_j}$ are absorbed in $\mathcal{E}\left(s^{-2}\right)$. Further, we approximate the product of exponential terms using (C.2) according to

$$
\prod_{i=1}^{I} \exp\left(\frac{\gamma_i^2}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) = \prod_{i=1}^{I} \left(1 + \frac{\gamma_i^2}{4s\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) = 1 + \frac{1}{4s} \sum_{i=1}^{I} \frac{\gamma_i^2}{\lambda_i} + \mathcal{E}\left(s^{-2}\right),
$$

where the last step is in analogy to (C.3).

The resulting approximation for $f(y)$ is

$$
\begin{aligned}
f(y) = \frac{2^{-I/2} \prod_{i=1}^{I} \lambda_i^{-1/2}}{2\pi\imath} &\exp\left(-\sum_{i=1}^{I} \frac{\gamma_i^2}{2}\right) \int_{-\imath\infty}^{\imath\infty} \exp(sy)\, s^{-I/2} \\
&\times \left(1 - \frac{1}{4s} \sum_{i=1}^{I} \lambda_i^{-1} + \mathcal{E}\left(s^{-2}\right)\right) \left(1 + \frac{1}{4s} \sum_{i=1}^{I} \frac{\gamma_i^2}{\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) \mathrm{d}s.
\end{aligned}
$$

Using again the argument from (C.3) we arrive at

$$
\begin{aligned}
f(y) = \frac{2^{-I/2} \prod_{i=1}^{I} \lambda_i^{-1/2}}{2\pi\imath} &\exp\left(-\sum_{i=1}^{I} \frac{\gamma_i^2}{2}\right) \int_{-\imath\infty}^{\imath\infty} \exp(sy)\, s^{-I/2} \\
&\times \left(1 + \frac{1}{4s} \sum_{i=1}^{I} \frac{\gamma_i^2 - 1}{\lambda_i} + \mathcal{E}\left(s^{-2}\right)\right) \mathrm{d}s.
\end{aligned}
$$

The final step is performing the inverse Fourier transformation and noticing that the error term $\mathcal{E}\left(s^{-2}\right)$ in the *frequency* domain translates to an error term $\mathcal{E}\left(y^2\right)$ in the *time* domain. Note that $\mathcal{E}\left(y^2\right)$ is in analogy to $\mathcal{E}\left(x^2\right)$, which means that the smallest exponent of y in the error term $\mathcal{E}\left(y^2\right)$ is 2. With this the approximation (5.5) for $f(y)$ for small positive values of y is obtained according to

$$f(y) = \frac{y^{(I/2)-1} \exp\left(-\sum_{i=1}^{I} \frac{\gamma_i^2}{2}\right)}{2^{I/2}\Gamma(I/2) \prod_{i=1}^{I} \lambda_i^{1/2}} \left(1 + \frac{y}{2I} \sum_{i=1}^{I} \frac{\gamma_i^2 - 1}{\lambda_i} + \mathcal{E}\left(y^2\right)\right).$$

This approximations works well as long as the term $\sum_{i=1}^{I} \frac{\gamma_i^2}{2}$ does not become too large, which causes numerical problems because of too small values for $f(y)$.

D Wishart Distribution

The Wishart distribution [66,67] is a generalization of the Chi-square distribution to multiple dimensions. It is a stochastic description for positive semidefinite random matrices.

Assume a real valued matrix $\mathbf{X}$ of dimensions $L \times N$. Each row l of this matrix corresponds to an independent realization of a length N random vector $\mathbf{x}_l$ following a zero mean jointly Gaussian probability distribution with $N \times N$ covariance matrix $\mathbf{\Sigma}$. The Wishart distribution is defined as the probability distribution of the positive semidefinite matrix

$$\mathbf{S} = \mathbf{X}^T\mathbf{X},$$

which is also known as the scatter matrix. The unbiased ML estimate of the covariance matrix $\mathbf{\Sigma}$ is simply $\mathbf{S}/L$. Note that this definition also applies to complex random vectors, i.e. to a complex valued matrix $\mathbf{X}$. By splitting the complex valued random vectors into length $2N$ real random vectors, the same steps as above can be applied.

In order to indicate that $\mathbf{S}$ is distributed according to a Wishart distribution the following notation is commonly used:

$$\mathbf{S} \sim W_N\left(\mathbf{\Sigma}, L\right),$$

where N denotes the dimensionality and L the number of degrees of freedom. Note that $W_1\left(1, L\right)$ is equal to a Chi-square distribution with L degrees of freedom.

The Wishart PDF for a positive definite matrix $\mathbf{M}$ with parameters N, L, and $\mathbf{\Sigma}$ is given by

$$f_W\left(\mathbf{M}\right) = \frac{\left|\mathbf{M}\right|^{\frac{L-N-1}{2}} \exp\left(-\frac{1}{2}\mathrm{Tr}\left(\mathbf{M}\mathbf{\Sigma}^{-1}\right)\right)}{2^{\frac{LN}{2}}\left|\mathbf{\Sigma}\right|^{N/2}\pi^{N(N-1)/4}\prod_{i=1}^{N}\Gamma\left(\frac{L-i+1}{2}\right)}.$$

E Expectation of the Product of Four Gaussian Random Variables

Let us consider four correlated Gaussian random variables X_1, X_2, X_3, and X_4, which are stacked into the random vector $\mathbf{x}$. The four-dimensional characteristic function of this MV Gaussian random vector is given by

$$\Psi(\imath\mathbf{w}) = \mathrm{E}\left[\exp\left(\imath\mathbf{w}^T\mathbf{x}\right)\right] = \exp\left(\imath\mathbf{m}^T\mathbf{w} - \frac{1}{2}\mathbf{w}^T\mathbf{S}\mathbf{w}\right),$$

where $\mathbf{m} = \mathrm{E}[\mathbf{x}]$ is the mean vector of $\mathbf{x}$ and $\mathbf{S} = \mathrm{E}\left[(\mathbf{x}-\mathbf{m})(\mathbf{x}-\mathbf{m})^T\right]$ is the covariance matrix of $\mathbf{x}$.

The expected value of the product of these four Gaussian random variables is given by

$$\mathrm{E}[X_1X_2X_3X_4] = (\imath)^4 \left.\frac{\partial^4\Psi(\imath\mathbf{w})}{\partial\mathbf{w}[1]\,\partial\mathbf{w}[2]\,\partial\mathbf{w}[3]\,\partial\mathbf{w}[4]}\right|_{\mathbf{w}[1]=\mathbf{w}[2]=\mathbf{w}[3]=\mathbf{w}[4]=0}.$$

After performing the partial derivatives and evaluating the resulting expression at $\mathbf{w}[1] = \mathbf{w}[2] = \mathbf{w}[3] = \mathbf{w}[4] = 0$, we obtain

$$\begin{aligned}\mathrm{E}[X_1X_2X_3X_4] = {} & [\mathbf{S}]_{1,2}[\mathbf{S}]_{3,4} + [\mathbf{S}]_{1,3}[\mathbf{S}]_{2,4} + [\mathbf{S}]_{1,4}[\mathbf{S}]_{2,3} + [\mathbf{S}]_{1,2}\mathbf{m}[3]\,\mathbf{m}[4] \\ & + [\mathbf{S}]_{1,3}\mathbf{m}[2]\,\mathbf{m}[4] + [\mathbf{S}]_{1,4}\mathbf{m}[2]\,\mathbf{m}[3] + [\mathbf{S}]_{2,3}\mathbf{m}[1]\,\mathbf{m}[4] \\ & + [\mathbf{S}]_{2,4}\mathbf{m}[1]\,\mathbf{m}[3] + [\mathbf{S}]_{3,4}\mathbf{m}[1]\,\mathbf{m}[2] + \mathbf{m}[1]\,\mathbf{m}[2]\,\mathbf{m}[3]\,\mathbf{m}[4].\end{aligned}$$

Acronyms

AIC	Akaike's information criterion
AoA	angle of arrival
CDF	cummulative distribution function
CIR	channel impulse response
CR	channel response
CU	central unit
ED	energy detection
EM	expectation maximization
LNA	low noise amplifier
LoS	line-of-sight
MAP	maximum a posteriori probability
ML	maximum likelihood
MV	multivariate
PAMP	power amplifier
PDF	probability density function
PDP	power delay profile
PL	path loss
PMF	probability mass function
PN	pseudo noise
RF	radio frequency
RMS	root mean square
RSS	received signal strength

RX	receiver
SNR	signal-to-noise ratio
TDoA	time difference of arrival
ToA	time of arrival
TX	transmitter
UWB	ultra-wideband

Notation

$\arg[x]$	phase angle of complex number x
c_0	speed of light in air (≈ 299702547 m/s)
$\imath$	$\sqrt{-1}$
$\mathrm{P}(H_i \vert H_j)$	probability of deciding for H_i given H_j is true
$(\cdot)^*$	complex conjugate
$*$	convolution
$\mathcal{D}_e$	average positioning error
$\triangleq$	definition
$\vert\mathbf{A}\vert$	determinant of matrix **A** (result is a scalar)
$\mathrm{diag}[\mathbf{A}]$	main diagonal of matrix **A** (result is a vector)
$\delta(t)$	Dirac delta function
$d_{i,j}$	distance between center points of regions i and j
$\hat{\Theta}$	estimate of parameter Θ
$\mathrm{E}[\cdot]$	expectation operator
$\Gamma(\cdot)$	Gamma function
$\Gamma'(\cdot)$	first derivative of Gamma function
H_m	m-th hypothesis
$\mathbf{a}^H$	hermitian transpose of column vector **a**
M	number of hypotheses, i.e. number of regions
$\mathbf{A}$	matrix
$[\mathbf{A}]_{i,j}$	j-th element in the i-th row of matrix **A**

n_{maxabs}	reference sample for temporal alignment based on maximum absolute value
N	length of location fingerprint vector
$\|\cdot\|$	L^2 norm, i.e. Euclidian norm
$\mathcal{P}_e$	total error probability
π	constant Pi (≈ 3.14159)
π_m	a priori probability of hypothesis H_m
$\text{Tr}[\mathbf{A}]$	trace of matrix **A** (result is a scalar)
$\mathbf{a}^T$	transpose of column vector **a**
a	column vector

Bibliography

[1] J. Warrior, E. McHenry, and K. McGee, "They know where you are - location detection," *IEEE Spectrum*, vol. 40, pp. 20–25, July 2003.

[2] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, pp. 70–84, July 2005.

[3] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: Possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Processing Magazine*, vol. 22, pp. 41–53, July 2005.

[4] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *WSNA '02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, (New York, NY, USA), pp. 88–97, ACM, 2002.

[5] H. Lücken, T. Zasowski, C. Steiner, F. Trösch, and A. Wittneben, "Location-aware adaptation and precoding for low complexity IR-UWB receivers," in *IEEE International Conference on Ultra-Wideband, ICUWB 2008*, vol. 3, pp. 31–34, September 2008.

[6] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proceedings of the IEEE*, vol. 97, pp. 404–426, February 2009.

[7] J. Reed, *An Introduction to Ultra Wideband Communication Systems*. Upper Saddle River, NJ, USA: Prentice Hall Press, 2005.

[8] H. Arslan, Z. N. Chen, and M.-G. D. Benedetto, *Ultra Wideband Wireless Communication*. Wiley-Interscience, 2006.

[9] M.-G. D. Benedetto, T. Kaiser, A. F. Molisch, I. Oppermann, C. Politano, and D. Porcino, *UWB Communication Systems: A Comprehensive Overview*. Hindawi Publishing Corporation, 2006.

[10] "Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems." First Report and Order, ET Doc. 98-153, FCC 02-48, Adopted: February 14, 2002, Released: April 22, 2002.

[11] F. Trösch, C. Steiner, T. Zasowski, T. Burger, and A. Wittneben, "Hardware aware optimization of an ultra low power UWB communication system," *IEEE International Conference on Ultra-Wideband, ICUWB 2007*, pp. 174–179, September 2007.

[12] V. Zeimpekis, G. M. Giaglis, and G. Lekakos, "A taxonomy of indoor and outdoor positioning techniques for mobile location services," *ACM SIGecom Exchanges*, vol. 3, no. 4, pp. 19–27, 2003.

[13] Z. Sahinoglu, S. Gezici, and I. Guvenc, *Ultra-wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols*. Cambridge University Press, 2008.

[14] S. Gezici and H. Poor, "Position estimation via ultra-wide-band signals," *Proceedings of the IEEE*, vol. 97, pp. 386–403, February 2009.

[15] S. Valaee, B. Champagne, and P. Kabal, "Localization of wideband signals using least-squares and total least-squares approaches," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1213 –1222, May 1999.

[16] K. Cheung, H. So, W.-K. Ma, and Y. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Transactions on Signal Processing*, vol. 52, pp. 1121 – 1130, April 2004.

[17] E. Larsson and D. Danev, "Accuracy comparison of LS and squared-range LS for source localization," *IEEE Transactions on Signal Processing*, vol. 58, pp. 916 –923, February 2010.

[18] I. Guvenc and Z. Sahinoglu, "Threshold-based ToA estimation for impulse radio UWB systems," *IEEE International Conference on Ultra-Wideband, ICU 2005*, pp. 420–425, September 2005.

[19] I. Guvenc, Z. Sahinoglu, and P. Orlik, "ToA estimation for IR-UWB systems with different transceiver types," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, pp. 1876–1886, June 2006.

[20] A. Rabbachin, I. Oppermann, and B. Denis, "GML ToA estimation based on low complexity UWB energy detection," *IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, p. 5, September 2006.

[21] A. Rabbachin, I. Oppermann, and B. Denis, "ML time-of-arrival estimation based on low complexity UWB energy detection," *IEEE International Conference on Ultra-Wideband, ICUWB 2006*, pp. 599–604, September 2006.

[22] D. Dardari, C.-C. Chong, and M. Win, "Threshold-based time-of-arrival estimators in UWB dense multipath channels," *IEEE Transactions on Communications*, vol. 56, pp. 1366–1378, August 2008.

[23] C. Steiner and A. Wittneben, "Robust time-of-arrival estimation with an energy detection receiver," in *IEEE International Conference on Ultra-Wideband, ICUWB 2009*, pp. 834–838, September 2009.

[24] J. D. Hol, F. Dijkstra, and T. B. Schön, "Tightly coupled UWB/IMU pose estimation," in *IEEE International Conference on Ultra-Wideband, ICUWB 2009*, pp. 688–692, September 2009.

[25] J.-Y. Lee and R. A. Scholtz, "Ranging in a dense multipath environment using an UWB radio link," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1677–1683, December 2002.

[26] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 672–681, March 2006.

[27] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, pp. 54–69, July 2005.

[28] H. Wymeersch, J. Lien, and M. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, pp. 427–450, February 2009.

[29] G. Sun, J. Chen, W. Guo, and K. J. R. Liu, "Signal processing techniques in network-aided positioning: A survey of state-of-the-art positioning designs," *IEEE Signal Processing Magazine*, vol. 22, pp. 12–23, July 2005.

[30] K. Kaemarungsi, *Design of indoor positioning systems based on location fingerprinting technique*. Ph.d. thesis, University of Pittsburgh, 2005.

[31] M. Brunato and R. Battiti, "Statistical learning theory for location fingerprinting in wireless LANs," *Computer Networks and ISDN Systems*, vol. 47, no. 6, pp. 825–845, 2005.

[32] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 1 ed., October 2007.

[33] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proceedings of IEEE INFOCOM*, (Tel Aviv, Israel), pp. 775–784, March 2000.

[34] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas, and J. Sievänen, "A probabilistic approach to WLAN user location estimation," *International Journal of Wireless Information Networks*, vol. 9, pp. 155–164, July 2002.

[35] C. Nerguizian, C. Despins, and S. Affès, "Geolocation in mines with an impulse response fingerprinting technique and neural networks," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 603–611, March 2006.

[36] C. Nerguizian and V. Nerguizian, "Indoor fingerprinting geolocation using wavelet-based features extracted from the channel impulse response in conjunction with an artificial neural network," in *IEEE International Symposium on Industrial Electronics, ISIE 2007*, pp. 2028–2032, June 2007.

[37] F. Althaus, F. Trösch, and A. Wittneben, "Geo-regioning in UWB networks," in *14th IST Mobile and Wireless Communications Summit*, (Dresden, Germany), June 2005.

[38] F. Althaus, F. Trösch, and A. Wittneben, "UWB geo-regioning in rich multipath environment," in *IEEE Vehicular Technology Conference, VTC 2005-Fall*, (Dallas, USA), pp. 1001–1005, September 2005.

[39] A. Chehri, P. Fortier, and P. M. Tardif, "Geolocation for UWB networks in underground mines," in *IEEE Annual Wireless and Microwave Technology Conference, WAMICON 2006*, p. 4, December 2006.

[40] W. Malik and B. Allen, "Wireless sensor positioning with ultrawideband fingerprinting," *First European Conference on Antennas and Propagation, EuCAP 2006*, p. 5, November 2006.

[41] R. Moses, D. Krishnamurthy, and R. Patterson, "A self-localization method for wireless sensor networks," *EURASIP Journal on Applied Signal Processing*, vol. 2003, pp. 348–358, March 2003.

[42] H. Lücken, C. Steiner, and A. Wittneben, "ML timing estimation for generalized UWB-IR energy detection receivers," in *IEEE International Conference on Ultra-Wideband, ICUWB 2009*, pp. 829–833, September 2009.

[43] C. Steiner and A. Wittneben, "Low complexity location fingerprinting with generalized UWB energy detection receivers," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1756–1767, March 2010.

[44] H. van Trees, *Detection, Estimation, and Modulation Theory*. Wiley, New York, 1968. Part I.

[45] N. Swangmuang and P. Krishnamurthy, "An effective location fingerprint model for wireless indoor localization," *Pervasive Mobile Computing*, vol. 4, no. 6, pp. 836–850, 2008.

[46] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2. Prentice Hall PTR, January 1998.

[47] J. G. Proakis, *Digital Communications*. McGraw-Hill, 4 ed., 2001.

[48] F. Althaus, F. Trösch, T. Zasowski, and A. Wittneben, "STS measurements and characterization," *PULSERS Deliverable D3b6a*, vol. IST-2001-32710 PULSERS, 2005.

[49] T. Zasowski, *A System Concept for Ultra Wideband (UWB) Body Area Networks*. Ph.d. thesis, ETH Zürich, No. 17259, 2007.

[50] C. Steiner and A. Wittneben, "UWB geo-regioning using multivariate channel statistics," in *IEEE International Conference on Communications 2007*, (Glasgow, Scottland), pp. 5763–5768, June 2007.

[51] C. Steiner, F. Althaus, F. Trösch, and A. Wittneben, "Ultra-wideband geo-regioning: A novel clustering and localization technique," *EURASIP Journal on Advances in Signal Processing*, p. 13, January 2008.

[52] A. Molisch, *Wireless Communications*. New York: Wiley-IEEE press, 2005.

[53] A. Molisch, "Ultra-wide-band propagation channels," *Proceedings of the IEEE*, vol. 97, pp. 353–371, February 2009.

[54] W. Braun and U. Dersch, "A physical mobile radio channel model," *IEEE Transactions on Vehicular Technology*, vol. 40, pp. 472–482, May 1991.

[55] R. Vaughan and J. Andersen Bach, *Channels, Propagation and Antennas for Mobile Communications*. London, U.K.: The Institiution of Electrical Engineers, 2003.

[56] D. Cassioli, M. Win, and A. F. Molisch, "The ultra-widebandwidth indoor channel: From statistical model to simulations," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1247–1257, August 2002.

[57] J. Keignart and N. Daniele, "Channel sounding and modelling for indoor UWB communications," in *International Workshop on Ultra Wideband Systems Proceedings, IWUWBS*, (Oulu, Finland), June 2003.

[58] P. Pagani, P. Pajusco, and S. Voinot, "A study of the ultra-wide band indoor channel: Propagation experiment and measurement results," in *International Workshop on Ultra Wideband Systems Proceedings, IWUWBS*, (Oulu, Finland), June 2003.

[59] U. G. Schuster and H. Bölcskei, "Ultrawideband channel modeling on the basis of information-theoretic criteria," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2464–2475, July 2007.

[60] U. G. Schuster, *Wireless Communication over Wideband Channels*. Ph.d. thesis, ETH Zürich, Series in Communication Theory, ISSN 1865-6765, 2009.

[61] H. Akaike, "Information theory and an extension of the maximum likelihood principle," *Breakthroughs in Statistics, S. Kotz and N.L. Johnson, Eds. New York, USA: Springer*, vol. 1, pp. 610–624, 1992.

[62] A. F. Molisch, K. Balakrishnan, D. Cassioli, C. Chong, S. Emami, A. Fort, J. Karedal, J. Kunisch, H. Schantz, U. Schuster, K. Siwiak, "IEEE 802.15.4a channel model - Final report," 2004.

[63] A. Molisch, "Ultrawideband propagation channels-theory, measurement, and modeling," *IEEE Transactions on Vehicular Technology*, vol. 54, pp. 1528–1545, September 2005.

[64] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, June 2005.

[65] A. Papoulis and U. S. Pillai, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 4 ed., 2002.

[66] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. 20 A, no. 1-2, pp. 32–52, 1928.

[67] T. Anderson, *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons, Inc., 3 ed., 2003.

[68] C. M. Bishop, *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1st ed., 1995.

[69] G. Abreu, F. Althaus, J. Montellet, A. Rabbachin, C. Steiner, F. Trösch, A. Wittneben, K. Yu, and T. Zasowski, "Signal processing techniques for positioning, tracking and identification," *PULSERS Deliverable D3b6b*, 2005.

[70] C. Steiner and A. Wittneben, "Cognitive interference suppression for low complexity UWB transceivers," in *IEEE International Conference on Ultra-Wideband, ICUWB 2008*, vol. 2, pp. 165–168, September 2008.

[71] H. Lücken, T. Zasowski, and A. Wittneben, "Synchronization scheme for low duty cycle UWB impulse radio receiver," in *IEEE International Symposium on Wireless Communication Systems 2008, ISWCS 2008*, pp. 503–507, October 2008.

[72] G. Tziritas, "On the distribution of positive-definite gaussian quadratic forms," *IEEE Transactions on Information Theory*, vol. 33, pp. 895–906, November 1987.

[73] A. D'Amico, U. Mengali, and E. Arias-de Reyna, "Energy-detection UWB receivers with multiple energy measurements," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2652–2659, July 2007.

[74] U. Grenander, H. O. Pollak, and D. Slepian, "The distribution of quadratic forms in normal variates: A small sample theory with applications to spectral analysis," *Journal of the Society for Industrial and Applied Mathematics*, vol. 7, no. 4, pp. 374–401, 1959.

[75] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, no. 1-2, pp. 199–201, 1960.

[76] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous Multivariate Distributions, Models and Applications*, vol. 1. Wiley, second ed., 2000.

[77] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1. Wiley, second ed., 1994.

[78] C. Steiner and A. Wittneben, "Clustering of wireless sensors based on ultra-wideband geo-regioning," in *Asilomar Conference on Signals, Systems, and Computers*, pp. 1345–1349, November 2007.

[79] G. Mclachlan and D. Peel, *Finite Mixture Models*. Wiley Series in Probability and Statistics, Wiley-Interscience, October 2000.

[80] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society Ser. B*, vol. 39, 1977.

[81] R. Redner and H. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, no. 2, pp. 195–239, 1984.

[82] J. A. Blimes, "A gentle tutorial of the EM algorithm and its application to parameter estimation for gaussian mixture and hidden markov models," *Technical Report, University of Berkeley, ICSI-TR-97-021*, 1997.

[83] D. McNamara, C. Pistorius, and J. Malherbe, *Introduction to the Uniform Geometrical Theory of Diffraction*. Artech House, 1990.

[84] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall PTR, 2 ed., December 2001.

[85] R. J.-M. Cramer, R. A. Scholtz, and M. Win, "Evaluation of an ultra-wide-band propagation channel," in *IEEE Transactions on Antennas and Propagation*, vol. 50, pp. 561–570, May 2002.

[86] R.-R. Lao, J.-H. Tarng, and C. Hsiao, "Transmission coefficients measurement of building materials for UWB systems in 3-10 GHz," in *IEEE Vehicular Technology Conference, VTC 2003-Spring*, vol. 1, pp. 11–14, April 2003.

[87] A. Richter, *Estimation of Radio Channel Parameters: Models and Algorithms*. Ph.d. thesis, Technische Universität Illmenau, 2005.

[88] A. Richter and R. Thomä, "Joint maximum likelihood estimation of specular paths and distributed diffuse scattering," in *IEEE Vehicular Technology Conference, VTC 2005-Spring*, vol. 1, pp. 11–15, June 2005.

[89] R. Zetik, J. Sachs, and R. Thomä, "Imaging of propagation environment by UWB channel sounding," in *COST273 Temporary Document TD(05) 058*, January 2005.

[90] R. Thomä, O. Hirsch, J. Sachs, and R. Zetik, "UWB sensor networks for position location and imaging of objects and environments," in *The Second European Conference on Antennas and Propagation, EuCAP 2007*, p. 9, November 2007.

[91] D. Daniels, *Ground Penetrating Radar*. Institution of Engineering and Technology, London, 2 ed., 2004.

[92] E. Kreyszig, *Advanced Engineering Mathematics*. John Wiley & Sons, 9 ed., 2006.

[93] H. V. Poor, *An introduction to signal detection and estimation*. Springer-Verlag New York, Inc., 2 ed., 1994.

[94] B. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 434–450, March 1999.

[95] J. Li and R. Wu, "An efficient algorithm for time delay estimation," *IEEE Transactions on Signal Processing*, vol. 46, pp. 2231–2235, August 1998.

[96] J. Högbom, "Aperture synthesis with a non-regular distribution of interferometer baselines," *Astronomy and Astrophysics Supplement*, vol. 15, pp. 417–426, 1974.

[97] L. G. Shapiro and G. C. Stockman, *Computer Vision*. Prentice Hall, January 2001.

[98] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, pp. 11–15, January 1972.

Curriculum Vitae

Name:	Christoph Steiner
Birthday:	March 11, 1980
Birthplace:	Graz, Austria

Education

10/2005-06/2010 **ETH Zurich, Switzerland**
PhD studies at the Communication Technology Laboratory, Department of Information Technology and Electrical Engineering

09/1999-05/2005 **Graz University of Technology, Austria**
Bachelor and Master studies in Telematics. Degree: Dipl.-Ing.

09/1990-06/1998 **Bundesrealgymnasium (BRG) Spittal a.d. Drau, Austria**
AHS-Matura

Experience

09/2005-10/2010 **ETH Zurich, Switzerland**
Research assistant at the Communication Technology Laboratory headed by Prof. Dr. Armin Wittneben

- Research in position location and wireless communication for UWB systems
- Project work and management in
  - industry projects for Phonak AG, Switzerland
  - European research project PULSERS
  - Swiss research project NCCR-MICS
- Teaching and presentation experience
- Supervision of master and semester theses
- Reviewer of international conference and journal submissions

09/2003-11/2003 **Carpathian Institute of Informatics and Enterprising Ivano-Frankivsk, Ukraine**
IAESTE scholarship

07/2002-09/2002 **Joanneum Research Graz, Austria**
Trainee in the computer graphics laboratory

07/2001-09/2001 **Infineon Technologies Graz, Austria**
Trainee in the electronics laboratory

Awards

04/2006 Master's thesis award from the Austrian Electrotechnical Association

Publications

- Journal Papers

  1. **Low Complexity Location Fingerprinting with Generalized UWB Energy Detection Receivers**
     C. Steiner and A. Wittneben, *IEEE Transactions on Signal Processing*, Mar. 2010

  2. **Ultra-Wideband Geo-Regioning: A Novel Clustering and Localization Technique**
     C. Steiner, F. Althaus, F. Trösch, and A. Wittneben, *EURASIP Journal on Advances in Signal Processing*, Jan. 2008

- Conference and Workshop Papers

  1. **UWB Positioning with Virtual Anchors and Floor Plan Information**
     P. Meissner, C. Steiner, and K. Witrisal, *Workshop on Positioning, Navigation and Communication*, WPNC 2010, Dresden, Germany, Mar. 2010

  2. **Robust Time-of-Arrival Estimation with an Energy Detection Receiver**
     C. Steiner and A. Wittneben, *IEEE International Conference on Ultra-Wideband*, ICUWB 2009, Vancouver, Canada, Sept. 2009

  3. **ML Timing Estimation for Generalized UWB-IR Energy Detection Receivers**
     H. Lücken, C. Steiner, and A. Wittneben, *IEEE International Conference on Ultra-Wideband*, ICUWB 2009, Vancouver, Canada, Sept. 2009

  4. **Cognitive Interference Suppression for Low Complexity UWB Transceivers**
     C. Steiner and A. Wittneben, *IEEE International Conference on Ultra-Wideband*, ICUWB 2008, Hanover, Germany, Sept. 2008, invited paper

  5. **Location-aware Adaptation and Precoding for Low Complexity IR-UWB Receivers**
     H. Lücken, T. Zasowski, C. Steiner, F. Trösch, and A. Wittneben, *IEEE International Conference on Ultra-Wideband*, ICUWB 2008, Hanover, Germany, Sept. 2008, invited paper

  6. **Neural network based Geo-Regioning**
     L. Leins and C. Steiner, *IEEE International Conference on Ultra-Wideband*, ICUWB 2008, Hanover, Germany, Sept. 2008

  7. **Ultra Low Power UWB Modem Design: Experimental Verification and Performance Evaluation**
     C. Steiner, H. Lücken, T. Zasowski, F. Trösch, and A. Wittneben, *Union Radio Scientifique Internationale*, URSI 2008, Chicago, USA, Aug. 2008, invited paper

  8. **Clustering of Wireless Sensors based on Ultra-Wideband Geo-Regioning**
     C. Steiner and A. Wittneben, *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Nov. 2007

9. **Hardware Aware Optimization of an Ultra Low Power UWB Communication System**
   F. Trösch, C. Steiner, T. Zasowski, T. Burger, and A. Wittneben, *IEEE International Conference on Ultra-Wideband*, ICUWB 2007, Singapore, Sept. 2007
10. **On the Interference Robustness of Ultra-Wideband Energy Detection Receivers**
    C. Steiner and A. Wittneben, *IEEE International Conference on Ultra-Wideband*, ICUWB 2007, Singapore, Sept. 2007
11. **UWB Geo-Regioning using Multivariate Channel Statistics**
    C. Steiner and A. Wittneben, *IEEE International Conference on Communications*, ICC 2007, Glasgow, Scotland, June 2007
12. **On the Performance of UWB Geo-Regioning**
    C. Steiner, F. Althaus, and A. Wittneben, *The Seventh IEEE International Workshop on Signal Processing Advances in Wireless Communications*, SPAWC 2006, Cannes, France, July 2006
13. **UWB Geo-Regioning - Algorithm and Performance**
    F. Althaus, C. Steiner, and A. Wittneben, *Workshop on Positioning Navigation and Communication*, WPNC 2006, Hanover, Germany, Mar. 2006
14. **Multiuser Interference Modeling and Suppression for a Multichannel Differential IR-UWB System**
    C. Steiner and K. Witrisal, *IEEE International Conference on Ultra-Wideband*, ICU 2005, Zurich, Switzerland, Sept. 2005