Position Estimation in Networks with Sparse Infrastructure

Dem Promotionsausschuss der
Technischen Universität Hamburg

zur Erlangung des akademischen Grades

Doktor-Ingenieur (Dr.-Ing.)

genehmigte Dissertation

von
Rico Mendrzik

aus
Hamburg, Deutschland

2020

Vorsitzender des Prüfungsausschusses
Prof. Dr.rer.nat. Turau

1. Gutachter
Prof. Dr.-Ing. Gerhard Bauch

2. Gutachter
Prof. Dr. Henk Wymeersch

Datum der mündlichen Prüfung
22. November 2019

Declaration

I, Rico Mendrzik, hereby declare that I have written this doctoral thesis independently, and that I have not made use of any aid other than those acknowledged in this thesis. Neither this thesis, nor any other similar work, has been previously submitted to any examination board.

Hamburg, February 13, 2020

Rico Mendrzik

Abstract

This thesis studies the problem of position estimation in wireless networks with sparse infrastructure from a stochastic perspective. To narrow down the subjects in that wide space, the important terms of the previous sentence will be deciphered in more detail. Position estimation refers to the ability to conclude the geometric location of a network node from a number of noisy measurements. In this thesis, only self-positioning is considered, i.e. nodes attempt to estimate their own locations. In wireless networks, the noisy measurements are derived from the radio signals that are received from one or more transmitting nodes. Typical measurements are distances or angles to the transmitting nodes. Some nodes, so-called anchors, have global knowledge of their positions. Anchors are usually considered as part of the infrastructure of a wireless network. Measurements with respect to anchors translate the relative position information into a global reference frame. If receiving nodes with unknown positions, so-called agents, attempt to estimate their positions unambiguously, they generally require measurements to multiple anchors. Networks that employ only a small number of anchors per unit area are classified as networks with sparse infrastructure. Simple positioning methods such as trilateration or triangulation fail in these networks and more sophisticated positioning strategies are required. Most of these strategies require tools from the domain of statistical signal processing. In that domain, the stochastic models of the measurement errors are used to determine position information that inherently contains the uncertainties of the positions of agents. These uncertainties are quantified by probability distributions.

Two approaches have been considered in this thesis. Both enable unambiguous positioning in wireless networks with sparse infrastructure, and they rely heavily on statistical signal processing. Particularly, the framework of belief propagation was considered in both approaches. The first approach that was studied compensates the lack of anchors by *cooperation* among agents. In particular, agents derive cooperative distance measurements from the signals received from other agents in addition to those from anchors. Cooperative measurements must be considered carefully, since other agents – in contrast to anchors – have imperfect position information. In the second approach that was studied, large antenna arrays are used to overcome the lack of infrastructure by deriving numerous measurements from the received signal of a single anchor. These measurements include angular measurements like the angle-of-arrival and angle-of-departure. Moreover, every multipath component of the received signal can be harnessed in these approaches, providing even more measurements. With such a large number of measurements per anchor, agents can position themselves using only a *single anchor*.

The contributions of this thesis are summarized in what follows. Advances in the field of cooperative positioning have been made on the algorithmic frontier. A method has been developed which allows agents to confine their locations to small areas. This method is coined polygon outer-approximation, since it constrains the sets of potential positions of

agents to convex polygons. These convex geometric constraints are valuable since they restrict the positions of agents to small subsets of the entire search space. Embedding the constraints into position estimation problems assists the actual estimators by steering their focus to the relevant regions. The usage of constraints is shown to improve the performance of cooperative position estimators in terms of accuracy, computational complexity, and speed of convergence. In single-anchor positioning, theoretical and algorithmic advances have been made. Analytical studies of the estimation problems in that field have revealed that many other parameters can be estimated on top of the positions of agents. Those parameters include agent states (orientation and clock offset with respect to the anchor) as well as the positions of reflecting or scattering objects in the agent's radio environment. The research presented in this thesis deciphered these extended estimation problems and derived fundamental insights into them. One major finding is that non-line-of-sight paths can increase the positioning accuracy if the antenna arrays of the anchor and agent as well as the signal bandwidth are large enough. Moreover, the viability of unambiguous positioning in the absence of the line-of-sight path was proven if enough non-line-sight paths exist. The theoretical part of the studies has been concluded by the derivation of the fundamental limit embodied by the Cramér-Rao lower bound. Numerical investigations of these bounds have revealed that the achievable accuracies of the extended estimation problems are only reasonable if all measurements (angles-of-arrival, angles-of-departure and times-of-arrival) are precise, i.e. the measurements contain small measurement errors. Driven by the theoretical insights, a belief propagation-based estimator has been developed that allows an agent to jointly estimate its position, orientation, and clock offset with respect to an anchor along with the positions of reflecting or scattering objects in its radio environment.

Scenarios where the infrastructure can be expected to be sparse include important future wireless networks. In many applications in the internet-of-things, data sinks (anchors) may be deployed sparsely and they will be vastly outnumbered by sensors (agents). Here, cooperative positioning approaches become relevant. Since most internet-of-things-sensors are hardware-constrained, low-complexity cooperative positioning algorithms are required. The polygon outer-approximation algorithm presented in this thesis can help to make cooperative approaches more tractable, since it adds almost no complexity but relaxes the estimation problem considerably. Fifth generation millimeter wave networks are another important area where the infrastructure can be expected to be sparse. There, base stations (anchors) are going to be deployed densely but their coverage overlap is going to be small and agents can only communicate with a single base station. Consequently, the results of the single-anchor positioning paradigm studied in this thesis become relevant.

Contents

List of Acronyms

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
5G	fifth generation
AOA	angle-of-arrival
AOD	angle-of-departure
AWGN	additive white Gaussian noise
BEB	bias error bound
BP	belief propagation
CCDF	complementary cumulative distribution function
CRLB	Cramér-Rao lower bound
D2D	device-to-device
EFIM	equivalent Fisher information matrix
FIM	Fisher information matrix
GNSS	global navigation satellite system
i.i.d.	independent identically distributed
KDE	kernel density estimate
LLS	linearized least squares
LOS	line-of-sight
LS	least squares
LTE	Long Term Evolution
MAP	maximum a posteriori
ML	maximum likelihood
mMIMO	massive multiple-input multiple-output
MMSE	minimum mean square error
mmWave	millimeter wave

NLOS	non-line-of-sight
NLS	nonlinear least squares
OEB	orientation error bound
OLOS	obstructed-line-of-sight
PBP	parametrized belief propagation
PDF	probability density function
PEB	position error bound
PHY	physical
PMF	probability mass function
POA	polygon outer-approximation
PRS	positioning reference signal
PSD	power spectral density
RMSE	root-mean-square error
SBP	sample-based belief propagation
SEB	scatterer error bound
SLAM	simultaneous localization and mapping
SNR	signal-to-noise-power ratio
SoO	signals of opportunity
TOA	time-of-arrival
TOF	time-of-flight
UAV	unmanned aerial vehicle
ULA	uniform linear array
UWB	ultrawideband
WiFi	wireless fidelity
WLS	weighted least squares

Nomenclature

Notations

$\|\cdot\|$	Euclidean norm
a	scalar
$\mathbf{a}$	vector
$\mathbf{A}$	matrix
$\mathcal{A}$	set
$\|\mathcal{A}\|$	cardinality of set $\mathcal{A}$
$\|\mathbf{A}\|$	determinant of matrix $\mathbf{A}$
$\hat{a}$	estimate of a
$\mathfrak{A}$	family of distributions
$a(y\|x)$	conditional probability density of y given x
$D(f\|\|g)$	divergence measure between probability distributions f and g
$\mathcal{R}_{N_s}(\cdot)$	particle representation with N_s particles
Rank$(\cdot)$	matrix rank
tr$(\cdot)$	matrix trace

Subscripts and Superscripts

H	transpose conjugate complex
k	particle index
l	iteration index
T	transpose

Scalar, Vector, Matrices, and Functions

α	orientation
c	speed of light

d	distance
ϵ	clock offset
$\boldsymbol{\eta}$	general purpose parameter vector
e	error
$\mathbf{J}$	Fisher information matrix
λ	eigenvalue
$\mu(\cdot)$	belief propagation message
m	mean value
$\mathcal{N}(\cdot; \mathbf{m}, \mathbf{\Sigma})$	Gaussian distribution with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{\Sigma}$
N_s	number of particles
$\mathbf{p}$	position
σ	standard deviation
$\mathbf{\Sigma}$	covariance matrix
τ	path delay
θ_{RX}	angle-of-arrival
θ_{TX}	angle-of-departure
w	particle weight
$\mathbf{z}$	measurement vector

Chapter 1

Introduction

This chapter introduces the problem of position estimation in wireless networks. It motivates the importance of obtaining accurate information regarding the positions of network nodes whose positions are unknown a priori. In that context, the related challenges and difficulties will be discussed and the approaches that are considered in this thesis will be outlined subsequently. The chapter will be concluded with a brief summary of the remainder of this thesis.

1.1 Motivation

The need for reliable and widely-available position information is ubiquitous. Knowledge regarding the relative or absolute location[1] is the foundation of numerous applications. The most prominent example which relies on position information is navigation. Already centuries ago, our ancestors had to determine their positions based on coarse maps and known landmarks in order to safely travel from one place to another. Hundreds of years ago, maritime traffic demanded the availability of position information to navigate a vessel over the ocean. Nowadays, autonomous vehicles (including cars, unmanned aerial vehicles (UAVs), and robots) rely so heavily on position information that they could not be operated without positioning capabilities. For instance, consider a UAV which is supposed to gather information from a multitude of spatially distributed sensors. In order to steer the UAV towards the sensors, the position of the UAV must be identifiable. Even outside of the domain of navigation – in a variety of applications – position information is either a prerequisite or at least provides considerable benefits. Some examples are location-aware network services or smart home applications. In those domains, position information is often a key ingredient. For instance, in so-called geo-fencing, a network grants access only to users inside a certain area. Geo-fencing can be used for cafes or shops to provide network access exclusively to their customers while denying access to users outside of their properties. In the context of the smart home, the lighting can be automatically turned on when a user enters a room or music can be re-routed to other speakers as the user moves between rooms. The spectrum of applications that benefits or explicitly requires position information seems almost infinite, but the availability of such information is often not given in reality.

[1] The terms *position* and *location* are used interchangeably. Similarly, the terms *positioning* and *localization* are treated as equivalent.

The following fundamental question arises: *What does it take to provide widely-available and reliable position information at reasonable costs?* The abstract term 'costs' will be clarified in a bit. But first, the fundamental idea of positioning will be reviewed Centuries ago, our ancestors were using landmarks like characteristic mountains to determine their locations. In the early age of maritime expeditions, the main tools for positioning were the eyes of the navigator, who had to recognize parts of the observed coastline. The accuracies of these ancient positioning techniques was admittedly very coarse, ranging from tens of meters to a few kilometers. In recent decades, the availability of highly accurate position information became integral for many applications. This gave rise to the installation of a number of radio-based positioning systems like global navigation satellite systems (GNSSs) and terrestrial positioning systems. While many aspects of radio-based positioning systems differ greatly from preceding systems, the fundamental concept used hundreds of years ago remains unchanged. In particular, radio-based positioning systems still consider observations with respect to a number of known landmarks/reference nodes (e.g., satellites or cellular base stations). The positions of these reference nodes are known and they are usually referred to as anchors. In order to obtain position-related information, anchors typically transmit known radio signals. The receiving nodes that aim to determine their own positions (agents) derive position-related observations from their received signals. Such observations include distances and angles relative to the anchors. Many positioning problems can be represented in the following unified way. Let the vector $\mathbf{z}$ denote the set of all observations (e.g, distances and angles to one or multiple anchors) and the vector $\boldsymbol{\eta}$ contains all parameters of interest (e.g., the positions of one or more agents). Then, the problem of position estimation can be formulated as an optimization problem that minimizes a cost function or maximizes a utility function. For instance,

$$\hat{\boldsymbol{\eta}} = \arg\min_{\boldsymbol{\eta}} c(\boldsymbol{\eta}, \mathbf{z}), \tag{1.1}$$

where $c(\boldsymbol{\eta}, \mathbf{z})$ is a suitable cost function. Many positioning problems can be translated into the abstract form in (1.1). With this definition, the term *costs* mentioned earlier can be specified. In particular, the costs of obtaining accurate position information are related to certain mathematical properties of the optimization problem in (1.1) which raise the following questions:

1. How many observations to how many anchors are necessary in order to obtain an unambiguous solution to the problem in (1.1)?

2. How difficult is it to solve the problem in (1.1) numerically?

Answering those questions will help to identify the important challenges in radio-based positioning. For instance, every measurement obtained comes at a certain cost (energy, system bandwidth, etc.) and every anchor that is deployed by the network is expensive. Anchors are usually part of the infrastructure of a wireless network and deploying them is costly in many regards. Setting up anchors and maintaining them requires manual labor. Moreover, anchors are usually more expensive in terms of their hardware and they also consume more energy than regular network nodes. From the cost and scalability perspectives, it becomes very attractive to deploy only a small number of anchors. On the other hand, anchors are the backbone of localization and deploying only a few of them reduces the positioning capabilities of the network. Hence there is a fundamental trade-off in radio-based positioning systems: availability and accuracy versus the sparsity of

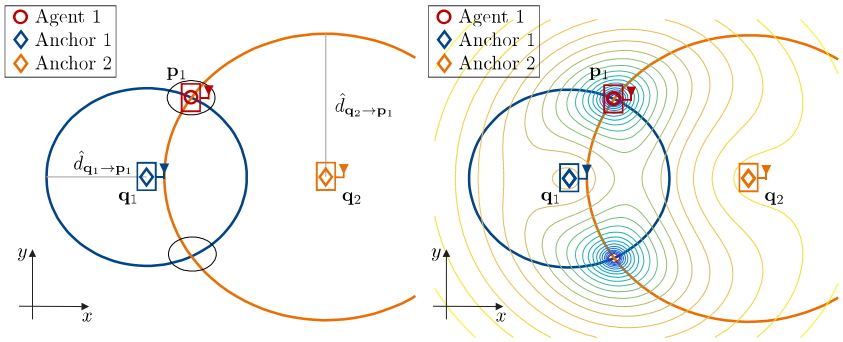

Figure 1.1: *Toy example* - an agent (red circle) attempts to determine its position using the distance estimates $\hat{d}_{\mathbf{q}_j\to\mathbf{p}_1}$, $j = 1, 2$ with respect to two anchor nodes (blue and orange diamonds, respectively). The topology is depicted on the left while a contour plot of the considered cost function is superimposed on the right. Costs are color-coded, where blue and yellow represent low and high costs, respectively. Unambiguous positioning is not possible since two intersection points (highlighted by black ellipses) are present. This can also be seen from the cost function which has two identical minima.

infrastructure. This trade-off raises the following question: *What are the possibilities to deploy networks with exceptional positioning capabilities and sparse infrastructure at the same time?* The research presented in this thesis addresses this question and attempts to find answers to it.

In order to visualize these challenges, a toy example will be used. It is depicted in Fig. 1.1. The topology is shown on the left, while a corresponding cost function is shown on the right. In this example, a single agent aims to determine its own location based on the distance estimates $\hat{d}_{\mathbf{q}_j\to\mathbf{p}_1}$ derived from the received signals of the two anchors. The number of anchors and observations in this example is too small for unambiguous positioning. Intuitively, each distance estimate $\hat{d}_{\mathbf{q}_j\to\mathbf{p}_1}$ translates into a circle that is centered at the known location of the respective anchor $\mathbf{q}_j$. The radius of the circle is given by the distance estimate. It can be observed that unambiguous positioning is not possible since the circles intersect in two points. Hence two possible locations remain. This observation is also confirmed by the corresponding cost function which has two distinct minima. Despite the simplicity of the previous example, it guides the way towards the answers to the question: *How can one reduce the number of anchors, while still being able to obtain accurate position information?* There are two abstract answers to this question:

1. Gather observations from other agents via cooperation.

2. Increase the number of observations per anchor.

The first answer relates to the field of *cooperative positioning*, while the second statement concerns the domain of *single-anchor positioning*. Both topics fall into the category of *positioning with sparse infrastructure* and they will be studied in this thesis. The concept and research challenges of the two approaches will be outlined in the following.

In cooperative positioning, agents do not only communicate with anchors (infrastructure) but also with their neighboring agents. In that way, each agent obtains additional

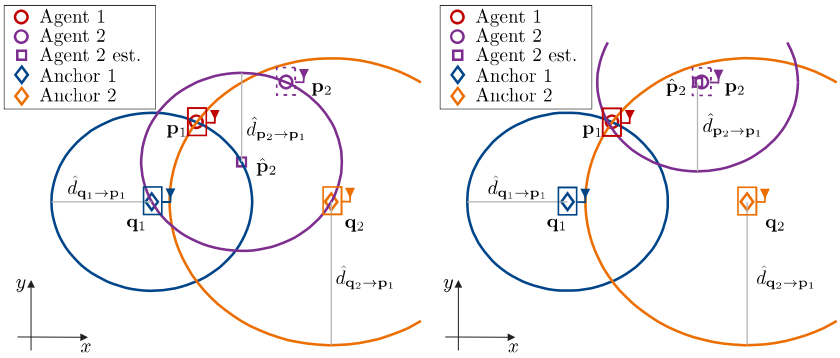

Figure 1.2: *Cooperative positioning* - agent 1 has obtained a cooperative distance estimate $\hat{d}_{\mathbf{p}_2\to\mathbf{p}_1}$ with respect to agent 2. If the position of agent 2 is known accurately (right), the cooperative distance estimate resolves the ambiguity of agent 1. If the position estimate of agent 2 is poor (left), the distance estimate is useless or potentially harmful to the positioning accuracy of agent 1.

(cooperative) observations. These observations must be treated more carefully than those from anchors. The reason is that other agents are not perfectly aware of their own positions. Hence the principal challenge in cooperative positioning is to encode the position uncertainty of agents. Only if their uncertainty is accounted for, cooperative observations can provide reliable position information. A small example which motivates the main problem of cooperative positioning is depicted in Fig. 1.2. The cooperating agent 2 is assumed to have a precise and poor position estimate in the right and left part of the figure, respectively. It can be concluded from the scenario depicted on the right that cooperating agents can act as anchors if their positioning accuracies are high and they are aware of that. Hence cooperation relaxes the requirements on the anchor density and increases the position accuracies of agents compared to non-cooperative networks. On the other hand, it can be seen that cooperating agents with imprecise position information can cause more harm than good. Especially, if cooperating neighbors are not aware or do not communicate their own position uncertainties, the positioning accuracies of surrounding agents can be degraded significantly. It can be concluded that encoding and communicating the position uncertainties of agents are the principal challenges in cooperative approaches. Moreover, the tractability of approaches which address these challenges is crucial for many practical applications.

Single-anchor positioning is an emerging field of research which is fueled by the millimeter wave (mmWave) massive multiple-input multiple-output (mMIMO) physical (PHY) layer proposal for fifth generation (5G) cellular communication systems. As the name implies, no more than one anchor is required to determine unambiguous location estimates. To facilitate this method, the agent must be able to determine angle and distance observations from the received signal. In contrast to previous transmission techniques, mmWave mMIMO enables the determination of the delays τ_j (hence the distances $d_j = c\tau_j$), angles-of-departure (AODs) $\theta_{\mathrm{TX},j}$, and angles-of-arrival (AOAs) $\theta_{\mathrm{RX},j}$ of all $j = 0, ..., J$ propagation paths of the received signal. This increases the number of position-related observations

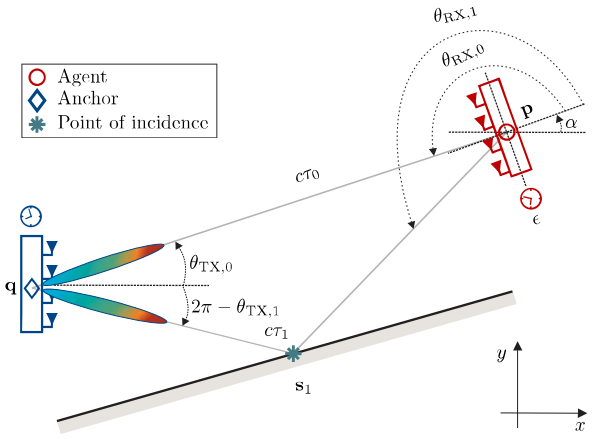

Figure 1.3: *Single-anchor positioning* - in mmWave mMIMO systems, the position-related parameters (angles and delays) of every propagation path of the radio signal can be extracted and employed for positioning. This requires estimating the reflection or scattering points $\mathbf{s}_j, \forall j$ (point of incidence), the orientation α, and possibly the clock offset ϵ. Similar to cooperative positioning, the uncertainty regarding these parameters must be encoded properly to obtain a reliable position estimate.

per anchor dramatically. Fig. 1.3 depicts an example where the agent is able to obtain accurate position information based on the received signal from a single anchor. To leverage the observations from each multipath component, the locations of the reflectors or scatterers[2] (green star in Fig. 1.3) must be estimated jointly with the position of the agent. Since the AOAs are measured in the local reference frame of the agent, the orientation α of the agent must be included in the estimation problem as well. A possible clock offset ϵ between the anchor and agent can also be included in the estimation problem in order to relax the synchronization requirements. The key challenge in single-anchor positioning resembles that of cooperative positioning. In particular, the uncertainty regarding certain parameters must be encoded properly. In contrast to cooperative positioning, it is not the uncertainty regarding the positions of other agents but rather the uncertainty of the points of incidence $\mathbf{s}_j, \forall j$, the orientation α, and the clock offset ϵ. Hence single-anchor positioning and cooperative positioning can be treated in a structurally similar way. Beyond this observation, many aspects of single-anchor positioning using mmWave signals with mMIMO technology are not yet fully understood. Therefore, it is another fundamental challenge to decompose this complicated estimation problem and reveal its mechanics, before developing estimators.

1.2 Thesis Organization

The body of this thesis contains three main parts. First, the fundamentals of radio-based positioning and statistical signal processing will be discussed in a tutorial-style fashion

[2]The term *point of incidence* will be used as a unifying expression for point of reflection or scattering.

(Chapters 2-4). Then, the research problems tackled in this thesis will be presented and their approaches will be described. Chapter 5 presents advances in cooperative positioning. Finally, the Chapters 6 and 7 depict the research conducted in the field of single-anchor positioning.

Chapter 2

This chapter describes some basic positioning principles. The difference between direct and indirect positioning will be explained. Subsequently, methods will be reviewed that extract the position-related parameters such as distances and angles from received radio signals. The canonical position estimation problem will be revisited and it will be discussed under which conditions it can be solved unambiguously. Particularly, the lateration and angulation approaches are reviewed which derive position estimates based on distance estimates and angular estimates, respectively. Examples are provided throughout this chapter to illustrate the theory.

Chapter 3

This chapter contains relevant material from the domain of statistical signal processing. Although this material will be discussed only briefly, it provides the tools which are required to understand the algorithms and analyses presented in subsequent chapters. First, a concise overview of the different representations of probability distributions is given. Subsequently, techniques, which frequently occur in statistical signal processing, will be discussed. In particular, importance sampling, resampling, and multiplication of probability densities will be treated. Throughout this chapter, a running example will be considered which is used to illustrate the abstract mathematical concepts.

Chapter 4

The tutorial part of this thesis will be concluded in this chapter by reviewing common belief propagation variants. Since the belief propagation algorithm is the cornerstone to solving all considered estimation problems in this thesis, its sample-based, parametrized, and discretized implementation will be explained. For the sake of illustration, the running example from Chapter 3 will be extended. It will be employed to visualize the theory of factor graphs as well as the different variants of belief propagation.

Chapter 5

This chapter contains the first major contribution of this thesis. It presents a novel cooperative positioning method that employs geometrical constraints in order to improve the performance of sample-based belief propagation in terms of accuracy, the speed of convergence, and computational complexity. The state of the art and the novel aspects of the presented method will be discussed first. Then, the system model will be described. Another running example will be introduced subsequently to highlight an important issue of sample-based belief propagation and to visualize the new methodology that tackles this issue. Particularly, a method called *polygon outer-approximation* will be introduced. This method determines geometrical constraints on the positions of all agents in a cooperative manner. How these constraints can be embedded in

the actual cooperative positioning problem will be discussed afterward. Then, a numerical analysis will be conducted that highlights the benefits of *constrained* cooperative position estimation. Concluding remarks will be given to close the chapter on cooperative positioning.

Chapter 6

The first part of the research regarding single-anchor positioning will be discussed in this chapter. Theoretical contributions will be presented that decipher the positioning capabilities of mmWave mMIMO systems. Before that, the topic will be introduced and the contributions will be outlined. The system model and the underlying assumptions will be discussed in the sequel. Then, two mmWave mMIMO positioning problems will be studied. Both analyses are conducted based on Fisher information. First, a position, orientation, and mapping problem will be examined, where the anchor and agent are synchronized. Secondly, the previous estimation problem will be extended by an unknown clock offset between the anchor and agent. A decomposition of the Fisher information matrix associated with the first problem will prove that non-line-of-sight (NLOS) paths increase the positioning accuracy in mmWave mMIMO systems. The Fisher information matrix associated with the second problem will reveal that joint positioning, orientation estimation, synchronization, and mapping becomes possible. A numerical study follows the theoretical analyses. This study depicts the performance limits in terms of accuracy and outlines the potential of mmWave mMIMO systems for single-anchor positioning. The results of this numerical study will be discussed to conclude the chapter.

Chapter 7

The theoretical insights from Chapter 6 will be leveraged in this chapter. In particular, a modular and unified estimation framework will be presented which can tackle a number of single-anchor positioning problems. The chapter starts with a review of existing estimators and with a list of contributions at the algorithmic forefront. Different estimation problems will be defined subsequently. Then, the estimation framework to tackle those problems will be presented based on the toughest of the presented estimation problems, where the anchor and agent are not synchronized. Particularly, a graphical representation of the problem will be derived. It will be outlined that the estimation framework can be adjusted to any other of the presented problems by simply adding or removing elements from the graphical representation. A sample-based variant of the belief algorithm will be proposed and evaluated numerically. Based on the numerical results, conclusions will be drawn and open problems will be identified.

Chapter 8

This chapter summarizes the thesis. The major issues of positioning in the context of sparse infrastructure are briefly reviewed and the presented solutions to these problems are recapitalized. The main contributions are highlighted in a compact form and future research directions are pointed out.

Chapter 2

Basic Principles of Radio-based Positioning

Radio-based positioning is a comprehensive topic. It includes an enormous number of aspects which makes a complete description impossible, considering the scope of this thesis. Nonetheless, the aim of this chapter is to give a broad but still concise overview of the aspects that will be important for this thesis. In doing so, a multitude of topics will be touched upon, starting with the definition of the general positioning problem. The fundamental concepts of direct and indirect positioning are differentiated, where the latter concept is in the focus. Therefore, strategies for distance and angle-of-arrival estimation based on the received signal are revisited. In order to transition from these position-related parameters to position estimates, lateration and angulation are recapitulated and the necessary conditions to obtain unambiguous position estimates are discussed.

2.1 Canonical Positioning Problem

Fig. 2.1 depicts the canonical positioning problem. Radio-based positioning relies on the transmission and reception of radio waves. The basic principle can be summarized as follows: One or multiple transmitters with known locations transmit radio signals $s_j(t), \forall j$ over a wireless channel. These signals are known to the receiver. Then, the receiver observes distorted versions of these signals. Based on the received signals $r_j(t), \forall j$, the receiver attempts to infer its location $\boldsymbol{\eta} \triangleq \mathbf{p}_1$.

Radio-based position is almost as old as the discovery of radio-waves itself. In 1904, the entrepreneur and inventor Christian Hülsmeyer invented the Telemobiloskop, an early version of a radar, to determine the position of vessels on waterways. Even though his groundbreaking idea is more than a hundred years old, its essence remains unchanged and is still used to solve today's positioning problems. What Hülsmeyer realized was that received radio waves contain spatial information, i.e. the received radio signals contain information about the relative positions of the transmitters and the receiver.

Most position estimation problems can be written as an optimum parameter estimation problem if all parameters of interest (e.g., the position of an agent) and all observations (e.g., samples from the received radio signals, distances, or angles) are given by $\boldsymbol{\eta}$ and $\mathbf{z}$, respectively:

$$\hat{\boldsymbol{\eta}} = \arg\max_{\boldsymbol{\eta}} p(\boldsymbol{\eta}|\mathbf{z}), \tag{2.1}$$

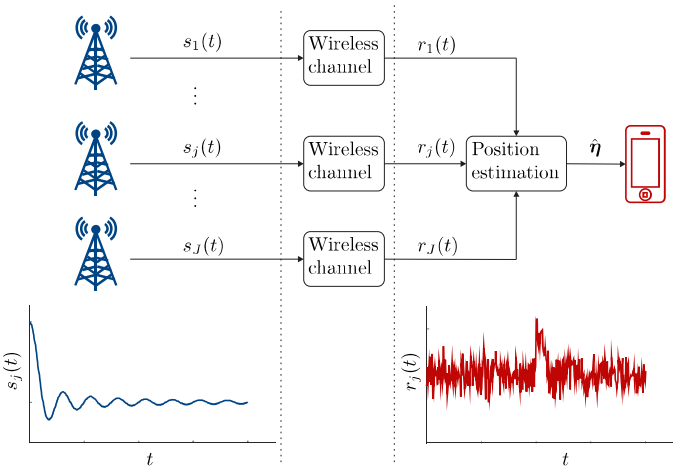

Figure 2.1: *Radio-based positioning* - known signals $s_j(t), \forall j$ are transmitted over the wireless channel and the receiver attempts to determine an estimate $\hat{\boldsymbol{\eta}}$ of its own location $\boldsymbol{\eta}$ based on the received signals $r_j(t), \forall j$.

where $p(\boldsymbol{\eta}|\mathbf{z})$ is the a posteriori distribution (short: *posterior distribution*). For simplicity and clarity of the explanations, it is assumed in this chapter that the parameter of interest is the position of a single agent in x and y directions, i.e. $\boldsymbol{\eta} \triangleq \mathbf{p}_1 = [p_{1,\mathrm{x}}, p_{1,\mathrm{y}}]^{\mathrm{T}}$. There are two fundamentally different paradigms in radio-based positioning; *direct* and *indirect* positioning [GP09]. Fig. 2.2 shows the difference between the two approaches graphically.

In direct positioning, the goal is to determine an estimate of the position of an agent directly based on the received signal. Since most state-of-the-art receivers are implemented digitally, it is assumed that Z samples have been obtained from all J received signals $r_j(t), j = 1, .., J$. The samples from the j^{th} signal are collected in the vector $\mathbf{r}_j = [r_{j,1}, ..., r_{j,Z}]^{\mathrm{T}}$. If all samples from all received signals are stacked in the observation vector $\mathbf{z} \triangleq [\mathbf{r}_1^{\mathrm{T}}, ..., \mathbf{r}_J^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{\eta} \triangleq \mathbf{p}_1$, the optimum direct positioning problem is given by (2.1).

Prominent examples of direct positioning schemes are variants of fingerprinting, where the received signals to multiple anchors nodes (e.g., wireless fidelity (WiFi) access points) are used to infer the position of an agent. The basic principle of fingerprinting is the following: Fingerprinting consists of two phases. First, a map with a finite grid of an environment is built (offline stage). For each point on the map, the received signals to all available anchor nodes are associated with this point and the map is stored in a database. This data base is typically known to all agents. If an agent attempts to estimate its position (online phase), the received signals at the agents are compared with the received signals from the map in the database. The position estimate is given by the point of the database whose received signals show the highest similarity to the actual received signals of the agent.

In contrast to direct positioning, position-related parameters like distances, angles-of-departure (AODs), and angles-of-arrival (AOAs) are determined in indirect positioning. They are estimated based on the samples from the received signals. These estimates are

then used to determine the position of an agent[1]. Assume that, e.g., the distance estimates $\hat{d}_j$ have been derived from all J received signals $r_j(t), j = 1, ..., J$. If the distance estimates are collected in the observation vector $\mathbf{z} \triangleq [\hat{d}_1, ..., \hat{d}_J]^{\mathrm{T}}$ and $\boldsymbol{\eta} \triangleq \mathbf{p}_1$ is considered, the maximum a posteriori (MAP)-optimum indirect positioning problem is given by (2.1). The previous example is a special case of indirect positioning, where only distance estimates are used. This case is called lateration. For the remainder of this thesis, only indirect positioning will be considered.

2.2 Estimation of Position-related Parameters

As discussed in the previous section, the position-related parameters provide the basis for indirect positioning. Therefore, this section is dedicated to the review of position-related parameter estimation. In particular, time-of-arrival (TOA) and AOA estimation are considered. It should be emphasized that this section does not claim completeness by any means, but is intended to provide the reader with a coarse intuition of position-related parameter estimation.

2.2.1 Time-of-Arrival

2.2.1.1 Basic Principle

The TOA of a signal traveling from one network node to another node can be used to infer the distance between those two nodes.

Intuition
Suppose that both nodes share the same clock and the transmitter includes the time instance of transmission in the package to be transmitted. When the package arrives at the receiver, the time instance of reception is measured. Then, the traveling time is given by

[1]From the position estimation perspective, the position-related parameters are observations. Strictly speaking, they are parameter estimates themselves. For the sake of differentiability, position-related parameters will be referred to as observations or measurements after Section 2.2.

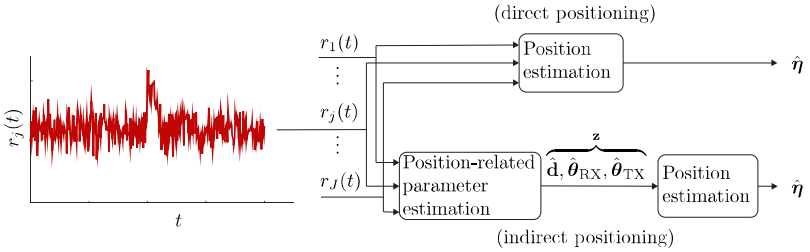

Figure 2.2: *Direct vs indirect positioning* - in direct positioning, the position is directly estimated from the received signal (e.g., by comparing the received signal with stored signals from a data base), while indirect positioning approaches first obtain estimates on position-related parameters $\mathbf{z}$ (e.g., distances $\hat{\mathbf{d}}$ and angles $\hat{\boldsymbol{\theta}}_{\mathrm{RX/TX}}$) based on which the position is estimated.

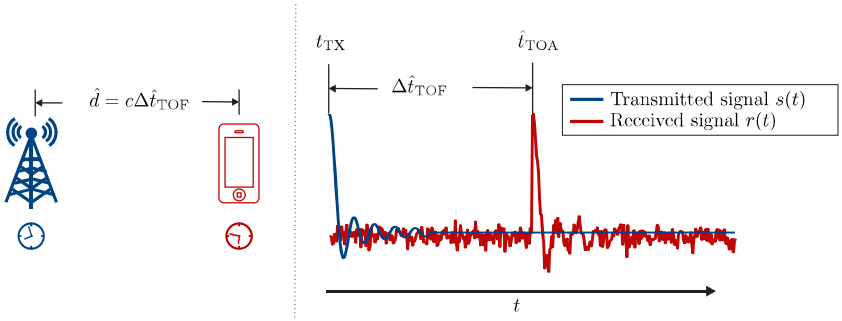

Figure 2.3: *TOA-based distance estimation* - a known signal $s(t)$ is transmitted over a noisy wireless channel and the receiver determines the TOA estimate $\hat{t}_{\mathrm{TOA}}$ based on the received signal $r(t)$. From this estimate and the transmission time t_{TX}, the receiver can infer the time-of-flight (TOF) $\Delta\hat{t}_{\mathrm{TOF}}$ and hence the distance $\hat{d}$.

the difference of the time instances of reception and transmission. Finally. the distance can be concluded from the traveling time and the speed of light.

Mathematical description

The transmitting node sends a known signal $s(t)$ over the wireless channel. The time required to travel from one node to the other is denoted by the TOF $\Delta t_{\mathrm{TOF}} = d/c$. where d is the distance between the nodes and c is the speed of light. The receiver cannot directly determine the TOF but rather estimates the time instance when the signal arrives. This time instance is called time-of-arrival and denoted by t_{TOA}. The TOA is then given in the receivers local time frame. In order to determine the TOF, the receiver must know when the signal was transmitted. Hence both nodes must have a common clock and the instance of transmission t_{TX} must be conveyed to the receiver. In that case, the receiver can determine a distance estimate based on the estimated TOA $\hat{t}_{\mathrm{TOA}}$, the known instance of transmission t_{TX}, and the speed of light c, i.e.

$$\hat{d} = d + e_d = c\Delta\hat{t}_{\mathrm{TOF}\cdot} = c \cdot (\hat{t}_{\mathrm{TOA}} - t_{\mathrm{TX}}), \tag{2.2}$$

where e_d denotes the distance estimation error. A schematic plot is depicted in Fig. 2.3. Note that the synchronization of transmitter and receiver is a critical issue since every nanosecond of offset will result in a systematic ranging error of $|\Delta d| \approx 0.3$ m. The synchronization assumption can be removed at the cost of increased ranging overhead if so-called two-way protocols are considered [LS02, ZLK07].

In the following, it will be explained how the receiver can obtain the TOA estimate $\hat{t}_{\mathrm{TOA}}$ using a simple correlation-based TOA estimator [DCF+09]. Notice that a huge variety of TOA estimators exists in the literature. For the sake of brevity, only the previously-mentioned approach is considered since this basic principle is prevalent in almost all other TOA estimators. The block-diagram of a correlation-based TOA estimator is depicted in Fig. 2.4. The receiver correlates the received signal[2] $r(t) = s(t - \Delta t_{\mathrm{TOF}}) + n(t)$ with the known transmitted pulse $s(t)$. i.e. $r_{rs}(\tau) = r(-\tau)^* s(\tau)$. The TOA estimate is given by

[2]Here, an additive white Gaussian noise (AWGN) channel is assumed.

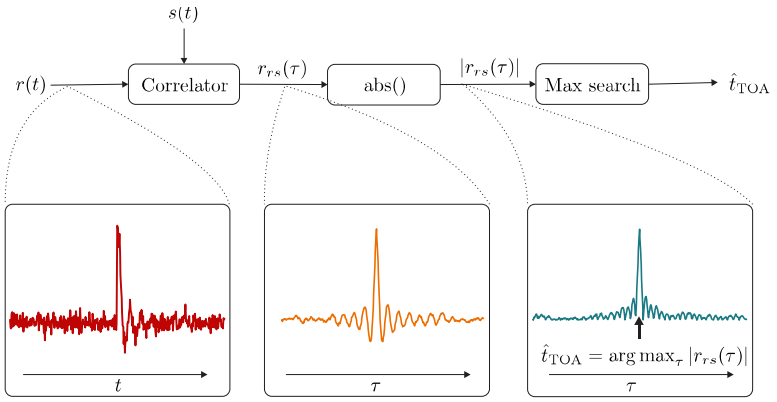

Figure 2.4: *TOA estimator* - the received signal $r(t)$ is correlated with the known transmit signal $s(t)$. The time instance with the highest correlation yields the TOA estimate $\hat{t}_{\mathrm{TOA}}$.

the time instance corresponding to the maximum absolute value of the correlated signal, i.e. $\hat{t}_{\mathrm{TOA}} = \arg\max_\tau |r_{rs}(\tau)|$. In an AWGN channel, this estimator yields a maximum likelihood (ML) estimate and its performance attains the Cramér-Rao lower bound (CRLB) for large signal-to-noise-power ratios (SNRs) [Tre02].

2.2.1.2 Main Sources of Error

The accuracy of the TOA estimates is mainly impacted by noise, multipath propagation, interference, and synchronization errors [DCF+09]. Noise perturbs the received signal which generally leads to a random shift of the highest peak of the correlated signal $|r_{rs}(\tau)|$. Hence the TOA estimate $\hat{t}_{\mathrm{TOA}}$ will be distorted in the presence of noise. Interference due to multipath propagation also causes estimation errors if the multipath components have a considerable delay spread. In multipath environments, more sophisticated TOA estimators than the one presented in this section are required [STK05]. Increasing the bandwidth makes the multipath components resolvable in the temporal domain which alleviates the effect of multipath propagation on the estimation error [PAK+05]. In-band and out-of-band interference from other systems also cause a degradation of the estimation accuracy. Moreover, recall that TOA estimation requires perfect synchronization between the transmitter and receiver. Even after synchronization, remaining clock offsets or clock drift lead to reduced TOA estimation accuracy. The distribution of the distance estimation errors depends on the environment and the actual estimators. It can be modeled as Gaussian distribution [YWI08, N. 03], exponential distribution [MB17, VWS12, GWSR11b], and Gaussian mixture distribution [WLW09]. The standard deviations of the distributions vary from a few centimeters [MPG15] to tens of meters [BHE00], depending mainly on the transmission environment, bandwidth, and SNR. The distribution of the error becomes important in the context of Section 5.4.

2.2.2 Angle-of-Arrival

2.2.2.1 Basic Principle

Fig. 2.5 depicts the schematic of a simple AOA estimator. Consider a transmitter with a single antenna and a receiver whose uniform linear array (ULA) consists of N_{RX} antennas.

Intuition
If an impinging plane wave arrives at the receiver under a certain angle, the wave has traveled different distances to each of the antenna elements. The actual differences depend on the antenna spacing Δd and the impinging angle of the wave, which is also called the angle-of-arrival. Due to the different path lengths, the phases of the received signals differ from antenna element to antenna element. The AOA is concluded based on these phase differences and the known antenna spacing Δd.

Mathematical Description
The array response vector is given by

$$\mathbf{a}(\theta_{\mathrm{RX}}) = \begin{bmatrix} 1 \\ \exp(j2\pi\Delta d/\lambda \sin(\theta_{\mathrm{RX}})) \\ \vdots \\ \exp(j2\pi\Delta d(N_{\mathrm{RX}}-1)/\lambda \sin(\theta_{\mathrm{RX}})) \end{bmatrix}. \tag{2.3}$$

The transmitter sends a known unit-energy signal $s(t)$ which is received on every receive antenna. Based on the received signals $r_i(t), i = 1, ..., N_{\mathrm{RX}}$ on all antennas, the receiver aims to estimate the AOA θ_{RX}. For simplicity, it is assumed that only the LOS path is present. The noisy observations on all antennas are collected in the vector $\mathbf{r}(t) = [r_1(t), ..., r_{N_{\mathrm{RX}}}(t)]^{\mathrm{T}} = \mathbf{a}(\theta_{\mathrm{RX}})s(t - \Delta t_{\mathrm{TOF}}) + \mathbf{n}$, where $\mathbf{n}$ the noise vector. For digital processing, it is assumed that Q samples of all individual signal $r_i, i = 1, ..., N_{\mathrm{RX}}$ with sampling period T_{sample} have been obtained. The q^{th} sample on the i^{th} antenna is denoted by $r_i(q)$ and the samples from all antennas at time qT_{sample} are collected in the vector $\mathbf{r}(q) = [r_1(q), ..., r_{N_{\mathrm{RX}}}(q)]^{\mathrm{T}}$. Furthermore, it is assumed that the TOF is known[3] (e.g., by employing an estimator as described in 2.2.1 prior to AOA estimation). To estimate the AOA, the received signal on every antenna is first correlated with the transmitted signal shifted by the TOF, i.e.

$$\tilde{\mathbf{a}} = \frac{1}{Q}\sum_{q=1}^{Q} \mathbf{r}(q)s^*(q - q_{\Delta t_{\mathrm{TOF}}}) \tag{2.4}$$

$$= \mathbf{a}\underbrace{\frac{1}{Q}\sum_{q=1}^{Q} |s^*(q - q_{\Delta t_{\mathrm{TOF}}})|^2}_{=1} + \frac{1}{Q}\sum_{q=1}^{Q} \mathbf{n}(q)s^*(q - q_{\Delta t_{\mathrm{TOF}}}), \tag{2.5}$$

where $q_{\Delta t_{\mathrm{TOF}}}$ is the time-discrete index in multiples of the sampling period corresponding to the TOF. Based on the correlation values $\tilde{\mathbf{a}}$, the receiver determines an estimate of the

[3] If the TOF is unknown, other AOA estimation strategies can be considered that estimate the absolute phase difference (induced by the TOF) jointly with the AOA.

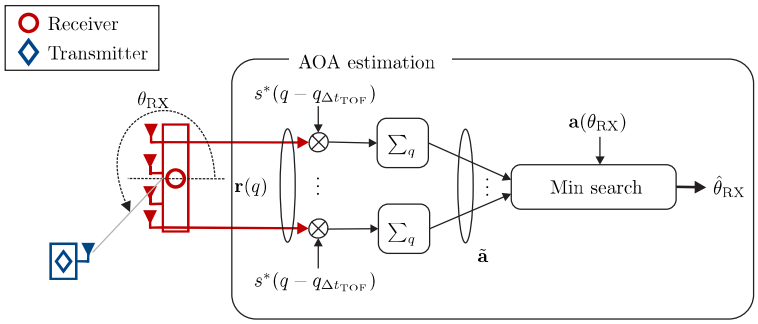

Figure 2.5: *AOA estimator* - the digital received signals $\mathbf{r}(q)$ from all antennas are processed in order to estimate the steering vector $\tilde{\mathbf{a}}$. The AOA estimate $\hat{\theta}_{\mathrm{RX}}$ is given by the angle of the steering vector $\mathbf{a}(\theta_{\mathrm{RX}})$ which minimizes the Euclidean distance to the estimated steering vector $\tilde{\mathbf{a}}$.

AOA by identifying the value of θ_{RX} for which $\tilde{\mathbf{a}}$ and $\mathbf{a}(\theta_{\mathrm{RX}})$ have the smallest Euclidean distance. i.e.

$$\hat{\theta}_{\mathrm{RX}} = \arg\min_{\theta_{\mathrm{RX}}} \|\tilde{\mathbf{a}} - \mathbf{a}(\theta_{\mathrm{RX}})\| . \tag{2.6}$$

Such an optimization problem could be solved, e.g., using a finite grid of angles $\boldsymbol{\theta}_{\mathrm{trial}} = [\theta_{1,\mathrm{trial}}, \cdots, \theta_{N_{\mathrm{trial}},\mathrm{trial}}]^{\mathrm{T}}$ and evaluating the expression $\left\|\tilde{\mathbf{a}} - \mathbf{a}(\theta_{i,\mathrm{trial}})\right\|, i = 1, ..., N_{\mathrm{trial}}$. The optimum angle $\hat{\theta}_{\mathrm{RX}}$ is then given by the angle $\theta_{i,\mathrm{trial}}$ for which the expression $\left\|\tilde{\mathbf{a}} - \mathbf{a}(\theta_{i,\mathrm{trial}})\right\|$ is minimized. Note that for very precise estimates, the granularity of the grid must be extremely fine.

2.2.2.2 Main Sources of Error

AOA estimates are impaired by the same sources discussed in the TOA section: Mainly by additive noise and multipath. The resulting AOA estimation errors are typically modeled as zero-mean Gaussian with standard deviation $\sigma_{\theta_{\mathrm{RX}}}$ ranging from sub-1° [SGD+18a] to a few degrees [AP04], depending on the number of antennas and the SNR.

2.3 Lateration

Lateration refers to distance-based position estimation, where the position of an agent is determined based on distinct distance estimates with respect to multiple anchors. A special case of lateration is so-called trilateration, where an agent determines its own position using the distance estimates to exactly three anchors.

2.3.1 Problem Formulation

Fig. 2.6 (left) depicts the lateration-based positioning problem with one agent and $M = 3$ anchors in the absence of noise. The position of the agent and the positions of the anchors are denoted by $\mathbf{p}_1$ and $\mathbf{q}_i, i = 1, ..., M$, respectively. As mentioned in Section 2.1, the MAP-optimum solution to positioning problems can be expressed according to (2.1). In

lateration, the observation vector $\mathbf{z}$ summarizes the distance estimates $\hat{d}_{\mathbf{q}_i \to \mathbf{p}_1}, i = 1, ..., M$ to all anchors, i.e.

$$\mathbf{z} = [\hat{d}_{\mathbf{q}_1 \to \mathbf{p}_1}, ..., \hat{d}_{\mathbf{q}_M \to \mathbf{p}_1}]^{\mathrm{T}}. \tag{2.7}$$

The parameter vector $\boldsymbol{\eta}$ contains only the position of the agent $\mathbf{p}_1$, i.e. $\boldsymbol{\eta} \triangleq \mathbf{p}_1$. The goal is to infer the position of the agent given all observations $\mathbf{z}$.

Intuition
In the absence of noise, solving this problem breaks down to finding the intersection of M circles which are respectively centered around the positions of the anchors with radii equal to the distances from the anchors to the agent. For general geometric constellations, the intersection of at least three yields a single point. Hence the sufficient condition to unambiguously solve the noise-free lateration problem is that the number of distance estimates to distinct anchors must be larger or equal to three [ZB11], i.e. $M \geq 3$. Notice that, in some special cases, the problem can be solved unambiguously; even if $M = 2$. An example is depicted in Fig. 2.6 (bottom).

2.3.2 Lateration in the Absence of Noise

When no noise is present, the distance estimates $\hat{d}_{\mathbf{q}_i \to \mathbf{p}_1}, i = 1, ..., M$ become equal to the true distances $d_{\mathbf{q}_i \to \mathbf{p}_1} = \left\| \mathbf{p}_1 - \mathbf{q}_i^* \right\|, i = 1, ..., M$, where $\mathbf{q}_i^*$ denotes the known position of the i^{th} anchor. Then, the positioning problem can be reformulated as a nonlinear system of equations, where the i^{th} distance induces the following equation:

$$d_{\mathbf{q}_i \to \mathbf{p}_1} = \sqrt{(p_{1,\mathrm{x}} - q_{i,\mathrm{x}}^*)^2 + (p_{1,\mathrm{y}} - q_{i,\mathrm{y}}^*)^2}, \quad \text{for } i = 1, ..., M. \tag{2.8}$$

Squaring both sides and rearranging the terms yields

$$\frac{(p_{1,\mathrm{x}} - q_{i,\mathrm{x}}^*)^2}{d_{\mathbf{q}_i \to \mathbf{p}_1}^2} + \frac{(p_{1,\mathrm{y}} - q_{i,\mathrm{y}}^*)^2}{d_{\mathbf{q}_i \to \mathbf{p}_1}^2} = 1, \tag{2.9}$$

which is the equation of a circle with mid-point $\mathbf{q}_i^*$ and radius $d_{\mathbf{q}_i \to \mathbf{p}_1}$. If M anchors are considered, the localization problem boils down to finding the intersection of M circles. An illustrative example is depicted in Fig. 2.6. This nonlinear system of equations can also be transformed into a linear system of equations by subtracting any equation i like that in (2.9) from all others ($j \neq i$). As a consequence, $M - 1$ linear equations result. These equations have the following form:

$$p_{1,\mathrm{x}}(q_{i,\mathrm{x}}^* - q_{j,\mathrm{x}}^*) + p_{1,\mathrm{y}}(q_{i,\mathrm{y}}^* - q_{j,\mathrm{y}}^*) = \frac{1}{2}\left(d_{\mathbf{q}_j \to \mathbf{p}_1}^2 - d_{\mathbf{q}_i \to \mathbf{p}_1}^2 + \|\mathbf{q}_i^*\|^2 - \left\|\mathbf{q}_j^*\right\|^2 \right), \tag{2.10}$$

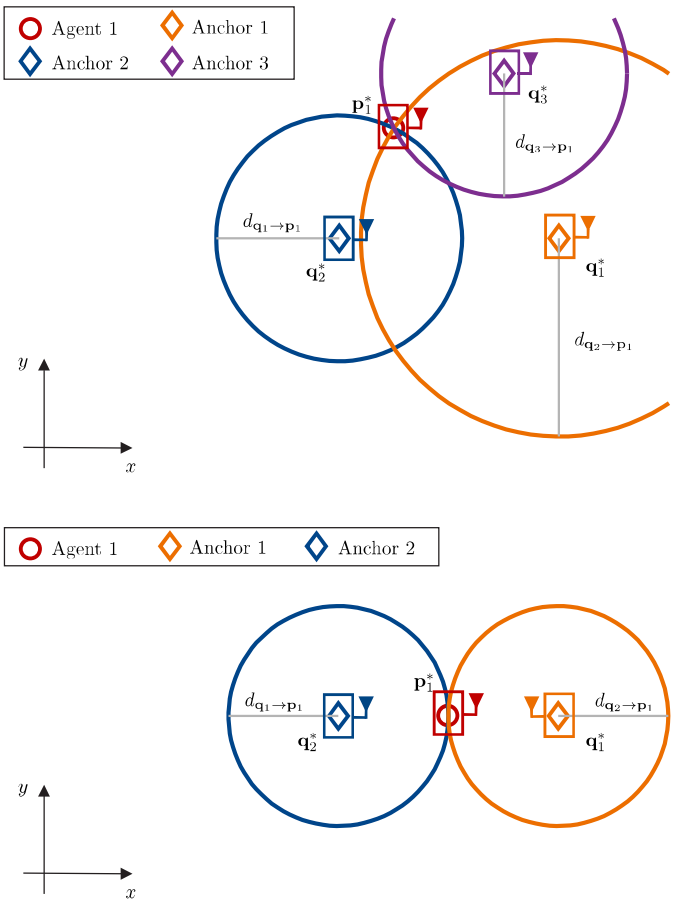

Figure 2.6: *Lateration examples* - lateration with $M = 3$ (top) and $M = 2$ (bottom) anchors. Generally, $M \geq 3$ anchors are necessary to obtain an unambiguous position estimate. In some special cases (bottom), unambiguous positioning is possible with only two anchors.

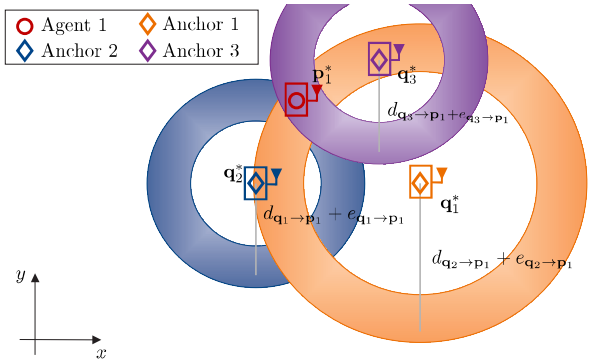

Figure 2.7: *Lateration in the presence of noise* - when noise corrupts the received signals, distance estimation errors $e_{\mathbf{q}_i\to\mathbf{p}_1}, \forall i$ result. Intuitively, the uncertainty caused by these errors results in annular shapes (loosely speaking: donuts). In the presence of noise, there is not unique solution to the lateration problem in 2.11. Hence a rigorous position estimation strategy like a least squares approach will be required.

for $j = 1, ...i-1, i+1, ..., M$. These equations can be summarized as the matrix vector product

$$\underbrace{\begin{bmatrix} q_{i,\mathrm{x}}^* - q_{1,\mathrm{x}}^* & q_{i,\mathrm{y}}^* - q_{1,\mathrm{y}}^* \\ \vdots & \vdots \\ q_{i,\mathrm{x}}^* - q_{i-1,\mathrm{x}}^* & q_{i,\mathrm{y}}^* - q_{i-1,\mathrm{y}}^* \\ q_{i,\mathrm{x}}^* - q_{i+1,\mathrm{x}}^* & q_{i,\mathrm{y}}^* - q_{i+1,\mathrm{y}}^* \\ \vdots & \vdots \\ q_{i,\mathrm{x}}^* - q_{M,\mathrm{x}}^* & q_{i,\mathrm{y}}^* - q_{M,\mathrm{y}}^* \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} p_{1,\mathrm{x}} \\ p_{1,\mathrm{y}} \end{bmatrix}}_{\mathbf{p}} = \underbrace{\frac{1}{2}\begin{bmatrix} d_{\mathbf{q}_1\to\mathbf{p}_1}^2 - d_{\mathbf{q}_i\to\mathbf{p}_1}^2 + \left\|\mathbf{q}_i^*\right\|^2 - \left\|\mathbf{q}_1^*\right\|^2 \\ \vdots \\ d_{\mathbf{q}_{i-1}\to\mathbf{p}_1}^2 - d_{\mathbf{q}_i\to\mathbf{p}_1}^2 + \left\|\mathbf{q}_i^*\right\|^2 - \left\|\mathbf{q}_{i-1}^*\right\|^2 \\ d_{\mathbf{q}_{i+1}\to\mathbf{p}_1}^2 - d_{\mathbf{q}_i\to\mathbf{p}_1}^2 + \left\|\mathbf{q}_i^*\right\|^2 - \left\|\mathbf{q}_{i+1}^*\right\|^2 \\ \vdots \\ d_{\mathbf{q}_M\to\mathbf{p}_1}^2 - d_{\mathbf{q}_i\to\mathbf{p}_1}^2 + \left\|\mathbf{q}_i^*\right\|^2 - \left\|\mathbf{q}_M^*\right\|^2 \end{bmatrix}}_{\mathbf{b}.} \tag{2.11}$$

Note that the system of equations in (2.11) is linear even though the problem is of nonlinear nature in its original form. Any two rows of the system of equations in (2.11) solve the positioning problem. Notice that two columns include the distances to three anchors. The system can be solved by elementary row operations or matrix inversion, i.e.

$$\mathbf{p} = \frac{1}{2}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}. \tag{2.12}$$

2.3.3 Lateration in the Presence of Noise

In the presence of noise, the system of equations in (2.11) does not have a unique solution, since the true distances are replaced by noisy distance estimates, i.e. $d_{\mathbf{q}_i\to\mathbf{p}_1} \to \hat{d}_{\mathbf{q}_i\to\mathbf{p}_1}$. The consequences are depicted graphically in Fig. 2.7. If noise is present, the positioning problem in (2.11) needs to be reformulated by replacing the vector $\mathbf{b}$ by $\hat{\mathbf{b}}$, where the j^{th} entry is defined as follows $[\hat{\mathbf{b}}]_j = \hat{d}_{\mathbf{q}_j\to\mathbf{p}_1}^2 - \hat{d}_{\mathbf{q}_i\to\mathbf{p}_1}^2 + \left\|\mathbf{q}_i^*\right\|^2 - \left\|\mathbf{q}_j^*\right\|^2$.

In general, the noise on the observations gives rise to the introduction of an optimality criterion like the least squares (LS) criterion:

$$\hat{\mathbf{p}}_1 = \arg\min_{\mathbf{p}_1} \sum_{i=1}^{M} \left(\hat{d}_{\mathbf{q}_i \to \mathbf{p}_1} - \sqrt{(p_{1,\mathrm{x}} - q^*_{i,\mathrm{x}})^2 + (p_{1,\mathrm{y}} - q^*_{i,\mathrm{y}})^2} \right)^2 . \tag{2.13}$$

Notice that the optimization problem in (2.13) is a nonlinear least squares (NLS) problem, whose solution generally requires numerical solvers like the *Levenberg-Marquardt* [Mar63] or *Trust-Region-Reflective* [Sor82] algorithm. However, using the linearized notation of the problem in (2.11), it can be shown that

$$\hat{\mathbf{p}}_1 = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\hat{\mathbf{b}} \tag{2.14}$$

provides an optimum estimate in the linearized least squares (LLS) sense [ZB11]. Generally, the linearized solution in (2.14) and the nonlinear solution obtained by, e.g, the Levenberg-Margquardt algorithm differ. A numerical example in the following subsection emphasizes this difference. Similar to the noise-free case, any two rows in (2.11) provide an LLS position estimate. However, notice that considering more than two distance estimates generally increases the positioning accuracy [ZB11].

If the statistics of the noise are known, the MAP criterion can be considered as defined in (2.1). In the following, the example of zero-mean Gaussian distance estimation noise is considered. For simplicity, it is assumed that the corresponding variances are identical, i.e. $\sigma_d^2 \triangleq \sigma^2_{d_{\mathbf{q}_i \to \mathbf{p}_1}} = \sigma^2_{d_{\mathbf{q}_j \to \mathbf{p}_1}}, \forall i, j$. In that case, it can be easily verified that the MAP criterion can be written as the LS criterion in (2.13). Assuming conditional independence of the distance estimates, the MAP distribution can be written as

$$f(\mathbf{p}_1|\mathbf{z}) \propto \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi\sigma_d^2}} \exp\left(-\frac{\left(\hat{d}_{\mathbf{q}_i \to \mathbf{p}_1} - \sqrt{(p_{1,\mathrm{x}} - q^*_{i,\mathrm{x}})^2 + (p_{1,\mathrm{y}} - q^*_{i,\mathrm{y}})^2} \right)^2}{\sigma_d^2} \right) f(\mathbf{p}_1), \tag{2.15}$$

where $f(\mathbf{p}_1)$ is the prior information on the position of agent 1. For simplicity of the argument, uniform prior information is assumed and hence $f(\mathbf{p}_1|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{p}_1)$ follows. Substituting (2.15) into (2.1) yields

$$\hat{\mathbf{p}}_1 = \arg\max_{\mathbf{p}_1} \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi\sigma_d^2}} \exp\left(-\frac{\left(\hat{d}_{\mathbf{q}_i \to \mathbf{p}_1} - \sqrt{(p_{1,\mathrm{x}} - q^*_{i,\mathrm{x}})^2 + (p_{1,\mathrm{y}} - q^*_{i,\mathrm{y}})^2} \right)^2}{\sigma_d^2} \right), \tag{2.16}$$

Taking the log of the argument in (2.16) and dropping the terms that are constant for the maximization over $\mathbf{p}_1$ yields

$$\hat{\mathbf{p}}_1 = \arg\max_{\mathbf{p}_1} \sum_{i=1}^{M} - \left(\hat{d}_{\mathbf{q}_i \to \mathbf{p}_1} - \sqrt{(p_{1,\mathrm{x}} - q^*_{i,\mathrm{x}})^2 + (p_{1,\mathrm{y}} - q^*_{i,\mathrm{y}})^2} \right)^2 . \tag{2.17}$$

By changing the signs of each term under the sum, the maximization is turned into a minimization. Comparing the resulting expression with (2.13) reveals that the LS formulation

has been obtained. In other words, under uniform priors and conditionally-independent distance estimates which are observed under zero-mean Gaussian noise with identical variance, the MAP solution to the lateration problem is identical to the NLS solution.

2.3.4 Numerical Example

Consider the topology shown in Fig. 2.6 (top). All position and distance-related quantities are given in meters. The coordinates of the anchors are $\mathbf{q}_1^* = [25, 35]^{\mathrm{T}}$, $\mathbf{q}_2^* = [10, 20]^{\mathrm{T}}$, and $\mathbf{q}_3^* = [30, 20]^{\mathrm{T}}$. The agent is located at $\mathbf{p}_1^* = [15, 30]^{\mathrm{T}}$. Hence the true distances between the anchors and agents are $d_{\mathbf{q}_1 \to \mathbf{p}_1} = 11.18$, $d_{\mathbf{q}_2 \to \mathbf{p}_1} = 11.18$, and $d_{\mathbf{q}_3 \to \mathbf{p}_1} = 18.03$. The standard deviation of the noise is $\sigma_d = 1$. The following distance estimates have been observed by the agent: $\mathbf{z} = [\hat{d}_{\mathbf{q}_1 \to \mathbf{p}_1}, \hat{d}_{\mathbf{q}_2 \to \mathbf{p}_1}, \hat{d}_{\mathbf{q}_3 \to \mathbf{p}_1}]^{\mathrm{T}} = [8.92, 11.71, 19.86]^{\mathrm{T}}$.

In the following, the cost function of the NLS problem in (2.13) is examined. The cost associated with the first term $(\hat{d}_{\mathbf{q}_1 \to \mathbf{p}_1} - ||\mathbf{p}_1 - \mathbf{q}_1||)^2$ of the cost function in (2.13) is depicted in dB-scale in Fig. 2.9 (top). This part corresponds to the costs implied by the distance estimate with respect to the first anchor. Observe that the lowest values are obtained on the circle which is centered at $\mathbf{q}_1$ with radius $\hat{d}_{\mathbf{q}_1 \to \mathbf{p}_1}$. Considering only a single distance estimate, all points on this circle have identical cost, i.e. no distinct minimum exists and the position of the agent cannot be deduced unambiguously. Considering all terms of the cost function in (2.13) (see Fig. 2.9 (bottom)), it can be seen that a single global minimum (asterisk in Fig. 2.9 (bottom)) exists. This global minimum is the solution to the NLS problem and can be determined by, e.g., the Levenberg-Marquardt algorithm. Note that the LLS solution ('x' in Fig. 2.9 (bottom)) differs from the NLS solution. In this example, the position estimates of the LLS and NLS problem are given by $\hat{\mathbf{p}}_{1,\mathrm{LLS}} = [13.56, 33.35]^{\mathrm{T}}$ and $\hat{\mathbf{p}}_{1,\mathrm{NLS}} = [15.26, 31.74]^{\mathrm{T}}$, respectively. Therefore, the position errors under the l2-norm are $e_{\mathbf{p}_{1,\mathrm{LLS}}} = \left\|\hat{\mathbf{p}}_{1,\mathrm{LLS}} - \mathbf{p}_1^*\right\| = 3.64$ and $e_{\mathbf{p}_{1,\mathrm{NLS}}} = \left\|\hat{\mathbf{p}}_{1,\mathrm{NLS}} - \mathbf{p}_1^*\right\| = 1.76$. Notice that, in this example, the optimum solution in the maximum a posteriori sense is directly given by $\hat{\mathbf{p}}_{1,\mathrm{MAP}} = \hat{\mathbf{p}}_{1,\mathrm{NLS}} = [15.26, 31.74]^{\mathrm{T}}$ as argued in Subsection 2.3.3. Note that the numerical values of the considered realization are solely provided for reproducibility.

2.4 Angulation

Angulation refers to angle-based position estimation, where the position of an agent is determined based on the AOA estimates with respect to multiple anchors. In contrast to lateration position estimation using angulation generally depends on the orientation of the antenna array of the agent.

2.4.1 Problem Formulations

Fig. 2.8 depicts the angulation-based positioning problem with one agent and $M = 3$ anchors in the absence of noise. The position of the agent and the positions of the anchors are denoted by $\mathbf{p}_1$ and $\mathbf{q}_i, i = 1, ..., M$, respectively. The orientation of the agent is denoted by α. It is assumed that the agent has obtained the AOAs $\theta_{\mathrm{RX},i}, i = 1, ..., M$ with respect to every anchor and these observations are summarized in the vector $\mathbf{z} = [\theta_{\mathrm{RX},1}, ..., \theta_{\mathrm{RX},M}]^{\mathrm{T}}$. As mentioned in Section 2.1, the MAP-optimum solution to positioning problems can be expressed according to (2.1). In the context of angulation, two conditions are differentiated; (i) known and (ii) unknown orientation of the agent.

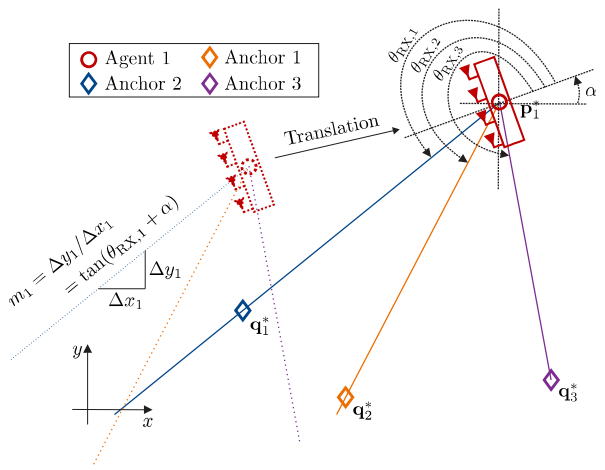

Figure 2.8: *Angulation with known orientation* - the localization problem can be interpreted as finding the translation of agent 1 such that the lines corresponding the AOAs intersect with all anchors $\mathbf{q}_i^*$.

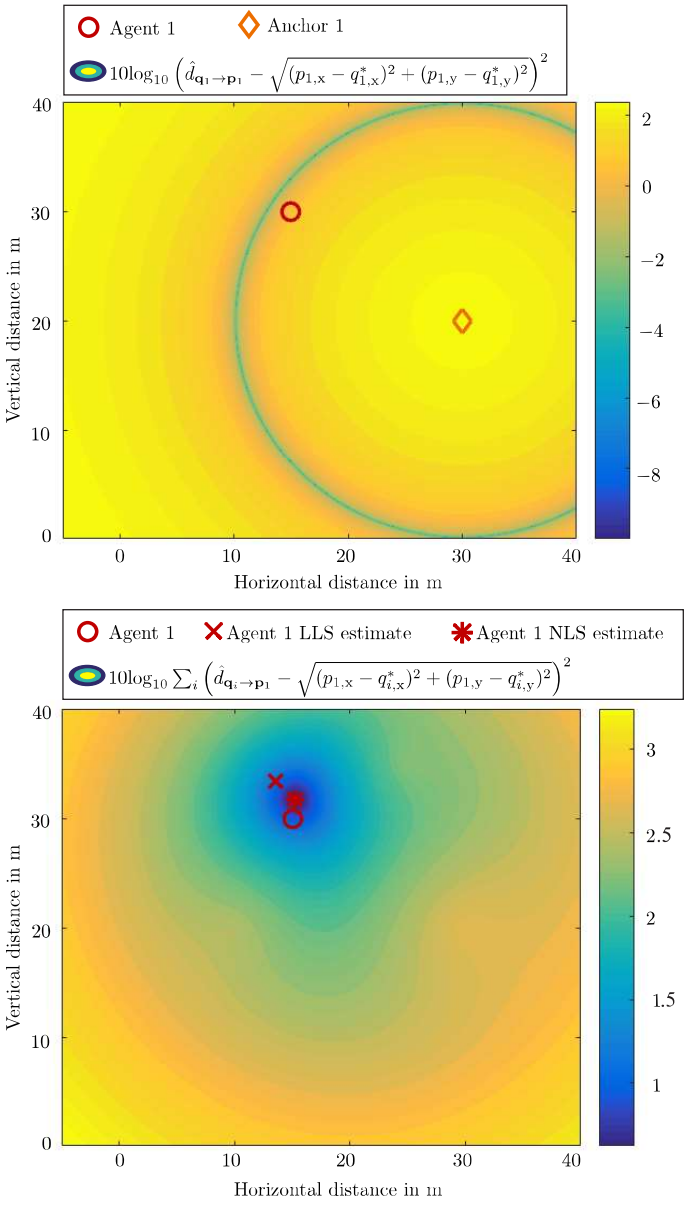

Figure 2.9: *Lateration cost functions* - costs in the LS sense in lateration: The costs related to a single distance estimate (top) show an annular shape and no unambiguous position estimate can be obtained. Considering the sum of the costs (bottom) related to three distance estimates. a distinct global minimum exists which is the solution to the NLS problem.

2.4.1.1 Known Orientation - Intuition

If the orientation of the agent is known, the parameter vector contains only the position of the agent, i.e. $\boldsymbol{\eta} \triangleq \mathbf{p}_1$. The goal is to infer the position of the agent given all observations $\mathbf{z}$. In the absence of noise, this problem can be interpreted as follows. From the perspective of the agent, a line with the angle $\theta_{\mathrm{RX},i} + \alpha$ is drawn for every AOA $\theta_{\mathrm{RX},i}$. Then, the goal is to find the translation of the agent such that all lines intersect with the positions of the respective anchors (see Fig. 2.8). Alternatively, from the perspective of the anchors, a line through each anchor is drawn whose slope is determined by the respective AOA $\theta_{\mathrm{RX},i}$ and orientation α. The intersection point of all lines represents the position of the agent (see Fig. 2.10). This scenario requires $M \geq 2$ anchors to determine an unambiguous position estimate.

2.4.1.2 Unknown Orientation - Intuition

In the case where the orientation is unknown, the estimation problem is more complicated, since the parameters of interest do not only include the position of the agent but also its orientation, i.e. $\boldsymbol{\eta} = [\mathbf{p}_1^{\mathrm{T}}, \alpha]^{\mathrm{T}}$. Since one more parameter has to be estimated compared to the scenario with known orientation, at least three anchors are required to unambiguously estimate the position and orientation of the agent. The angulation problem with unknown orientation can be interpreted in a similar way as in Fig. 2.10. Since the orientation is unknown, the position of an agent is given by the optimum translation *and* rotation. In other words, a line based on each AOA is drawn starting from the agent. Then, the agent is translated and rotated until all lines intersect with the respective position of an anchor.

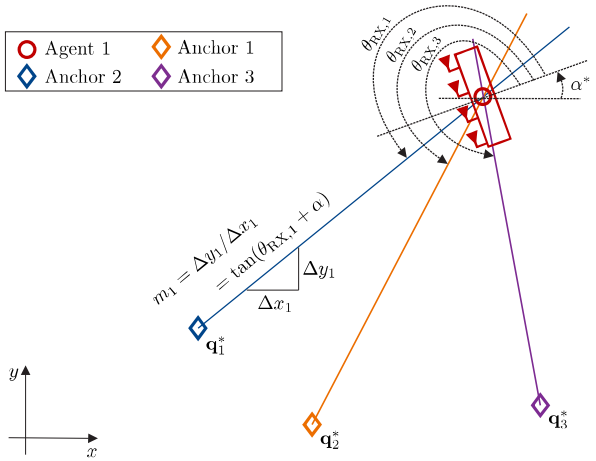

Figure 2.10: *Angulation with known orientation* - if the orientation of the agent is known, each AOA estimate gives a line which connects the respective anchor and the agent. The intersection of at least two lines provides an unambiguous position estimation.

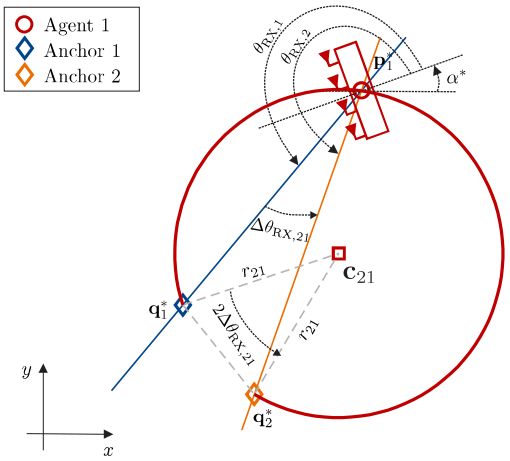

Figure 2.11: *Differential AOA* - when the orientation of the receiver is unknown, it can be eliminated from the estimation problem by considering the difference of any pair of AOA estimates. Such a pair enforces the position of the agent to be on the geometric locus depicted by the red arc.

2.4.2 Angulation in the Absence of Noise

2.4.2.1 Known Orientation - Mathematical Description

When no noise is present, the AOA estimates $\hat{\theta}_{\mathrm{RX},i}, i = 1, ..., M$ become equal to the true AOAs $\theta_{\mathrm{RX},i} = \operatorname{atan2}\left((q^*_{i,\mathrm{y}} - p_{1,\mathrm{y}})/(q^*_{i,\mathrm{x}} - p_{1,\mathrm{x}})\right) - \alpha$. As a result for the i^{th} AOA, a line through $\mathbf{q}_i^*$ with slope $m_i \triangleq \tan\left(\theta_{\mathrm{RX},i} + \alpha\right)$ is obtained (see Fig. 2.10), i.e.

$$p_{1,\mathrm{y}} = m_i(p_{1,\mathrm{x}} - q^*_{i,\mathrm{x}}) + q^*_{i,\mathrm{y}} \qquad \text{for } i = 1, ..., M. \tag{2.18}$$

Similar to the linearized lateration problem in (2.11), the angulation problem with known orientation can be formulated as a linear system of equations

$$\underbrace{\begin{bmatrix} m_1 & -1 \\ \vdots & \vdots \\ m_M & -1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} p_{1,\mathrm{x}} \\ p_{1,\mathrm{y}} \end{bmatrix}}_{\mathbf{p}} = \underbrace{\begin{bmatrix} m_1 q^*_{1,\mathrm{x}} - q^*_{1,\mathrm{y}} \\ \vdots \\ m_1 q^*_{M,\mathrm{x}} - q^*_{M,\mathrm{y}} \end{bmatrix}}_{\mathbf{b}}. \tag{2.19}$$

In the absence of noise. any two rows in (2.19) provide an unambiguous position estimate.

2.4.2.2 Unknown Orientation - Mathematical Description

If the orientation itself is unknown, $M = 3$ anchors are necessary to solve the extended positioning problem, where the orientation has to be determined as well. To circumvent joint position and orientation estimation, the extended positioning problem can be reformulated by considering the differential AOAs, e.g., $\theta_{\mathrm{RX},j} - \theta_{\mathrm{RX},i}, j \neq i$. Instead of considering the

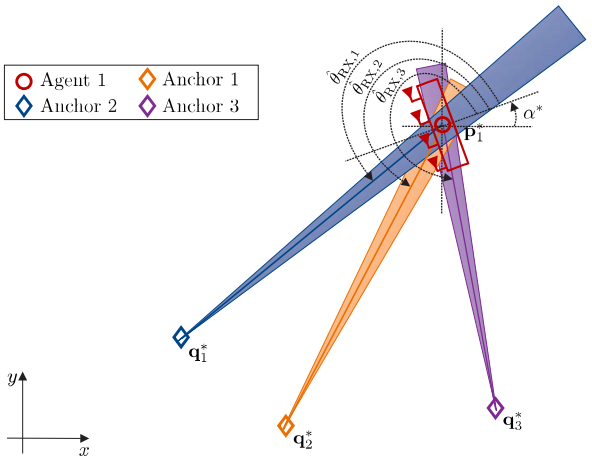

Figure 2.12: *Angulation in the presence of noise* - the lines formed by each AOA transform into cones if the AOA is observed the influence of noise.

absolute AOAs, all AOAs are related to one (arbitrary) AOA, e.g., the i^{th} AOA. In that way, the orientation is eliminated from the positioning problem. In particular,

$$\begin{aligned} \theta_{\mathrm{RX},j} - \theta_{\mathrm{RX},i} &= \operatorname{atan2}\left(\tfrac{q_{j,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{j,\mathrm{x}}^* - p_{1,\mathrm{x}}}\right) - \alpha - \operatorname{atan2}\left(\tfrac{q_{i,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{i,\mathrm{x}}^* - p_{1,\mathrm{x}}}\right) + \alpha, \quad \forall j \neq i \\ \Delta\theta_{\mathrm{RX},ji} &\triangleq \operatorname{atan2}\left(\tfrac{q_{j,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{j,\mathrm{x}}^* - p_{1,\mathrm{x}}}\right) - \operatorname{atan2}\left(\tfrac{q_{i,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{i,\mathrm{x}}^* - p_{1,\mathrm{x}}}\right), \quad \forall j \neq i \end{aligned} \tag{2.20}$$

Note that the system of equations in (2.20) is nonlinear. The relationship described in (2.20) for any pair of anchors j and i represents a geometric locus which circumscribes all positions $\mathbf{p}$ with the angle $\Delta\theta_{\mathrm{RX},ji}$ to the anchors j and i. An example is depicted in Fig. 2.11. In particular, this locus given by is an arc whose center and radius are $\mathbf{c}_{ji}$ and r_{ji}, respectively. Both quantities are computed from the known positions of the anchors and the differential angle $\Delta\theta_{\mathrm{RX},ji}$, using geometric relationships in an isosceles triangle (dashed triangle in Fig. 2.11). The intersection of any two loci gives an unambiguous position estimate. Hence considering any two of the $M-1$ equations formed in (2.20) make it possible to determine the position $\mathbf{p}_1$ of the agent. Recall that two such equations correspond to three AOA estimates. Due to the geometric similarity of the loci described by (2.20) and regular circles, any angulation problem with unknown orientation can be rewritten as a lateration problem [NN03a].

2.4.3 Angulation in the Presence of Noise

In the presence of noise, neither the system of equations in (2.19) nor in (2.20) has a unique solution, since the true AOAs are replaced by noisy AOA estimates, i.e. $\theta_{\mathrm{RX},i} \rightarrow \hat{\theta}_{\mathrm{RX},i}, \forall i$. Instead, an optimality criterion is introduced with respect to which the position of the agent is determined. For consistency, the LS criterion is considered again. Intuitively, the lines formed by the AOA estimates transform into cones as depicted in Fig. 2.12

2.4.3.1 Known Orientation

If the orientation α is known, the optimum solution in the LS sense is given by

$$\hat{\mathbf{p}}_1 = \arg\min_{\mathbf{p}} \sum_{i=1}^{M} \left(\hat{\theta}_{\mathrm{RX},i} - \underbrace{\operatorname{atan2}\left(\frac{q_{i,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{i,\mathrm{x}}^* - p_{1,\mathrm{x}}} \right) - \alpha}_{\theta_{\mathrm{RX},i}} \right)^2 . \tag{2.21}$$

Note that the optimization problem in (2.21) has a highly nonlinear objective function due to the inverse tangent. Generally, no closed-form solution to this problem exists and numerical solvers as mentioned in Subsection 2.3.3 have to be employed.

2.4.3.2 Unknown Orientation

If the orientation of the agent cannot be identified a priori, differential AOAs can be considered as indicated in (2.20). In that case, the LS position estimate is given by

$$\hat{\mathbf{p}}_1 = \arg\min_{\mathbf{p}} \sum_{i=1}^{M} \left(\Delta\hat{\theta}_{\mathrm{RX},ji} - \underbrace{\operatorname{atan2}\left(\frac{q_{j,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{j,\mathrm{x}}^* - p_{1,\mathrm{x}}} \right) + \operatorname{atan2}\left(\frac{q_{i,\mathrm{y}}^* - p_{1,\mathrm{y}}}{q_{i,\mathrm{x}}^* - p_{1,\mathrm{x}}} \right)}_{\Delta\theta_{\mathrm{RX},ji}} \right)^2 . \tag{2.22}$$

Like (2.21), the optimization problem in (2.22) is nonlinear and has no closed-form solution.

2.4.4 Numerical Example

Consider the topology shown in Fig. 2.8 with $M = 3$ anchors. All distance and position-related quantities are given in meters, while all angle-related quantities are given in degrees. The coordinates of the anchors are $\mathbf{q}_1^* = [20, 10]^{\mathrm{T}}$, $\mathbf{q}_2^* = [40, -15]^{\mathrm{T}}$, and $\mathbf{q}_3^* = [80, -10]^{\mathrm{T}}$. The agent is located at $\mathbf{p}^* = [70, 70]^{\mathrm{T}}$ with known orientation $\alpha^* = 20$. Hence the true AOAs between the anchors and agent are $\theta_{\mathrm{RX},1} = -84.80$, $\theta_{\mathrm{RX},2} = -64.44$, and $\theta_{\mathrm{RX},3} = -37.87$. The standard deviation of the noise is $\sigma_{\theta_{\mathrm{RX}}} = 1$. The following AOA estimates have been observed by the agent: $\mathbf{z} = [\hat{\theta}_{\mathrm{RX},1}, \hat{\theta}_{\mathrm{RX},2}, \hat{\theta}_{\mathrm{RX},3}]^{\mathrm{T}} = [-82.80, -63.44, -37.37]^{\mathrm{T}}$.

The cost associated with the third term of the cost function in (2.21) is depicted in dB-scale in Fig. 2.13 (top). As described by the functional relationship in (2.18), this term of the cost function produces the lowest costs along the line from anchor 3 to agent 1 with the angle $\hat{\theta}_{\mathrm{RX},3} + \alpha$. Considering only the AOA to anchor 3, all points on this line have identical cost and the position of the agent cannot be deduced unambiguously. Fig. 2.13 (bottom) depicts the costs associated with all three anchors. Considering all terms in (2.13), the cost function has a single global minimum (see Fig. 2.13 (bottom)) which is given by the LS solution. The obtained position estimate is $\hat{\mathbf{p}}_{1,\mathrm{NLS}} = [69.07, 66.18]^{\mathrm{T}}$ and hence the position error under the l2-norm is $e_{\mathbf{p}_{1,\mathrm{NLS}}} = \left\| \hat{\mathbf{p}}_{1,\mathrm{NLS}} - \mathbf{p}_1^* \right\| = 3.934$. Note that the numerical values are solely provided for reproducibility.

2.5 Summary

The basic principles of radio-based positioning have been revisited in this chapter, emphasizing the difference between direct and indirect positioning. Aspects of indirect positioning have been elaborated, since the research presented in later chapters focuses on this field. Indirect positioning refers to a two-step notion where position-related parameters like angles-of-arrival and distances are derived from the received signal in the first step. In the second step, the position of an agent is determined based on these parameters. To that end, a time-of-arrival and an angle-of-arrival estimator were discussed. Subsequently, the paradigms of lateration (distance-based positioning) and angulation (angle-based positioning) have been recapitulated. In particular, the general problem was revisited and optimization problems to obtain least-squares position estimates were formulated. Numerical examples have been used for illustration. The essential point is that in lateration and angulation an agent needs to fuse the information from multiple anchors in order to determine its own position unambiguously. These requirements may not always be met in practice, since deploying and maintaining a large number of anchors is costly. Hence lateration and angulation-based positioning fails in scenarios with sparse infrastructure. Therefore, alternative positioning solutions will be presented in the Chapters 5 - 7 which provide accurate position estimates even in the presence of very few anchors.

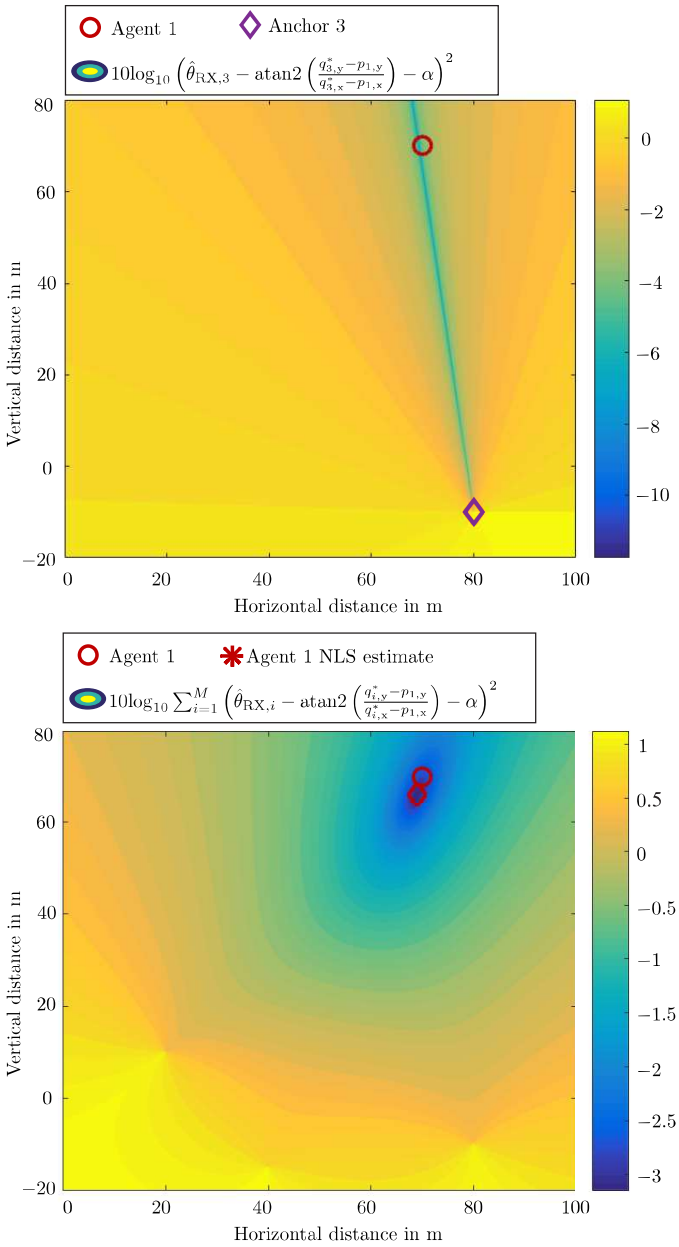

Figure 2.13: *Angulation cost functions* - cost in the LS sense in angulation with known orientation: The costs related to a single AOA estimate (top) are given by a conic shape. In that case, no unambiguous position estimate can be obtained. If three AOA estimates are considered the corresponding cost function (bottom) has a single global minimum, allowing for unambiguous position estimation.

Chapter 3

Theoretical Background on Statistical Signal Processing

The field of statistical signal processing has gained considerable attention in the past few decades. A number of different techniques, including the well-known *belief propagation* algorithm fall under the umbrella of statistical signal processing. Statistical signal processing takes a fundamentally different approach compared to many conventional, deterministic signal processing techniques. Unlike deterministic approaches, statistical signal processing does not try to optimize metrics such as the Euclidean distance, but rather probabilistic measures such as the likelihood function or posterior distribution are optimized. In that way, the uncertainties of the parameters to be estimated are inherently considered. Hence statistical signal processing has become extremely popular in many inference tasks, since it provides tractable, high-quality solutions for a variety of problems.

Irrespective of the considered application, optimum solutions to inference problems (like MAP) are desirable. However, if inference problems are high-dimensional, the joint posterior distributions are often hard to optimize. This is especially true if the posterior distribution is nonlinear and non-Gaussian. To circumvent estimation in high dimensional spaces, algorithms like belief propagation determine the lower dimensional marginal posterior distributions in a computationally efficient manner. Estimates are then derived from the low-dimensional marginals.

This chapter is devoted to recapitalizing a number of basic tools required for statistical signal processing. First, different representations of probability densities will be reviewed. Subsequently, two essential operations will be discussed: *sampling* from densities and *multiplication* of densities . Similar to the previous chapter, the goal of this chapter is not to provide a complete description of all aspects of statistical signal processing, but rather present a concise overview of the essential aspects that will be required in the upcoming chapters. A simplistic running example will be used throughout this chapter to illustrate the concepts to be discussed.

Running Example

Consider a simple one-dimensional (1D) positioning system which is depicted in Fig. 3.1. It consists of one agent and two anchors. The agent set is denoted by $\mathcal{A} = \{1\}$, while $\mathcal{B} = \{A, B\}$ is the set of all anchors. For notational simplicity, the position of a node is denoted by[a] $\eta_i, i \in \mathcal{A} \cup \mathcal{B}$. The true position of a node is distinguished by a superscript asterisk, e.g., η_1^* (agent 1) or η_A^* (anchor A). It is assumed that the agent can estimate the distance $\hat{d}_{i\to 1}, i \in \mathcal{B}$ with respect to both anchors. The noisy distance estimates are given by

$$\hat{d}_{i\to 1} = \underbrace{|\eta_i^* - \eta_1^*|}_{d_{i\to 1}} + e_{i\to 1}, \quad i \in \mathcal{B}$$

where $d_{i\to 1}$ and $e_{i\to 1}$ are the true distance and the distance estimation error between the agent and the i^{th} anchor, respectively. It is assumed that the distributions of both ranging errors $e_{i\to 1}, i \in \mathcal{B}$ are independent. Both distance estimates are collected in the vector $\mathbf{z} = [\hat{d}_{A\to 1}, \hat{d}_{B\to 1}]^{\mathrm{T}}$. Then, the joint likelihood function can be written as

$$f(\mathbf{z}|\eta_1, \eta_A^*, \eta_B^*) = f(\hat{d}_{A\to 1}|\eta_1, \eta_A) f(\hat{d}_{B\to 1}|\eta_1, \eta_B). \tag{3.1}$$

Consequently, the joint posterior distribution is

$$f(\eta_1, \eta_A^*, \eta_B^*|\mathbf{z}) \propto \underbrace{f(\hat{d}_{A\to 1}|\eta_1, \eta_A^*)}_{\triangleq f_1(\eta_1)} \underbrace{f(\hat{d}_{B\to 1}|\eta_1, \eta_B^*)}_{\triangleq f_2(\eta_1)} f_3(\eta_1). \tag{3.2}$$

where '$\propto$' means proportional up to a normalization constant and the perfect prior knowledge $\delta(\eta_i - \eta_i^*), i \in \mathcal{B}$ of anchors was absorbed in the likelihood functions $f_1(\eta_1)$ and $f_2(\eta_1)$. It should be stressed that $f_1(\eta_1)$ and $f_2(\eta_1)$ are shorthand notations for conditional distributions with varying parameter η_1. A uniform prior in the interval $[2, 7]$ is considered, i.e. $f_3(\eta_1) = \mathcal{U}(2, 7)$. The distance estimation errors $e_{A\to 1}$ and $e_{B\to 1}$ are assumed to be Gaussian and one-sided exponential, respectively. Hence the likelihood functions $f(\hat{d}_{A\to 1}|\eta_1, \eta_A^*)$ and $f(\hat{d}_{B\to 1}|\eta_1, \eta_B^*)$ are Gaussian and one-sided exponential, respectively.

Since 1D positioning is considered, the distance to an anchor suggests two possible locations: Left or right of the anchor. For instance, $\eta_1 = \eta_A^* + d_{A\to 1}$ or $\eta_1 = \eta_A^* - d_{A\to 1}$. For simplicity of explanation, it is assumed that the agent knows whether an anchor is located right or left of it.

The goal of the agent is to determine an optimum estimate in the minimum mean square error (MMSE) sense, i.e.

$$\hat{\eta}_{1,\mathrm{MMSE}} = \int \eta_1 f(\eta_1, \eta_A^*, \eta_B^*|\mathbf{z}) \mathrm{d}\eta_1, \tag{3.3}$$

where the integral is taken over the entire sample space (support) of η_1.

[a] Note that, in this thesis, the position of a node is generally denoted by $p_{i,\mathrm{x}}$ and $\mathbf{p}_i = [p_{i,\mathrm{x}}, p_{i,\mathrm{y}}]^{\mathrm{T}}$ in 1D and two-dimensional (2D), respectively. For simplicity of the notation, in the Chapters 3 and 4, η_i represents the position of the node i.

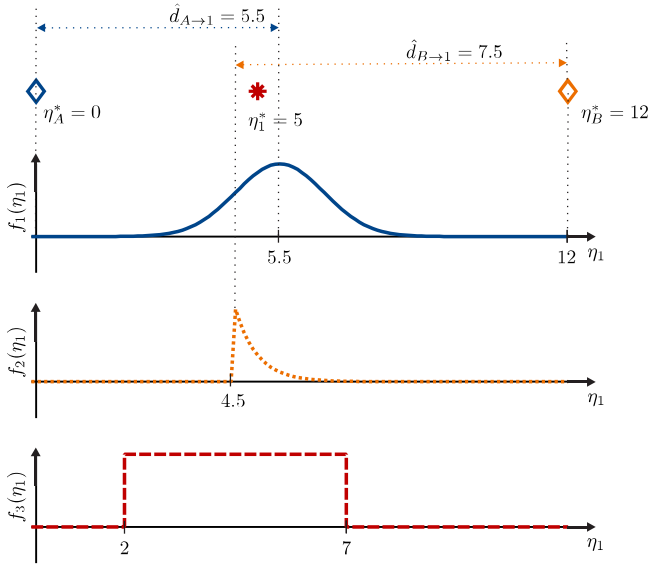

Figure 3.1: *Running example* - An agent attempts to determine an MMSE estimate of its position in 1D space based on the observed ranges $\mathbf{z} = [5.5, 7.5]^{\mathrm{T}}$. The agent (red asterisk) is located at $\eta_1^* = 5$ and the anchors A (blue diamond) and B (orange diamond) reside at $\eta_A^* = 0$ and $\eta_B^* = 12$, respectively. It is assumed that the ranging error of $\hat{d}_{A\to1}$ is zero-mean Gaussian distributed with standard deviation $\sigma = 1$, while the ranging error of $\hat{d}_{B\to1}$ is one-sided exponentially distributed with mean $m = 0.5$. The topology, densities $f_1(\eta_1)$ and $f_2(\eta_1)$, and the uniform prior distribution $f_3(\eta_1)$ are depicted.

3.1 Representations of Probability Distributions and Densities

Probability distributions govern the stochastic properties of random variables. Random variables can be broadly classified into discrete and continuous. In the former case, random variables have a finite sample space, i.e. the set of possible outcomes of a random experiment is limited. An obvious example is a regular dice whose sample space is given by the set $\Omega = \{1, 2, 3, 4, 5, 6\}$. If the dice is fair, its probability distribution[1] assigns the same probability ($\Pr(1) = \Pr(2) = \ldots = \Pr(6) = 1/6$) to each possible outcome. In contrast to discrete random variables, the sample space of continuous random variables is infinite, e.g, the set of real numbers, i.e. $\Omega = \mathbb{R}$. Examples of continuous random variables are temperatures, distances, positions, etc. Since an infinite number of possible outcomes exists, the probability of a particular outcome must be zero to assure that the probability of all outcomes adds up to one. As a consequence, continuous random variables are not

[1]The term probability distribution is used for both continuous and discrete random variables. In contrast, the expression probability mass function (PMF) refers exclusively to discrete random variables, while the word *density* implies that a continuous random variable is considered.

characterized by probability mass functions, but rather by continuous probability densities. Such densities can be represented in different ways. The following representations of continuous densities are discussed in the following:

- Discrete representations
  1. Discretized representation
  2. Sample-based representation
  3. Particle-based representation
- Parametrized representations

It should be noted that discretized, sample-based, and particle-based representations are not explicitly classified as discrete representations in the literature. However, due to the similarity of these methods, such a classification appears to be meaningful and is adopted in this thesis.

3.1.1 Discrete Representations of Probability Distribution

Discrete representations capture the essence of continuous distributions by using finite sets of points. In some cases, these points are also weighted. The goal of these points and weights is to approximate all statistical moments (such as mean and variance) of a distribution. The key difference between discretized, sample-based, and particle-based representations is the way how these points and possibly weights are chosen. The peculiarities of these representations will be explained in the following subsections.

3.1.1.1 Discretized Representation

A naive, but simple approach to represent a continuous distribution $f(\eta)$ is to deterministically discretize the entire sample space using N_s points. For illustration, the running example is considered. In particular, the Gaussian density $f_1(\eta_1)$ is discretized in the interval from [0,12] using $N_s = 10$ and $N_s = 50$ samples, respectively. The discretized and continuous probability density function (PDF) are depicted in Fig. 3.2.

Discretized representations consist of two sets:

1. a set of quantization points $\mathcal{H} = \{\eta^{(1)}, \ldots, \eta^{(N_s)}\}$ and
2. a set of weights $\mathcal{F} = \{f^{(1)}, \ldots, f^{(N_s)}\}$.

The distance $\Delta\eta$ between two adjacent quantization points is given by $\Delta\eta = |\eta^{(k-1)} - \eta^{(k)}|$ and the k^{th} quantization point is given by $\eta^{(k)} = k\Delta\eta$. The unnormalized weight $\tilde{f}^{(k)}$ of the quantization point $\eta^{(k)}$ is given as the functional value of $f(\eta^{(k)})$, i.e.

$$\tilde{f}^{(k)} = f(\eta^{(k)}). \tag{3.4}$$

For mathematical convenience, the weights are often normalized, i.e.

$$f^{(k)} = \frac{\tilde{f}^{(k)}}{\sum_{j=1}^{N_s} \tilde{f}^{(j)}} \tag{3.5}$$

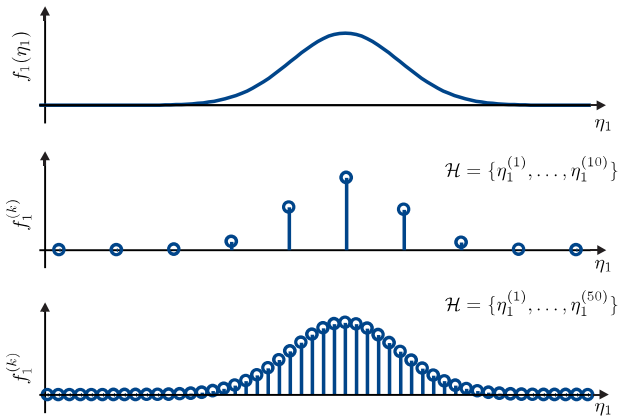

Figure 3.2: *Discretized representations of probability densities* - a discretized representation of the continuous density $f_1(\eta_1)$ is obtained by determining N_s equally-spaced quantization points $\eta_1^{(k)}, \forall k$ in the interval of interest and by determining the weight $f_1^{(k)}$ of each quantization point $\eta_1^{(k)}$.

The distribution is then approximated as a finite set of weights $\mathcal{F}$ and quantization points $\mathcal{H}$.

Note that employing this discretized representation requires that the continuous density $f(\eta)$ can be evaluated for all η in the considered interval. From the example in Fig. 3.2, it can be seen that the essential shape of the density can be recovered. However, the shortcomings of this approach can also be noticed: Many quantization points reside in regions of the sample space where the probability distribution has virtually no mass. In other words, quantization points are not efficiently utilized. For instance, whenever an estimate $\hat{\eta}$ is obtained from the distribution $f(\eta)$ (e.g., mean value). quantization points with relatively small weight contribute only little. In case of the mean value, this can be directly observed:

$$\hat{m}_\eta = \int \eta f(\eta) \mathrm{d}\eta \approx \sum_{k=1}^{N_s} f^{(k)} \eta^{(k)}. \tag{3.6}$$

It can be seen that $f^{(k)}$ directly weights the k^{th} quantization point $\eta^{(k)}$. If $f^{(k)}$ is extremely small, the corresponding quantization point does not change the estimation result notably and hence carries negligible information from the estimation perspective. This observation gives rise to a more sophisticated way of representing a continuous density using a finite number of points.

3.1.1.2 Sample-based Representation

Sample-based representations embody an alternative method of representing a continuous density in a discrete way. In contrast to the discretized representation, where the quantization points are chosen in a deterministic and uniform fashion, sample-based representations

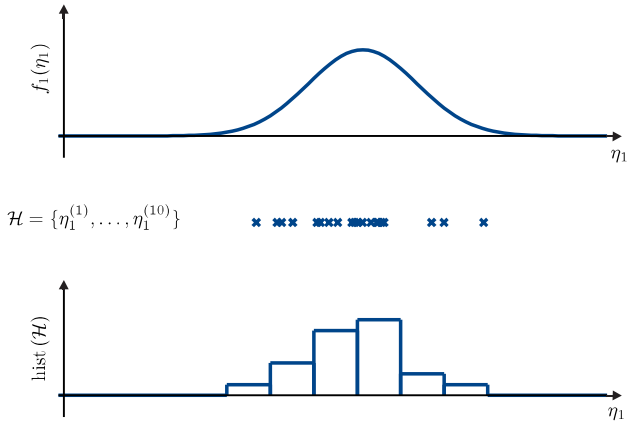

Figure 3.3: *Sample-based representation of probability densities* - a sample-based representation of $f_1(\eta_1)$ can be obtained by directly drawing samples $\eta_1^{(k)}$ from $f_1(\eta_1)$ if direct sampling is tractable. The histogram of all samples $\mathcal{H}$ suggests that $f_1(\eta_1)$ is well represented. since its shape resembles the density $f_1(\eta_1)$.

draw the quantization points randomly. Hence quantization points $\eta^{(k)}$ are referred to as samples and the actual set of samples $\mathcal{H} = \{\eta^{(1)}, \ldots, \eta^{(N_s)}\}$ differs from one realization to another. More precisely, the characteristics of the density $f(\eta)$ are represented by a set of samples which are directly drawn from this distribution. i.e. $\eta^{(k)} \sim f(\eta)$, where '$\sim$' is equivalent to: *drawn from*. Therefore, the inherent advantage of this approach over discretization is that samples reside in areas where the continuous probability density has significant probability mass. Compared to the previous example in (3.6) where the mean of η was estimated, each sample factors in with a weight of $1/N_s$ rather than $f^{(k)}$:

$$\hat{m}_\eta = \int \eta f(\eta) \mathrm{d}\eta \approx \sum_{k=1}^{N_s} \frac{1}{N_s} \eta^{(k)}. \tag{3.7}$$

To emphasize the differences of the discretized and sample-based representations, the example from Fig. 3.2 is revisited. A sample-based representation $\mathcal{H}$ of $f_1(\eta_1)$ with $N_s = 20$ samples, the corresponding histogram of the samples, and the continuous density are shown in Fig. 3.3. It can be observed that most samples reside in regions where the continuous density has significant probability mass (close to the mode of $f_1(\eta_1)$).

Even though, the sample-based representation has inherent advantages over the discretized representation, it is worth mentioning that this technique also has a major restriction. In particular, it is assumed that one can directly sample from the distribution $f(\eta)$ in a tractable manner. While this requirement is met for a small subset of distributions (e.g., Gaussian, exponential, uniform, etc.), direct sampling from arbitrary distributions is generally difficult. To bypass this drawback, the notion of importance sampling (as will be explained in Section 3.2.1.1) is often considered. Importance sampling is a key tool for

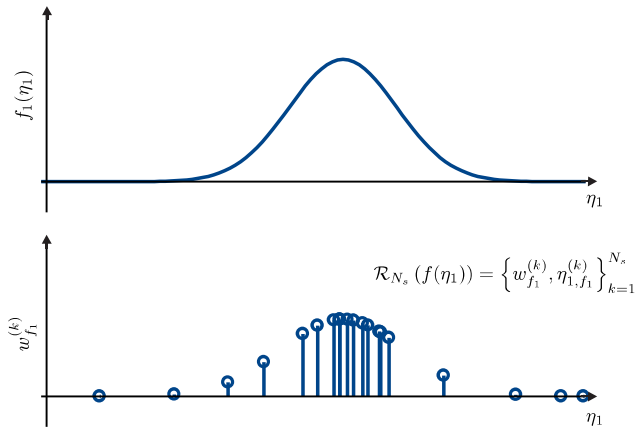

Figure 3.4: *Particle-based representations of probability densities* - a particle-based representation $\mathcal{R}_{N_s}(f_1(\eta_1))$ of $f_1(\eta_1)$ contains a set of samples $\{\eta_{1,f_1}^{(k)}\}_{k=1}^{N_s}$ with associated weights $\{w_{f_1}^{(k)}\}_{k=1}^{N_s}$.

obtaining particle-based representations of densities. This type of representation will be discussed in the following.

3.1.1.3 Particle-based Representation

A set of weighted samples is referred to as *particle representation* and given by

$$\mathcal{R}_{N_s}(f(\eta)) = \left\{w_f^{(k)}, \eta_f^{(k)}\right\}_{k=1}^{N_s}, \tag{3.8}$$

where $f(\eta)$ is the target distribution, N_s is the number of particles, $\eta_f^{(k)}$ is a random sample with associated weight[2] $w_f^{(k)}$. The combination of a sample and its weight is referred to as *particle*. In contrast to the discretized approach, the samples $\eta_f^{(k)}$ are not deterministic but random. If the weights are normalized, a particle $\{w_f^{(k)}, \eta_f^{(k)}\}$ can be interpreted as an hypothesis, i.e. $\eta_f^{(k)}$ occurs with probability $w_f^{(k)}$. An exemplary particle representation in the context of the running example is depicted in Fig. 3.4. For comparability, the same example as in the previous two subsections is revisited. It can be observed that the shape of the continuous distribution is preserved, while, in contrast to the discretized approach, the samples are not equally-spaced. Particle representations will play an essential role throughout many problems that are considered in this thesis. How they are obtained will be described in Section 3.2.1.1. It is worth mentioning that sample-based representations are special cases of particle-based representations, in which $w_f^{(k)} = 1/N_s, \forall k$.

[2]Section 3.2.1.1 will explain how the samples $\eta_f^{(k)}$ are drawn and how the weights $w_f^{(k)}$ are determined.

3.1.2 Parametrized Representation

Densities from the class of parametric distributions $f(\eta;\boldsymbol{\theta})$ can be fully described by a finite number of parameters $\boldsymbol{\theta}$. The most prominent member from the parametric family of distributions is probably the Gaussian distribution whose parameters are its mean and variance. In particular, a univariate Gaussian random variable η is completely described by its probability density function

$$f(\eta;m,\sigma^2)=\underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\eta-m)^2}{2\sigma^2}\right)}_{\triangleq\mathcal{N}(\eta;m,\sigma)}.$$

where m and σ^2 are the mean and variance, respectively. For illustration, consider the distribution $f_1(\eta_1)$ from the running example in Fig. 3.1. Parametrized representations of densities are by nature very compact, since the number of parameters that is required to fully describe them is generally small. In case of the univarite Gaussian distribution, two parameters are sufficient. Note, however, that it is not possible to represent arbitrary probability densities in a closed form using a finite set of parameters. As a matter of fact, most densities which occur in statistical signal processing cannot be described by parametrized distributions without errors. Hence alternative representations which capture the stochastic behavior of the corresponding random variables are needed.

3.1.3 Comparison of Representations

Different representations of continuous densities were considered. Their characteristics are summarized in the following table [LFS+12].

Table 3.1: Characteristics of different representation approaches

Representation	Size of Representation	Arbitrary Distributions
Parametric	small	no
Discretized	large	yes
Sample-based	medium	no
Particle-based	medium	yes

As discussed in Section 3.1.2, parametrized representations typically use only a small number of parameters in order to describe a density and hence the size of these representations is small. On the other hand, arbitrary distributions cannot be represented by parametrized representations without approximation error. Discretized representations are alternative ways of representing continuous densities. In contrast to parametrized representations, they use sets of points to capture the characteristics of the continuous density. Generally, their size is large when compared to parametrized representations. However, discretized and particle-based representations are capable of representing arbitrary distributions without error as the number of points approaches infinity.

3.2 Frequent Operations in Statistical Signal Processing

A number of different techniques and tools are frequently required in statistical signal processing. From the perspective of this thesis, the most relevant operations are sampling from arbitrary densities and multiplication of arbitrary densities. The former task requires methods like importance sampling and resampling, while the latter is related to operations like kernel density estimation.

3.2.1 Sampling

As seen in Section 3.1.1.3, it is often necessary to draw samples from an arbitrary probability density in order to obtain a sample or particle-based representation of a continuous density. This demands for a rigorous sampling method that can be applied to any distribution.

3.2.1.1 Importance Sampling

Importance sampling is a powerful sampling technique which allows one to determine a particle-based representation of any probability distribution $f(\eta)$ [DGA00]. As mentioned in the previous section, often it is intractable to directly obtain a set of samples from an arbitrary probability distribution $f(\eta)$. From a high level perspective, importance sampling provides a means to obtain a set of samples from a target distribution $f(\eta)$ that cannot be sampled from directly. Instead, it obtains samples from a *proposal distribution* $q(\eta)$ and attaches a weight to each sample. The weight accounts for the mismatch between the target distribution $f(\eta)$ and the proposal distribution $q(\eta)$. An example is depicted in Fig. 3.5. Importance sampling consists of two steps:

1. Draw N_s samples $\eta_f^{(k)}$ from the proposal distribution $q(\eta)$.

2. Determine the associated sample weights $w_f^{(k)}$.

As mentioned in Section 3.1.1.3, a weighted sample is referred to as *particle*. The weight of every particle can be seen as a correcting factor which accounts for the fact that the sample $\eta_f^{(k)}$ is not drawn from $f(\eta)$, but rather from $q(\eta)$. For clarity, an illustration using the running example will be considered in the following. Subsequently, a mathematical description of importance sampling will be provided.

Running example: Suppose that one wants to obtain a particle-based representation of the density $\tilde{f}(\eta_1) \triangleq f_1(\eta_1)f_2(\eta_1)$. For simplicity of sampling, the proposal distribution $q(\eta_1)$ is Gaussian with mean $m_q = 5$ and variance $\sigma_q^2 = 1$. Fig. 3.5 depicts the density $\tilde{f}(\eta_1)$ and proposal distribution $q(\eta_1)$ in the upper part of the figure. The $N_s = 20$ samples $\{\eta_{1,\tilde{f}}^{(k)}\}_{k=1}^{N_s}$ and particles $\{w_{\tilde{f}}^{(k)}, \eta_{1,\tilde{f}}^{(k)}\}_{k=1}^{N_s}$ are depicted in the lower part of the figure.

In a general sampling problem, let η be a continuous random variable and the continuous density $f(\eta)$ should be approximated by a set of N_s samples. Assuming that $f(\eta)$ cannot be sampled directly, the goal is to obtain a particle representation $\mathcal{R}_{N_s}(f(\eta))$ of $f(\eta)$ using importance sampling. To obtain a particle representation of $f(\eta)$, first, N_s

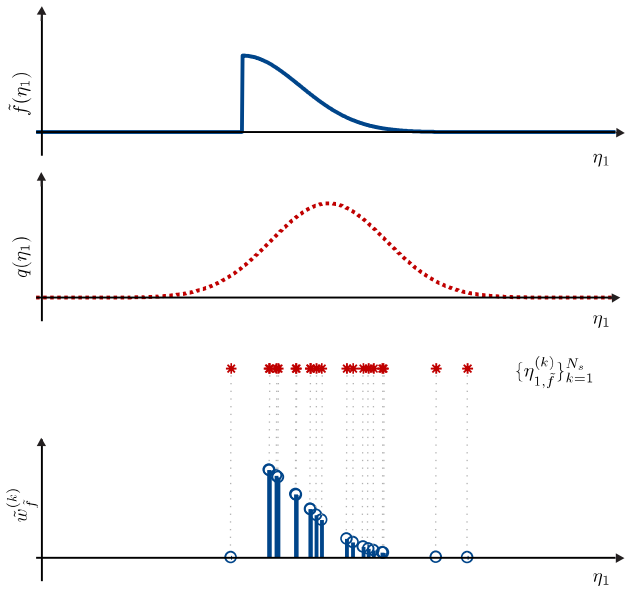

Figure 3.5: *Importance sampling* - the continuous density $\tilde{f}(\eta_1) = f_1(\eta_1)f_2(\eta_1)$ (solid blue line) is approximated by a set of particles $\mathcal{R}_{N_s}(\tilde{f})\{w_{\tilde{f}}^{(k)}, \eta_{1,\tilde{f}}^{(k)}\}_{k=1}^{N_s}$ (blue Dirac impulses). In order to obtain the set of particles, a proposal distribution $q(\eta_1)$ (red dotted line) is employed to obtain a set of samples $\{\eta_{1,\tilde{f}}^{(k)}\}_{k=1}^{N_s}$ (red asterisks). The unnormalized weight $\tilde{w}_{\tilde{f}}^{(k)}$ of each sample is determined using (3.9).

independent identically distributed (i.i.d.) samples $\{\eta_f^{(k)}\}_{k=1}^{N_s}$ are drawn from a suitable proposal distribution $q(\eta)$. i.e.

$$\eta_f^{(k)} \sim q(\eta).$$

Regarding the choice of the proposal distribution, there are almost no constraints. In particular, the proposal distribution can have arbitrary shape as long as sampling is tractable and its support dominates the support of $f(\eta)$. i.e. $f(\eta) > 0 \Rightarrow q(\eta) > 0, \forall \eta$. To account for the mismatch between $f(\eta)$ and $q(\eta)$, a weight $w_f^{(k)}$ is associated with every sample $\eta_f^{(k)}$. The unnormalized weight is computed as follows:

$$\tilde{w}_f^{(k)} = f(\eta^{(k)})/q(\eta^{(k)}). \tag{3.9}$$

For mathematical convenience, all weights are normalized such that they add up to one, i.e.

$$w_f^{(k)} = \frac{\tilde{w}_f^{(k)}}{\sum_{j=1}^{N_s} \tilde{w}_f^{(j)}}. \tag{3.10}$$

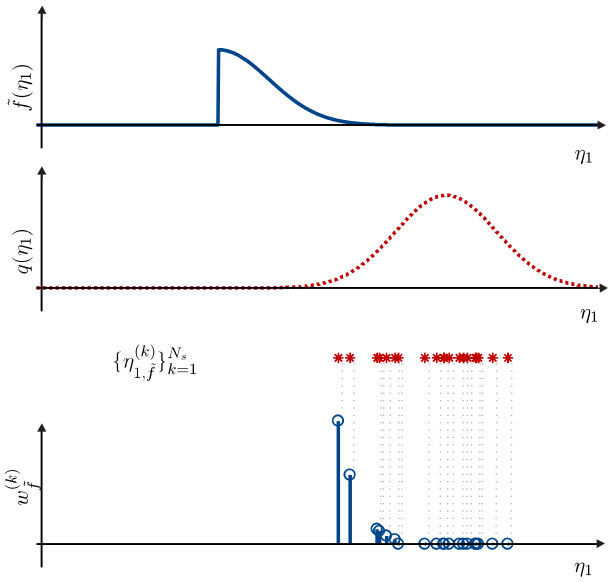

Figure 3.6: *Importance sampling (mismatched proposal)* - if the proposal distribution $q(\eta_1)$ (dashed red line) has considerable probability mass in regions where the target distribution $\tilde{f} = f_1(\eta_1)f_2(\eta_1)$ (solid blue line) has negligible mass, a weight $w_{\tilde{f}}^{(k)}$ of zero or close to zero will be assigned to many samples (red asterisks). In that case, for a finite number of samples N_s, the particle representation $\{w_{\tilde{f}}^{(k)}, \eta_{1,\tilde{f}}^{(k)}\}_{k=1}^{N_s}$ (blue Dirac impulses) does not represent the target distribution well.

As can be seen in the previous example in Fig. 3.5, importance sampling yields a set of particles which closely represents the target distribution $f(\eta)$ in a discrete fashion. As mentioned earlier, the only constraint regarding the proposal distribution $q(\eta)$ is that its support dominates the support of the target distribution $f(\eta)$. If this condition is fulfilled, a particle representation obtained by importance sampling describes the target distribution $f(\eta)$ arbitrarily close as $N_s \to \infty$ [KTB11].

However, for a finite number of particles, there are some practical implications that are related to the choice of the proposal distribution. Ideally, the samples $\eta_f^{(k)}$ generated by the proposal distribution $q(\eta)$ reside in regions of the sample space, where the target distribution $f(\eta)$ has significant probability mass. This proposition has the following intuition. If a sample $\eta_f^{(k)}$ was drawn from a region where $f(\eta)$ has negligible mass, $f(\eta_f^{(k)}) \approx 0$ and hence $w_f^{(k)} \approx 0$. In other words, from the estimation perspective, the particle does not carry information, since its weight is almost zero. If the number of samples is finite, samples should be used as effectively as possible, meaning that the number of samples with weights close to zero should be small. This idea can be translated into the following property of the proposal distribution. The proposal distribution should resemble the target

distribution closely, i.e. the proposal distribution $q(\eta)$ should have significant probability mass in regions where the target distribution $f(\eta)$ has significant probability mass. If the proposal distribution $q(\eta)$ shows significant probability mass in regions where $f(\eta)$ has virtually no mass, a relatively large number of particles with weights close to zero will be generated. In that case, a large number of samples is required to accurately approximate $f(\eta)$ by a set of particles. Having a large number of particles is generally undesirable, since the computational complexity of inference algorithms increases as the number of particles goes up. An example of a poorly chosen proposal distribution is depicted in Fig. 3.6. In that figure, the example from 3.5 is revisited with a different proposal distribution $q(\eta_1)$. In particular, the proposal distribution is Gaussian with mean $m_q = 8$ and variance $\sigma_q^2 = 1$. Observe that compared to the proposal distribution used in Fig. 3.5 only the mean value has changed, but the effect on the approximation accuracy is detrimental. It can be seen from Fig. 3.6 that the majority of samples has a weight of (almost) zero, making them uninformative from the estimation perspective. More importantly, the region with the largest probability mass (close to the mode of $\tilde{f}(\eta_1)$) is not covered at all.

From this example, one can deduce that the proposal distribution $q(\eta)$ should resemble the target distribution $f(\eta)$ closely to ensure that the samples $\eta^{(k)}$ reside in areas where $f(\eta)$ has significant probability mass. A priori, it is not known where the target distribution $f(\eta)$ has significant probability mass and hence it is difficult to choose a suitable proposal distribution $q(\eta)$. As a major part of the contribution of this thesis, a novel method will be introduced in Section 5.4 which confines the relevant regions where $f(\eta)$ has significant probability mass. Based on these confined regions, advanced proposal distributions are derived. A rigorous description of the method will be provided in section 5.4.

3.2.1.2 Resampling

In the previous subsection, it was discussed how one can obtain a particle representation of a probability density. From the examples in the Figs. 3.5 and 3.6, it became evident that particles can have greatly different weights and often some particles have weights that are close to zero. In fact, in many applications, including positioning, weights can differ by multiple orders of magnitude. Recall that samples can be seen as hypotheses and their weights as the corresponding probabilities. From that perspective, one might be tempted to retain only the subset of the most relevant hypothesis and discard extremely unlikely hypotheses. One incentive could be to reduce the overall number of particles in order to reduce the computational complexity. Alternatively, for a constant number of particles, one may wish to focus the computational resources only on particles with significant weights. Both of these goals give rise to a selection strategy of particles which chooses particles according to their weights. Different approaches can be pursued. The simplest and most obvious method would be to introduce a likelihood threshold and simply discard all particles below the threshold and renormalize the weights of the retained particles. This admittedly simple scheme has two shortcomings: (i) It is generally difficult to set the threshold and (ii) the number of remaining particles is difficult to control.

Resampling is a more rigorous selection scheme which determines the set of retained samples in a stochastic manner. In particular, a particle representation $\mathcal{R}_{N_s}(f(\eta)) = \{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$ is treated as a discrete distribution, where $\Pr(\eta_f^{(k)}) = w_f^{(k)}$. From this distribution, N_r samples $\eta_f^{(k)^*}$ are drawn resulting in a new particle representation with equal weights, i.e. $\{1/N_r, \eta_f^{(k)^*}\}_{k=1}^{N_r}$. Intuitively, on average, low-weighted particles are

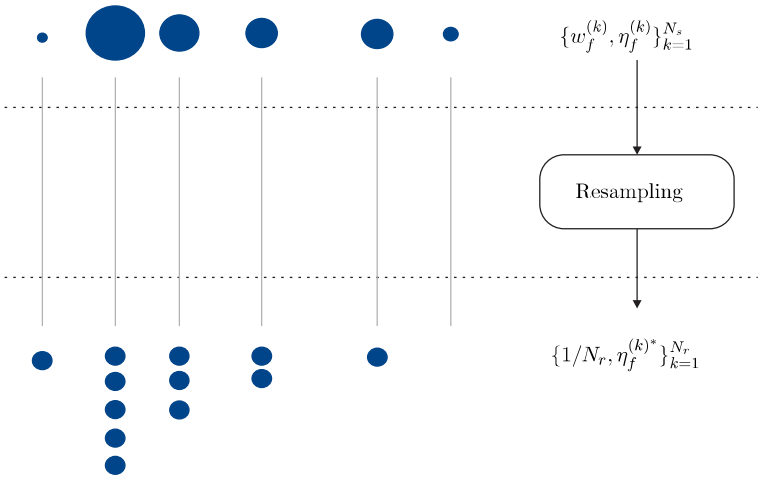

Figure 3.7: *Resampling* - a particle representation $\{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$ with $N_s = 6$ samples is considered. Samples $\eta_f^{(k)}$ are illustrated by disks whose size depends on their weights (high weight equals large disk). The particle representation $\{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$ is resampled resulting in a new particle representation with $N_r = 12$ samples whose weight is identical, i.e. $1/N_r$. Particles with large weight $w_f^{(k)}$ are reproduced multiple times while one particle with a low weight (rightmost particle) was erased.

discarded and multiple copies of high-weighted particles are generated in a stochastic sense. From an evolutionary angle, resampling can be interpreted as a version of the *survival of the fittest*, where high weights imply high fitness values. For example, consider a sample $\eta_f^{(l)}$ that has an associated weight of $w_f^{(l)} = 10^{-8}$. To obtain on average one copy of this sample, $N_r = 10^8$ random samples need to be drawn. In contrast, consider a sample $\eta_f^{(j)}, j \neq l$ that has an associated weight of $w_f^{(j)} = 0.1$. Drawing $N_r = 10^8$ random samples from $\{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$, would result in (on average) 10^7 copies of the sample $\eta_f^{(j)}$. For clarification, a simple example is depicted in Fig. 3.7.

In order to pursue this selection strategy, one has to be able to draw samples from an arbitrary discrete distribution. For clarity, the following example is considered. A particle representation $\mathcal{R}_{N_s}(f(\eta))$ of the distribution $f(\eta)$ has been obtained and the samples and weights are given by $\{\eta_f^{(1)}, \ldots, \eta_f^{(4)}\} = \{0.9, 1.5, 2.1, 2.8\}$ and $\{w_f^{(1)}, \ldots, w_f^{(4)}\} = \{0.1, 0.5, 0.3, 0.1\}$. If the particle representation is interpreted as a probability mass function, the numerical values of the cumulative distribution function (CDF) are given as $F(1) = 0.1$, $F(2) = 0.1 + 0.5 = 0.6$, $F(3) = 0.6 + 0.3 = 0.9$, and $F(4) = 0.9 + 0.1 = 1$. Note that the argument of the CDF is the sample index. In order to resample $\mathcal{R}_{N_s}(f(\eta))$, N_r i.i.d. samples $u^{(k)}, k = 1, \ldots, N_r$ are drawn from a uniform distribution $\mathcal{U}(0, 1)$. Now the drawn values $u^{(k)}$ are compared to the CDF. In particular, if $u^{(k)} < F(1)$, a copy of the first sample $\eta_f^{(1)}$ is generated. If $F(1) \leq u^{(k)} < F(2)$. a copy of the second sample is generated, etc. Fig. 3.8 depicts an example of this procedure for

$\mathbf{u} = [u^{(1)}, ..., u^{(N_r)}]^{\mathrm{T}} = [0.23, 0.72, 0.41]^{\mathrm{T}}$. In this example, $N_r = 3$ samples are drawn from $\mathcal{R}_{N_s}(f(\eta))$. In particular, $\eta_f^{(1)^*} = \eta_f^{(2)}$, $\eta_f^{(2)^*} = \eta_f^{(3)}$, and $\eta_f^{(3)^*} = \eta_f^{(2)}$. Note that the outcome of resampling is random, since the entries of $\mathbf{u}$ are random. That is, the set of retained values $\{1/N_r, \eta_f^{(k)^*}\}_{k=1}^{N_r}$ varies even if resampling is executed repeatedly with the same set of input samples $\{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$.

From the previous example. it becomes evident that, for finite N_r, resampling reduces the diversity of the samples $\eta_f^{(k)}$, since samples with small weights are usually erased. In other words. the cardinality of the set of distinct retained samples is generally smaller than the cardinality of the original set of distinct samples. On the other hand, the set of relevant samples (samples with large weight) is increased, since multiple copies of the relevant samples are retained with high probability. In summary, resampling leads to a particle representation which concentrates on the particles with large weights at the expense of reduced diversity of the samples. After resampling, all particles have the same weights. Typically, though not necessarily. the number of samples after resampling is reduced, i.e. $N_r \leq N_s$.

3.2.2 Multiplication of Probability Densities

Frequently in statistical signal processing, densities have to be multiplied in order to fuse information from multiple observations. In fact, multiplication of densities is such an integral operation that almost all applications rely on it. For instance, consider the posterior distribution from the running example in (3.2) which is given as the product of two likelihood functions $f(\hat{d}_{i\to 1}|\eta_1, \eta_i^*), i \in \{A, B\}$ and the prior $f_3(\eta_1)$. Without multiplication of densities, neither MMSE nor MAP estimates on the position η_1 can be obtained. Multiplication of arbitrary densities is not necessarily a simple task. The actual multiplication method depends on the representations of the distributions to be multiplied, as well as the representation of the product. Numerous combinations of representations are possible. For brevity, the discussion of the different combinations in this chapter is restricted to the following:

1. Discretized input - discretized output
2. Particle-based input - particle-based output
3. Parametrized input - parametrized output

The product $f(\eta) \triangleq \prod_{i=1}^{I} f_i(\eta)$ of an arbitrary number of densities $f_i(\eta)$ can be recursively defined as the product of two densities, i.e.

$$f(\eta) \triangleq \tilde{f}_{I-1}(\eta) \cdot f_I(\eta), \tag{3.11}$$

where $\tilde{f}_i(\eta) \triangleq \tilde{f}_{i-1}(\eta) f_i(\eta)$ with $\tilde{f}_1(\eta) = f_1(\eta)$. Therefore, only pairwise multiplication is considered in this section. For notational convenience, subscripts are dropped and distributions are distinguished by different letters. In particular, the following problem is considered

$$h(\eta) = f(\eta) \cdot g(\eta). \tag{3.12}$$

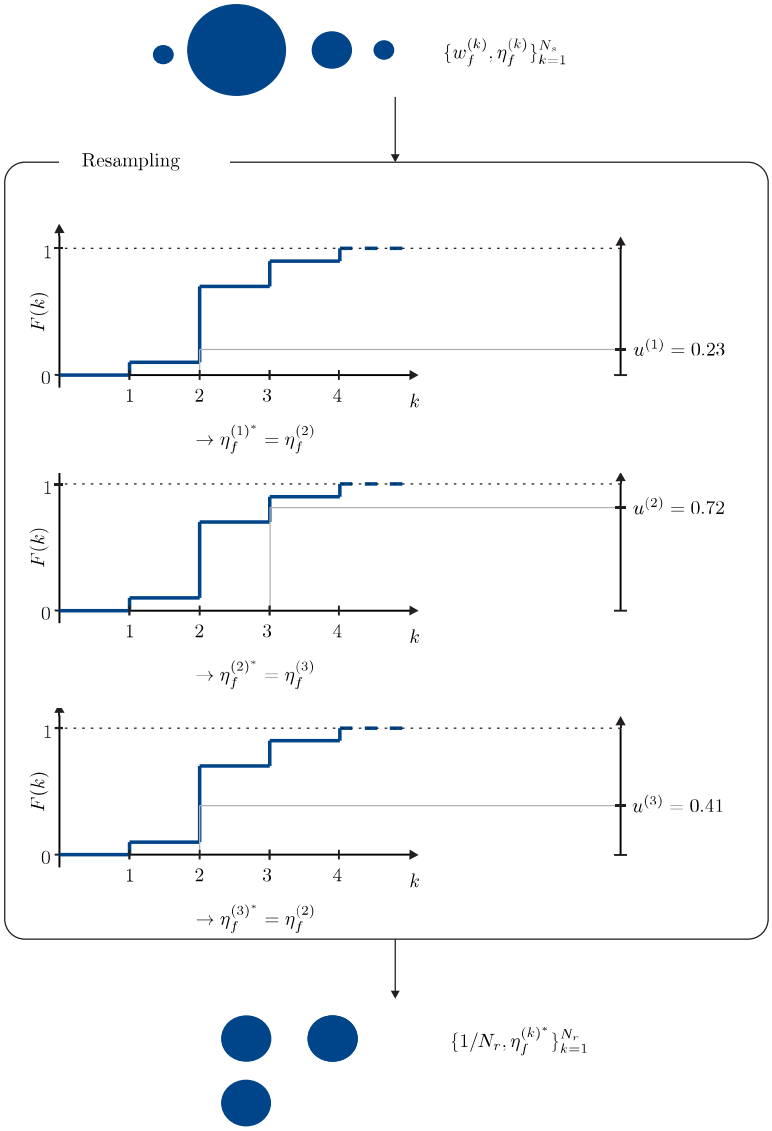

Figure 3.8: *Resampling example* - the particle representation $\mathcal{R}_{N_s}(f(\eta))$ with samples $\{\eta_f^{(1)}, \ldots, \eta_f^{(4)}\} = \{0.9, 1.5, 2.1, 2.8\}$ and weights $\{w_f^{(1)}, \ldots, w_f^{(4)}\} = \{0.1, 0.5, 0.3, 0.1\}$ is resampled using $N_r = 3$ samples. For the given sequence $\{u^{(1)}, u^{(2)}, u^{(3)}\} = \{0.23, 0.72, 0.41\}$, samples and weights of the resampled particle representation are given by $\{\eta_f^{(1)^*}, \eta_f^{(2)^*}, \eta_f^{(3)^*}\} = \{1.5, 2.1, 1.5\}$ and $\{w_f^{(1)^*}, w_f^{(2)^*}, w_f^{(3)^*}\} = \{1/N_r, 1/N_r, 1/N_r\}$.

3.2.2.1 Multiplication of Discretized Densities

Consider that $f(\eta)$ and $g(\eta)$ have been discretized as explained in Section 3.1.1.1. It is assumed that the quantization points are identical, i.e. $\mathcal{H} = \mathcal{H}_g = \mathcal{H}_f = \{\eta^{(1)}, \ldots, \eta^{(N_s)}\}$. Two sets of weights have been obtained, i.e. $\mathcal{F} = \{f^{(1)}, \ldots, f^{(N_s)}\}$ and $\mathcal{G} = \{g^{(1)}, \ldots, g^{(N_s)}\}$. If the weights of both sets are stacked in the respective vectors $\mathbf{f} = [f^{(1)}, \ldots, f^{(N_s)}]^{\mathrm{T}}$ and $\mathbf{g} = [g^{(1)}, \ldots, g^{(N_s)}]^{\mathrm{T}}$, the vector of the weights of the product is obtained by the taking the point-wise product of $\mathbf{f}$ and $\mathbf{g}$, i.e.

$$\mathbf{h} = [h^{(1)}, \ldots, h^{(N_s)}]^{\mathrm{T}} = [f^{(1)} \cdot g^{(1)}, \ldots, f^{(N_s)} \cdot g^{(N_s)}]^{\mathrm{T}}. \tag{3.13}$$

The discretized representation of the product of $f(\eta)$ and $g(\eta)$ is finally given by the set of quantization points $\mathcal{H}$ and the set of weights $\{h^{(1)}, \ldots, h^{(N_s)}\}$. Observe that any pair of densities can be multiplied in that fashion, as long as both densities $f(\eta)$ and $g(\eta)$ can be evaluated in the same quantization points. Notice that the weights $\mathbf{h}$ do not add up to one, even if the weights $\mathbf{f}$ and $\mathbf{g}$ are normalized. Thus, an additional normalization step is necessary.

3.2.2.2 Multiplication of Particle-based Densities

If the area of nonparametric methods is considered, densities are often given as particle representation. For the purpose of illustration, consider the previous example, where the density $\tilde{f}(\eta_1) \triangleq f_1(\eta_1) f_2(\eta_1)$ shall be computed. Here, the goal is to obtain a particle representation $\mathcal{R}_{N_s}\left(\tilde{f}(\eta_1)\right) = \{w_{\tilde{f}}^{(k)}, \eta_{1,\tilde{f}}^{(k)}\}_{k=1}^{N_s}$. In this example, it becomes evident that the multiplication is not straightforward if both densities f_1 and f_2 are given as particle representations, i.e. $\mathcal{R}_{N_s}(f_1(\eta_1)) = \{w_{f_1}^{(k)}, \eta_{1,f_1}^{(k)}\}_{k=1}^{N_s}$ and $\mathcal{R}_{N_s}(f_2(\eta_1)) = \{w_{f_2}^{(k)}, \eta_{1,f_2}^{(k)}\}_{k=1}^{N_s}$. This can be directly seen from Fig. 3.10. since, in contrast to the discretized case, the samples of the particles $\eta_{1,f_1}^{(k)}$ and $\eta_{1,f_2}^{(k)}$ are not identical. Because the samples $\eta_{1,f_1}^{(k)}$ and $\eta_{1,f_2}^{(k)}$ from both sets were drawn independently, they will be distinct with probability one.

To avoid further notational clutter, the abstract problem $h(\eta) = f(\eta)g(\eta)$ is considered again. The particle representations of $f(\eta)$ and $g(\eta)$ are given by $\mathcal{R}_{N_s}(f(\eta)) = \{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$ and $\mathcal{R}_{N_s}(g(\eta)) = \{w_g^{(k)}, \eta_g^{(k)}\}_{k=1}^{N_s}$, respectively. In order to multiply the particle representations of the densities, continuous estimates of the particle representations (here: $\hat{f}(\eta)$ and $\hat{g}(\eta)$) are first obtained using kernel density estimation (see the grey box at the bottom of the page). After the kernel density estimates $\hat{f}(\eta)$ and $\hat{g}(\eta)$ of the particle representations $\mathcal{R}_{N_s}(f(\eta))$ and $\mathcal{R}_{N_s}(g(\eta))$ are obtained, the product $\hat{h}(\eta) = \hat{f}(\eta)\hat{g}(\eta)$ can be evaluated for arbitrary values of η. Subsequently, random samples $\{\eta_h^{(k)}\}_{k=1}^{N_s}$ are drawn from a suitable proposal distribution $q(\eta)$, i.e. $\eta_h^{(k)} \sim q(\eta)$. Finally, the weights of the output particle representations are determined as the product of the continuous densities which are evaluated in the positions of the samples, i.e. $\tilde{w}_h^{(k)} = \hat{f}(\eta_h^{(k)})\hat{g}(\eta_h^{(k)})$. Fig. 3.9 shows a block diagram of the procedure. The individual parts of the block diagram will be discussed in more detail in the following paragraphs.

Obtaining a Kernel Density Estimate of a Distribution

Intuition:
A kernel density estimate (KDE) of a particle representation $\mathcal{R}_{N_s}(f(\eta))$ coats every particle with a continuous kernel $K(\eta;\boldsymbol{\theta}_K)$ like a Gaussian kernel. Kernels are centered at the respective positions of every particle $\eta_f^{(k)}$. Every kernel is scaled by the weight $w_f^{(k)}$ of the particle. Then, the superposition of all kernels yields the continuous kernel density estimate $\hat{f}(\eta)$. An illustration in the context of the running example is depicted in the upper part of Fig. 3.10.

Mathematical description:
Given a particle representation $\mathcal{R}_{N_s}(f(\eta)) = \{w_f^{(k)}, \eta_f^{(k)}\}_{k=1}^{N_s}$ of distribution $f(\eta)$, a kernel density estimate using Gaussian kernels can be obtained according to

$$\hat{f}(\eta) = \sum_{k=1}^{N_s} w_f^{(k)} \mathcal{N}(\eta; \eta_f^{(k)}, \hat{\sigma}^2_{\mathrm{KDE}}), \tag{3.14}$$

where $\mathcal{N}(\eta; \eta_f^{(k)}, \hat{\sigma}^2_{\mathrm{KDE}})$ is the Gaussian kernel with mean $\eta_f^{(k)}$ and variance $\hat{\sigma}^2_{\mathrm{KDE}}$. Observe that the mean of the kernel is directly given by the position of the particle $\eta_f^{(k)}$. However, the variance $\hat{\sigma}^2_{\mathrm{KDE}}$ of the kernel has to be estimated based on the properties of the particle representation $\mathcal{R}_{N_s}(f(\eta))$ [Bot07]. The field of kernel density estimation treats the problem of estimating $\hat{\sigma}^2_{\mathrm{KDE}}$ [Bot07]. Popular algorithms are, e.g., least squares cross-validation and the improved plug-in method [Bot07]. It is worthwhile to note that the choice of the kernel function does not affect estimation accuracy considerably, while the choice of the variance $\hat{\sigma}^2_{\mathrm{KDE}}$ has a major impact on the accuracy [Sil86].

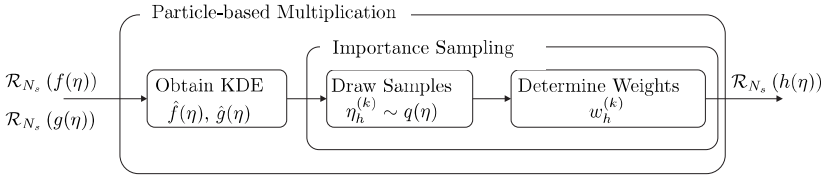

Figure 3.9: *Block diagram of particle-based multiplication of densities* - the product of the KDEs $\hat{f}(\eta)$ and $\hat{g}(\eta)$ is determined via importance sampling.

In order to obtain a particle representation of the product of $\hat{f}(\eta)$ and $\hat{g}(\eta)$, importance sampling as described in Section 3.2.1.1 will be employed. First, random samples $\eta_h^{(k)}, k = 1, ..., N_s$ are drawn from a proposal distribution $q(\eta)$. These samples are then weighted according to

$$\tilde{w}_h^{(k)} = \frac{\hat{h}\left(\eta_h^{(k)}\right)}{q\left(\eta_h^{(k)}\right)} = \frac{\hat{f}\left(\eta_h^{(k)}\right)\hat{g}\left(\eta_h^{(k)}\right)}{q\left(\eta_h^{(k)}\right)}. \tag{3.15}$$

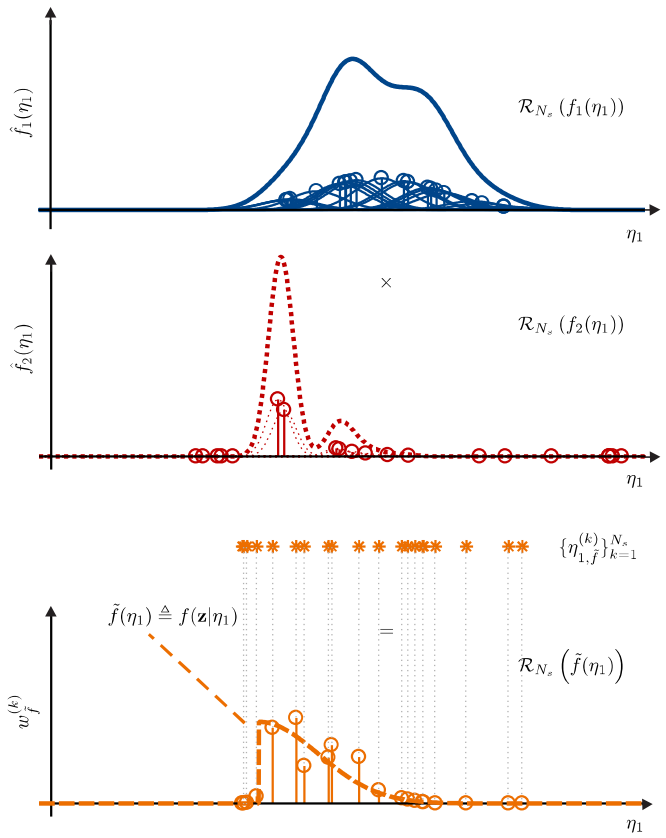

Figure 3.10: *Particle-based multiplication of densities* - a particle representation $\mathcal{R}_{N_s}(\tilde{f}(\eta_1))$ of the product of two particle representations $\mathcal{R}_{N_s}(f_1(\eta_1))$ and $\mathcal{R}_{N_s}(f_2(\eta_1))$ is obtained by (i) obtaining the KDEs $\hat{f}_1(\eta_1)$ and $\hat{f}_2(\eta_1)$ first and (ii) performing importance sampling subsequently. It can be seen that the exact (true) density $\tilde{f}(\eta_1)$ (orange dash-dotted line) is well approximated by the particle representation $\mathcal{R}_{N_s}(\tilde{f}(\eta_1))$.

Note that the weights are not normalized. The particle representation of $h(\eta)$ is then given as $\mathcal{R}_{N_s}(h(\eta)) = \{w_h^{(k)}, \eta_h^{(k)}\}_{k=1}^{N_s}$, where $w_h^{(k)} = \tilde{w}_h^{(k)} / \sum_j \tilde{w}_h^{(j)}, \forall k$ are the normalized weights. The steps of the particle-based multiplication (except for normalization) are visualized in Fig. 3.10 in the context of the running example.

Since particle-based multiplication employs importance sampling, the same observations regarding the proposal distributions apply here. In particular, the proposal distribution $q(\eta)$ needs to be chosen such that it has significant probability mass where the product of densities $h(\eta) = f(\eta)g(\eta)$ has considerable mass. Notice that this is a *chicken-or-the-egg* problem, since one usually does not know where the product $h(\eta) = f(\eta)g(\eta)$ has significant probability mass before evaluating it. To overcome this problem, a major contribution of this work is a strategy which confines the regions of considerable probability mass. The

method is coined polygon outer-approximate and is presented in Section 5.4.

3.2.2.3 Multiplication of Parametrized Densities

If the input densities $f(\eta;\boldsymbol{\theta}_f)$ and $g(\eta;\boldsymbol{\theta}_g)$ are given as parametrized densities and the output density $h(\eta;\boldsymbol{\theta}_h)$ should be a parametrized density of a certain class, multiplication becomes more complicated compared to the discretized case. This can be easily verified if the running example in Fig. 3.1 is considered, where the densities $f_1(\eta_1)$ and $f_2(\eta_1)$ are Gaussian and one-sided exponential, respectively. Note that the product of the two densities is neither Gaussian, nor one-sided exponential. If the output density $h(\eta;\boldsymbol{\theta}_h)$ is from the class of Gaussian distributions, an approximation error is inevitable. It is worth noting that the error will depend on the choice of the measure. This observation gives rise to the introduction of a criteria that describes how the mismatch been $h(\eta;\boldsymbol{\theta}_f)$ and $f(\eta;\boldsymbol{\theta}_f)\cdot g(\eta;\boldsymbol{\theta}_g)$ is measured (a so-called *divergence measure*) and how the parameters $\boldsymbol{\theta}_h$ are determined.

On a more general scale: Multiplying arbitrary parametrized densities $f(\eta;\boldsymbol{\theta}_f)$ and $g(\eta;\boldsymbol{\theta}_g)$ (even if they are of the same type) generally yields a density $h(\eta)$ which is neither of the type of f nor of the type of g. Only very few exceptions exist. The most prominent examples are Gaussian distributions, whose product is again Gaussian [KF09]. In particular, if $f(\eta;m_f,\sigma_f^2)=\mathcal{N}(\eta;m_f,\sigma_f^2)$ and $g(\eta;m_g,\sigma_g^2)=\mathcal{N}(\eta;m_g,\sigma_g^2)$, $h(\eta;m_h;\sigma_h^2)=\mathcal{N}(\eta;m_h,\sigma_h^2)$ where m_h and σ_h^2 can be determined in a closed form:

$$m_h=\frac{m_f\sigma_f^2+m_g\sigma_g^2}{\sigma_f^2+\sigma_g^2}\qquad\text{and}\qquad\sigma_h^2=\frac{1}{1/\sigma_f^2+1/\sigma_g^2}. \tag{3.16}$$

For non-Gaussian distributions $f(\eta;\boldsymbol{\theta}_f)$ and $g(\eta;\boldsymbol{\theta}_f)$, tools from the field of variational inference are often employed to determine $h(\eta;\boldsymbol{\theta}_h)$ as a distribution parametrized in $\boldsymbol{\theta}_h$. Though, the details of the individual methods differ, all methods share a common approach which will be explained in the following.

Intuition:
Generally, parametrized distributions from a certain family (like the Gaussian family) cannot represent products of densities exactly but only approximately. The goal of parametrized message multiplications is to find a parametrized distribution which resembles the actual product as closely as possible. Along this road, the family of distributions must be chosen. Then, the criterion of divergence (e.g., Kullback-Leibler divergence) must be chosen. In other words, it has to be decided how the closeness of distributions is measured. Finally, the optimum set of parameters must be determined such that the highest resemblance/lowest divergence results. This procedure is depicted in a block diagram in Fig. 3.11 and in the context of the running example in Fig. 3.12.

Mathematical description:
First, the family of output distributions $\mathfrak{H}$ must be chosen. Then, the optimum member $h(\eta;\boldsymbol{\theta}_{h,\text{opt}})$ of the family $\mathfrak{H}$ has to be determined. Optimality refers to a predefined divergence measure like the Kullback-Leibler divergence. Hellinger distance, χ^2-distance, etc. This divergence measure is minimized in order to find the parameters $\boldsymbol{\theta}_{h,\text{opt}}$ of the

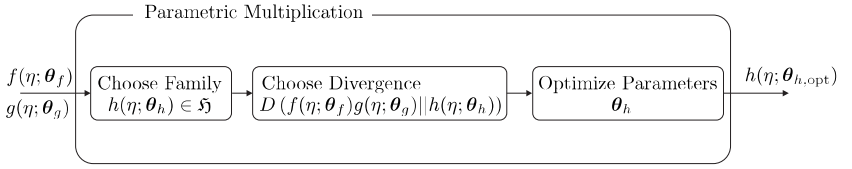

Figure 3.11: *Block diagram of parametric multiplication of densities* - by choosing a family of distributions $\mathfrak{H}$ and a divergence measure $D(\cdot||\cdot)$, parametric densities $f(\eta;\boldsymbol{\theta}_f)$ and $g(\eta;\boldsymbol{\theta}_g)$ can be multiplied by optimizing the parameters of $h(\eta;\boldsymbol{\theta}_h)$ such that the divergence $D(f(\eta;\boldsymbol{\theta}_f)g(\eta;\boldsymbol{\theta}_g)||h(\eta;\boldsymbol{\theta}_h))$ is minimized.

optimum member $h(\eta;\boldsymbol{\theta}_{h,\mathrm{opt}}) \in \mathfrak{H}$. A rigorous mathematical description goes as follows:

$$\boldsymbol{\theta}_{h,\mathrm{opt}} = \underset{\boldsymbol{\theta}_h}{\arg\min}\, D\left(f(\eta;\boldsymbol{\theta}_f)\cdot g(\eta;\boldsymbol{\theta}_g)||h(\eta;\boldsymbol{\theta}_h)\right). \tag{3.17}$$

where $D\left(\cdot||\cdot\right)$ is the divergence between the actual product $f(\eta;\boldsymbol{\theta}_f)\cdot g(\eta;\boldsymbol{\theta}_g)$ and the approximated product $h(\eta;\boldsymbol{\theta}_h)$.

For the purpose of clarity, the running example from Fig. 3.1 is revisited. In particular, the density $\tilde{f}(\eta_1) = f_1(\eta_1;\boldsymbol{\theta}_{f_1})f_2(\eta_1;\boldsymbol{\theta}_{f_2})$ should be computed, which is the product of the Gaussian density from anchor A, $f_1(\eta_1) = \mathcal{N}(\eta_1; m_{f_1}, \sigma^2_{f_1})$, and the one-sided exponential density from anchor B, $f_2 = \mathcal{E}(\eta_1; m_{f_2})$, where $\mathcal{E}(\cdot; m)$ denotes the one-sided exponential distribution with mean m. Observe that the density $\tilde{f}(\eta_1)$ is non-Gaussian, but should be approximated by a member $h(\eta_1; m_h, \sigma^2_h)$ of the family $\mathfrak{H}$ of Gaussian distributions. In this particular example, the variational problem can be written as

$$\boldsymbol{\theta}_{h,\mathrm{opt}} = \underset{\boldsymbol{\theta}_h}{\arg\min}\, D\left(f_1(\eta_1; m_{f_1}, \sigma^2_{f_1})\cdot f_2(\eta_1; m_{f_2})||h(\eta_1;\boldsymbol{\theta}_h)\right). \tag{3.18}$$

Since any member $h(\eta_1)$ of the family of univariate Gaussian distributions $\mathfrak{H}$ has only two parameters (mean m_h and variance σ^2_h), the problem can be reformulated

$$\{m_{h,\mathrm{opt}}, \sigma^2_{h,\mathrm{opt}}\} = \underset{\boldsymbol{\theta}=\{m_h,\sigma^2_h\}}{\arg\min}\, D\left(f_1(\eta_1; m_{f_1}, \sigma^2_{f_1})\cdot f_2(\eta_1; m_{f_2})||h(\eta_1; m_h, \sigma^2_h)\right). \tag{3.19}$$

It can be shown that the optimum parameters (under the Kullback-Leibler divergence) $m_{h.\mathrm{opt}}$ and $\sigma^2_{h,\mathrm{opt}}$ of a Gaussian distribution are given by the mean and variance of $f_1(\eta_1; m_{f_1}, \sigma^2_{f_1})\cdot f_2(\eta_1; m_{f_2}, \sigma^2_{f_2})$, respectively [Min05]. Mathematically that means:

$$m_{h,\mathrm{opt}} = \int \eta_1 f_1(\eta_1; m_{f_1}, \sigma^2_{f_1})\cdot f_2(\eta_1; m_{f_2}, \sigma^2_{f_2}), \mathrm{d}\eta_1, \tag{3.20}$$

$$\sigma^2_{h,\mathrm{opt}} = \int (\eta_1 - m_{h,\mathrm{opt}})^2 f_1(\eta_1; m_{f_1}, \sigma^2_{f_1})\cdot f_2(\eta_1; m_{f_2}, \sigma^2_{f_2})\mathrm{d}\eta_1. \tag{3.21}$$

Hence a very simple result can be concluded: If any product of densities is approximated by a member of the family of Gaussian distributions, the optimum parameters of the Gaussian distribution are given by the mean and variance of the product of densities which can be

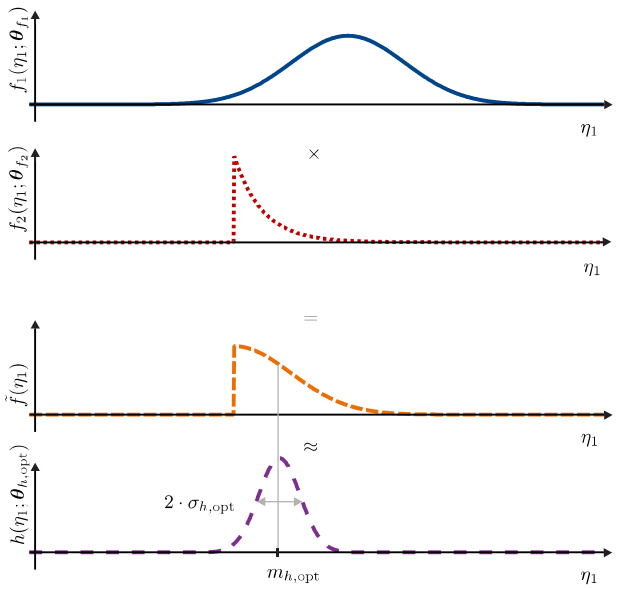

Figure 3.12: *Parametrized multiplication of densities* - the product $\tilde{f}(\eta_1) = f_1(\eta_1;\boldsymbol{\theta}_{f_1})f_2(\eta_1;\boldsymbol{\theta}_{f_2})$ of a Gaussian density $f_1(\eta_1;\boldsymbol{\theta}_{f_1})$ and a one-sided exponential density $f_2(\eta_1;\boldsymbol{\theta}_{f_2})$ is neither Gaussian nor exponentially distributed. In order to obtain a parametrized representation, $\tilde{f}(\eta_1)$ is approximated by the optimum member $h(\eta_1;\boldsymbol{\theta}_{h,\mathrm{opt}})$ of the Gaussian family of distributions. The parameters $m_{h,\mathrm{opt}}$ and $\sigma^2_{h,\mathrm{opt}}$ are chosen such that the KLD between $\tilde{f}(\eta_1)$ and $h(\eta_1;\boldsymbol{\theta}_h)$ is minimized.

determined, for instance, by means of numerical integration. Fig. 3.12 visualizes the Gaussian distribution $\mathcal{N}(\eta_1; m_{h,\mathrm{opt}}, \sigma^2_{h,\mathrm{opt}})$ that optimally approximates the density $\tilde{f}(\eta_1)$.

It should be noted, however, that it is generally difficult to determine the optimum set of parameters if the target density h is non-Gaussian and/or the divergence measure is not given by the Kullback-Leibler divergence. In the general case, the optimization problem in (3.17) cannot be solved in a closed form. Often, the resulting optimization problems are nonlinear and non-convex which makes it difficult to find globally optimum estimates $\hat{\boldsymbol{\theta}}_h$.

3.3 Summary

In statistical signal processing, the representations of probability densities play an importance role. Popular ways of representing densities are the following.

- Discrete representations
    1. Discretized representation

2. Sample-based representation
3. Particle-based representation

- Parametrized representations

Discrete representations capture the characteristics of a density by a finite set of points and possibly weights. Discretized representations perform a uniform quantization of the sample space and distributions are represented by deterministically chosen quantization points and the associated weights. In contrast to the discretized case, sample-based and particle-based representations employ sets of random samples to capture the properties of continuous densities. In addition, particle-based descriptions also assign a weight to each sample. Parametrized representations describe densities by a finite (typically small) set of parameters. The most prominent example is the Gaussian distribution whose parameters are given by its mean and variance.

Essential operations in statistical signal processing are (i) sampling from arbitrary distributions and (ii) multiplication of arbitrary densities.
Importance sampling is a sampling technique which allows one to draw samples from a target distribution that cannot be sampled directly. In the first step of that process, samples are drawn from a so-called proposal distribution for which sampling is tractable. Subsequently, a weight is associated with every sample in order to account for the mismatch between the target distribution and the proposal distribution. For numerical stability, weights are usually normalized as the last step of the procedure. The outcome of importance sampling is a so-called particle representation which consists of a set of samples and weights.
Multiplication of arbitrary densities can be facilitated in different ways, depending on the representations of the distributions to be multiplied. In general, the multiplication of densities is a computationally expensive operation. For instance, multiplying parametrized distributions requires solving optimization problems in which the optimum set of parameters of the product of densities is determined. It can be summarized that particle-based representations are difficult to multiply, since they often require obtaining kernel density estimates of the particle representations and employ importance sampling subsequently.

Chapter 4

Variants of Belief Propagation

This chapter revisits different variants of belief propagation which is an algorithm that is frequently used in Bayesian inference problems. For that reason. Bayesian inference problems are briefly revisited. Subsequently, factor graphs are introduced as a means of structuring such inference problems. This chapter is concluded by recapitalizing different variants of the belief propagation algorithm for continuous variables. In particular, discretized, parametric and sample-based belief propagation are discussed. For the sake of illustration, the running example from Chapter 3 is extended as follows.

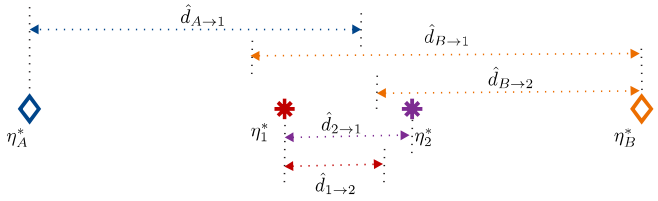

Figure 4.1: *Extended running example* - The agents 1 and 2 attempt to localize using the distance estimates $\{\hat{d}_{A\to1}, \hat{d}_{B\to1}, \hat{d}_{B\to2}\}$ with respect to the anchors A and B, as well as the cooperative distance estimates $\{\hat{d}_{1\to2}, \hat{d}_{2\to1}\}$. Recall that the true positions of anchors $\eta_i^*, i \in \{A, B\}$ are known while the true positions of agents $\eta_j^*, j \in \{1, 2\}$ are unknown.

Running Example - Extension

The extended topology is depicted in Fig. 4.1. In the extended running example an additional agent with location $\eta_2^* = 7.5$ is considered. It is assumed that agent 1 and 2 cooperate in the sense that they have obtained the noisy cooperative distance estimates $\hat{d}_{1\to2} = 2$ and $\hat{d}_{2\to1} = 2.5$. In addition, agent 2 has obtained a range estimate $\hat{d}_{B\to2} = 5$ with respect to anchor B. The distributions of the ranging errors $e_{1\to2}$, $e_{2\to1}$, and $e_{B\to2}$ are assumed to be zero-mean Gaussian with variance $\sigma^2 = 1$. Considering the additional observations, the vector of all measurements is given by $\mathbf{z} = [\hat{d}_{A\to1}, \hat{d}_{B\to1}, \hat{d}_{B\to2}, \hat{d}_{1\to2}, \hat{d}_{2\to1}]^{\mathrm{T}}$. All assumptions and numerical values regarding agent 1 from the original running example are also considered in the extended example. Assuming conditional independent observations, the joint likelihood $f(\mathbf{z}|\underbrace{\eta_1, \eta_2, \eta_A^*, \eta_B^*}_{\triangleq\boldsymbol{\eta}})$ can be written as follows

$$\begin{aligned} f(\mathbf{z}|\boldsymbol{\eta}) = &\underbrace{f(\hat{d}_{A\to1}|\eta_1, \eta_A)}_{f_1(\eta_1)} \underbrace{f(\hat{d}_{B\to1}|\eta_1, \eta_B)}_{f_2(\eta_1)} \\ &\underbrace{f(\hat{d}_{B\to2}|\eta_2, \eta_B)}_{f_3(\eta_2)} \underbrace{f(\hat{d}_{2\to1}|\eta_1, \eta_2)}_{f_4(\eta_1,\eta_2)} \\ &\underbrace{f(\hat{d}_{1\to2}|\eta_1, \eta_2)}_{f_5(\eta_1,\eta_2)}. \end{aligned} \tag{4.1}$$

Hence the joint posterior of the extended problem is given by

$$\begin{aligned} f(\boldsymbol{\eta}|\mathbf{z}) \propto &\underbrace{f(\hat{d}_{A\to1}|\eta_1, \eta_A^*)}_{f_1(\eta_1)} \underbrace{f(\hat{d}_{B\to1}|\eta_1, \eta_B^*)}_{f_2(\eta_1)} \\ &\underbrace{f(\hat{d}_{B\to2}|\eta_2, \eta_B^*)}_{f_3(\eta_2)} \underbrace{f(\hat{d}_{2\to1}|\eta_1. \eta_2)}_{f_4(\eta_1,\eta_2)} \\ &\underbrace{f(\hat{d}_{1\to2}|\eta_1, \eta_2)}_{f_5(\eta_1,\eta_2)} f_6(\eta_1) f_7(\eta_2). \end{aligned} \tag{4.2}$$

where the prior distribution of agent 1 and 2 are given by $f_6(\eta_1) = f_7(\eta_2) = \mathcal{U}(2, 7)$ and the priors $\delta(\eta_i - \eta_i^*), i \in A, B$ of the anchors A and B have been absorbed in the densities f_1, f_2, and f_3. The goal of the extended problem is to determine the marginal posterior distributions $f(\eta_1|\mathbf{z})$ and $f(\eta_2|\mathbf{z})$ of the agents based on which the position estimates $\hat{\eta}_1$ and $\hat{\eta}_2$ should be obtained like in (3.3).

4.1 Bayesian Inference

In Bayesian inference, parameters such as the position of a node are treated as realizations of random variables with known prior distributions. Many problems in the field of positioning like cooperative positioning or multipath-assisted positioning can be formulated

as multidimensional Bayesian inference problems. Given a set of observations (e.g., distances and angles) the goal of Bayesian inference is to infer the unknown parameters (like positions and orientations of agents) based on the statistics of the observations and the prior information on the parameters. For a general problem, where the set of parameters $\boldsymbol{\eta} = [\boldsymbol{\eta}_1^{\mathrm{T}}, ..., \boldsymbol{\eta}_N^{\mathrm{T}}]^{\mathrm{T}}$ shall be inferred based on the observations $\mathbf{z}$, the optimum estimate $\hat{\boldsymbol{\eta}}_{\mathrm{MAP}}$ in the a posteriori sense is given by

$$\hat{\boldsymbol{\eta}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\eta}} f(\boldsymbol{\eta}|\mathbf{z}). \tag{4.3}$$

Using Bayes rule, the posterior distribution $f(\boldsymbol{\eta}|\mathbf{z})$ can be rewritten as

$$f(\boldsymbol{\eta}|\mathbf{z}) = \frac{f(\mathbf{z}|\boldsymbol{\eta})f(\boldsymbol{\eta})}{f(\mathbf{z})} \propto f(\mathbf{z}|\boldsymbol{\eta})f(\boldsymbol{\eta}), \tag{4.4}$$

where $f(\mathbf{z}|\boldsymbol{\eta})$ is the joint likelihood function and $f(\mathbf{z})$ is the likelihood of observation $\mathbf{z}$. Observe that $f(\mathbf{z})$ does not depend on $\boldsymbol{\eta}$ and thus, $f(\mathbf{z})$ can be disregarded for maximization. Hence the optimization problem can be rewritten as

$$\hat{\boldsymbol{\eta}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\eta}} f(\mathbf{z}|\boldsymbol{\eta})f(\boldsymbol{\eta}). \tag{4.5}$$

Intuitively speaking, the maximum a posteriori (MAP) estimate is given by the point $\hat{\boldsymbol{\eta}}_{\mathrm{MAP}}$ which is the best explanation of the observations $\mathbf{z}$ while taking the prior knowledge on $\boldsymbol{\eta}$ into account.

Often the problem described by (4.5) is difficult to solve numerically if the dimension of $\boldsymbol{\eta}$ is large. In many situations, it is more tractable to determine estimates $\hat{\boldsymbol{\eta}}_i$ on the lower-dimensional marginal posterior distributions $f(\boldsymbol{\eta}_i|\mathbf{z})$, rather than the joint posterior distribution $f(\boldsymbol{\eta}|\mathbf{z})$, i.e.

$$\hat{\boldsymbol{\eta}}_{i,\mathrm{MAP}} = \arg\max_{\boldsymbol{\eta}_i} f(\boldsymbol{\eta}_i|\mathbf{z}). \tag{4.6}$$

Note that the marginal posteriors $f(\boldsymbol{\eta}_i|\mathbf{z})$ have reduced dimensions, e.g.

$$\dim(\boldsymbol{\eta}_1) = \dim(\boldsymbol{\eta}) - \dim([\boldsymbol{\eta}_2^{\mathrm{T}}, ..., \boldsymbol{\eta}_N^{\mathrm{T}}]^{\mathrm{T}}), \tag{4.7}$$

which makes it easier to obtain marginal MAP estimates rather than joint MAP estimates.

Mathematically, the marginal posterior of $\boldsymbol{\eta}_i$ is determined by integrating the joint posterior over all variables other than $\boldsymbol{\eta}_i$, i.e.

$$\begin{aligned} f(\boldsymbol{\eta}_i|\mathbf{z}) &= \int \dots \int \int \dots \int f(\boldsymbol{\eta}|\mathbf{z}) \mathrm{d}\boldsymbol{\eta}_1 \dots \mathrm{d}\boldsymbol{\eta}_{i-1} \mathrm{d}\boldsymbol{\eta}_{i+1} \dots \mathrm{d}\boldsymbol{\eta}_N \\ &\triangleq \int f(\boldsymbol{\eta}|\mathbf{z}) \mathrm{d}{\sim}\boldsymbol{\eta}_i, \end{aligned} \tag{4.8}$$

where ${\sim}\boldsymbol{\eta}_i$ denotes the set of all parameters except for $\boldsymbol{\eta}_i$. Calculating the marginals as described in (4.8) is referred to as *naive marginalization*. If the variables $\boldsymbol{\eta}_i, \forall i$ of the joint posterior show some structure (conditional independence), naive marginalization performs certain calculations repeatedly [Wym07], leading to a waste of computational resources. The belief propagation algorithm leverages the structure of the variables in order to determine the marginal posterior distributions in a computationally efficient way,

i.e. without redoing calculations [Wym07]. In order to reveal the structure of variables, the concepts of *factorizations* and *factor graphs* will be introduced in the next section.

4.2 Factorizations and Factor Graphs

As mentioned in the previous section, densities can be often factorized. For clarity, the joint posterior distribution $f(\boldsymbol{\eta}|\mathbf{z})$ of the running example is revisited:

$$f(\boldsymbol{\eta}|\mathbf{z}) \propto f_1(\eta_1)f_2(\eta_1)f_3(\eta_2)f_4(\eta_1,\eta_2)f_5(\eta_1,\eta_2)f_6(\eta_1)f_7(\eta_2). \tag{4.9}$$

The posterior is given as the product of seven factors; the priors and the densities $f_i, i = 1, ..., 5$ and the priors f_6, f_7. Observe that factorizations are not unique, since another factorization results by simply lumping together the last two factors:

$$f(\eta_1|\mathbf{z}) \propto f_1(\eta_1)f_2(\eta_1)f_3(\eta_2)f_4(\eta_1,\eta_2)f_5(\eta_1,\eta_2)\tilde{f}(\eta_1,\eta_2) \tag{4.10}$$

where $\tilde{f}(\eta_1,\eta_2) \triangleq f_6(\eta_1)f_7(\eta_2)$. Since different factorizations exist, the structure of the variables can be either hidden or revealed. Generally, it can be noted that the more factors a factorization has, the more independence among variables is revealed. This can be directly seen if (4.9) and (4.10) are compared. (4.9) implies that the priors are independent, while such an observation does not follow from (4.10).

The most general factorization of a density $f(\boldsymbol{\eta} = [\boldsymbol{\eta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\eta}_N^{\mathrm{T}}]^T)$ is given by

$$f(\boldsymbol{\eta}) = \prod_{m=1}^{M} f_m(\mathcal{S}_m), \tag{4.11}$$

where $\mathcal{S}_m \subseteq \{\boldsymbol{\eta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\eta}_N^{\mathrm{T}}\}$ is the m^{th} variable subset and $f_m(\cdot) \geq 0$ is a real-valued function. The set $\mathcal{S}_m$ is also referred to as the neighborhood of f_m and summarizes all nodes that are part of the argument of f_m. For instance, if the joint posterior distribution of the running example in (4.2) is considered, it can be seen that $\mathcal{S}_1 = \{\eta_1\}$, while $\mathcal{S}_4 = \{\eta_1, \eta_2\}$. There exists always a trivial factorization with one factor $f_1(\cdot) \triangleq f(\cdot)$ and $\mathcal{S}_1 \triangleq \{\boldsymbol{\eta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\eta}_N^{\mathrm{T}}\}$.

Factorizations of densities can be relatively convoluted, especially if the number of variables is large and many factors exist. As an alternative to the mathematical description, factorizations can be represented by graphical models, including so-called *factor graphs*. Graphical representations are generally easier to comprehend for humans. There is a one-to-one mapping between factorizations and factor graphs [Wym07]. For a given factorization, it is generally straightforward to construct the corresponding factor graph. Once constructed, various properties of the factorization can be obtained by simply looking at the graph. For instance, the concept of connectedness can be directly seen from the graph by a simple inspection of the graph, i.e. if no parts of the graph are isolated, the graph is connected.

In the literature, mainly two types of factor graphs exist; conventional and normal (Forney-style) factor graphs [Wym07]. In this thesis. only conventional factor graphs are considered. The following recipe describes how a factor graph is drawn based on a given factorization as in (4.11).

Construction of Factor Graphs

The construction of factor graphs is illustrated by an example in Fig. 4.2.

1. For every variable $\boldsymbol{\eta}_i, i = 1, \ldots, N$, draw a circle (variable node) with the corresponding label $\boldsymbol{\eta}_i$.
2. For every factor $f_m, m = 1, \ldots, M$, draw a rectangle (factor node) with the corresponding label f_m.
3. For every factor f_m, draw $|\mathcal{S}_{f_m}|$ edges which connect the factor node f_m with the respective variable nodes $\boldsymbol{\eta}_i \in \mathcal{S}_{f_m}$.

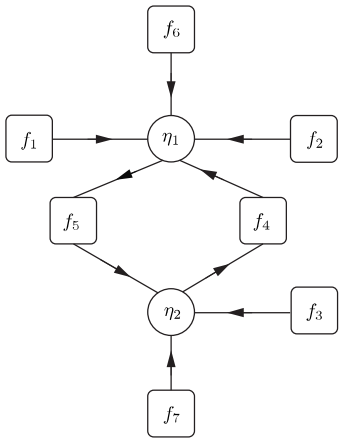

Figure 4.2: *Factor graph example* - the factor graph of the posterior distribution $f(\boldsymbol{\eta}|\mathbf{z})$ in (4.9) is depicted. Factor nodes are given by rectangles, while variable nodes are represented by circles.

The factor graph of the posterior distribution of the extended running example is depicted in Fig. 4.2. It is worth mentioning that factor graphs are in general undirected graphs, meaning that edges have no direction. In the context of belief propagation, however, it is often useful to add directions to edges as will be seen in the following section. It should also be stressed that the factor graph in Fig. 4.2 does not contain the variable nodes of the anchors A and B, since they are not random and their positions are perfectly known. These variables can be absorbed in the factor nodes to which they are connected. For instance,

$$f_1(\eta) \triangleq f(\hat{d}_{A\to 1}|\eta_1, \eta_A)\delta(\eta_A - \eta_A^*) = f(\hat{d}_{A\to 1}|\eta_1, \eta_A^*). \tag{4.12}$$

4.3 Belief Propagation

Motivation:
Often it is desirable to obtain estimates on the joint posterior distribution, but the optimization task in (4.3) may be intractable. Especially, if the dimensionality of the variables to be estimated is high, the problem in (4.3) is difficult or even impossible to solve. In order to obtain estimates on the unknown variables (like the positions of agents), the lower-dimensional marginal posterior distributions are frequently considered, since it is easier to determines estimates. These marginals could be determined one by one according to (4.8), but if there are dependencies between the variables, many intermediate calculations are done repeatedly. Belief propagation is a popular tool which determines all marginal posterior distributions in a computationally efficient way. Based on these marginals, the ultimate goal can be achieved: Estimating the values of the unknown variables. In the following, belief propagation (BP) will be described in a brief manner using the running example.

Belief propagation leverages factorizations, i.e. it makes use of the conditional (in)dependencies of the random variables, allowing for an efficient computation of the marginal posterior distributions [Wym07]. To obtain the marginal posterior distributions, messages (densities) are passed along the edges of an underlying factor graph [Wym07]. At factor nodes, the outgoing messages are determined using filtering operations of the incoming messages, while the outgoing messages at variable nodes are determined as the product of the incoming messages. In cycle-free graphs, belief propagation computes the exact marginal posterior distributions, while the marginal posterior distributions are only approximated in cyclic graphs [Wym07]. Approximate marginal posterior distributions are also called *beliefs*. If the factor graph contains cycles, beliefs are computed iteratively until a convergence criterion is met [Wym07]. In general, convergence is not guaranteed and difficult to prove in general [YFW05]. However, several theoretical studies partially justify the application of belief propagation to graphs with cycles [JW03, WF01, YFW05]. In this thesis, only cyclic graphs are considered, since the factor graphs of the localization problems in this thesis are all cyclic.

Belief propagation is a message passing scheme in which messages are passed along the edges of an underlying factor graph. Generally, belief propagation is initialized at the leaf nodes of the graph, i.e. at nodes which have only a single edge[1]. Then, the sequence of message updates is determined by the so-called message schedule or simply *schedule*. At any node in the graph, outgoing messages are updated based on their incoming messages. Recall that any factor graph consists of two types of nodes; factor and variable nodes. Thus, two generic operations are performed.

1. Factor nodes: message filtering is performed.

2. Variable nodes: message multiplication is implemented.

For the sake of clarity, the belief propagation algorithm will be revisited in the context of the extended running example. It is important to note that in localization, the messages that are computed and passed along the edges of the factor graph have geometrical interpretations. These properties will be highlighted during the course of this section whenever it is appropriate.

[1] If the factor graph does not contain leaf nodes, factor nodes with uninformative (uniform) priors can be attached to variable nodes in order to assure the existence of leaf nodes.

Initialization:
As can be seen in Fig. 4.2, the factors f_1, f_2, f_3, f_6, and f_7 are the leaf nodes. Here, the following messages are sent and visualized in the factor graph in Fig. 4.3. The superscript (l) denotes the l^{th} iteration. Observe that all messages during the initialization process

Table 4.1: Messages transmitted during the initialization.

Leaf Node	Transmitted Message
f_1	$\mu^{(0)}_{f_1\to\eta_1}(\eta_1) = f_1(\eta_1) = f(\hat{d}_{A\to1}\|\eta_1, \eta^*_A)$
f_2	$\mu^{(0)}_{f_2\to\eta_1}(\eta_1) = f_2(\eta_1) = f(\hat{d}_{B\to1}\|\eta_1, \eta^*_B)$
f_3	$\mu^{(0)}_{f_3\to\eta_2}(\eta_2) = f_3(\eta_2) = f(\hat{d}_{B\to2}\|\eta_2, \eta^*_B)$
f_6	$\mu^{(0)}_{f_6\to\eta_1}(\eta_1) = f_6(\eta_1) = \mathcal{U}(2,7)$
f_7	$\mu^{(0)}_{f_7\to\eta_2}(\eta_2) = f_7(\eta_2) = \mathcal{U}(2,7)$

are either densities that are related to anchors who have perfect knowledge regarding their own position or prior information of the positions of agents. For instance, the message $\mu^{(1)}_{f_1\to\eta_1} = f(\hat{d}_{A\to1}|\eta_1, \eta^*_A)$ provides information on the position of agent 1 based on the known position η^*_A of the anchor and the observed distance $\hat{d}_{A\to1}$. Its essential information is that the position of agent 1 is somewhere around the position of anchor A with the distance $\hat{d}_{A\to1}$. In this context, 'somewhere' accounts for the uncertainty of the observed distance $\hat{d}_{A\to1}$ which is encoded by the ranging error model, i.e. the observed distance $\hat{d}_{A\to1}$ contains a ranging error with a certain distribution. Subsequently, message multiplication is performed.

Iterative Message Update
It can be seen from Fig. 4.2 that the factor graph contains cycles (e.g., $\eta_1 \to f_5 \to \eta_2 \to$

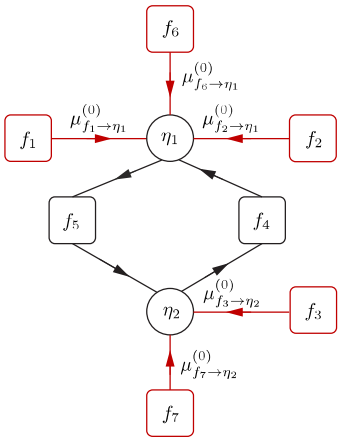

Figure 4.3: *Transmitted messages during the initialization process* - leaf nodes and transmitted messages (see also Tab. 4.1) are marked in red.

$f_4 \to \eta_1$). Hence the exact marginals cannot be determined using belief propagation, but rather approximate marginals are determined in an iterative fashion. Every iteration consists of message multiplication at all variable nodes and subsequent message filtering at all factor nodes. This procedure is repeated N_{BP} times. In the context of the running example, the variable nodes η_1 and η_2 determine the outgoing messages in the l^{th} iteration as the product of the incoming messages (message multiplication). Message flow is indicated by black arrows in Fig. 4.2. For clarification, the incoming messages and the outgoing message are highlighted in the factor graph in Fig. 4.4. Observe that directed messages are assumed, so that the variable node η_1 computes only a single outgoing message $\mu^{(l)}_{\eta_1 \to f_5}(\eta_1)$ to the factor node f_5. The directionality is a peculiarity of cooperative localization and was originally introduced in [WLW09] as a means to reduce the amount of intrinsic information passed along the edges of the graph. The outgoing message of η_1 is given by

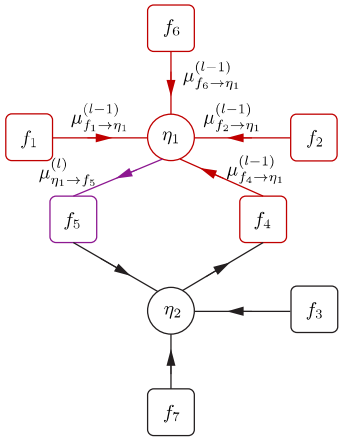

Figure 4.4: *Message multiplication in the variable node η_1* - incoming messages are marked in red, while the outgoing message is marked in magenta.

$$\mu^{(l)}_{\eta_1 \to f_5}(\eta_1) = \prod_{j \in \mathcal{S}_{\to 1}} \mu^{(l-1)}_{f_j \to \eta_1}(\eta_1), \tag{4.13}$$

where $\mathcal{S}_{\to 1} = \{1, 2, 4, 5, 6\}$. Note that if no message has been received on an edge (*here*: the message $\mu^{(l-1)}_{f_4 \to \eta_1}(\eta_1)$ is missing), a uniform distribution over the entire sample space is assumed.

It is worthwhile to spare a thought on the interpretation of the outgoing message $\mu^{(l)}_{\eta_1 \to f_5}(\eta_1)$ in (4.13). The outgoing message is computed based on the product of all incoming messages $\mu^{(l-1)}_{f_j \to \eta_1}(\eta_1), \forall j \in \mathcal{S}_{\to 1}$. Each of the incoming messages $\mu^{(l-1)}_{f_j \to \eta_1}(\eta_1)$ can be seen as the opinion of a distinct expert on the position of agent 1. For instance, $\mu^{(l-1)}_{f_1 \to \eta_1}(\eta_1)$ is the opinion of anchor A regarding η_1 based on the observed distance $\hat{d}_{A \to 1}$. Since the product of all incoming messages is taken in (4.13), the information from all experts is combined, considering their uncertainty.

After the outgoing messages from the variable nodes η_1 and η_2 are determined, message filtering is performed at the factor nodes f_4 and f_5 (see Fig. 4.5), e.g.,

$$\mu^{(l)}_{f_5\to\eta_2}(\eta_2) \propto \int f_5(\eta_1,\eta_2)\cdot\mu^{(l)}_{\eta_1\to f_5}(\eta_1)\mathrm{d}\eta_1 = \int f(\hat{d}_{1\to2}|\eta_1,\eta_2)\cdot\mu^{(l)}_{\eta_1\to f_5}(\eta_1)\mathrm{d}\eta_1, \quad (4.14)$$

where $\mu^{(l)}_{\eta_1\to f_5}(\eta_1)$ is computed according to (4.13). It is, again, worthwhile to consider the interpretation of this operation. Intuitively, the operation in (4.14) generates an expert opinion in the previously elaborated sense. The message $\mu^{(l)}_{f_5\to\eta_2}(\eta_2)$ contains information on the position of agent 2 considering the observed distance estimate $\hat{d}_{1\to2}$ *as well as* the position uncertainty of agent 1 which is encoded by the incoming message $\mu^{(l)}_{\eta_1\to f_5}(\eta_1)$. In the extreme case that agent 1 has virtually no uncertainty, $\mu^{(l)}_{\eta_1\to f_5}(\eta_1) \approx \delta(\eta_1-\eta_1^*)$ and hence the integral in (4.14) is approximately given by $\mu^{(l)}_{f_5\to\eta_2}(\eta_2) \approx f(\hat{d}_{1\to2}|\eta_1^*,\eta_2)$. Now assume that agent 1 has more position uncertainty in the sense that two positions $\eta_1^\dagger$ and η_1' are possible with equal likelihood. Then, the outgoing message would be $\mu^{(l)}_{f_5\to\eta_2}(\eta_2) \approx 1/2 f(\hat{d}_{1\to2}|\eta_1^\dagger,\eta_2) + 1/2 f(\hat{d}_{1\to2}|\eta_1',\eta_2)$. Extending this idea by adding more and more candidate positions regarding η_1, the solution to the integral in (4.14) is obtained. Note, though, that the likelihood of all candidate solutions η_1 is generally not equal, but it is quantified by $\mu^{(l)}_{\eta_1\to f_5}(\eta_1)$. Irrespective of the actual uncertainty on the position of agent 1, the outgoing message $\mu^{(l)}_{f_5\to\eta_2}(\eta_2)$ to agent 2 can always be interpreted as the expert opinion of node 1 on the position of agent 2 by taking the observed distance $\hat{d}_{1\to2}$ into consideration. However, the information on the position of agent 2 that is carried by the opinion of the expert depends on how well the position of agent 1 is known.

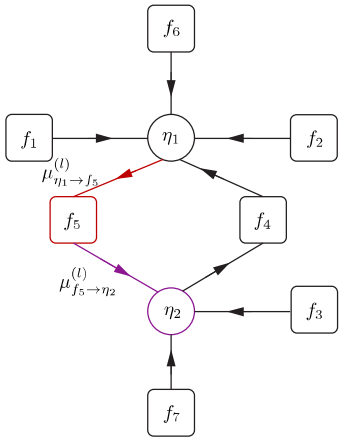

Figure 4.5: *Message filtering in the factor node* f_5 - the incoming message is marked in red, while the outgoing message is marked in magenta.

The l^{th} iteration is concluded by computing the beliefs. Assuming the directionality described above, the belief in the l^{th} iteration is given by

$$b_{\eta_1}^{(l)}(\eta_1) \triangleq \prod_{j \in \mathcal{S}_{\to 1}} \mu_{f_j \to \eta_1}^{(l)}(\eta_1). \tag{4.15}$$

In a general setting with non-directive edges, the belief is defined as the product of all incoming messages

$$b_{\eta_1}^{(l)}(\eta_1) \propto \prod_{j \in \mathcal{S}_{\to 1}} \mu_{f_j \to \eta_1}^{(l)}(\eta_1). \tag{4.16}$$

Remark: Note that, in general, the integral in (4.14) cannot be solved in a closed form. To perform belief propagation for continuous random variables, the filtering and multiplication operation must be adjusted such that (i) the integrals appearing in the filtering process in (4.14) and (ii) the products of densities as in (4.13) and (4.15) can be solved with reasonable computational effort. In the following, three variants of belief propagation will be discussed; discretized, parametrized, and sample-based belief propagation. These variants approximate all densities in belief propagation by discretized, parametrized, and particle-based representations, respectively. In the following subsections, the adjustments to the filtering and multiplication operation are presented if the densities are represented in the aforementioned ways. For clarity of the explanation, the densities in (4.13) and (4.14) are considered.

4.3.1 Discretized Belief Propagation

In this variant of belief propagation, all messages are discretized using N_s quantization points as described in Section (3.1.1.1). The quantization points are chosen such that they are identical for all messages, i.e. $\mathcal{H} = \{\eta^{(1)}, \ldots, \eta^{(N_s)}\}$. In a practical implementation, these points must be known to all agents in the network (here: agent 1 and 2). Instead of the continuous messages (densities), only the weights of the quantization points of the messages are passed along the edges of the factor graph.

Message Multiplication

The multiplication of messages becomes a trivial operation, since only the weights of the respective messages are directly multiplied. The procedure is visualized in the context of the extended running example and it is depicted in Fig. 4.6. Mathematically, the unnormalized weights $\tilde{f}_{\mu_{\eta_1 \to f_5}}^{(k,l)}, \forall k$ of the product of messages are given by

$$\tilde{f}_{\mu_{\eta_1 \to f_5}}^{(k,l)} = \prod_{j \in \mathcal{S}_{\to 1}} \mu_{f_j \to \eta_1}^{(l-1)}(\eta_1 = \eta^{(k)}) = \prod_{j \in \mathcal{S}_{\to 1}} f_{\mu_{f_j \to \eta_1}}^{(k,l-1)}, \tag{4.17}$$

where $f_{\mu_{f_j \to \eta_1}}^{(k,l-1)}, \forall k$ are the normalized weights of the incoming message $\mu_{f_j \to \eta_1}^{(l-1)}(\eta_1 = \eta^{(k)})$. Consequently, the message to any factor node is given by a set of weights and the quantization points. An illustration is depicted in Fig. 4.6. Recall the discussion on the interpretation of the messages from the previous section. It was argued that every incoming message $\mu_{f_j \to \eta_1}^{(l-1)}(\eta_1)$ can be seen as an opinion of different experts on the position of agent 1. The uncertainties of the experts are given by the shapes and the spreads of the

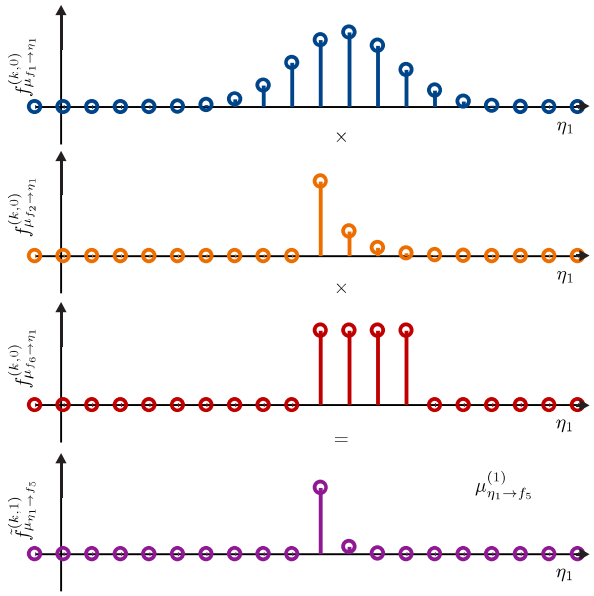

Figure 4.6: *Discretized message multiplication* - the message $\mu^{(1)}_{\eta_1\to f_5}$ from variable node η_1 to factor node f_5 in the first iteration using discretized belief propagation is depicted.

densities in the respective iterations. An example is depicted in Fig. 4.6. Combining the opinions of all experts by multiplication yields the belief on the position of agent 1.

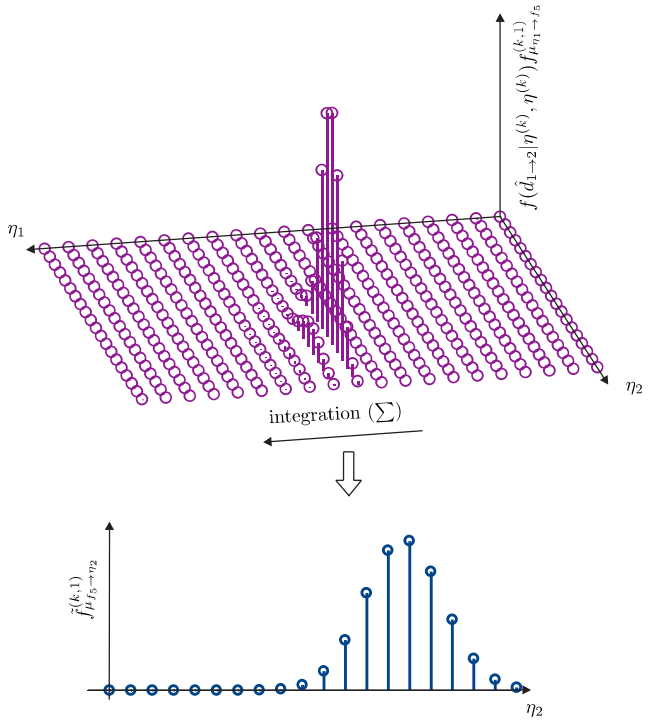

Figure 4.7: *Discretized message filtering* - the message $\mu^{(1)}_{f_5\to\eta_2}$ from factor node f_5 to the variable node η_2 in the first iteration using discretized belief propagation is shown.

Message Filtering

Since the incoming message is given by a set of weights and quantization points, the integral in (4.14) can be solved by means of numerical integration such as the rectangle or trapezoidal rule of integration [DR07]. An illustration in the context of the extended running example is depicted in Fig. 4.7. Using the rectangle rule, the integral in (4.14) simplifies to

$$\mu^{(l)}_{f_5\to\eta_2}(\eta_2) = \Delta\eta \sum_{k=1}^{N_s} f(\hat{d}_{1\to 2}|\eta_1 = \eta^{(k)}, \eta_2) f^{(k,l)}_{\mu_{\eta_1\to f_5}}, \tag{4.18}$$

where $\Delta\eta$ is the Euclidean distance between two quantization points and $f^{(k,l)}_{\mu_{\eta_1\to f_5}} = \tilde{f}^{(k,l)}_{\mu_{\eta_1\to f_5}} / \sum_j \tilde{f}^{(j,l)}_{\mu_{\eta_1\to f_5}}$. The unnormalized weights corresponding to $\mu^{(l)}_{f_5\to\eta_2}(\eta_2)$ are obtained by evaluating (4.18)

in the quantization points $\eta_2 = \eta^{(k)}, \forall k$, i.e.

$$\tilde{f}^{(k,l)}_{\mu_{f_5\to\eta_2}} = \mu^{(l)}_{f_5\to\eta_2}(\eta_2 = \eta^{(k)}) = \Delta\eta \sum_{k=1}^{N_s} f(\hat{d}_{1\to 2}|\eta^{(k)}, \eta^{(k)}) f^{(k,l)}_{\mu_{\eta_1\to f_5}}. \tag{4.19}$$

Observe that, in principle, the quantization points $\eta^{(k)} \in \mathcal{H}$ of η_1 and η_2 can be chosen differently. However, for simplicity, it is assumed that they are identical.
Recalling the discussion on the interpretation of filtered messages and considering Fig. 4.18, it can be seen how integrating over all possible positions of η_1 yields the message which is the expert opinion of agent 1 on the position of agent 2.

It is worthwhile to mention that the complexity of discretized belief propagation is dominated by the filtering operation in (4.19). This operation scales linearly with the number of quantization points and exponential with the dimension of the random variables to be integrated, i.e. $\mathcal{O}(N_s^{d+1})$, where d is the dimension of the random variable. In this 1D positioning example, $d = 1$. However, for two-dimensional (2D) and three-dimensional (3D) positioning, message filtering has cubic $\mathcal{O}(N_s^3)$ and polynomial complexity $\mathcal{O}(N_s^4)$, respectively. It can be summarized that, in discretized belief propagation, message multiplication is straightforward, while message filtering shows prohibitive complexity for 2D or 3D positioning applications.

4.3.2 Parametrized Belief Propagation

In parametrized belief propagation (PBP), messages and beliefs are approximated by members of the family of parametrized distributions such as Gaussian distributions. The goal is to obtain a more compact representation when compared to the discretized approach. The representation is more compact. because parametrized densities can generally be represented by small sets of parameters. Moreover, parametrized belief propagation also tackles the problem of computational complexity by making the computations in belief propagation more tractable. On the other hand, many messages in belief propagation cannot be described arbitrarily closely by such simple distributions like the Gaussian distribution and approximation errors cannot be avoided. Hence the estimation performance is often degraded compared to exact message computation.

Most of the required elements of parametrized message computation have already been discussed in Section 3.2.2.3. The general approach of parametrized belief propagation is to determine the optimum sets of parameters for every message such that all true (exact) messages are approximated as closely as possible. In order to pursue this strategy. two decisions must be taken:

1. The family $\mathfrak{H}$ of distributions must be chosen.

2. The divergence measure between the true (exact) message μ and the approximated message $\hat{\mu}$ must be chosen.

Both decisions must be taken in a way that the resulting optimization problems become tractable while the approximation errors are kept relatively small. Message approximation of most parametrized belief propagation schemes can be expressed by the following unifying approach.

$$\hat{\mu} = \arg\min_{h\in\mathfrak{H}} D(\mu||h), \tag{4.20}$$

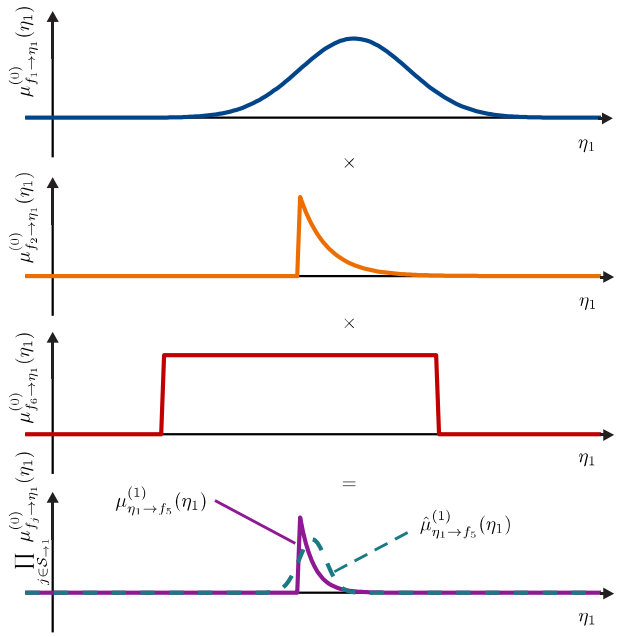

Figure 4.8: *Parametrized message multiplication* - the exact message $\mu^{(1)}_{\eta_1\to f_5}(\eta_1)$ (solid magenta line) from variable node η_1 to factor node f_5 in the first iteration is approximated by a Gaussian density $\hat{\mu}^{(1)}_{\eta_1\to f_5}(\eta_1)$ (dashed green line) whose parameters are optimally (in the sense of the minimum Kullback-Leibler divergence) chosen. It can be observed that the approximating distribution $\hat{\mu}^{(1)}_{\eta_1\to f_5}(\eta_1)$ does not represent the exact distribution $\mu^{(1)}_{\eta_1\to f_5}(\eta_1)$ closely, since, e.g., it assigns probability mass in regions where the true distribution does not have any mass.

where $\mathfrak{H}$ and $D(\cdot||\cdot)$ are the chosen family of distributions and divergence measure, respectively. Prominent examples from the literature choose the Kullback-Leibler divergence $D_{\mathrm{KL}}(\mu||h)$ as a measure of dissimilarity of the distributions μ and h [Min05, VWS12, MSW13]. The family $\mathfrak{H}$ of distributions must be chosen such that the properties (like the shape or number of modes) of the exact messages can be well represented. This is generally a hard task, since the shape of the exact messages depends strongly on the problem at hand. Hence the family of distributions is often tailored to individual problems and not applicable to other cases.

In the following, the two major operations in belief propagation are discussed when parametrized message approximation is applied. For that purpose, the running example is considered again and, for simplicity, it is assumed that all messages are approximated by members of the Gaussian family of distributions. In that case, all messages can simply be represented by their means and variances. Note that, in general, Gaussian distributions are not appropriate for general applications and more complex densities are required

which make the optimization problems more difficult than those described in the following.

Message Multiplication
For consistency, the product of messages from (4.13) is considered. An illustration of the parametric multiplication procedure is depicted in Fig. 4.8. Note that if the Kullback-Leibler divergence $D_{\mathrm{KL}}(\cdot||\cdot)$ is used and the Gaussian family of distributions is considered, the optimum parameters $m^{(l)}_{\eta_1 \to f_5}$ and $(\sigma^{(l)}_{\eta_1 \to f_5})^2$ are simply given by the mean and variance of the product of messages, i.e.

$$m^{(l)}_{\eta_1 \to f_5} = \int \eta_1 \cdot \prod_{j \in \mathcal{S}_{\to 1}} \mu^{(l-1)}_{f_j \to \eta_1}(\eta_1) \mathrm{d}\eta_1, \tag{4.21}$$

$$(\sigma^{(l)}_{\eta_1 \to f_5})^2 = \int (\eta_1 - m^{(l)}_{\eta_1 \to f_5})^2 \cdot \prod_{j \in \mathcal{S}_{\to 1}} \mu^{(l-1)}_{f_j \to \eta_1}(\eta_1) \mathrm{d}\eta_1. \tag{4.22}$$

Observe that the general optimization problem in (4.20) breaks down to solving the two integrals in (4.21) and (4.22), respectively. Only in very few special cases, closed-form solutions of the integrals exist. Generally, numerical integration is required.

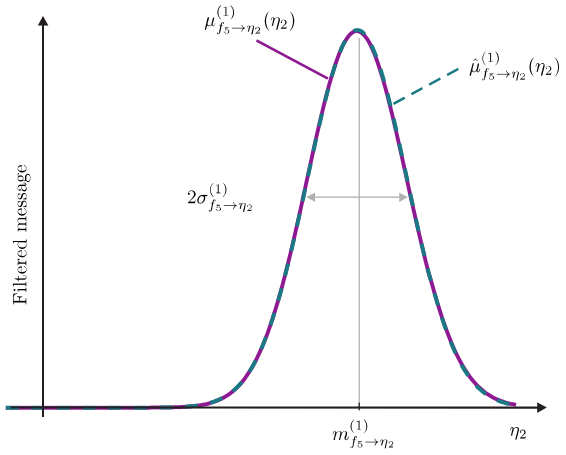

Figure 4.9: *Parametrized message filtering* - the exact message $\mu^{(1)}_{f_5 \to \eta_2}(\eta_2)$ (solid magenta line) and the approximated message $\hat{\mu}^{(1)}_{f_5 \to \eta_2}(\eta_2)$ (dashed green line) from factor node f_5 to the variable node η_2 in the first iteration are depicted. It can be observed that the approximation error is relatively small when the exact distribution has a shape that can be well represented by a member of the chosen family $\mathfrak{H}$ of distributions.

Message Filtering
Recall the filtering operation in (4.14). The exact filtered message $\mu^{(l)}_{f_5 \to \eta_2}(\eta_2)$ is given by

the integral over the product of the incoming message $\mu^{(l)}_{\eta_1 \to f_5}(\eta_1)$ and the function of the factor node $f_5(\eta_1, \eta_2) = f(\hat{d}_{1\to 2}|\eta_1, \eta_2)$. Suppose that, similar as in message multiplication, the filtered message $\mu^{(l)}_{f_5 \to \eta_2}(\eta_2)$ is approximated by a member of the Gaussian family of distributions and the Kullback-Leibler divergence is chosen as divergence measure. In that case, the parameters of the approximated filtered message are given by

$$m^{(l)}_{f_5 \to \eta_2} = \int \eta_2 \cdot \int f(\hat{d}_{1\to 2}|\eta_1, \eta_2) \cdot \mu^{(l)}_{\eta_1 \to f_5}(\eta_1) \mathrm{d}\eta_1 \mathrm{d}\eta_2, \tag{4.23}$$

$$(\sigma^{(l)}_{f_5 \to \eta_2})^2 = \int (\eta_2 - m^{(l)}_{f_5 \to \eta_2})^2 \cdot \int f(\hat{d}_{1\to 2}|\eta_1, \eta_2) \cdot \mu^{(l)}_{\eta_1 \to f_5}(\eta_1) \mathrm{d}\eta_1 \mathrm{d}\eta_2. \tag{4.24}$$

An illustration in the context of the extended running example is depicted in Fig. 4.9. Note that the messages in Fig. 4.8 and 4.9 have the same interpretation as discussed in the previous subsection.

4.3.3 Sample-based Belief Propagation

In contrast to discretized and parametrized belief propagation, all messages in sample-based belief propagation (SBP) are represented by particle representations. The goal of the particle representations is to describe the exact messages as accurately as possible with a finite number of particles. As discussed in Section 3.1.3, the messages in SBP are relatively comprehensive compared to parametrized messages. Yet, they are generally more compact than messages in discretized belief propagation. Recall that a particle representation consists of a set of samples with associated weights, where the samples are drawn from a proposal distribution. As discussed in Section 3.2.1.1, the choice of the proposal distribution has a crucial impact on how well a distribution can be approximated. This aspect will be stressed throughout the entire section as a motivation for the method that will be presented in the next chapter.

In contrast to discretized belief propagation, where message multiplication is a computationally simple operation, the multiplication of messages in SBP can embody the computational bottleneck of this approach. On the other hand, message filtering becomes lightweight in terms of computations. Both operations (message multiplications and filtering) will be described in the following paragraphs.

Message Multiplication

Since all messages in SBP are given as particle representation, straightforward multiplication is not possible, as discussed in Section 3.2.2.2. For consistency, the running example is considered. Here, the product in (4.13) can be written as the product of particle representations $\mathcal{R}_{Ns}(\cdot)$, i.e.

$$\mu^{(l)}_{\eta_1 \to f_5}(\eta_1) = \prod_{j \in \mathcal{S}_{\to 1}} \mathcal{R}_{Ns}\left(\mu^{(l-1)}_{f_j \to \eta_1}(\eta_1)\right). \tag{4.25}$$

To enable multiplication, continuous kernel density estimates $\dot{\mu}^{(l-1)}_{f_j \to \eta_1}(\eta_1)$ of the discrete particle-based messages $\mathcal{R}_{Ns}\left(\mu^{(l-1)}_{f_j \to \eta_1}(\eta_1)\right)$ are obtained. In this thesis only Gaussian kernels are considered. If Gaussian kernels are used, all continuous kernel density estimates

$\hat{\mu}^{(l-1)}_{f_j\to\eta_1}(\eta_1)$ are Gaussian mixture densities (short: *Gaussian mixtures*) with N_s components. Note that the product of $|\mathcal{S}_{\to 1}|$ Gaussian mixtures with N_s components is again a Gaussian mixture. However, the number of components increases exponentially with the number of messages, i.e. $\mathcal{O}(N_s^{|\mathcal{S}_{\to 1}|})$ which makes exact multiplication intractable. In order to circumvent the issue of prohibitive complexity, importance sampling is used, as described in Section 3.2.1.1. In particular, the exact product of Gaussian mixtures $\prod_{j\in\mathcal{S}_{\to 1}}\hat{\mu}^{(l-1)}_{f_j\to\eta_1}(\eta_1)$ is approximated by a particle representation $\mathcal{R}_{N_s}\left(\mu^{(l)}_{\eta_1\to f_5}(\eta_1)\right)$, whose samples are drawn from the proposal distribution $q_1(\eta_1)$. An illustration of this process is depicted in Fig. 4.10. In particular, multiplication of particle representations requires the followings steps.

1. Draw N_s samples[2] $\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}$ from a suitable proposal distribution $q_1(\eta_1)$.

2. Determine the unnormalized weight $\tilde{w}^{(k,l)}_{\mu_{\eta_1\to f_5}}$ of every sample based on the kernel density estimates (KDEs) of the incoming messages and the proposal distribution.

3. Normalize all weights $\tilde{w}^{(k,l)}_{\mu_{\eta_1\to f_5}}, \forall k$.

The second and third steps are elaborated in more detail in the following. After N_s samples $\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}$ are drawn. the weights of the particles have to be determined. In particular, the ratio of the product of the kernel density estimates of the messages $\prod_{j\in\mathcal{S}_{\to 1}}\hat{\mu}^{(l-1)}_{f_j\to\eta_1}(\eta_1)$ and the proposal distribution $q_1(\eta_1)$ is evaluated in every sample $\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}$, i.e.

$$\tilde{w}^{(k,l)}_{\mu_{\eta_1\to f_5}} \propto \frac{\prod_{j\in\mathcal{S}_{\to 1}}\hat{\mu}^{(l-1)}_{f_j\to\eta_1}(\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}})}{q_1(\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}})}. \tag{4.26}$$

Note that the complexity of adjusting the weights scales with $\mathcal{O}(|\mathcal{S}_{\to 1}|N_s^2)$ which makes SBP a complex algorithm if N_s is large. The unnormalized weights are then normalized as follows:

$$w^{(k,l)}_{\mu_{\eta_1\to f_5}} = \frac{\tilde{w}^{(k,l)}_{\mu_{\eta_1\to f_5}}}{\sum_j \tilde{w}^{(j,l)}_{\mu_{\eta_1\to f_5}}}. \tag{4.27}$$

Finally the particle representation of the product of messages is given by

$$\mathcal{R}_{N_s}\left(\mu^{(l)}_{\eta_1\to f_5}(\eta_1)\right) = \left\{w^{(k,l)}_{\mu_{\eta_1\to f_5}}, \eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}\right\}_{k=1}^{N_s} \tag{4.28}$$

Note that the particle representation of the product of messages is often resampled as described in Section 3.2.1.2 in order to focus on the relevant particles.

Remark: It is worthwhile to recall that the proposal distribution $q_1(\eta_1)$ should be chosen such that it resembles the product of messages $\prod_{j\in\mathcal{S}_{\to 1}}\hat{\mu}^{(l-1)}_{f_j\to\eta_1}(\eta_1)$ closely. In that way, most particles will have a significant weight and only a small number of particles is sufficient in order to represent the product of messages accurately. Unfortunately, the product $\prod_{j\in\mathcal{S}_{\to 1}}\hat{\mu}^{(l-1)}_{f_j\to\eta_1}(\eta_1)$ cannot be known before it is actually computed, which makes

[2]Note that an for an unambiguous notation, samples require subscripts to indicate from which message they originate.

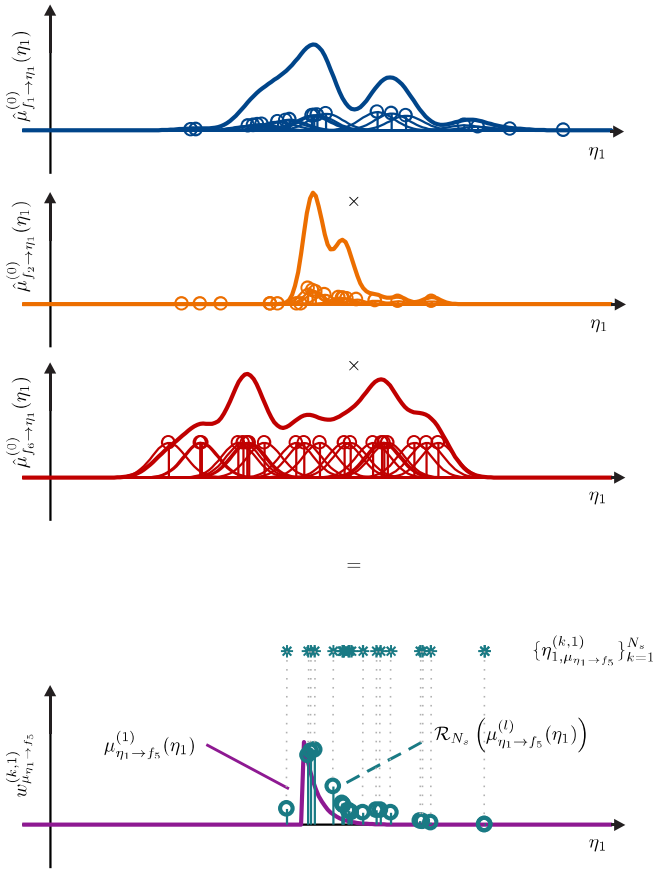

Figure 4.10: *Sample-based message multiplication* - the product of the incoming messages, which are given as particle representations, is determined by applying importance sampling to the product of kernel density estimates $\hat{\mu}^{(0)}_{f_j\to\eta_1}(\eta_1), \forall j$. First, N_s samples $\eta^{(k,1)}_{1,\mu_{\eta_1\to f_5}}$ (green asterisk) are drawn from the proposal distribution $q_1(\eta_1)$ (distribution $q_1(\eta_1)$ not shown). The corresponding weights $w^{(k,1)}_{\mu_{\eta_1\to f_5}}$ are determined subsequently by evaluating the quotient in (4.26) and normalizing afterwards.

the choice of $q_1(\eta_1)$ a *chicken or the egg* problem. In Section 5.4, a method called *polygon outer-approximation* (POA) is presented that allows for rigorous choices of the proposal distributions by confining the sets of possible locations of agents to subsets of the entire sample space.

Message Filtering

In contrast to discretized belief propagation, the filtering operation that is performed at

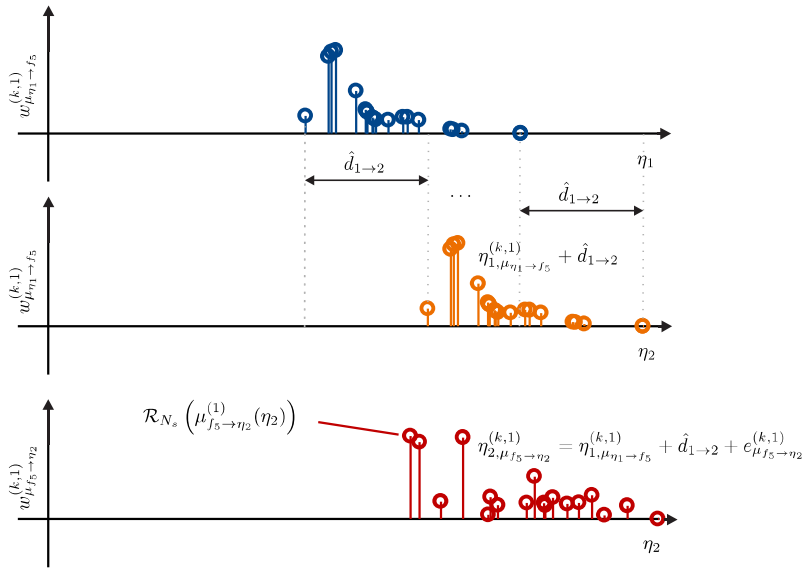

Figure 4.11: *Sample-based message filtering* - the outgoing particle representation $\mathcal{R}_{N_s}\left(\mu^{(1)}_{f_5\to\eta_2}(\eta_2)\right) = \{w^{(k,1)}_{\mu_{f_5\to\eta_2}}, \eta^{(k,1)}_{2,\mu_{f_5\to\eta_2}}\}_{k=1}^{N_s}$ is determined based on the incoming particle representation $\mathcal{R}_{N_s}\left(\mu^{(1)}_{\eta_1\to f_5}(\eta_1)\right) = \{w^{(k,1)}_{\mu_{\eta_1\to f_5}}, \eta^{(k,1)}_{1,\mu_{\eta_1\to f_5}}\}_{k=1}^{N_s}$ and the density $f(\hat{d}_{1\to 2}|\eta_1,\eta_2)$. In this context, all samples $\eta^{(k,1)}_{1,\mu_{\eta_1\to f_5}}, \forall k$ of the incoming message are first shifted by the distance estimate $\hat{d}_{1\to 2}$ (middle). Then, a random noise term $e^{(k,1)}_{\mu_{f_5\to\eta_2}}$ is added to each sample which is distributed according to the measurement noise (bottom). The weights of the outgoing message are given by the weights of the incoming message, i.e. $w^{(k,1)}_{\mu_{f_5\to\eta_2}} = w^{(k,1)}_{\mu_{\eta_1\to f_5}}$. In other words, the particles in the graph at the bottom are obtained by shifting the particles in the graph in the middle randomly according to $e^{(k,1)}_{\mu_{f_5\to\eta_2}}$.

factor nodes can be simple if the incoming and outgoing messages are given as particle representations. In particular, the message in (4.14) can be rewritten as follows if the incoming message $\mu^{(l)}_{\eta_1\to f_5}(\eta_1)$ is given by the particle representation $\mathcal{R}_{N_s}\left(\mu^{(l)}_{\eta_1\to f_5}(\eta_1)\right) = \{w^{(k,l)}_{\mu_{\eta_1\to f_5}}, \eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}\}_{k=1}^{N_s}$:

$$\begin{aligned}\mu^{(l)}_{f_5\to\eta_2}(\eta_2) &\propto \int f_5(\eta_1,\eta_2)\cdot\sum_{k=1}^{N_s} w^{(k,l)}_{\mu_{\eta_1\to f_5}}\delta(\eta_1-\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}})\mathrm{d}\eta_1 \\ &= \sum_{k=1}^{N_s} w^{(k,l)}_{\mu_{\eta_1\to f_5}} \underbrace{f(\hat{d}_{1\to 2}|\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}},\eta_2)}_{f_5(\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}},\eta_2)},\end{aligned} \tag{4.29}$$

where $\delta(\cdot)$ is the Dirac impulse. In that case, a particle representation $\mathcal{R}_{N_s}\left(\mu^{(l)}_{f_5\to\eta_2}(\eta_2)\right)$ of the outgoing message can be obtained by direct sampling, i.e., for every sample $\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}$, a sample $\eta^{(k,l)}_{2,\mu_{f_5\to\eta_2}}$ is drawn from the density $f(\hat{d}_{1\to 2}|\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}},\eta_2)$ and the weight is directly copied, i.e. $\eta^{(k,l)}_{2,\mu_{f_5\to\eta_2}} \sim f(\hat{d}_{1\to 2}|\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}},\eta_2)$ and $w^{(k,l)}_{\mu_{f_5\to\eta_2}} = w^{(k,l)}_{\mu_{\eta_1\to f_5}}$.

This procedure is visualized in the context of the running example in Fig. 4.11. Here, sampling from $f(\hat{d}_{1\to 2}|\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}},\eta_2)$ simplifies to shifting the incoming sample $\eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}}$ by[3] $\hat{d}_{1\to 2}+e^{(k,l)}_{\mu_{f_5\to\eta_2}}$, where $e^{(k,l)}_{\mu_{f_5\to\eta_2}}$ is a random sample which is distributed according to the measurement noise. Hence an outgoing sample $\eta^{(k,l)}_{2,\mu_{f_5\to\eta_2}}$ is given by

$$\eta^{(k,l)}_{2,\mu_{f_5\to\eta_2}} = \eta^{(k,l)}_{1,\mu_{\eta_1\to f_5}} + \hat{d}_{1\to 2} + e^{(k,l)}_{\mu_{f_5\to\eta_2}}, \tag{4.30}$$

where $e^{(k,l)}_{\mu_{f_5\to\eta_2}}$ is distributed according to the measurement error of $\hat{d}_{1\to 2}$. Loosely speaking, an incoming sample is shifted by the observed distance estimate $\hat{d}_{1\to 2}$ plus a random error term $e^{(k,l)}_{\mu_{f_5\to\eta_2}}$ that simulates the measurement noise.

Remark: It is worth noting that there are also other variants of sample-based belief propagation than the one that was presented in this subsection. The presented variant performs well (in terms of the estimation accuracy) if no prior information regarding the estimation parameters is available. In the presence of prior information, an alternative variant of sample-based belief propagation can be employed to perform inference. This variant will be explained in Section 7.3.3.2.

4.4 Summary

In Bayesian inference, the goal is to determine an estimate of a set of parameters from a collection of observations (measurements) based on the stochastic models of the observations and the prior information on the parameters. If the set of parameters is large, the joint posterior distribution is high-dimensional which makes it difficult to obtain optimum estimates (e.g., MMSE or MAP). Thus, the lower-dimensional marginal posteriors are often considered in order to estimate the parameters. A popular method of determining the marginal posteriors in a computationally efficient way is belief propagation which is a message passing algorithm. In belief propagation, messages are passed along the edges of an underlying factor graph. Messages are updated using two main operations; message filtering and message multiplication. Performing these operations for continuous parameters is often difficult, since they involve solving integrals that do not have closed-form solutions. In order to deal with these integrals, different variants of belief propagation have been discussed. Particularly, discretized, parametrized, and sample-based belief propagation have been considered. Note that this list is by no mean exhaustive.

Discretized belief propagation determines quantized versions of the exact messages. If the number of quantization points is large, the marginal posteriors are estimated accurately. However, discretized belief propagation suffers from a high complexity and is only suited for small estimation problems.

[3] Note that generally a sample is shifted by $\pm\hat{d}_{1\to 2}+e^{(k,l)}_{\mu_{f_5\to\eta_2}}$, since agent 2 could be left or right of agent 1. For simplicity, it was assumed earlier that nodes know whether they are left or right of each other.

In parametrized belief propagation, all messages are approximated by members from a chosen family of distributions (like the family of Gaussian distributions). A message is determined by solving the optimization problem which finds the best set of parameters such that the divergence between the true and approximated messages is minimized. The divergence is typically measured by the Kullback-Leibler divergence, but generally any divergence measure (e.g., the Hellinger distance) can be chosen. Parametrized belief propagation suffers from the fact that solving the optimization problems is only tractable for a relatively small number of families and divergence measures. Due to these restrictions, the approximated messages can deviate significantly from the exact messages which, in turn, degrades the estimation accuracy.
Messages in sample-based belief propagation are given by particle representations, i.e. samples with associated weights. A particle can be interpreted as a hypothesis of a parameter value whose probability is given by the weight. Drawing these samples is usually difficult, because it is not known a priori where a message will have significant probability mass. Hence many samples need to be drawn in order to assure that some of the samples reside in the important regions of the sample space. The challenge of sampling will be addressed in Section 5.4.

Chapter 5

Cooperative Positioning

In most radio systems with localization capabilities such as global navigation satellite systems (GNSSs), current Long Term Evolution (LTE) systems or wireless fidelity (WiFi) systems, *self-localization*[1] of agents is based on the reception of radio signals from a set of anchor nodes with known locations. Cooperative (sometimes also called *collaborative*) localization[PAK+05, WLW09, ZB11], however, also leverages direct inter-agent communications to support localization. The main idea of cooperative localization is that agents help each other determining their positions. In that way, cooperative localization can provide improvements both in terms of coverage (agents at the edge of the coverage area can act as anchors) and accuracy (inter-agent communications increase the number of available measurements) compared to non-cooperative systems [BWV18].

Despite the potential advantages of cooperative localization approaches, implementations of cooperative approaches are not yet adopted by state-of-the-art communication or localization systems. In fact, not one of the aforementioned systems has implemented cooperative localization strategies, despite the fact that some systems already provide the capabilities for cooperation. LTE is a good example. Since LTE release 14, device-to-device (D2D) communication modes are included in the standard which allow for a straightforward implementation of cooperative localization schemes. Although many theoretical studies have proven the power of cooperative approaches, only very few practical deployments exist. One of the most compelling reasons is that many cooperative localization algorithms still suffer from exhaustive complexity burdens. This is especially true in scenarios with very sparse infrastructure, since the position uncertainties of agents are particularly high and powerful algorithms are necessary to account for those uncertainties. A major goal of the work in this thesis is to develop cooperative localization methods that explicitly tackle the complexity issue of state-of-the-art cooperative localization algorithms.

This chapter starts with a brief background on cooperative localization schemes within which a variety of existing approaches will be discussed and open challenges will be identified. This discussion is followed by the specification of the system model that will be considered throughout the chapter. Section 5.3 will present a running example, in which an exemplary cooperative network is studied. The running example will be used to explain sample-based belief propagation for cooperative localization to point out the difficulties of cooperative localization and to visualize the proposed localization method. It will also be used throughout all sections of this chapter to explain complicated problems. In Section 5.4, the cornerstone of the proposed cooperative localization approach will be presented.

[1]The term *self-localization* refers to the problem of agents attempting to determine their own positions. For brevity, henceforth the terms *positioning* and *localization* will be used in place of *self-positioning*.

This section deals with an essential problem in cooperative localization: *Where to search for the positions of agents?*. As one of the main contributions of the thesis, a novel cooperative algorithm named polygon outer-approximation (POA) will be presented. POA is a distributed and low-complexity algorithm which constrains the locations of agents to convex polygons. With these constraints, cooperative localization is revisited and reformulated as a *constrained* inference problem. The polygonal constraints help reducing the notoriously high computational burden of state-of-the-art cooperative localization algorithms. A combination of POA and sample-based belief propagation – a two-phase localization approach – will be presented in Section 5.5. Finally, an exhaustive numerical analysis of the proposed method will be conducted in Section 5.6 and conclusions will be drawn in Section 5.7.

5.1 Introduction

Cooperative localization is a problem that has been gaining considerable attention for already more than one decade. The benefits of collaboration among agents have been investigated both theoretically and algorithmically. Most theoretical studies are based on the analysis of the Fisher information associated with the corresponding cooperative localization problem. In this domain, different aspects are investigated such as the fundamental nature of cooperation and the scaling with the network size [SWW10, GGM11], power allocation [DSW15, DSW18], anchor placement [JR06], and neighbor selection [DW12]. Such insights are helpful as they often reveal guidelines on how to deploy cooperative networks with exceptional localization capabilities. Even beyond the scope of network design, the fundamental insights from the Fisher information analyses usually assist the development of algorithms. As a consequence, many different algorithms have been presented over the last decade. These algorithms can be broadly classified according to the following metrics: Algorithms are usually distinguished by their type of signal processing. For instance, the nature of signal processing can be deterministic or stochastic. Alternative classifications such as centralized versus distributed processing[2] can be considered. Classifications according to the measurement types are also possible. For example, range-based versus range-free algorithms can be differentiated. In this thesis, range-based and distributed algorithms are considered and the discussed approaches will be distinguished according to their nature of signal processing, i.e. deterministic or stochastic.

5.1.1 State of the Art

The class of deterministic approaches encloses algorithms such as (weighted) least squares [GCW07, SRL02, NJSW15, WCM+11], geometric approaches [GWSR11a, GWSR11b, GWG13], multi-hop methods [NN01, NN03b, Nic04], and many others. Deterministic approaches often struggle to account for the uncertainties regarding the positions of agents. Hence they perform poorly in scenarios with sparse infrastructure, where the position uncertainties are high. Stochastic approaches inherently take the position uncertainties of agents into consideration. Almost all stochastic inference-based approaches can be traced

[2] In distributed processing, agents perform local computations of measurements and share the results with their neighbors, while centralized approaches gather all measurements at a central computation entity that processes the entire data and disseminates the results afterwards.

back to some variants of message passing (belief propagation) algorithms. In the context of localization, messages are probability densities regarding the positions of agents. This fact ensures that the uncertainties of the positions of agents are accounted for since probability distributions may have multiple modes (corresponding to multiple possible positions) and each mode can accommodate a certain spread of the probability mass (corresponding to a confidence level regarding the possible position). In contrast to many other algorithms, message passing-based algorithms often attain the fundamental limits of cooperative localization[SW10a, SWW10]. The range of stochastic algorithms that is considered in this thesis is restricted to the class of belief propagation-based approaches due to their superior localization capabilities in the context of sparse infrastructure.

Existing belief propagation-based approaches can be broadly classified in the set of (i) parametric [LFS+12, EMH+17, EBHS13, VWS12, LWWK14, YWE+16]. (ii) sample-based [SIFW03, IFMW05, SII+10, SZ13, SZ09, LWG+18], and (iii) hybrid[3] [MHW13, ZCW+17] approaches. In parametric belief propagation, the true beliefs and messages are approximated by parametrized distributions as outlined in Section 4.3.2. The parameters are determined by minimizing some divergence metric between the true belief or message and the approximating distribution. The main drawback of PBP is that the parameter determination is only feasible for certain families of distributions [Min01]. Hence complex distributions cannot be represented arbitrarily closely since the class of applicable distributions is relatively small. In [LFS+12], the approximating distributions are members of the exponential family and the Kullback-Leibler divergence is minimized to determine the parameters. In contrast to [LFS+12], [VWS12, EMH+17, EBHS13, LWWK14, YWE+16] consider only a subset of the exponential family of distributions: Gaussian distributions. In particular, the beliefs are approximated by Gaussian distributions which inherently imply unimodality of the beliefs (meaning a single possible location). Unimodality cannot be assumed for networks with sparse infrastructure, unless informative prior information regarding the positions of the agents is available. In most practical scenarios, however, no prior information is available which renders the approaches in [VWS12, EMH+17, EBHS13, LWWK14, YWE+16] infeasible for the problem considered in this thesis.

In sample-based belief propagation, particle representations (sets of samples with associated weights) are used to approximate the true beliefs and messages as described in Section 4.3.3. The samples are drawn from so-called proposal distributions [SIFW03, IFMW05, SII+10, SZ13, SZ09, LWG+18]. In [SIFW03, SII+10], every agent considers an arbitrary incoming message as its proposal distribution. Since all messages are already given by sets of samples and weights, the samples are readily available and do not have to be drawn explicitly (a more detailed explanation will be provided in Section 5.3.2). Other proposal distributions are presented in [SZ13, SZ09]. [SZ13] proposes mixture importance sampling with reference particles. In this approach, the samples are not taken from a single incoming message but from all incoming messages. To enhance robustness, a certain percentage of the samples is sprinkled uniformly over the considered area. Auxiliary sampling, where the proposal distribution is augmented with an auxiliary variable, was also proposed in [SZ13]. The auxiliary variable takes information from anchors into account, resulting in samples that are concentrated in the areas close to the true locations of nodes. However, this approach requires a high anchor density. Boxed importance sampling is presented in [SZ09], where the supports of the proposal distributions are constrained heuristically by rectangles and samples are drawn within these rectangles. A region-sampling approach is

[3]Hybrid approaches employ tools from both the parametric and the nonparametric domain.

proposed in [LWG+18], where samples are drawn from regions that are close to the true locations of nodes. In order to determine these regions, however, the approach requires prior knowledge on the positions of the agents and a map of the environment. Both prior and map information are generally not available to the agents which restricts the applicability of the approach in [LWG+18] to relatively special cases.

Hybrid belief propagation approaches employ both particle representations and parametrized distributions to represent the true distributions which occur in belief propagation. Beliefs are represented by particles in [MHW13], while other messages are approximated by parametrized distributions. A similar approach is chosen in [ZCW+17]. Messages are represented by members of the exponential family, while beliefs are computed based on particle representations.

In summary, the belief propagation approaches in the literature suffer from either (i) high complexity and/or (ii) restricted applicability, since many assumptions are not necessarily satisfied in practical scenarios. The exhaustive computational burden of SBP approaches originates from the fact that the necessary number of samples can be prohibitively high. PBP approaches, on the other hand, usually have relaxed computational requirements but they generally demand prior knowledge on the positions of agents which is often unavailable in practice. Thus, both SBP and PBP can benefit substantially from a method which provides constraints on the search space of possible positions of agents. Such constraints can be interpreted as generated prior information which would render both SBP and PBP more practical from the computational perspective.

5.1.2 Contribution

Despite the effort of previous works, cooperative positioning in scenarios with sparse infrastructure remains to be a hard problem from the computational perspective. Especially for low-cost devices, state-of-the-art cooperative positioning algorithms often turn out to be computationally intractable. The essential finding in this chapter is that constraints on the positions of agents can be leveraged in order to relax the computational burden of cooperative positioning. To that end, a two-phase approach is proposed:

1. First, a low-complexity and distributed algorithm called POA is employed to determine sets of polygonal constraints $\{\mathbf{p}_j | \mathbf{A}_j \mathbf{p}_j \leq \mathbf{b}_j\}$ on the positions of all agents based on the distance estimates $\hat{d}_{i \to j}, \forall j, i \in \mathcal{S}_{\to j,\mathrm{NLOS}}$ of non-line-of-sight (NLOS) and obstructed-line-of-sight (OLOS)[4] links[5].

2. Secondly, these constraints are used to determine proposal distributions $q_j(\mathbf{p}_j), \forall j$ with concentrated probability mass in the vicinity of the true locations of agents. These proposal distributions are harnessed in order to ease the inference problem that is considered for cooperatively estimating the positions of the agents.

[4] In this thesis, the line-of-sight (LOS) condition means that only the direct path exists but no indirect multipath components are present. NLOS refers to the channel condition, where the direct LOS path exists and multiple indirect propagation paths are present as well. The NLOS condition can also include the case, where the LOS path is visibly blocked but the radio wave penetrates a LOS-blocking obstacle. OLOS describes channel conditions where multiple propagation paths are present, but the LOS path is absent (obstructed).

[5] To avoid notational clutter, the term *OLOS* is not mentioned in the subscripts of variables that refer to NLOS and OLOS links.

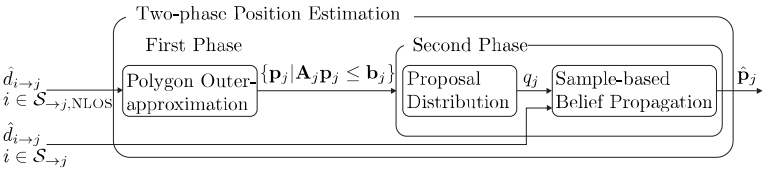

Figure 5.1: *Block diagram of the proposed two-phase position estimation approach* - distance estimates $\hat{d}_{i\to j}, i \in \mathcal{S}_{\to j,\mathrm{NLOS}}$ from NLOS and OLOS links are used to constrain the positions $\mathbf{p}_j, \forall j$ of agents via polygon outer-approximation. These constraints $\{\mathbf{p}_j|\mathbf{A}_j\mathbf{p}_j \leq \mathbf{b}_j\}$ restrict the proposal distributions $q_j, \forall j$ used by sample-based belief propagation.

Fig. 5.1 depicts the high-level overview of the proposed two-phase approach. In the first phase, the distributed and cooperative POA algorithm derives sets of constraints $\{\mathbf{p}_j|\mathbf{A}_j\mathbf{p}_j \leq \mathbf{b}_j\}, \forall j$ on the positions of the agents. These constraints confine the possible location $\mathbf{p}_j$ of every agent j and they are used as prior knowledge in the second phase. In this thesis, sample-based belief propagation is considered during the estimation phase due to its superior positioning accuracy compared to other cooperative position estimators. The constraints are employed in order to determine smart proposal distributions.

It is important to note that the second phase in the proposed two-phase approach does not necessarily have to be SBP. Polygon outer-approximation can be expected to also improve the performance of other cooperative positioning methods such as least-squares or parametric belief propagation. Possible applications are as follows: The constraints imposed by POA in the first phase can be used to initialize the numerical solver of a least-squares estimator or to constrain the range of the parameters in parametrized belief propagation. Since the positioning accuracy of SBP is generally higher than that of other methods, only SBP will be considered during the analysis of the two-phase approach.

The following list contains the main contributions of this chapter:

- A distributed algorithm called polygon outer-approximation is developed that constrains the positions of agents tightly.
- Based on these constraints, a novel type of proposal distribution for sample-based belief propagation is introduced that inherently allows for drawing samples in the vicinity of the true locations of agents.
- Simulations show that increased localization accuracy, quicker convergence, and reduced computation time can be achieved when compared to state-of-the-art cooperative positioning approaches.

5.2 System Model

The problem of cooperative localization is considered in which agents attempt to estimate their own positions in a distributed fashion based on the distance measurements with respect to neighboring anchors *and* agents. First, the network topology will be described. The measurement model will be elaborated subsequently. The chapter is concluded with the definition of the position estimation criterion.

5.2.1 Network Topology

The position of agent j is denoted by $\boldsymbol{\eta}_j \triangleq \mathbf{p_j} \triangleq [p_{\mathrm{x},j}, p_{\mathrm{y},j}]^{\mathrm{T}}$. It is assumed that all nodes are static. All agents can determine range estimates $\hat{d}_{i\to j}$ to all neighboring nodes $i \in \mathcal{S}_{\to j}$ inside the communication range r_{com}. In particular, the index set $\mathcal{S}_{\to j}$ of neighbors of agent j is given by $\mathcal{S}_{\to j} = \{\forall i \neq j \mid \|\mathbf{p}_i - \mathbf{p}_j\|_2 \leq r_{\mathrm{com}}\}$. Anchor coverage is assumed to be sparse, meaning that the number of anchor indices in $\mathcal{S}_{\to j}$ is small and the majority of agents can only range with a single anchor. An exemplary topology with sparse infrastructure is depicted in Fig. 5.2. Anchors and agents are indicated by green diamonds and red circles, respectively. The communication ranges r_{com} of the anchors are visualized by black circles. Note that most agents in Fig. 5.2 are connected to one or two anchors only, making conventional distance-based positioning as described in Section 2.3 infeasible. With the help of cooperation among agents, however, accurate positioning becomes possible as will be shown in Section 5.6.

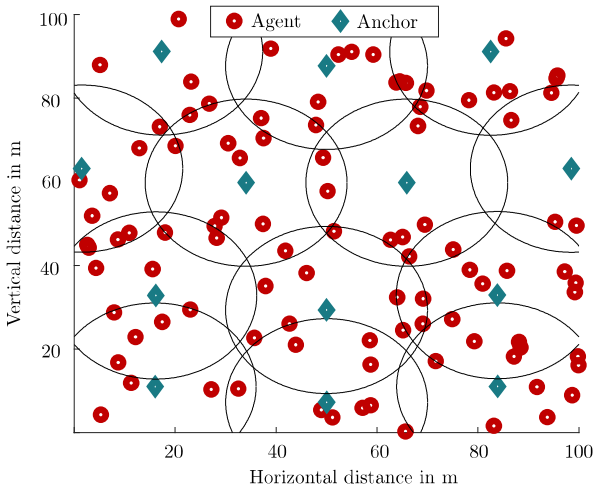

Figure 5.2: *Exemplary topology with sparse anchor coverage (infrastructure)*- anchors and agents are depicted as diamonds and circles, respectively. Black circles indicate the coverage regions of anchors.

5.2.2 Measurement Model

Distance estimates are given by

$$\hat{d}_{i\to j} = \|\mathbf{p}_i - \mathbf{p}_j\| + e_{i\to j}, \tag{5.1}$$

where $\|\mathbf{p}_i - \mathbf{p}_j\|$ is the true Euclidean distance between node i and node j. The ranging error is denoted by $e_{i\to j}$. Distance estimates can be obtained by employing estimators such as the one presented in Section 2.2.1. All distance measurements $\hat{d}_{i\to j}$ are collected in the joint measurement vector $\mathbf{z}$ in an arbitrary order.

The ranging error generally depends on the receiver characteristics, including thermal noise, the distance estimation algorithm, and the environment of transmission. In the indoor environment, NLOS and OLOS propagation are especially critical [LS02, DCF+09, JDW08, FDMW06]. NLOS and OLOS propagation induce positive biases into distance estimates [BWV18]. Hence the ranging error depends on the LOS condition of the link between a pair of nodes. In particular, a reasonable model is

$$e_{i\to j} = \begin{cases} n_{i\to j} & \text{LOS} \\ b_{i\to j} + n_{i\to j} & \text{NLOS and OLOS,} \end{cases} \tag{5.2}$$

where $n_{i\to j}$ is the error introduced by thermal noise and $b_{i\to j}$ is the bias introduced by NLOS and OLOS propagation.

Recall the definition of the LOS, NLOS, and OLOS conditions. LOS means that only the direct path exists. NLOS refers to the case where the LOS path and at least one multipath component are present. Note that this case also includes the scenario, where the LOS path is received after piercing an obstacle and at least one indirect path exists. Finally, OLOS refers to the situation, where the LOS path is completely blocked and at least one multipath component is received. The thermal noise error $n_{i\to j}$ is typically modeled as zero-mean Gaussian distributed. It is worthwhile to emphasize that both NLOS and OLOS lead to positive biases of range estimates. On the one hand, in the case of NLOS propagation, it is obvious from physics that the excess delay due to the reduced speed of the radio wave inside the LOS-blocking material causes a delayed arrival of the radio wave (compared to the regular LOS condition). On the other hand, in the case of OLOS propagation, the first arriving path is not the direct path, meaning that the radio wave has traveled longer compared to the time required to travel the direct path. Hence, in both cases, the time-of-flight will certainty be longer than the time-of-flight of the direct path. Consequently, the time-of-flight will be over-estimated which also over-estimates the distance. In summary, there is a general consensus in the literature that both NLOS and OLOS propagation lead to positively biased range estimates [BWV18]. Different models are used for $b_{i\to j}$ like Rayleigh distributed, uniformly distributed or one-sided exponentially distributed [GC09, YFGZ13]. Especially in the context of ultrawideband (UWB) signals, $n_{i\to j}$ typically takes small values due to the large bandwidth of the signal [MPG15]. Hence the bias $b_{i\to j}$ is generally much larger than then the error $n_{i\to j}$ due to the thermal noise if the link is in NLOS or OLOS condition, i.e. $b_{i\to j} \gg n_{i\to j}$. Consequently, the ranging errors of NLOS and OLOS links can be approximated by $e_{i\to j} \approx b_{i\to j}$. In other words, for NLOS and OLOS links, positive ranging errors are observed almost surely and negative ranging errors are extremely unlikely – if they occur at all [MGWW10, WMGW12]. Thus, for NLOS and OLOS links, the model used in [MB17, VWS12, GWSR11b] is employed which assumes one-sided non-negative exponentially distributed ranging errors. For LOS links, zero-mean Gaussian distributed ranging errors are assumed as used in [BWV18]. Therefore, the distribution of the ranging error is given as

$$f(e_{i\to j}) = \begin{cases} \mathcal{N}(e_{i\to j}; 0, \sigma^2_{i\to j}) & \text{LOS} \\ \mathcal{E}(e_{i\to j}; \lambda_{i\to j}) & \text{NLOS and OLOS,} \end{cases} \tag{5.3}$$

where $\mathcal{N}(\cdot; 0, \sigma^2_{i\to j})$ is the Gaussian distribution with zero mean and variance $\sigma^2_{i\to j}$ and $\mathcal{E}(\cdot; \lambda_{i\to j})$ is the one-sided non-negative exponential distribution with mean $1/\lambda_{i\to j}$. For simplicity, it is assumed that the parameters of links are identical, i.e. $\sigma^2 = \sigma^2_{i\to j}, \forall i, j$ and

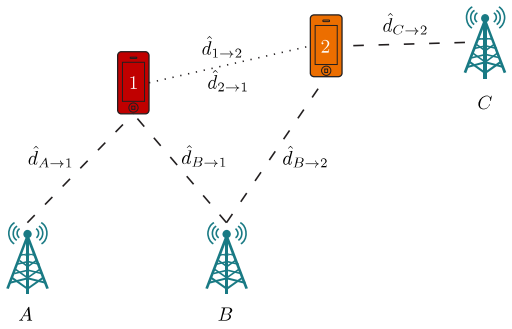

Figure 5.3: *Topology of the running example* - the agents 1 and 2 attempt to estimate their own positions based on the distance estimates with respect to the anchors A, B, and C (dashed lines), as well the cooperative distance estimates (dotted line).

$\lambda = \lambda_{i\to j}, \forall i, j$.

Remark: Note that the POA algorithm that will be described in Section 5.4 relies on the assumption of non-negative ranging errors to determine constraints on the positions of agents. It should be stressed, however, that the POA algorithm does not explicitly assume a model regarding the measurement errors. The only embedded assumption is that the ranging errors of NLOS and OLOS links are positive and that the condition of the links can be identified reliably. These assumptions will be discussed in detail in Section 5.4.3, where a robust version of the POA algorithm will be presented. LOS identification will be discussed, and the effects of miss-classified LOS condition will be reviewed.

5.2.3 Position Estimation

The global goal of every agent j in the network is to determine an estimate $\hat{\mathbf{p}}_j$ of its own position based on all distance measurements $\hat{d}_{i\to j}, \forall i, j$ taken in the network according to an optimality criterion like the MAP or MMSE criterion. Let $\mathbf{z}$ be the vector of all distance measurements $\hat{d}_{i\to j}, \forall i, j$, then the marginal MMSE estimate $\hat{\mathbf{p}}_{j,\mathrm{MMSE}}$ is given by

$$\hat{\mathbf{p}}_{j,\mathrm{MMSE}} = \int \mathbf{p}_j f(\mathbf{p}_j|\mathbf{z}) \mathrm{d}\mathbf{p}_j, \tag{5.4}$$

where $f(\mathbf{p}_j|\mathbf{z})$ is the marginal posterior distribution of the position of agent j. The marginal posterior distributions of all nodes are approximated by belief propagation as outlined in Chapter 4. A detailed description of SBP for cooperative localization will be given in the context of the running example in Section 5.3.

5.3 Running Example

The exemplary topology shown in Fig. 5.3 is used to visualize the concept of SBP in the context of cooperative localization. Starting from the joint posterior distribution, the factor graph will be derived. Based on this graph, the message passing procedure will be examined mathematically. For the convenience of the reader, mathematical operations

will be visualized based on the topology of the running example. Finally, challenges of the estimation procedure will be outlined.

5.3.1 Joint Posterior Distribution and Factor Graph

Let $\boldsymbol{\eta}$ be the parameter vector of the positions of all nodes in the running example, i.e. $\boldsymbol{\eta} = [\mathbf{p}_1^{\mathrm{T}}, \mathbf{p}_2^{\mathrm{T}}, \mathbf{p}_A^{\mathrm{T}}, \mathbf{p}_B^{\mathrm{T}}, \mathbf{p}_C^{\mathrm{T}}]^{\mathrm{T}}$. For the sake of differentiability, agents are enumerated by Arabic numbers, while anchors are enumerated by alphabetic characters. It is worthwhile to mention that the parameter vector can be reduced by the positions of the anchors, since their values are known. On the other hand, the positions of the anchors can be considered as part of the parameter vector by imposing a Dirac distribution as prior, e.g., $f(\mathbf{p}_A) = \delta(\mathbf{p}_A - \mathbf{p}_A^*)$, where $\mathbf{p}_A^*$ is the true (known) position of the anchor. In this chapter, the latter notion is used. Moreover, all distance measurements are collected in the vector $\mathbf{z} = [\hat{d}_{A\to1}, \hat{d}_{B\to1}, \hat{d}_{B\to2}, \hat{d}_{C\to2}, \hat{d}_{2\to1}, \hat{d}_{1\to2}]^{\mathrm{T}}$. Note that the distance estimates in $\mathbf{z}$ have arbitrary order and any other combination is also valid. It is assumed that the distance estimates are independent conditioned on the positions of the involved nodes and that the prior distributions are also independent. Then, the joint posterior $f(\boldsymbol{\eta}|\mathbf{z})$ can be broken up into the following factors:

$$\begin{aligned} f(\boldsymbol{\eta}|\mathbf{z}) \propto & \prod_j f_j(\mathbf{p}_j) \prod_{i\in\mathcal{S}_{\to j}} f_{ij}(\hat{d}_{i\to j}|\mathbf{p}_i, \mathbf{p}_j) \\ = & f_{A1}(\hat{d}_{A\to1}|\mathbf{p}_A, \mathbf{p}_1) f_{B1}(\hat{d}_{B\to1}|\mathbf{p}_B, \mathbf{p}_1) f_{21}(\hat{d}_{2\to1}|\mathbf{p}_2, \mathbf{p}_1) \\ & \times f_{B2}(\hat{d}_{B\to2}|\mathbf{p}_B, \mathbf{p}_2) f_{C2}(\hat{d}_{C\to2}|\mathbf{p}_C, \mathbf{p}_2) f_{12}(\hat{d}_{1\to2}|\mathbf{p}_1, \mathbf{p}_2) \\ & \times f_A(\mathbf{p}_A) f_B(\mathbf{p}_B) f_C(\mathbf{p}_C) f_1(\mathbf{p}_1) f_2(\mathbf{p}_2), \end{aligned} \tag{5.5}$$

where $f_{ij}(\hat{d}_{i\to j}|\mathbf{p}_i, \mathbf{p}_j)$ is the likelihood function of the measurement $\hat{d}_{i\to j}$ between node i and j, and $f_j(\mathbf{p}_j)$ is the prior of node j. The index set $\mathcal{S}_{\to j}$ contains all neighbors[6], e.g., $\mathcal{S}_{\to1} = \{2, A, B\}$. For notational convenience, the set $\mathcal{A}_{\to j}$ is introduced which denotes the set of all anchors inside the communication range of node j, e.g., $\mathcal{A}_{\to1} = \{A, B\} \subset \mathcal{S}_{\to1}$.

Based on the factorization in (5.5), it is straightforward to sketch the corresponding factor graph. Recall that the notion of factor graphs was explained in Section 4.2. Fig. 5.4 depicts the factor graph of the factorization in (5.5). Variable nodes are shown as circular vertices, factor nodes as rectangular vertices, and BP messages are indicated by arrows. Wireless links are depicted by dotted (inter-agent) and dashed (anchor-agent) lines according to Fig. 5.4. Solid edges determine intra-agent messages. Observe that, for compactness of the graph, the factor nodes of the priors are not enclosed in Fig. 5.4. Generally, it is assumed that agents have no prior knowledge regarding their positions which can be expressed by a uniform (uninformative) prior. Observe that no messages are sent towards the anchor nodes, since their positions are perfectly determined. From Fig. 5.4, it can be seen that the factor graph contains a cycle ($\mathbf{p}_1 \to f_{12} \to \mathbf{p}_2 \to f_{21} \to \mathbf{p}_1$). Hence the marginal posteriors cannot be determined exactly but only approximately when using belief propagation.

[6]Nodes within the communication range are loosely referred to as neighbors.

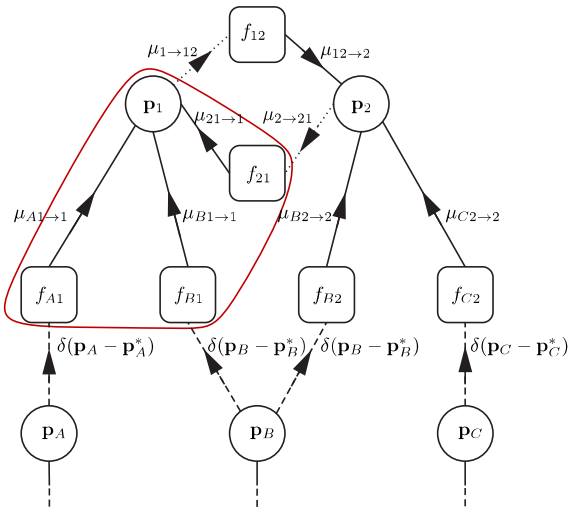

Figure 5.4: *Factor graph of the running example corresponding to the factorization in* (5.5) - the following shorthand notation is used: $\mu_{i\to ij} \triangleq \mu_{\mathbf{p}_i\to f_{ij}}$ and $\mu_{ij\to j} \triangleq \mu_{f_{ij}\to\mathbf{p}_j}$. Operations to be performed by vertices inside of the red frame are determined locally at agent 1.

5.3.2 Sample-based Belief Propagation for Cooperative Positioning

A comprehensive overview of the belief propagation algorithm was already provided in Section 4.3. Consequently, only the peculiarities of SBP in the context of cooperative positioning will be highlighted and exemplary messages will be visualized.

It is helpful to develop an intuition of the messages that are sent along the edges of the factor graph. Messages from variable nodes to factor nodes contain the information regarding the position of the sending variable node. For instance, the message $\mu^{(l)}_{1\to12}(\mathbf{p}_1)$ encloses the information of the position of node 1 along with its position uncertainty in iteration l. However, messages from factor nodes to variable nodes enclose position information regarding the receiving node under consideration of the distance estimates to the sending nodes and their positioning knowledge. For instance, $\mu^{(l)}_{12\to2}(\mathbf{p}_2)$ provides information regarding the position of node 2. considering the distance estimate $\hat{d}_{1\to2}$ conditioned on the position information of node 1. Illustrations that support this intuition will be provided within the next few paragraphs. To simplify the notation, the iteration superscript l is dropped whenever appropriate.

Initialization:
At iteration $l = 0$, message passing starts at the anchor nodes since they are the only nodes with informative prior information. In particular, every anchor broadcasts its position, e.g., $\mathbf{p}_A^*$. For the sake of explanations, the message passing procedure will be elaborated mostly for agent 1, but it is analogous for agent 2. The received anchor positions are then fed

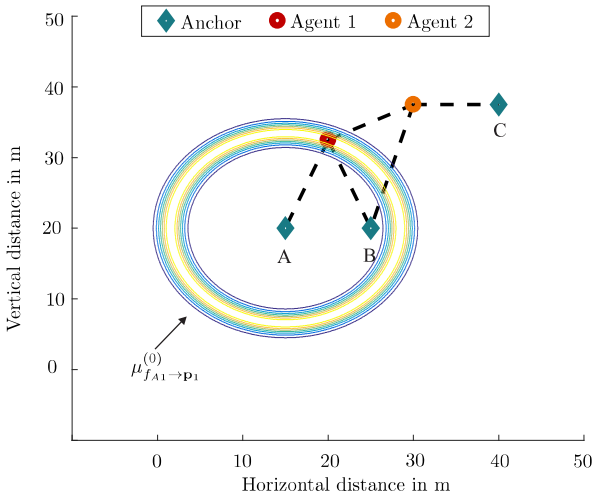

Figure 5.5: *Contour plot of* $\mu^{(0)}_{f_{A1}\to\mathbf{p}_1}$ - the message from factor node f_{A1} to variable node $\mathbf{p}_1$.

into the factor nodes of the respective distance estimates. For instance, agent 1 receives the positions $\mathbf{p}_A^*$ and $\mathbf{p}_B^*$ from the anchors A and B, respectively. These positions are used to determine the messages $\mu_{f_{i1}\to\mathbf{p}_1}, i \in \mathcal{A}_{\to 1}$. Recall that messages from factor nodes to variables nodes are obtained using message filtering, i.e.

$$\mu^{(l)}_{f_{ij}\to\mathbf{p}_j}(\mathbf{p}_j) = \int f_{ij}(\hat{d}_{i\to j}|\mathbf{p}_i,\mathbf{p}_j)\mu^{(l)}_{\mathbf{p}_i\to f_{ij}}(\mathbf{p}_i)\mathrm{d}\mathbf{p}_i. \tag{5.6}$$

The integral in (5.6) cannot be solved in a closed form for arbitrary incoming messages $\mu^{(l)}_{\mathbf{p}_i\to f_{ij}}(\mathbf{p}_i)$. Notice, however, that all messages $\mu^{(l)}_{\mathbf{p}_i\to f_{ij}}(\mathbf{p}_i)$ from anchors $i \in \mathcal{A}_{\to j}$ are Dirac distributions as outlined earlier and shown in Fig. 5.4. In this special case, the integral in (5.6) vanishes and the outgoing messages $\mu^{(l)}_{f_{ij}\to\mathbf{p}_j}(\mathbf{p}_j), i \in \mathcal{A}_{\to j}$ are given by

$$\mu^{(l)}_{f_{ij}\to\mathbf{p}_j}(\mathbf{p}_j) = f_{ij}(\hat{d}_{i\to j}|\mathbf{p}_i^*,\mathbf{p}_j), i \in \mathcal{A}_{\to j}. \tag{5.7}$$

Assuming that the link between the nodes i and j is in LOS condition and assuming the ranging error model in (5.2), the message is given by

$$\mu^{(l)}_{f_{ij}\to\mathbf{p}_j}(\mathbf{p}_j) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\hat{d}_{i\to j} - \left\|\mathbf{p}_i^* - \mathbf{p}_j\right\|^2)}{2\sigma^2}\right). \tag{5.8}$$

This message is depicted in the context of the running example for exemplary values of $\hat{d}_{i\to j}$ and σ in Fig. 5.5. Recall the intuition on that message which was provided at the beginning of this section. The message $\mu_{f_{A1}\to\mathbf{p}_1}(\mathbf{p}_1)$ provides information regarding the position of agent 1. In particular, with the highest likelihood, agent 1 is on any point of

the circle around the anchor A with radius $\hat{d}_{A\to 1}$. Note, however, that other radii than $\hat{d}_{A\to 1}$ are also possible, but they have reduced likelihood. In that way, the uncertainty of the measurement error $e_{i\to j}$ is accounted for. This uncertainty translates into a position uncertainty of node 1 via (5.8), leading to the doughnut-like shape in Fig. 5.8.

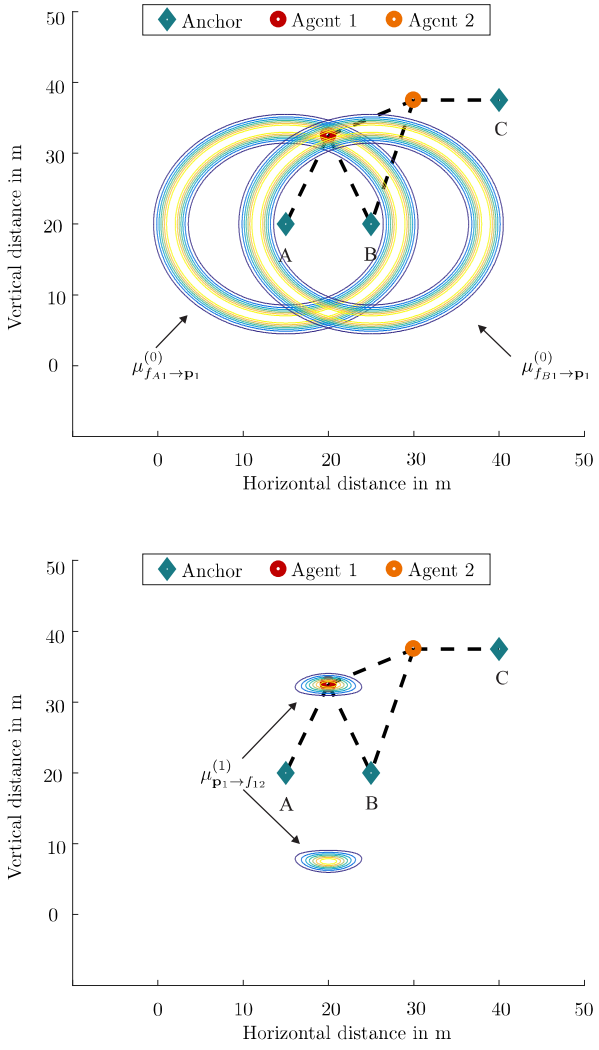

Figure 5.6: *Top* - contour plot of the messages $\mu^{(0)}_{f_{i1}\to \mathbf{p}_1}, i = \{A, B\}$ to be multiplied. *Bottom* - contour plot of the outgoing message $\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$ from variable node $\mathbf{p}_1$ to factor node f_{12}.

Iterative Message Update ($l = 1$):
Recalling the topology of the running example, it becomes evident that agent 1 receives the messages $\mu^{(0)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}_1), i \in \mathcal{A}_{\to 1} = \{A, B\}$. Since no message is received from agent 2 during the 1st iteration, the outgoing message $\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$ is given as the product of the incoming anchor-messages $\mu^{(1)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}_1), i \in \mathcal{A}_{\to 1} = \{A, B\}$, i.e.

$$\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1) = \prod_{i\in\mathcal{A}_{\to 1}} \mu^{(0)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}_1) = \mu^{(0)}_{f_{A1}\to\mathbf{p}_1}(\mathbf{p}_1) \times \mu^{(0)}_{f_{B1}\to\mathbf{p}_1}(\mathbf{p}_1). \tag{5.9}$$

The outgoing messages of variable nodes like in (5.9) also constitute the beliefs of the agents, e.g. $b^{(1)}_{\mathbf{p}_1}(\mathbf{p}_1) \triangleq \mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$. Observe that the closed-form solution in (5.8) allows for an evaluation of the messages $\mu^{(1)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}_1)$, $i \in \{A, B\}$ at arbitrary values of $\mathbf{p}_1$. The product of the two messages, however, cannot be easily described in a closed form[7]. Visually, this can be understood easily by considering Fig. 5.6 in which the incoming messages $\mu^{(0)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}_1), i \in \{A, B\}$ (top) and the outgoing message $\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$ (bottom) are depicted. The outgoing message has two modes whose shapes cannot be described without errors by tractable parametric distributions. The application of nonparametric tools offers a solution in which the modes are represented by particles. In particular, importance sampling with a suitable proposal distribution $q_1(\mathbf{p}_1)$ can be applied in order to obtain a particle representation $\mathcal{R}_{Ns}\left(\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)\right)$ of the message $\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$. Sampling and weighting is depicted in Fig. 5.7. Given a set of samples $\{\mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}}\}_{k=1}^{N_s} \sim q_1(\mathbf{p}_1)$, the weights $w^{(k,1)}_{\mu_{\mathbf{p}_1\to f_{12}}}$ of the particle representation are given by

$$w^{(k,1)}_{\mu_{\mathbf{p}_1\to f_{12}}} = \prod_{i\in\mathcal{A}_{\to 1}} \frac{\mu^{(0)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}})}{q_1(\mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}})}. \tag{5.10}$$

Note that it is not trivial to find a *good* proposal distribution $q_1(\mathbf{p}_1)$. In theory (for an infinite number of samples) any distribution $q_1(\mathbf{p}_1)$ whose support dominates the support of $\mu_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$ results in a particle representation that approximates the true message in (5.9) arbitrarily closely. In practice, however, poorly chosen proposal distributions will greatly affect the accuracy of the approximation. For a finite and *practical*[8] number of samples, a poorly chosen proposal distribution can lead to unsatisfactory approximation accuracy. This issue will be addressed very briefly in the following and more elaborately in Section 5.4 as a motivation for the proposed polygon outer-approximation algorithm.

Generally, a continuous distribution can be approximated relatively accurately if the samples of the approximating particle representation reside in areas where the distribution to be approximated has significant probability mass. In the context of message multiplication, the challenge of identifying a good proposal distribution refers back to the question: *How can one draw samples that reside in areas where the product of messages has significant probability mass without computing the product explicitly?*. Notice that this is a

[7] In this context, a closed-form solution refers to a solution which does not include the multiplication of the messages $\mu^{(l)}_{f_{i1}\to\mathbf{p}_1}(\mathbf{p}_1)$.

[8] Practical sample sizes range from hundreds to thousands of samples.

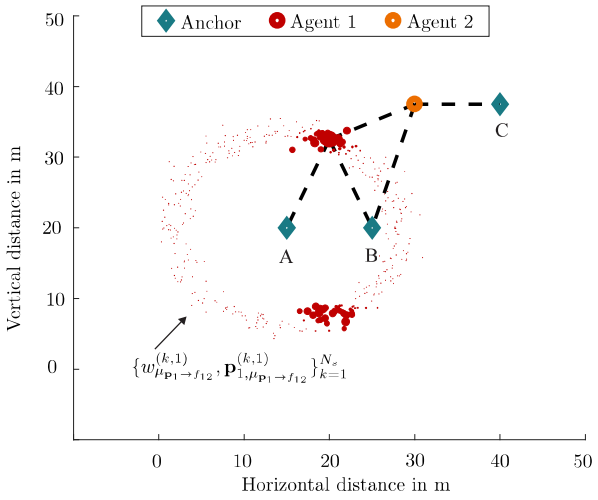

Figure 5.7: *Particle representation* $\mathcal{R}_{Ns}\left(\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)\right) = \{w^{(k,1)}_{\mu_{\mathbf{p}_1\to f_{12}}}, \mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}}\}_{k=1}^{N_s}$ *of the outgoing message* $\mu_{\mathbf{p}_1\to f_{12}}$ - samples are indicated by red dots and their weights are encoded by the size of the respective dots.

chicken-and-the-egg problem. A common way of drawing samples [LFS+12] is depicted in Fig. 5.7. Here, the samples $\mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}}, \forall k$ are drawn based on the following intuition. Given the position of an anchor, say anchor A, agent 1 must reside somewhere on the circle around $\mathbf{p}_A$ with radius $\hat{d}_{A\to 1} + e_{A\to 1}$, where $e_{A\to 1}$ is an error term that emulates the distance estimation error. Loosely speaking, the samples are drawn on the ring of the likelihood function $f(\hat{d}_{A\to 1}|\mathbf{p}_1, \mathbf{p}^*_A)$. Mathematically, a sample is generated by drawing two auxiliary samples $e^{(k,1)}_{A\to 1} \sim f(e_{A\to 1})$ and $\alpha^{(k,1)}_{A\to 1} \sim \mathcal{U}(0, 2\pi)$ and transforming them according to

$$\mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}} = \mathbf{p}_A + (\hat{d}_{A\to 1} + e^{(k,1)}_{A\to 1}) \begin{bmatrix} \cos(\alpha^{(k,1)}_{A\to 1}) \\ \sin(\alpha^{(k,1)}_{A\to 1}) \end{bmatrix}. \tag{5.11}$$

Observe that the majority of particles in Fig. 5.7 have negligible weights and hence they will be erased with high probability by the subsequent resampling step as elaborated in Section 3.2.1.2. An essential insight of this thesis is that finding good proposal distributions matters in practice as will be shown in Section 5.6. For instance, smart proposal distributions allow for (i) a reduction of the number of particles without reducing the estimation accuracy or (ii) enable higher estimation accuracy for a given number of samples. After resampling, the incoming message $\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$ to the factor node f_{12} is a particle representation with equally weighted particles, i.e.

$$\mathcal{R}_{Ns}\left(\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)\right) = \left\{1/N_s, \mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}}\right\}_{k=1}^{N_s}, \tag{5.12}$$

where the samples[9] $\mathbf{p}_{1,\mu_{\mathbf{p}_1\to f_{12}}}^{(k,1)}$ are determined via resampling.

Then, the resampled incoming message $\mu_{\mathbf{p}_1\to f_{12}}^{(1)}(\mathbf{p}_1)$ must be processed in order to compute the outgoing message $\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)$. This procedure is visualized in Fig. 5.8. In particular, the integral in (5.6) must be solved, i.e.

$$\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)\int f_{12}(\hat{d}_{1\to 2}|\mathbf{p}_1,\mathbf{p}_2)\mu_{\mathbf{p}_1\to f_{12}}^{(1)}(\mathbf{p}_1)\mathrm{d}\mathbf{p}_1. \tag{5.13}$$

Let $\mathbf{p}_{1,\mu_{\mathbf{p}_1\to f_{12}}}^{(k,1)}$ denote the k^{th} resampled particle. Then, the integral simplifies to

$$\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)\approx\frac{1}{N_s}\sum_{k=1}^{N_s}f_{12}(\hat{d}_{1\to 2}|\mathbf{p}_{1,\mu_{\mathbf{p}_1\to f_{12}}}^{(k,1)},\mathbf{p}_2). \tag{5.14}$$

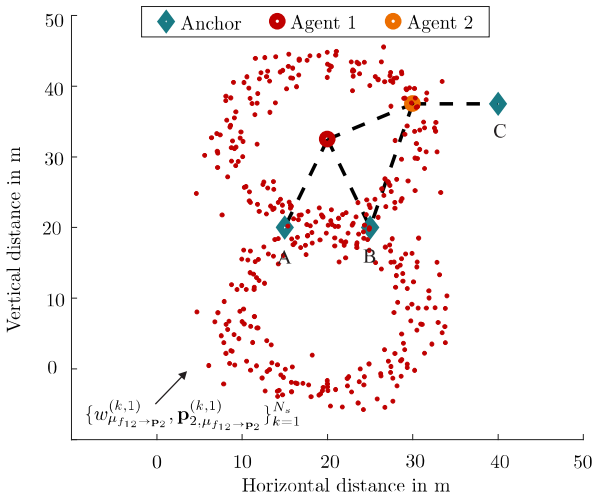

Figure 5.8: *Particle representation* $\mathcal{R}_{Ns}\left(\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)\right)=\{w_{\mu_{f_{12}\to\mathbf{p}_2}}^{(k,1)},\mathbf{p}_{2,\mu_{f_{12}\to\mathbf{p}_2}}^{(k,1)}\}_{k=1}^{N_s}$ *of the outgoing message* $\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)$ - samples are depicted by red dots.

In order to obtain a particle representation $\mathcal{R}_{Ns}\left(\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)\right)$ of the filtered message $\mu_{f_{12}\to\mathbf{p}_2}^{(1)}(\mathbf{p}_2)$ in (5.14), one sample is drawn for every likelihood-term in the sum in (5.14). This can be achieved by drawing the two auxiliary samples $e_{1\to 2}^{(k,1)}\sim f(e_{1\to 2})$ and $\alpha_{1\to 2}^{(k,1)}\sim\mathcal{U}(0,360)$ for every sample $\mathbf{p}_{1,\mu_{\mathbf{p}_1\to f_{12}}}^{(k,1)}$. The resulting particle representation

[9]Note that this is a slight abuse of the notation, since the sample $\mathbf{p}_{1,\mu_{\mathbf{p}_1\to f_{12}}}^{(k,1)}$ refers to the k^{th} original and resampled particle at the same time. This was done to avoid further notational clutter.

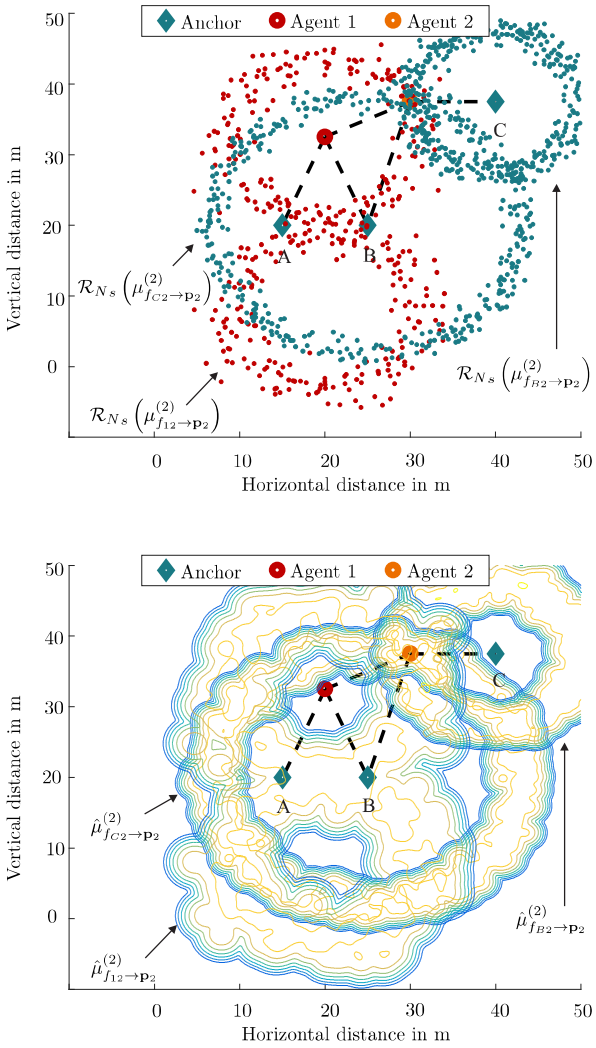

Figure 5.9: *Top* - particle representations $\mathcal{R}_{Ns}\left(\mu^{(2)}_{f_{i2}\to\mathbf{p}_2}(\mathbf{p}_2)\right) = \{w^{(k,2)}_{\mu_{f_{i2}\to\mathbf{p}_2}}, \mathbf{p}^{(k,2)}_{2,\mu_{f_{i2}\to\mathbf{p}_2}}\}_{k=1}^{N_s}$. *Bottom* - contour plot of the kernel density estimates $\hat{\mu}^{(2)}_{f_{i2}\to\mathbf{p}_2}(\mathbf{p}_2)$ of the messages.

$\mathcal{R}_{Ns}\left(\mu^{(1)}_{f_{12}\to\mathbf{p}_2}(\mathbf{p}_2)\right)$ is given by

$$\mathcal{R}_{Ns}\left(\mu^{(1)}_{f_{12}\to\mathbf{p}_2}(\mathbf{p}_2)\right) = \left\{w^{(k,1)}_{\mu_{f_{12}\to\mathbf{p}_2}}, \mathbf{p}^{(k,1)}_{2,\mu_{f_{12}\to\mathbf{p}_2}}\right\}_{k=1}^{N_s}, \tag{5.15}$$

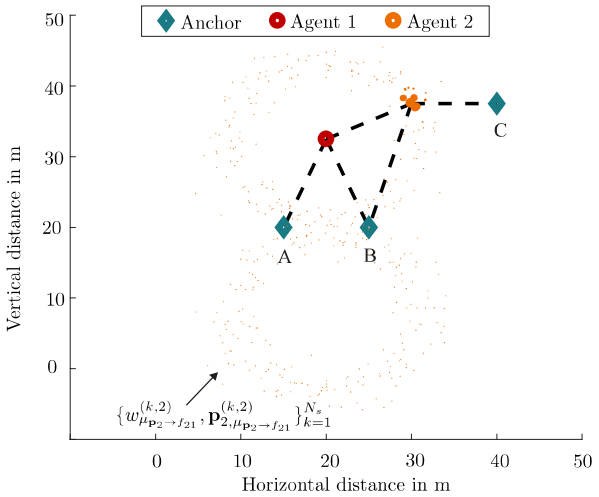

Figure 5.10: *Particle representation* $\mathcal{R}_{N_s}\left(\mu^{(2)}_{\mathbf{p}_2\to f_{21}}(\mathbf{p}_2)\right) = \{w^{(k,2)}_{\mu_{\mathbf{p}_2\to f_{21}}}, \mathbf{p}^{(k,2)}_{2,\mu_{\mathbf{p}_2\to f_{21}}}\}_{k=1}^{N_s}$ *of the outgoing message* $\mu^{(2)}_{\mathbf{p}_2\to f_{21}}(\mathbf{p}_2)$ - note that the position of agent 2 is unambiguously identified with the help of agent 1 and the anchor B and C, since the message $b^{(2)}_{\mathbf{p}_2}(\mathbf{p}_2)\mu^{(2)}_{\mathbf{p}_2\to f_{21}}(\mathbf{p}_2)$ is unimodal.

where $\mathbf{p}^{(k,1)}_{2,\mu_{f_{12}\to\mathbf{p}_2}}$ is computed according to (5.11). A visualization can be found in Fig. 5.8. Since the incoming message $\mu^{(1)}_{\mathbf{p}_1\to f_{12}}(\mathbf{p}_1)$ has two modes (see Fig. 5.7), two circles of particles result (see Fig. 5.8). After all filtered messages are computed, the first iteration terminates.

Iterative Message Update ($l \geq 2$):
In the second iteration, agent 1 and 2 compute the products of all incoming messages, including the cooperative messages $\mu_{f_{12}\to\mathbf{p}_2}(\mathbf{p}_2)$ and $\mu_{f_{21}\to\mathbf{p}_1}(\mathbf{p}_1)$. Note that, in contrast to the first iteration, the incoming messages from anchors *and* the other agent are received. For instance, from the perspective of agent 2, the following product must be computed (see Fig. 5.9)

$$\mu^{(l)}_{\mathbf{p}_2\to f_{21}}(\mathbf{p}_2) = \prod_{i\in\mathcal{S}_{\to 2}} \mu^{(l-1)}_{f_{i2}\to\mathbf{p}_2}(\mathbf{p}_2) = \underbrace{\mu^{(l-1)}_{f_{B2}\to\mathbf{p}_2}(\mathbf{p}_2)\mu^{(l-1)}_{f_{C2}\to\mathbf{p}_2}(\mathbf{p}_2)}_{\text{anchor info}}\underbrace{\mu^{(l-1)}_{f_{12}\to\mathbf{p}_2}(\mathbf{p}_2)}_{\text{agent info}}. \tag{5.16}$$

Bear in mind that anchor-messages $\mu^{(l-1)}_{f_{ij}\to\mathbf{p}_j}(\mathbf{p}_j), i \in \mathcal{A}_{\to j}$ do not change within different iterations, since the anchor positions are perfectly determined, i.e. $\mu^{(l-1)}_{f_{B2}\to\mathbf{p}_2}(\mathbf{p}_2) = \mu^{(0)}_{f_{B2}\to\mathbf{p}_2}(\mathbf{p}_2), \forall l$. Moreover, they are given in the closed form presented in (5.7). However, all incoming messages $\mu^{(l-1)}_{f_{ij}\to\mathbf{p}_j}(\mathbf{p}_j), i \in \mathcal{S}_{\to j}\backslash\mathcal{A}_{\to j}$ from other agents are given as particle representations. Hence straightforward message multiplication is not possible, since the positions of the samples will be distinct with probability one. In that case, kernel density

estimates $\hat{\mu}^{(l-1)}_{f_{i2}\to\mathbf{p}_2}(\mathbf{p}_2), i \in \mathcal{S}_{\to 2}\backslash\mathcal{A}_{\to 2}$ of the particle representations must be obtained as described in section 3.2.2.2. An illustration[10] is shown in Fig. 5.9. Using importance sampling, a particle representation of the outgoing message $\mu^{(l)}_{\mathbf{p}_2\to f_{21}}(\mathbf{p}_2)$ can be obtained by first drawing samples $\mathbf{p}^{(k,l)}_{2,\mu_{\mathbf{p}_2\to f_{21}}}$ from a suitable proposal distribution $q_2(\mathbf{p}_2)$ and weighting them according to

$$w^{(k.l)}_{\mu_{\mathbf{p}_2\to f_{21}}} = \frac{\prod_{i\in\mathcal{S}_{\to 2}} \hat{\mu}^{(l-1)}_{f_{i2}\to\mathbf{p}_2}(\mathbf{p}^{(k,l-1)}_{2,\mu_{\mathbf{p}_2\to f_{12}}})}{q_2(\mathbf{p}^{(k,l-1)}_{2,\mu_{\mathbf{p}_2\to f_{12}}})}. \tag{5.17}$$

An illustration of the resulting particle representation $\mathcal{R}_{Ns}\left(\mu^{(l)}_{\mathbf{p}_2\to f_{21}}(\mathbf{p}_2)\right)$ (for $l = 2$) based on the proposal distribution from [LFS+12] is depicted in Fig. 5.10. Notice, again, that with a poorly chosen proposal distribution, many of the particles have insignificant weights and do not contribute considerably to the result of the position estimation.

After computing the outgoing messages $\mu^{(l)}_{\mathbf{p}_2\to f_{21}}$ and $\mu^{(l)}_{\mathbf{p}_1\to f_{12}}$ (not shown) and resampling them, the filtering operation in (5.13) must be executed for all resampled messages in order to complete the l^{th} iteration. In this example, unimodality of the beliefs is achieved after two iterations and hence the positions of agents can be determined unambiguously. For larger networks, however, more iterations can be necessary to achieve unimodality of the beliefs of all agents. In general, the procedure of iterative message updates is repeated N_{BP} times or until convergence is achieved. The main take-away message from this section is that finding good proposal distributions is of crucial importance for efficient position estimation in practice as will be shown in Section 5.6.

5.4 Polygon Outer-Approximation

A common difficulty of most sample-based estimators is answering the question: *Which regions of the sample space contain significant probability mass?*. The answer to this question is certainly non-trivial, but it is a vital one as pointed out numerous times throughout this thesis. While in theory (for infinite sample sizes), the requirements imposed on proposal distributions are very loose, practical estimators (with hundreds to thousands of samples) demand smart proposal distributions in order to utilize the available samples as efficiently[11] as possible. The idea that will be elaborated in the remainder of this section allows for efficient sampling by imposing geometric constraints on the sample space. These constraints are embodied by convex polygons giving rise to the name of the approach: *polygon outer-approximation*.

5.4.1 Conception of Constraints

In radio-based localization, constraints emerge very naturally due to the radio environment. Range estimates from NLOS and OLOS links are inherently biased and generally overestimate the true distances between pairs of nodes as discussed in Section 5.2.2. The

[10] For simplicity of the argument, it is assumed that all messages (from agents *and* anchors) are given as particle representations. However, in practice, the closed-form solutions of anchor messages are used rather than kernel density estimates.

[11] In the context of importance sampling, efficiency refers to the ratio of particles with significant weights over all particles.

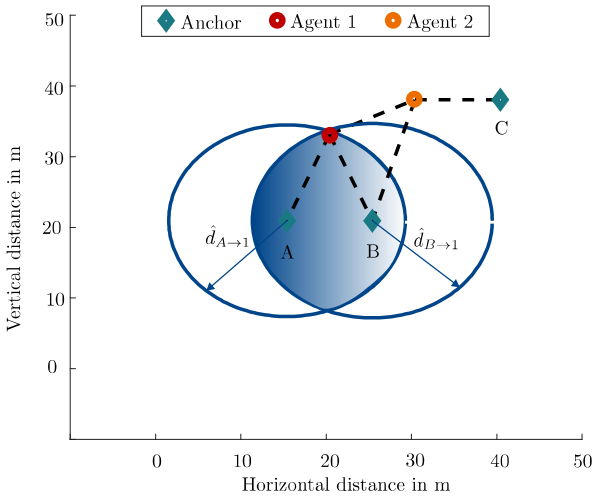

Figure 5.11: *Basic principle* - if the links to the anchors A and B are in NLOS or OLOS condition, agent 1 resides inside the disks that are bounded by the blue circles. Thus, the position of agent 1 is confined to the intersection of these disks (blue-shaded, eye-shaped region).

intuitive implication is the following: Assume that the link between agent 1 and anchor A of the running example is in NLOS condition. Then, the range estimate $\hat{d}_{A\to 1}$ overestimates the distance between the two nodes and it can be concluded that the position of agent 1 must be somewhere inside the disk with radius $\hat{d}_{A\to 1}$ centered at $\mathbf{p}_A^*$ (see Fig. 5.11). Mathematically, this constraint can expressed as

$$\hat{d}_{A\to 1} \geq \|\mathbf{p}_1 - \mathbf{p}_A^*\| \,. \tag{5.18}$$

Bear in mind that (5.18) assumes that the positive bias $b_{A\to 1}$ due to NLOS or OLOS propagation dominates the error due to thermal noise $n_{A\to 1}$, i.e. $b_{A\to 1} \gg n_{A\to 1}$. As long as this assumption holds, a natural constraint from the radio environment gives rise to a constraint on the sample space. In other words, it would be unreasonable to draw samples $\mathbf{p}_1^{(k)}$ that reside outside of the disk described in (5.18), since agent 1 must be inside of it. Suppose that the link to anchor B is also in NLOS or OLOS condition. These constraints on the position of agent 1 become tighter since agent 1 must reside within the intersection (eye-shaped region in Fig. 5.11) of the two individual constraints (disks).

The immediate consequence is the following: Suppose the LOS state of a link can be detected reliably[12], then, the NLOS and OLOS links can be used to impose geometric

[12]LOS identification can be done using various statistics of the received signal. Particularly, the delay spread of the received signal can be used to classify the state of the link. Numerous detectors have been presented in the literature [VB07, MGWW10, GCWI07, BHM98]. The rates for correct classification of links depend on the scenario as well as the classification method. Correct classification rates of 80% and larger have been achieved by the approach in [GCWI07] based on simulated data. Correct classification rates of 85% and higher were obtained by the method presented in [MGWW10] based on real-world experiments. A discussion on the implications of miss-classification will be conducted in Section 5.4.3.

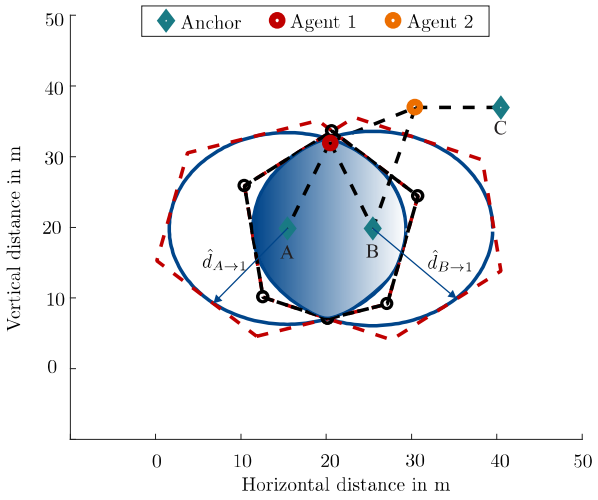

Figure 5.12: *Concept of polygon outer-approximation* - the two constrained regions (disks) are outer-approximated by polygons and the intersection of the polygons (black dash-dotted polygon) is computed via closed-form operations.

constraints on the locations of agents. Once the constraints are determined, they can be used as inputs to the position estimator. For instance, sample-based estimators can concentrate all samples in the constrained region of the sample space.

Observe that the intersections of arbitrary disks can take arbitrary shapes that are generally (i) difficult to compute and (ii) impractical to parametrize. In order to bypass these two complications, the two disks can be outer-approximated by polygons as shown in Fig. 5.12. Seemingly, the use of polygons instead of circles over-estimates the feasible area. At the same time, though, polygons are easy to intersect numerically. The vertices of the intersected polygon can be simply computed as the intersections of two straight lines. Determining the intersection of two straight lines can be solved in closed-form. Hence computing the intersections of polygons becomes tractable and the complication (i) is eliminated. Moreover, the intersections of convex polygons are again convex polygons which removes the hurdle (ii) regarding the representation.

For simplicity, the previous discussion was focused on the NLOS or OLOS links between agents and anchors. The idea of constraining the sample space can also be extended to cooperative distance estimates on NLOS and OLOS links between agents. In contrast to anchors, the positions of agents are not known a priori and the constraints need to be encoded differently compared to anchors. The main idea of including constraints based on agent-agent links will be described in what follows. An agent's location is not given by a single point but rather by a region (such as the polygon covering the eye-shaped region in Fig. 5.12). Extending this region by the respective range estimate to another agent, the extended region materializes into a new constraint for the other agent. A detailed explanation of the extension will be provided in the following subsection.

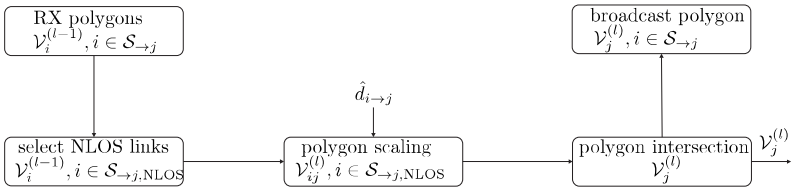

Figure 5.13: *POA flow diagram* - after receiving polygons $\mathcal{V}_i^{(l-1)}, i \in \mathcal{S}_{\to j}$ from all neighbors, every agent j selects the polygons $\mathcal{V}_i^{(l-1)}, i \in \mathcal{S}_{\to j,\mathrm{NLOS}}$ from NLOS and OLOS links. These polygons are scaled by the distance estimates $\hat{d}_{i\to j}$ and intersected subsequently in order to compute the updated polygon $\mathcal{V}_j^{(l)}$. Finally, all agents broadcast their polygons $\mathcal{V}_j^{(l)}, \forall j$ and repeat the previous steps during the next iteration.

5.4.2 Algorithm Blueprint

In this subsection, the polygon outer-approximation algorithm will be described in detail using the running example to graphically visualize the operations. The pseudo-code for the algorithm is shown in Algorithm 1 and a flow diagram is depicted in Fig. 5.13.

POA is an iterative algorithm that incrementally reduces the sizes of the constrained regions (polygons). The first iteration differs from all subsequent iterations and will be described in the following subsection. From the second iteration onwards, every agent j receives the polygons $\mathcal{V}_i^{(l-1)},\ \forall i \in \mathcal{S}_{\to j}$ from all neighbors. Then, agent j selects the polygons from the subset $\mathcal{S}_{\to j,\mathrm{NLOS}} \subseteq \mathcal{S}_{\to j}$ of neighbors whose links are in NLOS or OLOS condition. These polygons $\mathcal{V}_i^{(l-1)},\ \forall i \in \mathcal{S}_{\to j,\mathrm{NLOS}}$ are extended by the range estimates (see Fig. 5.16). This operation is called *polygon scaling*. The resulting (scaled) polygons are denoted by $\mathcal{V}_{ij}^{(l)}$. In order to update its own polygon, every agent j determines the intersection of the scaled polygons $\mathcal{V}_{ij}^{(l)},\ \forall i \in \mathcal{S}_{\to j,\mathrm{NLOS}}$, which results in the convex polygon $\mathcal{V}_j^{(l)}$. This operation is called *polygon intersection*. The intersected polygons $\mathcal{V}_j^{(l)}, \forall j$ are broadcasted and the procedure repeats. After a sufficient number N_{FS} of iterations, the sizes of the polygons converge. The polygon $\mathcal{V}_j^{(N_{\mathrm{FS}})}$ that each agent has obtained tightly confines the location of the respective agent j. Using the terminology from the optimization community, the phrase *feasible set* will often be used in place of the constrained region. The first iteration as well as the operations *polygon intersection* and *polygon scaling* will be described in the following subsections.

5.4.2.1 Anchor Polygon Processing

Recall that it is assumed that agents do not have prior knowledge regarding their positions. Hence only the information from anchors is considered in the first iteration. To start the procedure, all anchors broadcast their positions $\mathbf{p}_i^*, \forall i$. Let $\mathcal{A}_{\to j,\mathrm{NLOS}} \subseteq S_{\to j}$ denote the set of anchors in the communication range of agent j with positively biased range estimates $\hat{d}_{i\to j}, i \in \mathcal{A}_{\to j,\mathrm{NLOS}}$. Due to the positive ranging errors, the position of agent j is somewhere inside the disk with radius $\hat{d}_{i\to j}$ and center $\mathbf{p}_i^*$. Agent j's position can be confined to the convex feasible set

$$\mathcal{B}_{ij} = \left\{ \mathbf{p}_j \middle| \, \|\mathbf{p}_j - \mathbf{p}_i^*\| \leq \hat{d}_{i\to j} \right\}. \tag{5.19}$$

Algorithm 1 Polygon Support Outer-Approximation

given $\mathbf{z}$
for $l = 1$ to N_{FS} **do**
  nodes $j = 1$ to N **do** in parallel
    if $l = 1$ **then** (*first iteration*)
      receive $\mathbf{p}_i^*, \forall i \in \mathcal{A}_j$
      for $i \in \mathcal{A}_{\to j,\mathrm{NLOS}}$ **do**
        $\mathcal{V}_{ij}^{(1)} = \textit{anchor polygon processing}(\mathbf{p}_i^*, \hat{d}_{i\to j})$
        - see **Algorithm 2** -
      end for
      $\mathcal{V}_j^{(1)} = \textit{polygon intersection}(\mathcal{V}_{ij}^{(1)}\ \forall i \in \mathcal{A}_{\to j,\mathrm{NLOS}})$
      - see **Algorithm 3** -
    else (*subsequent iterations*)
      receive $\mathcal{V}_i^{(l-1)},\ \forall i \in \mathcal{S}_{\to j}$
      for $i \in \mathcal{S}_{\to j,\mathrm{NLOS}}$ **do**
        $\mathcal{V}_{ij}^{(l)} = \textit{polygon scaling}(\hat{d}_{i\to j},\ \mathcal{V}_i^{(l-1)})$
        - see **Algorithm 4** -
      end for
      $\mathcal{V}_j^{(l)} = \textit{polygon intersection}(\mathcal{V}_{ij}^{(l)},\ \forall i \in \mathcal{S}_{\to j,\mathrm{NLOS}})$
      - see **Algorithm 3** -
    end if
    broadcast $\mathcal{V}_j^{(l)}$
  nodes end
end for

To obtain a polygon $\mathcal{V}_{ij}^{(1)}$, this disk is outer-approximated by a polygon with N_E edges. This operation is coined *anchor polygon processing* and it is visualized in Fig. 5.14. Mathematically, a convex polygon with N_E edges is described by an ordered list of N_E vertices or, alternatively, by the intersection of N_E halfspaces. Both representations will be used during the description of the POA. For now, the description using vertices is considered.

To generate the vertices of an anchor polygon, the disk with radius $\hat{d}_{i\to j}$ is considered (see Fig. 5.14). For mathematical convenience, the disk is assumed to be in the origin of the coordinate system. Obtaining the vertices of the anchor polygon can be done efficiently in polar coordinates. Subsequently, the vertices are transformed into Cartesian coordinates and shifted from the origin to the position $\mathbf{p}_i^*$ of the anchor. The anchor polygon is denoted by $\mathcal{V}_{ij}^{(1)}$ (see Fig. 5.14).

In more details, N_E points are generated around the origin with uniform angular spacing $\alpha = 2\pi/N_\mathrm{E}$ and fixed radius

$$r_{ij} = \frac{\hat{d}_{i\to j}}{\cos(\alpha/2)}. \tag{5.20}$$

The list of vertices in polar coordinates is given by $\tilde{\mathbf{v}}_{ij,1}^{(1)} = [\alpha_0, r_{ij}]^\mathrm{T}$, $\tilde{\mathbf{v}}_{ij,2}^{(1)} = [\alpha + \alpha_0, r_{ij}]^\mathrm{T}$, ..., $\tilde{\mathbf{v}}_{ij,N_\mathrm{E}}^{(1)} = [(N_\mathrm{E}-1)\alpha + \alpha_0, r_{ij}]^\mathrm{T}$, where α_0 is an arbitrary angular offset. Every vertex is then transformed into Cartesian coordinates and shifted by the position $\mathbf{p}_i^*$ of anchor i. Let $\bar{\mathbf{v}}_{ij,m}^{(1)} \triangleq \mathit{pol2cart}(\tilde{\mathbf{v}}_{ij,m}^{(1)})$ denote the m^th vertex in Cartesian coordinates. Then, the shifted vertices are given by

$$\mathbf{v}_{ij,m}^{(1)} = \bar{\mathbf{v}}_{ij,m}^{(1)} + \mathbf{p}_i^*, \quad m = 1, ..., N_\mathrm{E}. \tag{5.21}$$

The ordered list of vertices $\mathbf{v}_{ij,m}^{(1)}$ constitutes the polygon $\mathcal{V}_{ij}^{(1)} = \left\{\mathbf{v}_{ij,1}^{(1)}, \mathbf{v}_{ij,2}^{(1)}, ..., \mathbf{v}_{ij,N_\mathrm{E}}^{(1)}\right\}$. Note that $\mathcal{V}_{ij}^{(1)}$ is always convex by construction.

Algorithm 2 Anchor Polygon Processing

given $\mathbf{p}_i^*, \hat{d}_{i\to j}$
vertex angular spacing $\alpha = \frac{2\pi}{N_\mathrm{E}}$
vertex radius $r_{ij} = \frac{\hat{d}_{i\to j}}{\cos(\alpha/2)}$
for $m = 1$ to N_E **do**
    compute vertex $\tilde{\mathbf{v}}_{ij,m}^{(1)} = [\alpha_0 + (m-1)\cdot\alpha, r_{ij}]^\mathrm{T}$
    $\bar{\mathbf{v}}_{ij,m}^{(1)} = \mathit{pol2cart}(\tilde{\mathbf{v}}_{ij,m}^{(1)})$
    shift vertex $\mathbf{v}_{ij,m}^{(1)} = \bar{\mathbf{v}}_{ij,m}^{(1)} + \mathbf{p}_i^*$
end for
determine polygon $\mathcal{V}_{ij}^{(1)} = \left\{\mathbf{v}_{ij,1}^{(1)}, \mathbf{v}_{ij,2}^{(1)}, ..., \mathbf{v}_{ij,N_\mathrm{E}}^{(1)}\right\}$

5.4.2.2 Polygon Intersection

After agent j has obtained all anchor polygons $\mathcal{V}_{ij}^{(1)}$, $\forall i \in \mathcal{A}_{\to j,\mathrm{NLOS}}$, it intersects these polygons to obtain $\mathcal{V}_j^{(1)}$. If an agent is not connected to any anchor with an NLOS or

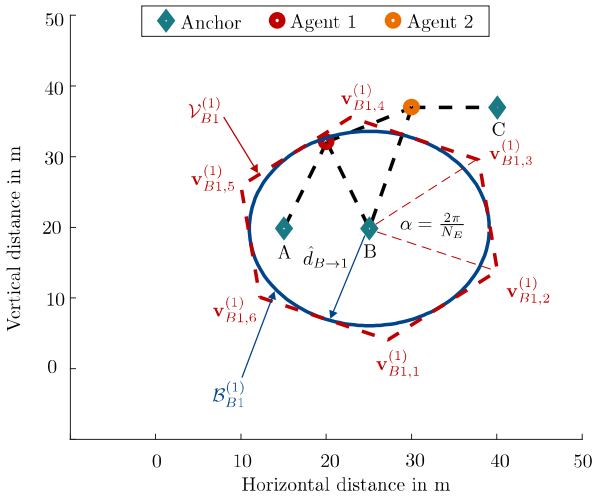

Figure 5.14: *Anchor polygon* - the feasible set $\mathcal{B}_{B1}$ is outer-aproximated by a polygon $\mathcal{V}_{B1}^{(1)}$ with $N_\mathrm{E} = 6$ edges. The vertices of $\mathcal{V}_{B1}^{(1)}$ are spaced with fixed radius $\hat{d}_{B\to1}/\cos(\alpha/2)$ and uniform angle $\alpha = 2\pi/N_\mathrm{E}$ around $\mathbf{p}_B^*$.

OLOS link, its polygon will be the entire plane which is considered for localization, i.e.

$$\mathcal{V}_j^{(1)} = \{[x_\mathrm{min}, y_\mathrm{min}]^\mathrm{T}, [x_\mathrm{max}, y_\mathrm{min}]^\mathrm{T}, [x_\mathrm{max}, y_\mathrm{max}]^\mathrm{T}, [x_\mathrm{min}, y_\mathrm{max}]^\mathrm{T}\}.$$

The polygon intersection operation is a successive application of the Sutherland-Hodgman algorithm [SH74]. In order to compute the intersection of $|\mathcal{A}_{\to j,\mathrm{NLOS}}|$ convex polygons, two polygons are intersected at a time using the Sutherland-Hodgman algorithm. Subsequently, the resulting polygon is intersected with one of the $|\mathcal{A}_{\to j,\mathrm{NLOS}}| - 2$ remaining polygons. The order of intersections is arbitrary. The procedure is repeated until all polygons are intersected. An example of *polygon intersection* is depicted in Fig. 5.15. The pseudo-code for polygon intersection is given in Algorithm 3. Note that beyond the first iteration ($l > 1$), every agent j will also take the polygons of neighboring agents into account and the number of polygons to be intersected increases from $|\mathcal{A}_{\to j,\mathrm{NLOS}}|$ to $|\mathcal{S}_{\to j,\mathrm{NLOS}}|$.

Sutherland-Hodgman Algorithm

The intersection of two polygons can be determined efficiently using the Sutherland-Hodgman algorithm. The Sutherland-Hodgman algorithm selects a convex clipping polygon and a subject polygon; the choice of the clipping polygon is arbitrary if both polygons are convex. All vertices of the subject polygon are added to an input list. The vertices of this list are updated as the edges of the clipping polygon are considered subsequently. Suppose that $\mathcal{V}_{ij}^{(l)}$ is the clipping polygon and it is given by a set of vertices $(\mathbf{v}_{ij,m}^{(l)}, m = 1, ..., N_\mathrm{E})$ or by the intersection of the halfspaces

$$\left\{\mathbf{p}_j | \mathbf{a}_{\mathcal{V}_{ij}^{(l)},m}^\mathrm{T} \mathbf{p}_j \leq c_{ij,m}^{(l)}\right\} \quad m = 1, ..., N_\mathrm{E}, \tag{5.22}$$

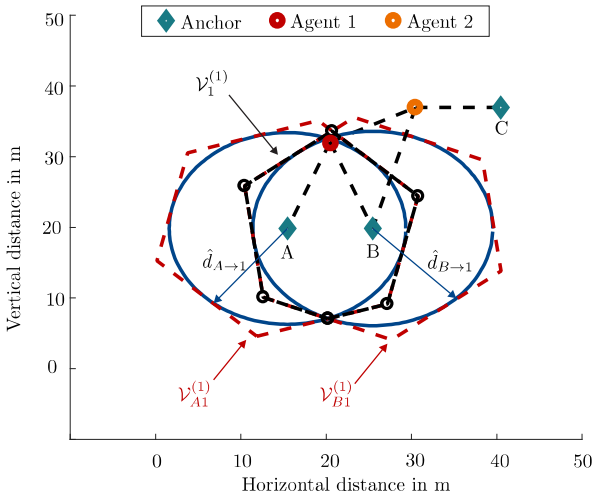

Figure 5.15: *Polygon intersection* - the intersection of the polygons $\mathcal{V}_{A1}^{(1)}$ and $\mathcal{V}_{B1}^{(1)}$ yields the polygon, $\mathcal{V}_1^{(1)}$, which tightly outer-approximates the feasible set (eye-shaped region). Recall that in the first iteration only anchor polygons are considered. Note that increasing the number N_{E} of polygon edges tightens the outer-approximation.

with outward normal vector $\mathbf{a}_{\mathcal{V}_{ij}^{(l)},m} \perp (\mathbf{v}_{ij,m}^{(l)} - \mathbf{v}_{ij,(m+1)\mathrm{mod}N_{\mathrm{E}}}^{(l)})$ and $c_{ij,m}^{(l)} = \mathbf{a}_{\mathcal{V}_{ij}^{(l)},m}^{\mathrm{T}} \mathbf{v}_{ij,m}^{(l)}$. Considering the representation with halfspaces, polygon intersection starts by selecting an arbitrary halfspace $\left\{ \mathbf{p}_j | \mathbf{a}_{\mathcal{V}_{ij}^{(l)},m}^{\mathrm{T}} \mathbf{p}_j \leq c_{ij,m}^{(l)} \right\}$ of the clipping polygon. All vertices of the subject polygon that reside inside the halfspace are added to an output list. Vertices outside of the halfspace are not added. Then, the edges of the subject polygon are traversed. New vertices are added to the output list if an edge of the subject polygon intersects with the hyperplane $\left\{ \mathbf{p}_j | \mathbf{a}_{\mathcal{V}_{ij}^{(l)},m}^{\mathrm{T}} \mathbf{p}_j = c_{ij,m}^{(l)} \right\}$. After the entire subject polygon has been traversed, the next halfspace $\left\{ \mathbf{p}_j | \mathbf{a}_{\mathcal{V}_{ij}^{(l)},n}^{\mathrm{T}} \mathbf{p}_j \leq c_{ij,n}^{(l)} \right\}, n = (m+1)\mathrm{mod}N_{\mathrm{E}}$ of the clipping polygon is considered. The output list of the previous halfspace m constitutes the input list to the next halfspace n. After every halfspace of the clipping polygon has been considered, the Sutherland-Hodgman algorithm terminates and the resulting output list yields the intersection of the subject and clipping polygon.

It is worthwhile mentioning that the intersection of two convex sets is again convex [BV04], i.e. polygon intersection preserves convexity. Also note that the number $\tilde{N}_{\mathrm{E}}$ of vertices of the resulting polygon $\mathcal{V}_j^{(l)}$ generally differs from the original number N_{E} of vertices. Empirical observations have shown that the number $\tilde{N}_{\mathrm{E}}$ of vertices after intersection is typically smaller than the original number N_{E} of vertices and the original number of vertices is only rarely exceeded. Numerical results which support this claim are provided in Section 5.6.

Remark: The polygon intersection operation is identical in every iteration. However, in

the first iteration only anchor polygons $\mathcal{V}_{ij}^{(1)}$, $\forall i \in \mathcal{A}_{\rightarrow j,\mathrm{NLOS}}$ are considered, while polygons from anchors and other agents $\mathcal{V}_{ij}^{(l)}$ $\forall i \in \mathcal{S}_{\rightarrow j,\mathrm{NLOS}}$ are considered in subsequent iterations.

Algorithm 3 Polygon Intersection

given $\mathcal{V}_{ij}^{(l)}$ $\forall i \in \mathcal{S}_{\rightarrow j,\mathrm{NLOS}}$
select index of the first neighbor $i = \mathcal{S}_{\rightarrow j,\mathrm{NLOS}}(1)$
initialize $\mathcal{V}_{j}^{(l)} = \mathcal{V}_{ij}^{(l)}$
for $v = 2$ to $|\mathcal{S}_{\rightarrow j,\mathrm{NLOS}}|$ **do**
    select index of the next neighbor
    $i = \mathcal{S}_{\rightarrow j,\mathrm{NLOS}}(v)$
    intersect two polygons according to [SH74]
    $\mathcal{V}_{j}^{(l)} = \textit{Sutherland-Hodgman}(\mathcal{V}_{j}^{(l)}, \mathcal{V}_{ij}^{(l)})$
end for

5.4.2.3 Polygon Scaling

From the second iteration onward, neighboring agents are also considered. In contrast to an anchor, an agent i cannot confine its location to its exact position. The position of agent $i \in \mathcal{S}_{\rightarrow j,\mathrm{NLOS}} \backslash \mathcal{A}_{\rightarrow j,\mathrm{NLOS}}$ is confined to the polygon $\mathcal{V}_{i}^{(l-1)}, l \geq 2$. In order to generate an additional constraint, the polygon $\mathcal{V}_{i}^{(l-1)}$ of agent i has to be extended by the range estimate $\hat{d}_{i\rightarrow j}$ (see Fig. 5.16). To ensure that agent j resides inside the extended polygon $\mathcal{V}_{ij}^{(l)}$, all edges of the polygon $\mathcal{V}_{i}^{(l-1)}$ have to be shifted by $\hat{d}_{i\rightarrow j}$ in the direction of the outward pointing normal vector $\mathbf{a}_{\mathcal{V}_{i}^{(l-1)},m}$ (orange arrows in Fig. 5.16) of the respective edge m. Recall that convex polygons can be represented by halfspaces (gray dashed lines in Fig. 5.16) according to (5.22). The scaled polygon is given by the intersection of the shifted halfspaces (dotted blue lines in Fig. 5.16). These shifted halfspaces are obtained by manipulating the right-hand side of all inequalities in (5.22). Since $c_{i,m}^{(l-1)} = \mathbf{a}_{\mathcal{V}_{i}^{(l-1)},m}^{\mathrm{T}} \mathbf{v}_{i,m}^{(l-1)}$, where $\mathbf{v}_{i,m}^{(l-1)}$ is any point on the corresponding hyperplane of $\mathcal{V}_{i}^{(l-1)}$, the shifted halfspace is obtained by shifting $\mathbf{v}_{i,m}^{(l-1)}$ to

$$\bar{\mathbf{v}}_{i,m}^{(l)} = \hat{d}_{i\rightarrow j} \cdot \mathbf{a}_{\mathcal{V}_{i}^{(l-1)},m} + \mathbf{v}_{i,m}^{(l-1)}. \tag{5.23}$$

The point $\bar{\mathbf{v}}_{i,m}^{(l)}$ in (5.23) yields a point on the m^{th} shifted hyperplane. Hence the m^{th} shifted halfspace is given by

$$\left\{ \mathbf{p}_j | \mathbf{a}_{\mathcal{V}_{i}^{(l-1)},m}^{\mathrm{T}} \mathbf{p}_j \leq \bar{c}_{i,m}^{(l)} = \mathbf{a}_{\mathcal{V}_{i}^{(l-1)},m}^{\mathrm{T}} \bar{\mathbf{v}}_{i,m}^{(l)} \right\}. \tag{5.24}$$

The scaled polygon $\mathcal{V}_{ij}^{(l)}$ is fully described by the set of shifted halfspaces.

In order to obtain a list of vertices of the polygon $\mathcal{V}_{ij}^{(l)}$, which is required by the subsequent *polygon intersection* operations, the intersections of adjacent hyperplanes have to be determined, i.e., for the m^{th} hyperplane, the intersection points with the $(m-1)^{\mathrm{th}}$ and $(m+1)^{\mathrm{th}}$ hyperplane are computed. These intersection points (orange circles in Fig. 5.16)

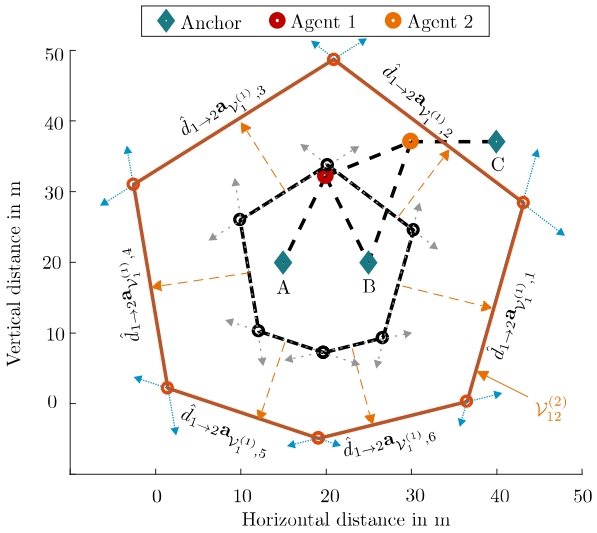

Figure 5.16: *Polygon scaling* - the polygon $\mathcal{V}_1^{(1)}$ (black dash-dotted lines) of agent 1 is received by agent 2 and it is scaled by the distance estimate $\hat{d}_{1\to 2}$. The received polygon $\mathcal{V}_1^{(1)}$ is given by $N_{\mathrm{E}} = 6$ halfspaces (gray dashed arrows), with outward normal vectors $\mathbf{a}_{\mathcal{V}_1^{(1)},i}, i = 1, ..., 6$ (dashed orange arrow). To scale the received polygon, each halfspace is shifted by $\hat{d}_{1\to 2}$ in the direction of its normal vector. The intersections of adjacent shifted halfspaces (blue dotted arrows) yield the vertices (orange circles) of the scaled polygon $\mathcal{V}_{12}^{(2)}$ (solid orange lines).

constitute the vertices of the scaled polygon $\mathcal{V}_{ij}^{(l)}$. Algorithm 4 shows the pseudo-code for polygon scaling.

5.4.3 Discussion on the Robustness

Note that the proposed polygon outer-approximation method does not fail, as long as the NLOS or OLOS links are detected correctly and the NLOS or OLOS bias $b_{i\to j}$ in (5.2) is larger than the error $n_{i\to j}$ due to thermal noise. i.e. $b_{i\to j} \geq -n_{i\to j}$. Whenever these two assumptions hold, the range estimates over-estimate the true distances and POA produces polygons which are guaranteed to contain the positions of the respective agents. In the unlikely cases of miss-classification of the LOS condition or if the ranging error of an NLOS link is negative, there is no guarantee that the POA algorithm determines polygons within which agents reside with certainty.

If the range estimate of a link under-estimates the true distance significantly and this link is used in the POA algorithm due to LOS miss-classification, the polygon intersection step in Algorithm 3 may return an empty set. This would lead to a failure of the POA algorithm. In order to improve the robustness of POA, additional precautions can be taken. In particular, a positive constant κ could be added to all range estimates to avoid failure of POA due to miss-classification of the LOS condition or large negative noise spikes. In other

Algorithm 4 Polygon Scaling

given $\mathcal{V}_i^{(l-1)}, \hat{d}_{i\to j}, i \in \mathcal{S}_{\to j,\mathrm{NLOS}} \backslash \mathcal{A}_{\to j,\mathrm{NLOS}}$
for $m = 1$ to N_E **do**
  shift hyperplane according to (5.23)
  $\bar{\mathbf{v}}_{i,m}^{(l)} = \hat{d}_{i\to j} \cdot \mathbf{a}_{\mathcal{V}_i^{(l-1)},m} + \mathbf{v}_{i,m}^{(l-1)}$
  determine shifted halfspace according to (5.24)
  $\left\{\mathbf{p}_j | \mathbf{a}^\mathrm{T}_{\mathcal{V}_i^{(l-1)},m} \mathbf{p}_j \leq \bar{c}_{i,m}^{(l)} = \mathbf{a}^\mathrm{T}_{\mathcal{V}_i^{(l-1)},m} \bar{\mathbf{v}}_{i,m}^{(l)}\right\}$
end for
determine intersection of adjacent hyperplanes
$\left\{\mathbf{p}_j | \mathbf{a}^\mathrm{T}_{\mathcal{V}_i^{(l-1)},m} \mathbf{p}_j = \bar{c}_{i,m}^{(l)}\right\} \cap \left\{\mathbf{p}_j | \mathbf{a}^\mathrm{T}_{\mathcal{V}_i^{(l-1)},m+1} \mathbf{p}_j = \bar{c}_{i,m+1}^{(l)}\right\}$
$\to \mathbf{v}_{ij,m}^{(l)},\ m = 1, ..., N_\mathrm{E}$
determine scaled polygon $\mathcal{V}_{ij}^{(l)} = \left\{\mathbf{v}_{ij,1}^{(l)}, ..., \mathbf{v}_{ij,N_\mathrm{E}}^{(l)}]^\mathrm{T}\right\}$

words, one could add a safety margin to each range estimate in order to assure that the distance estimates used for POA over-estimate the true distances. To over-estimate the true distance with high probability, the safety margin κ could be chosen based on the standard deviation σ of the ranging error n due to thermal noise, e.g. $\kappa = 3\sigma$. The rationale behind this choice is the following: Suppose that a LOS link is miss-classified as a NLOS link. In that case, the distance estimate of this link could under-estimate the true distance due to the effect of thermal noise. If the ranging error n is zero-mean Gaussian distributed with standard deviation σ, the probability that the sum of the distance estimate $\hat{d} = d + n$ and safety margin κ over-estimates the true distance is given by $\Pr(\hat{d} + \kappa > d) = 1 - \mathrm{erfc}(\kappa/\sigma)$. For $\kappa = 3\sigma$, this probability is $\Pr(\hat{d} + 3\sigma > d) \approx 1 - 2.2 \cdot 10^{-5}$. The introduction of such a safety margin improves the integrity of the POA algorithm. Note, however, that such a safety margin leads to polygons of increased size. Since, in case of UWB, the standard deviation of the ranging error due to thermal noise is typically small [MPG15] (few centimeters), the increase of the sizes of the polygons is insignificant, but the robustness benefits substantially from the safety margin. This will be shown numerically in Section 5.6.

5.5 Two-Phase Localization

The importance of good proposal distributions has been outlined numerous times throughout this thesis. Especially for practical sample sizes, smartly chosen proposal distributions matter. In the previous subsection, a framework has been presented that turns adverse effects such as NLOS and OLOS propagation into useful constraints on the sample space. These constraints provide valuable information regarding the positions of nodes and have the potential to improve the performance of position estimators considerably.

The approach that is considered in this thesis consists of a two-phase approach as depicted in Fig. 5.1. In the first phase, POA is employed in order to derive the polygonal constraints on the positions of the agents. The resulting constraints $\mathcal{V}_j^{(N_\mathrm{FS})}, \forall j$ are fed into the position estimator together with all range measurements $\hat{d}_{i\to j}, \forall i, j$. The position estimator considered in this thesis is a sample-based belief propagation estimator as described

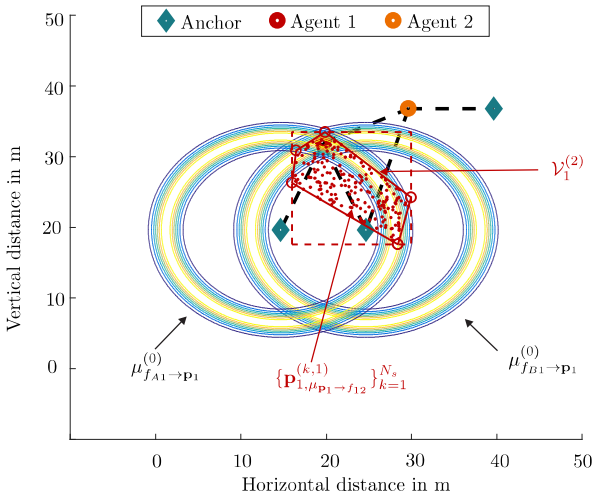

Figure 5.17: *Proposal distribution* - samples $\{\mathbf{p}^{(k,1)}_{1,\mu_{\mathbf{p}_1\to f_{12}}}\}_{k=1}^{N_s}$ drawn according to the proposed proposal distribution $q_1(\mathbf{p}_1)$. With constrained proposal distributions, the relevant region (close to the true location) can be confined tightly, leading to an efficient utilization of the samples.

in section 5.3.2. It is worth noting that this choice is by no means unique. Synergies with other estimators such as weighted least squares (WLS) or parametric belief propagation estimators can be expected.

In case of the considered sample-based belief propagation (SBP) estimator, the constraints are embedded by constraining the proposal distributions from which the samples are drawn. As discussed in Section 5.3.2, it is a common choice to take the samples from one of the incoming BP messages. These samples are usually not concentrated in the regions around the true locations of nodes. By introducing constrained proposal distributions whose supports are given by the respective polygons determined by POA, the samples are inherently in the regions of the sample space with significant probability mass. Hence most particles end up having considerable weights which leads to increased estimation accuracy compared to state-of-the-art proposal distributions. By comparing Fig. 5.7 and 5.17, it can be seen that the constraints on the proposal distribution allow for a better utilization of the samples.

The polygons provide constraints on the sample space, but no favored region within the constrained area can be concluded. Hence it is proposed to draw the samples uniformly within the polygons. Mathematically, the j^{th} proposal distribution is given by

$$q_j(\mathbf{p}_j) = \begin{cases} \frac{1}{A_{P,j}^{(N_{\text{FS}})}}, & \mathbf{p}_j \in \mathcal{V}_j^{(N_{\text{FS}})} \\ 0, & \text{otherwise}, \end{cases} \tag{5.25}$$

where $\mathcal{V}_j^{(N_{\text{FS}})}$ is determined as described in Section 5.4 and $A_{P,j}^{(N_{\text{FS}})}$ is the area of the j^{th} polygon. In order to draw samples uniformly inside a polygon, acceptance-and-rejection

sampling is used. In particular, samples are drawn uniformly from a rectangle which comprises the entire polygon $\mathcal{V}_j^{(N_{\mathrm{FS}})}$. A sample $\mathbf{p}_{j,\mu_{\mathbf{p}_j \to f_{ij}}}^{(k,l)}$ is accepted if $\mathbf{p}_{j,\mu_{\mathbf{p}_j \to f_{ij}}}^{(k,l)} \in \mathcal{V}_j^{(N_{\mathrm{FS}})}$. The area of the rectangle should be small in order to achieve a high acceptance rate. The acceptance rate is given by the ratio of the areas of the polygon and rectangle

$$R_a = \frac{A_{P,j}^{(N_{\mathrm{FS}})}}{A_{R,j}^{(N_{\mathrm{FS}})}}, \tag{5.26}$$

where $A_{R,j}^{(N_{\mathrm{FS}})}$ is the area of the rectangle. Consequently. N_s/R_a samples have to be drawn on average in order to obtain N_s accepted samples.

Fig. 5.17 depicts an example in which samples are drawn from the constrained proposal distribution $q_1(\mathbf{p}_1)$. The figure shows the polygon $\mathcal{V}_1^{(N_{\mathrm{FS}})}$ of agent 1 (solid red line), the rectangle for acceptance-and-rejection sampling (red dashed box), the samples $\left\{\mathbf{p}_{1,\mu_{\mathbf{p}_1 \to f_{12}}}^{(k,l)}\right\}_{k=1}^{N_s}$ (red dots), and the messages from anchor A and B (contour plots). In the next section, an exhaustive numerical analysis will be presented in which both the POA algorithm will be analyzed in isolation and the combination of the POA and SBP algorithms will be investigated.

5.6 Numerical Results

This section contains the numerical evaluation of the proposed two-phase localization framework. The section is clustered into three subsections. In the first subsection, the simulation setup and performance metrics will be discussed. Subsequently, the POA algorithm will be analyzed. The results of the subsection will be used to determine meaningful parameters for computing the constrained proposal distribution (like the number N_{E} of polygon edges and the number N_{FS} of POA iterations). The third subsection focuses on the analysis of the proposed two-phase approach.

5.6.1 Simulation Setup and Performance Measures

As reference topology, a common topology from the literature [VWS12, WLW09, LFS+12] is used. This topology emulates a large-scale ultra-wideband network in a 100m x 100m plane, with 100 uniformly distributed agents and 13 fixed anchors. A circular communication range of $r_{\mathrm{com}} = 20$ m is assumed. An exemplary scenario is depicted in Fig. 5.2. The NLOS or OLOS state of a link is randomly chosen according to a Bernoulli distribution with NLOS and OLOS probability ρ_{NLOS}. During the analysis, different values of ρ_{NLOS} are considered, ranging from $\rho_{\mathrm{NLOS}} = 0.6$ to $\rho_{\mathrm{NLOS}} = 1$. This range reflects typical indoor scenarios. The ranging errors are distributed according to (5.3) with the following parameters: Mean ranging error $1/\lambda = 0.38$ m of NLOS and OLOS links and standard deviation $\sigma = 0.05$ m of LOS links.

During the first part of the analysis, the key metric will be the size of the constrained region which is measured in m^2. Regarding the analysis of the two-phase approach, the localization accuracy will be measured in terms of the average localization error. Additionally, the complementary cumulative distribution function (CCDF) of the localization errors will be investigated. The localization error e_j of agent j is measured according to

$e_j = \|\hat{\mathbf{p}}_{j,\mathrm{MMSE}}^{(l)} - \mathbf{p}_j^*\|$, where $\hat{\mathbf{p}}_{j,\mathrm{MMSE}}^{(l)}$ is the MMSE estimate of the belief in the l^{th} iteration. For statistical significance, $N_{\mathrm{MC}} = 200$ random network topologies (positions of agents vary randomly, while anchors are fixed) are considered and position estimates are collected at every iteration for every agent.

In order to assess the complexity of the sample-based BP algorithm, two analyses will be conducted. First, the number of operations required for each variant of the algorithm is considered according to the big-$\mathcal{O}$ notion. Then, the actual computation time which is required will be measured. All computations of the simulation are performed on an Intel i7-5820k desktop CPU which was exclusively dedicated to the simulation. Three different aspects of the computation time are considered; the time required to determine the polygons, the time required to determine converged location estimates, and the accumulated time. Note that the time which is required to determine converged location estimates depends on the number of iterations required to achieve convergence[13].

5.6.2 Polygon Outer-Approximation

The following aspects are investigated:

1. The influence of the number N_{E} of polygon vertices on the area of the polygon and the speed of convergence.
2. The dependence of the polygon area on the number of NLOS and OLOS links.
3. The computation time of the algorithm.
4. The robustness to negative range estimates.

For comparison, the ellipse outer-approximation algorithm which was presented in [GWG13, GWSR11a, Gho11] is considered as a benchmark.

Aspect 1): Fig. 5.18 depicts the polygon area in m^2 against the number N_{FS} of iterations for a fixed NLOS and OLOS probability of $\rho_{\mathrm{NLOS}} = 1$. Recall that it is desirable to have small polygons in order to tightly constrain the positions of agents. It can be obtained from Fig. 5.18 that the sizes of the polygons decrease, as the number N_{E} of vertices increases. Increasing the number of vertices beyond $N_{\mathrm{E}} > 16$ adds no additional area reduction. Since additional vertices increase the complexity and communication overhead, but they do not provide smaller polygons, $N_{\mathrm{E}} = 16$ vertices will be considered for all further analysis. In addition, the largest area reduction is achieved in the 2^{nd} iteration. The reduction of polygon area from the 2^{nd} to the 3^{rd} iteration is only minor. Using more than $N_{\mathrm{FS}} = 3$ iterations does not decrease the polygon area notably. In order to limit the effort of deriving the polygons that are required for the constrained proposal distributions, $N_{\mathrm{FS}} = 3$ iterations are considered.

Aspect 2): The average polygon and ellipse area are depicted versus the NLOS and OLOS probability in Fig. 5.19. Three curves are shown: The ellipse-based, the polygon-based, and the robust polygon-based approach. In case of the robust polygon-based approach, a safety margin of $\kappa = 3\sigma$ is added to every range estimate before performing POA. Recall that only NLOS and OLOS links are used to confine the locations of nodes. For instance, if the NLOS and OLOS probability is $\rho_{\mathrm{NLOS}} = 0.6$, only 60% (on average) of the

[13]The following definition of convergence is used. Convergence is achieved if the average localization error does not change by more than 5% in the successive iteration.

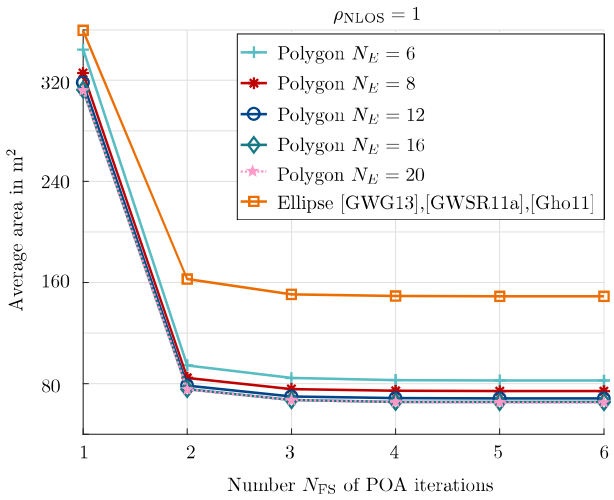

Figure 5.18: *Average polygon/ellipse area versus the number of iterations* - It becomes evident that after a few iterations (2-3) the average polygon and ellipse area does not decrease significantly. Moreover, a small number of polygon edges ($N_{\mathrm{E}} = 16$) is sufficient to achieve tight polygons.

links are available to determine the polygons or ellipses. It can be seen that the average sizes of the polygons and ellipses increase as the NLOS and OLOS probability decreases. The reason is that fewer links are used to confine the locations of nodes. It is important

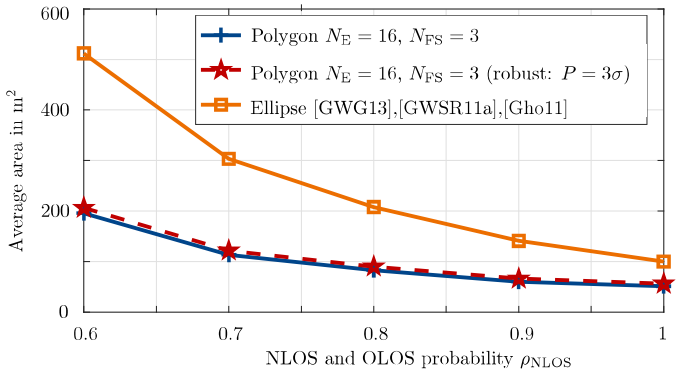

Figure 5.19: *Average polygon/ellipse area versus the NLOS and OLOS probability* - as the NLOS and OLOS probability decreases, fewer links are used to confine the positions of nodes and the average area increases. The increase of the constrained area of ellipses is steeper than that of polygons, meaning that the superiority of polygons over ellipses increases with decreasing ρ_{NLOS}.

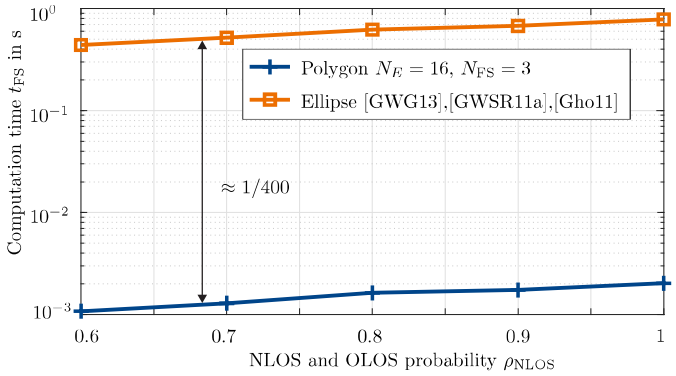

Figure 5.20: *Computation time versus the NLOS probability* - the time that an agent requires to compute its polygon is only a fraction of the time required for ellipses ($\approx 1/400$). As the number of NLOS and OLOS links increases, an agent has to perform more operations leading to increasing computation time t_{FS}.

to note that the average area of ellipses shows a steeper increase compared to polygons as the number of NLOS and OLOS links decreases. For instance, for $\rho_{\mathrm{NLOS}} = 0.6$, ellipses are approximately three times larger than polygons, while, for $\rho_{\mathrm{NLOS}} = 1$, ellipses are only two times larger than polygons. Another important aspect is that the increase of the average polygon size due to the safety margin of $\kappa = 3\sigma$ is only marginal. Hence increasing the robustness of POA comes at a small price in terms of the area gain.

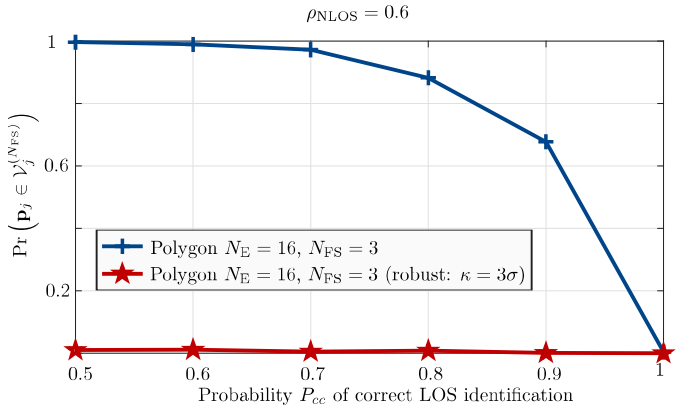

Figure 5.21: *Failure probability of POA versus the probability of correct NLOS and OLOS identification* - the probability that POA fails due to miss-classification of the NLOS and OLOS condition is moderately low. The robust version of POA with a safety margin of $\kappa = 3\sigma$ is very resilient to range measurements with negative ranging errors which are included in the algorithm due to miss-classification of the NLOS and OLOS condition.

Aspect 3): Fig. 5.20 depicts the computation time that an agent needs to determine its polygon or ellipse. The computation time is shown versus the NLOS and OLOS probability ρ_{NLOS}. The computation of polygons is substantially faster when compared to ellipses. The reason is that the computation of ellipses requires solving convex optimization problems, while the operations required to compute polygons are much simpler (like determining the intersection of two lines). For instance, if $\rho_{\text{NLOS}} = 0.6$, an agent needs approximately $t_{\text{FS}} = 1$ ms to compute its polygon, while approximately $t_{\text{FS}} = 400$ ms are required to determine its ellipse. It can also be seen that the computation times of both approaches increase with increasing NLOS and OLOS probability, since more NLOS links have to be taken into consideration.

Aspect 4): Fig. 5.21 depicts the probability that agents do not reside inside of their polygons against the probability of correct identification of the LOS condition. Classifications of links are performed in a Bernoulli-random fashion, i.e. a link is either classified correctly as LOS with probability P_{cc} or incorrectly classified as LOS with probability $(1 - P_{cc})$. Two curves are depicted; one for the standard POA algorithm and one for the robust POA algorithm with a safety margin of $\kappa = 3\sigma$. It can be observed that almost all agents reside inside of their polygons regardless of the number of classification errors. Hence the robust POA algorithm shows high resilience against miss-classification of the LOS condition. On the other hand, the standard version of POA is relatively susceptible to incorrect LOS detections. More and more agents reside outside of their polygons if the probability of correct NLOS and OLOS identification decreases. The following conclusion can be drawn from these observations: In practical deployments, where the NLOS and OLOS states are unknown and have to be detected, the robust variant of the POA algorithm should be considered. It provides high resilience to incorrect NLOS and OLOS identifications (see Fig. 5.21) and increases the constrained area only moderately (see Fig. 5.19).

5.6.3 Novel Proposal Distribution for Belief Propagation

The impact of the proposal distributions with polygonal constraints on sample-based BP is studied in this subsection. For comparability, the following benchmarks are considered

1. [SZ13]: SBP with sampling based on an auxiliary variable ('SBP aux.')
2. [LFS$^+$12]: SBP with sampling based on the incoming message with the lowest entropy ('SBP min.')
3. [MB17]: SBP with sampling based on elliptical constraints ('SBP ell.')
4. [LFS$^+$12]: parametric belief propagation ('PBP')
5. [SRL02]: weighted least squares ('WLS')

Key metrics of the analysis are (i) the speed of convergence, (ii) localization accuracy, (iii) the computation time, and (iv) the number of samples. For now, it is assumed that the NLOS and OLOS probability is $\rho_{\text{NLOS}} = 1$, i.e. all links are in NLOS condition. Lower NLOS probabilities resulting in mixed LOS/NLOS and OLOS conditions are discussed in Subsection 5.6.3.4. In the first part of the analysis, $N_{\text{s}} = 1000$ particles are used in all SBP variants.

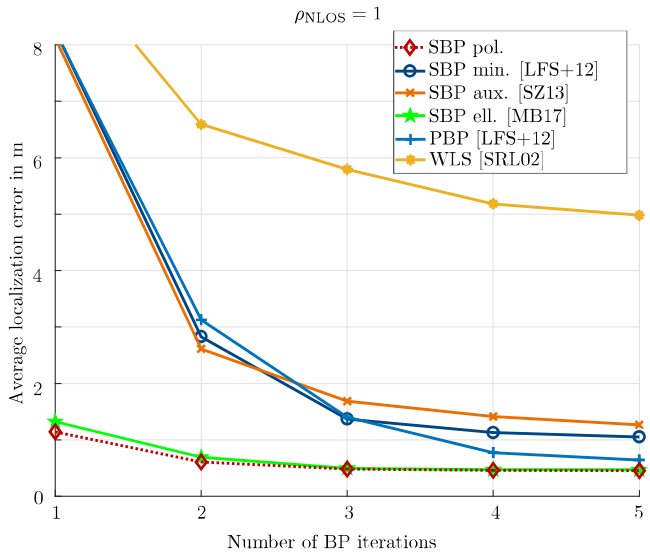

Figure 5.22: *Average localization error* - the average localization error is depicted against the number of iterations. The proposed SBP variant converges significantly faster compared to most baseline algorithms and shows the highest localization accuracy.

5.6.3.1 Convergence and Accuracy

Fig. 5.22 depicts the average localization error against the number of iterations. The proposed variant of SBP which employs polygonal constraints is termed 'SBP pol.'. From Fig. 5.22, two benefits of the polygonal constraints can be inferred. First, fast convergence ($N_{\mathrm{BP}} = 3$ iterations) is achieved. By employing the polygonal constraints on the proposal distributions, prior information is generated that accelerates convergence compared to the baseline approaches. Only the ellipse-based approach from [MB17] can compete, since the ellipses generate similar prior information compared to the proposed approach. Secondly, the proposed method achieves the highest localization accuracy among all competitors. Hence there is a two-fold benefit of incorporating the proposal distribution presented in Section 5.5. Those two advantages can be traced back to the following reasons. Quick convergence is achieved, since the polygonal constraints already restrict the possible locations and, after the first iteration, the beliefs are already concentrated in the regions close to the true locations of agents. High localization accuracy is achieved, since samples reside in the important regions of the sample space. Unlike in [SZ13, LFS$^+$12], many particles have significant weights and contribute to the location estimate.

The CCDF of the localization error allows for a more insightful analysis of the accuracy than the average localization error. Fig. 5.23 depicts the CCDFs of the localization errors after convergence. Most baseline approaches ([SZ13] 'SBP aux.', [LFS$^+$12] 'SBP min.', and [LFS$^+$12] PBP) require $N_{\mathrm{BP}} = 5$ iterations to converge, while [SRL02] 'WLS' needs ten iterations (not shown in Fig. 5.22). The proposed approach (SBP 'pol.') and [MB17] 'SBP ell.' converge after $N_{\mathrm{BP}} = 3$ iterations. It can be observed from Fig. 5.23 that the CCDF

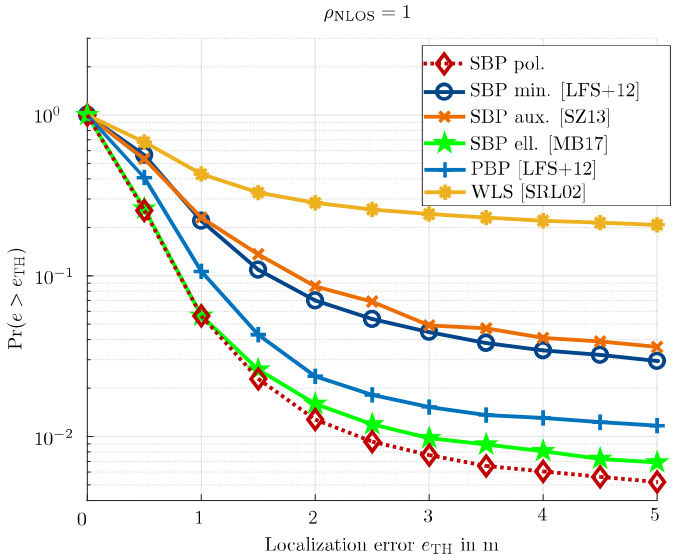

Figure 5.23: *CCDFs of the localization errors of all considered approaches* - SBP with polygonal constraints outperforms all benchmark methods in terms of the localization accuracy.

	Algorithms					
	SBP min.	SBP aux.	SBP ell.	PBP	WLS	SBP pol.
t_{FS} in s	-	-	0.7850	-	-	0.0020
t_{conv} in s	16.7189	16.8082	10.2384	0.6253	0.0618	10.9575
t_{c} in s	16.7189	16.8082	11.0234	0.6253	0.0618	10.9575

Table 5.1: *Computation time* - average time per node to achieve convergence.

of 'SBP pol.' decreases rapidly in the regime of small errors and especially, large errors can be mitigated better compared to the baseline approaches. With the help of polygonal constraints, all baseline approaches are outperformed. SBP with elliptical constrains is the strongest competitor in terms of accuracy.

5.6.3.2 Computation Time

Table 5.1 depicts the average accumulated computation time t_{c} per agent. The accumulated computation time t_{c} is broken up into the time t_{FS} that is required to compute the polygons or ellipses and the time t_{conv} required to achieve convergence with the respective localization algorithm t_{conv}. The sum of both quantities yields the total computation time $t_{\mathrm{c}} = t_{\mathrm{FS}} + t_{\mathrm{conv}}$. The convergence observations from the previous discussion are employed to determine t_{conv}.

In terms of computation time, WLS shows the lowest cost followed by PBP and the proposed polygon-based SBP. The following conclusions can be drawn

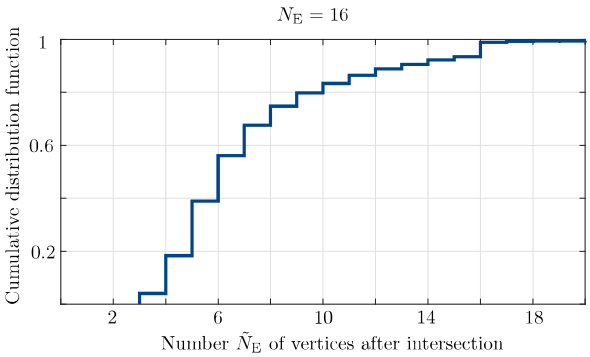

Figure 5.24: *CDF of the number $\tilde{N}_E$ of vertices after polygon intersection* - after intersecting polygons with N_E edges, the resulting polygons have usually fewer edges, i.e. $\tilde{N}_E < N_E$.

1. The computation time for POA is almost negligible compared to that of SBP.
2. The increased speed of convergence of SBP with polygonal constraints reduces computation time t_c significantly compared to the existing SBP variants.
3. Compared to WLS and PBP, the average accumulated computation time t_c of the proposed approach is still relatively high[14].

To gain some more insight into the first observation, it is worthwhile to review the cornerstones of the POA algorithm. Note that all operations in that algorithm can be solved in a closed form. Polygon scaling and anchor polygon processing (Algorithm 2 and 4, respectively) scale linearly with the number of edges, i.e. $\mathcal{O}(N_E)$. In terms of computation, the most demanding part of the algorithm is polygon intersection (Algorithm 3). The Sutherland-Hodgman algorithm scales quadratic in the number N_E of edges [SH74]. Since it intersects only pairs of polygons, the intersection of $|\mathcal{S}_{\to j,\mathrm{NLOS}}|$ polygons requires to execute the Sutherland-Hodgman algorithm $|\mathcal{S}_{\to j,\mathrm{NLOS}}|-1$ times. Recall, though, that the number $\tilde{N}_E$ of edges of intersected polygons depends on the input polygons and it cannot be generalized. Thus, the number of operations required to intersect $|\mathcal{S}_{\to j}|$ polygons cannot be quantified for arbitrary polygons. However, it has been observed empirically that the number $\tilde{N}_E$ of edges of intersected polygons is typically smaller than the number N_E of edges of the polygons to be intersected (see Fig. 5.24). Hence $\tilde{N}_E \approx N_E$ usually overestimates the number of computations. With this assumption, a conservative scaling of $\mathcal{O}(|\mathcal{S}_{\to j,\mathrm{NLOS}}|N_E^2)$ is considered for polygon intersection. Overall, POA scales quadratic in the number of edges and linearly with the number of neighbors, i.e. $\mathcal{O}(|\mathcal{S}_{\to j,\mathrm{NLOS}}|N_E^2)$. Considering that $N_E \ll N_S$, it is obvious that SBP is much more costly in terms of computations than polygon outer-approximation.

[14]It is worth mentioning that no parallel computing was implemented in the simulations to reduce the dependencies on the hardware. However, considerable speed-ups due to parallel computing can be expect for particle-based approaches, since the computations for all particles can be performed in parallel. The actual gain due to parallel computing depends on the hardware architecture on which the algorithms are implemented.

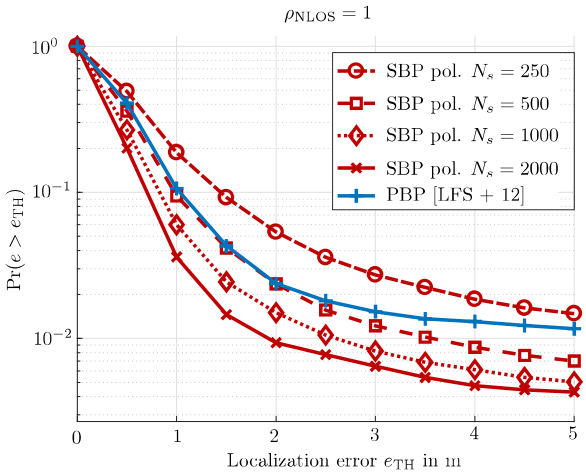

Figure 5.25: *CCDFs of the localization errors for different sample sizes* - reducing the number N_{s} of samples reduces the localization accuracy moderately if polygon-based SBP is considered. With only $N_s = 500$ samples, the PBP approach from [LFS+12] is still outperformed in terms of accuracy.

Observation 2) makes intuitively sense, since reducing the number N_{BP} of iterations reduces the complexity of SBP linearly. The proposed polygon-based and ellipse-based SBP algorithms have reduced computational costs compared to the other SBP benchmarks [LFS+12, SZ13], because both require fewer iterations to converge. In particular, a reduction of approximately 40% ($N_{\mathrm{BP}} = 3$ iterations versus $N_{\mathrm{BP}} = 5$ iterations) is achieved. Note that t_{conv} of the proposed polygon-based SBP approach is slightly higher compared to ellipse-based SBP [MB17]. The reason is that in [MB17] samples are drawn directly, while acceptance-and-rejection sampling is used in order to draw samples that reside inside the polygons.

Finally, it is worthwhile to mention that, although the computation time of the proposed SBP variant is reduced compared to the approaches in [LFS+12, SZ13], the overall computationally complexity is still relative high which reflects the fact that SBP remains a computationally intensive algorithm. On the other hand PBP is less computational complex compared to all SBP variants and embodies an attractive alternative compared to SBP. Yet, PBP shows slightly lower estimation accuracy than the proposed SBP variant.

5.6.3.3 Number of Samples

Fig. 5.25 depicts the CCDFs of the localization errors as various sample sizes are considered. The convergence observations from Subsection 5.6.3.1 are considered, i.e. polygon-based SBP converges within three iterations and PBP requires five iterations to converge. Two observations can be made:

- PBP can perform better than SBP in terms of accuracy if the number of samples is small.

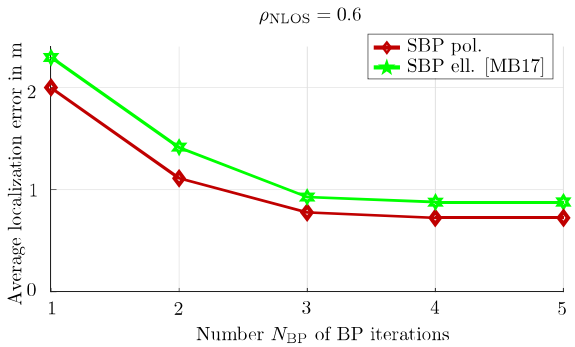

Figure 5.26: *Average localization error* - the proposed polygon-based algorithm outperforms the ellipse-based approach in terms of localization accuracy.

- SBP achieves higher positioning accuracy than PBP if the number of samples is sufficiently large.

For large sample sizes ($N_{\mathrm{s}} \geq 500$), SBP performs better than PBP in the regimes of both small and large errors. Bear in mind, though, that larger sample sizes result in higher computation times. Considering the results from Fig. 5.25 and the previous subsection, the following conclusion regarding the accuracy-computation trade-off can be drawn. For systems that do not aim to maximize the localization accuracy, it can be sufficient to choose PBP for cooperative network localization, since PBP is cheaper in terms of computations. If more computational resources are available, however, polygon-based SBP allows for an increase of the localization accuracy. The achievable localization accuracy of SBP can go far beyond what is possible with PBP if the number of samples is large. Compared to other SBP variants, polygon-based SBP provides a considerably improved accuracy-computation trade-off, i.e. higher localization accuracy and lower computational costs are achieved.

5.6.3.4 Mixed LOS and NLOS Condition

It has been observed in Subsection 5.6.2 that the sizes of the feasible sets increase as the NLOS and OLOS probability decreases. This, in turn, affects the accuracy of the polygon-based and ellipse-based SBP algorithms. The reason is that the sample density in the regions close to the true locations decreases if the number N_{s} of samples remains unchanged. To study the quantitative effects of the increased feasible sets on the localization accuracy, a scenario with a relatively low NLOS and OLOS probability of $\rho_{\mathrm{NLOS}} = 0.6$ is assumed. Fig. 5.26 depicts the average localization error versus the number N_{BP} of iterations. It can be observed that due to the increased polygons and ellipses both approaches converge slower compared to the case if all links are NLOS and OLOS (see Fig. 5.22). In particular, one additional iteration (four in total) is required to achieve convergence. The reason is that the polygons or ellipses provide less prior knowledge, since their sizes are increased. In addition, the advantage of polygons over ellipses becomes more apparent, since the gap between the two approaches widens. In case of $\rho_{\mathrm{NLOS}} = 1$, the average localization errors of polygon-based and ellipse-based SBP are approximately the same (see Fig. 5.22). On the other hand, if $\rho_{\mathrm{NLOS}} = 0.6$, the average localization error of polygon-based SBP is

approximately 20% lower than that of ellipse-based SBP. The reason is that the sizes of the ellipses increase faster than those of polygons as the NLOS and OLOS probability decreases (see Fig. 5.19). As a result, the proposed POA-based approach shows lower localization errors when compared to the ellipse-based approach. This can also be seen from the CCDFs of the localization errors which are depicted in Fig. 5.27. The proposed approach outperforms the ellipses-based approach in the regime of medium to large errors ($e > 0.5$ m).

In conclusion, in practical scenarios, where not all links are in NLOS and OLOS condition, the polygon-based approach performs better than the ellipse-based approach in terms of accuracy, since the polygons grow slower than ellipses as the NLOS and OLOS probability reduces.

5.7 Conclusions

This chapter has discussed the problem of cooperative localization in scenarios with sparse anchor coverage. It has been pointed out that the uncertainties regarding the positions of agents are relatively high if the infrastructure is deployed sparsely. In order to obtain accurate position estimates in such scenarios, the considered localization algorithm must be able to account for these uncertainties. Even though stochastic algorithms meet this requirement, their complexity is often prohibitive. However, the complexity of many of these algorithms could be greatly relaxed if the area of search was constrained to small regions. The obvious question is: *How can such constraints be determined?*

In many practical scenarios, physical constraints are inherently imposed by the propagation environment due to effects like NLOS and OLOS propagation. To that end, a method has been proposed in this chapter which leverages these physical constraints and transforms them into geometrical constraints on the positions of agents. These geometrical constraints are given by convex polygons which give rise to the name of the method: *polygon outer-approximation*. Polygon outer-approximation confines the positions of all agents to polygons in a distributed and iterative manner. It is computationally cheap, since all

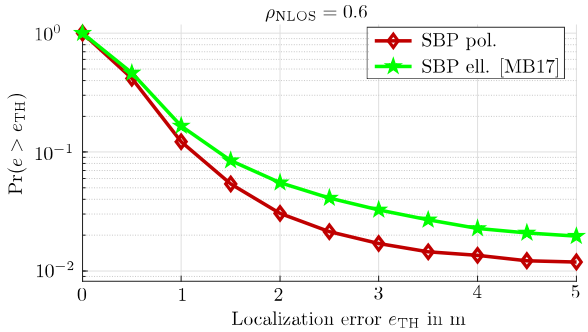

Figure 5.27: *CCDFs of the localization errors for mixed LOS/NLOS condition* - due to the smaller constrained regions of polygons compared to ellipses, SBP achieves higher localization accuracy in the regime of $e_{\mathrm{th}} > 0.5m$.

operations can be computed in a closed form. Consequently, it is a quick and easy way of restricting the area of search.

It was shown through simulations that these constraints can significantly improve the performance of sample-based belief propagation. For instance, they can (i) help to reduce the complexity burden considerably, (ii) speed up convergence, (iii) and increase the localization accuracy. In particular, the proposed polygon-constrained sample-based belief propagation method shows the best complexity-accuracy trade-off among all studied sample-based belief propagation variants. It is worth noting, though, that the overall complexity of the proposed approach may still be too high for low-cost and low-complexity devices. For such devices, parametric belief propagation may be more suitable, since it offers lower complexity at the costs of reduced accuracy. Whenever more computational resources are available, the proposed polygon-based approach is preferable, as it offers superior localization accuracy with tractable computational complexity.

However, it should be emphasized that the sizes of polygons increase steadily as the NLOS and OLOS probability decreases. Thus, in scenarios with extremely low NLOS probability, polygons can be expected to have relatively large sizes. In such cases, sample-based belief propagation with polygonal constraints can loose many of the advantages (fast convergence and high accuracy) mentioned above. Since a relatively large number of NLOS and OLOS links is required in order to determine tight constraints on the locations of agents, the proposed polygon-based cooperative position estimator is particularly suited for localization in indoor environments. Nonetheless, extensions of the proposed POA algorithm (similar to its robust variant) can be applied to include also LOS links. If a positive safety margin, which is large enough to compensate any negative ranging errors, is added to all links, LOS links can also be used to impose constraints on the positions of agents. Such an extension renders POA applicable to scenarios with many LOS links as well.

Chapter 6

Single-Anchor Positioning Part I: Theoretical Insights

It was shown in the previous section that cooperative approaches allow for accurate positioning in scenarios with sparse infrastructure. The enabling feature of cooperation is that every agent has access to a larger number of distinct[1] measurements (distance estimates) when compared to non-cooperative techniques. A natural way of extending this idea is to ask: *How can one increase the number of distinct measurements per anchor?* Ideally, the number of measurements is large enough so that every agent needs to communicate only with a single anchor. One way of obtaining multiple distinct measurements per anchor is to employ large antenna arrays at the transmitter and receiver which allow for accurate angular estimation. It turns out that fifth generation (5G) cellular communication networks bear a huge positioning potential. Especially the millimeter wave (mmWave) massive multiple-input multiple-output (mMIMO) proposal in 5G has been shown to have outstanding positioning capabilities [SGD+18b, ASZA+18b, GGD18a, MWBAS18]. In fact, these studies have revealed that single-anchor positioning becomes possible.

This chapter analyzes the positioning potential of mmWave mMIMO systems theoretically by means of Fisher information. First, a brief introduction will be given and the state of the art will be depicted in Section 6.1. Subsequently, the system model will be discussed in Section 6.2. The following two important cases will be distinguished: (i) Synchronized and (ii) non-synchronized transmitter and receiver clocks. Analytical Fisher information studies for both cases will be conducted in the Sections 6.3 and 6.4 in order to reveal the positioning potential of mmWave mMIMO systems. Numerical examples will be given in Section 6.5 in order to emphasize the findings of the theoretical analysis. Section 6.6 will summarize the results of this chapter.

6.1 Introduction and State of the Art

This section gives a brief introduction to the topic of single-anchor positioning in 5G mmWave mMIMO systems. The state of the art is described subsequently.

[1]The word *distinct* in this context refers to measurements that are obtained under different geometries. For instance, distance measurements with respect to neighboring nodes that reside in distinct positions are called distinct.

6.1.1 Introduction

Accurate radio-based positioning in networks with sparse infrastructure is a notoriously hard challenge. One way of compensating for the lack of anchors is to allow for cooperation among agents as described in Chapter 5. Another way of dealing with low anchor density is based on the concept of single-anchor positioning. Single-anchor positioning is an appealing concept, since, in most networks, coverage is usually assured by at least one anchor. Suppose that no coverage holes exist in a network, then agents could estimate their positions at any point in space without the need for cooperative approaches or multiple anchors. As attractive as the concept sounds, contemporary networks do not provide the capabilities for single-anchor positioning, since the position information (in the Fisher sense) that a single anchor can provide is insufficient for localization. In future 5G mmWave mMIMO networks, however, the tide can turn because a single anchor can provide a considerably larger amount of position information compared to anchors in current radio systems. One might ask: *Where does the potential from 5G mmWave mMIMO systems come from?* A rigorous answer to that question will be given during the course of the chapter.

As a first intuition, it is worthwhile to recall some basic knowledge from the field of radio-based positioning: (i) Large bandwidth is a catalyst for accurate delay or distance estimation [CWW07] and (ii) large antenna arrays are the gateway to precise angle estimation [LCP05, LSS09, HSZ+16a]. Future mmWave mMIMO systems will feature both aspects, since large blocks of bandwidth are available in the mmWave bands and the short wave length of mmWave signals allows for the deployment of a massive number of antennas in relatively small areas. As a consequence, accurate delay and angle estimates will be widely available in mmWave mMIMO systems. Hence 5G has the potential to provide precise position estimates based on a single anchor. Note that the positioning aspect is just a byproduct of 5G systems whose main purpose of deployment is to advance current cellular communication systems. Therefore, the expected cost[2] (in terms of deployment, maintenance, etc.) of positioning in 5G systems will potentially be lower than that of contemporary positioning systems. At the same time, though, the positioning accuracy can be expected to be relatively high. A conceptual cost-accuracy diagram is depicted in Fig. 6.1.

6.1.2 State of the Art

Very recently, single-anchor positioning has been gaining considerable attention. Research in this field has picked up a lot of momentum after early studies [WSGD+17, GWL+17, KHC+17, SGD+18a, ASZA+18a] have outlined the potential of mmWave mMIMO systems for single-anchor positioning. Even though the number of multipath components in mmWave mMIMO systems is usually small, a key observation was that every multipath component has the potential to increase the achievable positioning accuracy [SGD+18a]. This observation deviates from the common belief that multipath propagation is an adverse effect which impairs the precision of location estimates, when no prior information regarding the location of the corresponding points of incidence of non-line-of-sight (NLOS) paths is available [HSZ+16b, SW10a, SWW10, SW10b, WSD18]. The reason for this paradigm shift is that the large sizes of the envisioned antenna arrays in combination with the vast bandwidth envisioned in mmWave mMIMO systems provide substantial amounts of Fisher information. As a result, multipath components become resolvable in the temporal and

[2]This is a vague definition of cost which will only be used to indicate the potential of 5G systems.

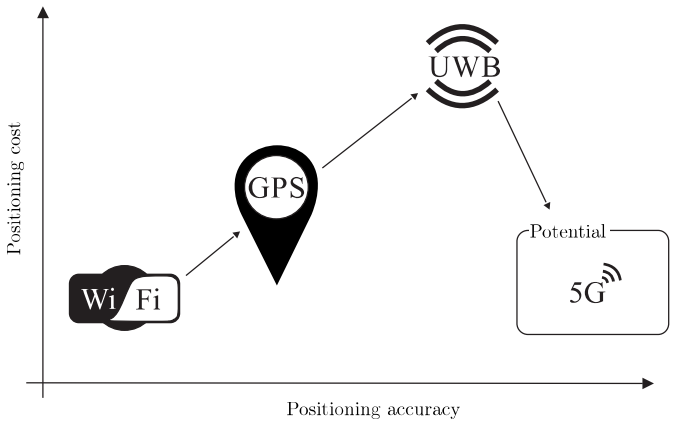

Figure 6.1: *Abstract cost-accuracy relationship of radio-based positioning* - WiFi is ubiquitous and hence can be used for position estimation at low costs. The achievable accuracy, however, is relatively low. GPS is also widely available and the achievable accuracy is higher than that of WiFi. UWB systems have great positioning capabilities, but they are usually deployed in an ad hoc fashion, rendering them costly. 5G has the potential to offer both high positioning accuracy and low positioning cost.

angular domains and hence the angle-of-arrival (AOA), angle-of-departure (AOD), and delays of every NLOS path can be estimated precisely [CW18, YZZ18]. Yet, the impact of multipath components was only studied numerically and there is a clear lack of the theoretical understanding.

Considering the large number of observations (triplets of AOA, AOD, and delay for every path) which a single anchor can generate, the problem of position estimation has been reevaluated and extended by parameters such as the orientation of the agent. The fundamental limits of position and orientation estimation using mmWave mMIMO in 5G have been recently investigated in [GGD18b, SGD$^+$18a, ASZA$^+$18a]. In [GGD18b], a single-anchor localization scheme was considered for indoor scenarios. The Fisher information matrix (FIM)[3] of the position and orientation parameters as well as the FIM of the NLOS parameters were presented. Based on these FIMs, the position error bound (PEB) and orientation error bound (OEB) were derived numerically for different array configurations. It was shown in [GGD18b] that increasing the number of antenna array elements increases the localization accuracy. In [SGD$^+$18a], the FIM of all observable channel parameters was presented. Using the geometric relationship of the channel parameters and the position and orientation-related parameters, the FIM of the position and orientation-related parameters was derived in closed-form. Moreover, the PEB and the OEB were determined numerically, and algorithms which attain the previously determined bounds were also presented. It was shown numerically that even in the absence of the line-of-sight (LOS) path, positioning with reasonable accuracy is possible. In [ASZA$^+$18a], fundamental limits of position and orientation estimation for uplink and downlink in 3D-space were presented. The FIM of the channel parameters was derived in a closed form similar to [SGD$^+$18a], which provided

[3]The achievable accuracy of an estimation problem can be linked to the corresponding FIM via the information inequality.

the FIM of the 2D channel parameters. Moreover, the structure of this FIM was analyzed and it was shown to become block diagonal when the bandwidth and the number of receive and transmit antennas are sufficiently large. In contrast to [SGD+18a], which considered uniform linear arrays, [ASZA+18a] presented the derivation of the PEB and the OEB in closed-form for arbitrary antenna array structures. The PEB and the OEB were derived in a similar manner compared to [SGD+18a]. In addition, the influence of different array types on the PEB and the OEB was investigated. Moreover, differences in the uplink and downlink were also considered in [ASZA+18a].

It is worthwhile to mention that the idea of exploiting multipath information has already been investigated before the advent of 5G [LMR+15, WML+16, GJW+16]. Preceding studies have focused on ultrawideband (UWB) signals [LMR+15, WML+16] and signals of opportunity (SoO) [GJW+16], which embody a class of signals from systems like Long Term Evolution (LTE) and wireless fidelity (WiFi). However, these approaches only exploit delay estimates and thus, they are subject to certain constraints. For instance, mobility of the agent is required to obtain unambiguous estimates, a large number of time steps can be necessary to achieve convergence and consequently, the computational complexity of the algorithms can be burdensome.

6.1.3 Contribution

Two fundamental estimation problems will be studied in this chapter. First, the role of NLOS paths will be investigated analytically. The studies in [SGD+18a, ASZA+18a] have only shown numerically that NLOS paths can increase the positioning accuracy. Important aspects such as the impact of the underlying geometry are still unclear. To that end, the equivalent Fisher information matrix (EFIM) which corresponds to the position and orientation estimation problem in a multipath environment will be derived and decomposed in a closed form. The results of the decomposition will reveal the impact of NLOS paths on the position and orientation estimation problem. Secondly, an extended problem will be analyzed. In the extended problem, it will be shown that an agent can determine its own position, orientation, *and* clock offset[4] based on a single transmission received from an anchor. Simultaneously, the transmitter can obtain a map of the points of incidence in its radio environment. The Fisher information matrix corresponding to this problem will be derived analytically and the respective performance bounds will be computed numerically. The following list contains the main contributions to the first and second problem:

1. Analysis of the role of NLOS paths in mmWave mMIMO
    - An expression for the EFIM of the position and orientation will be derived and it will be shown that this EFIM can be written as the sum of rank one matrices, where each NLOS component contributes a distinct rank one matrix to the EFIM.
    - It will be shown that the contribution from each NLOS component increases the position and orientation information content in the EFIM and thus reduces the PEB and OEB.
    - It will be revealed that NLOS components provide significant position and orientation information if and only if AOA, AOD, and delay can be estimated accurately.

[4]The terms *clock offset* and *clock bias* are used interchangeably.

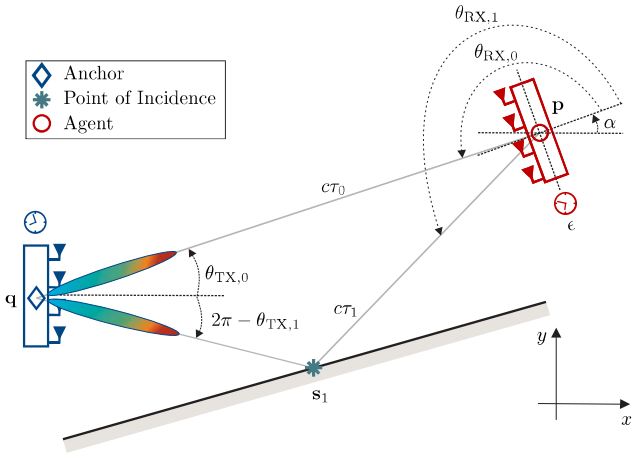

Figure 6.2: *Geometry of the scenario* - an agent with unknown position $\mathbf{p}$, orientation α, and clock offset ϵ attempts to localize itself and determine its orientation using the received signal from the anchor. The anchor has known location $\mathbf{q}$ and orientation ϕ. Single-bounce NLOS paths and a direct path are considered.

- The amount and direction[5] of information will be presented in a closed form, showing its relation to the geometry.

2. Analysis of the estimation capabilities in mmWave mMIMO
   - The FIM of the position, orientation, and clock offset estimation problem will be presented in a closed form.
   - The feasibility of joint position, orientation, clock offset, and map estimation will be shown numerical by showing that the corresponding FIM is invertible. It will be proven that the estimation of the aforementioned quantities becomes feasible if the LOS path and at least one NLOS path exist.

6.2 System Model

In this section, the geometry of the considered problem will be described. Subsequently, the transmitter, channel, and receiver models will be specified.

6.2.1 Geometry

Fig. 6.2 depicts the studied scenario. An agent is considered which aims to estimate its own location $\mathbf{p}$ and orientation α in 2D space, based on the downlink signal received from the anchor. If the agent and anchor are not precisely time-synchronized, the agent also

[5]Since Fisher information is a multidimensional concept, information can be represented as eigenvalue-eigenvector pairs of the (E)FIM [SW10a, SWW10]. The eigenvalue is referred to as the *gain* or *intensity* of information and the angle of the eigenvector as the *direction of information*.

attempts to estimate its clock offset ϵ with respect to the clock of the anchor. The position $\mathbf{q}$ and orientation ϕ of the anchor are perfectly known to the agent. This information can be piggybacked in the header of a package without increasing the overhead notably or via dedicated signals: so-called positioning reference signals (PRSs). Even though the analysis is focused on downlink position and orientation estimation, the estimation could also be carried out in the uplink as presented in [ASZA+18a]. The anchor and agent are both equipped with an array of N_{TX} transmit antennas and N_{RX} receive antennas, respectively. Their arrays have arbitrary but known geometry. The centroids of the anchor's and agent's arrays are located at the positions[6] $\mathbf{q} = [q_{\mathrm{x}}, q_{\mathrm{y}}]^{\mathrm{T}}$ and $\mathbf{p} = [p_{\mathrm{x}}, p_{\mathrm{y}}]^{\mathrm{T}}$, respectively. The known orientation of the anchor's array is denoted by ϕ, while α denotes the unknown orientation of the array of the agent.

6.2.2 Transmitter Model

It is assumed that mmWave signals in combination with massive MIMO technology are employed. For the sake of the analysis, all signals and systems are assumed in the equivalent base band. The transmitter transmits the signal $\tilde{\mathbf{s}}(t) \triangleq \sqrt{E_{\mathrm{s}}}\mathbf{F}\mathbf{s}(t)$, where E_{s} denotes the energy per symbol, $\mathbf{F} \triangleq [\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_{N_{\mathrm{B}}}]$ is a precoding matrix with N_{B} transmitted beams, and $\mathbf{s}(t) \triangleq [s_1(t), s_2(t), ..., s_{N_{\mathrm{B}}}(t)]^{\mathrm{T}}$ is the vector of pilot signals. The pilot signal of the l^{th} beam is given by

$$s_l(t) \triangleq \sum_{m=0}^{N_{\mathrm{p}}-1} d_{l,m} p(t - mT_{\mathrm{s}}), \tag{6.1}$$

where N_{p} denotes the number of pilot symbols per beam, T_s is the symbol duration, and $d_{l,m}$ is the m^{th} unit energy pilot symbol which is transmitted over the l^{th} beam with the unit energy pulse $p(t)$. The pilot symbols $d_{l,m}, \forall l, m$ are assumed to be known to the agent. The l^{th} column of $\mathbf{F}$ contains a directional beam pointing towards the azimuth angle $\theta_{\mathrm{BF},l}$

$$\mathbf{f}_l(\theta_{\mathrm{BF},l}) \triangleq \frac{1}{\sqrt{N_{\mathrm{B}}}} \mathbf{a}_{\mathrm{TX},l}(\theta_{\mathrm{BF},l}). \tag{6.2}$$

Here, $\mathbf{a}_{\mathrm{TX},l}(\cdot)$ is the unit-norm array response vector given by [Tre02]

$$\mathbf{a}_{\mathrm{TX},l}(\theta_{\mathrm{BF},l}) \triangleq \frac{1}{\sqrt{N_{\mathrm{TX}}}} \exp(-j\boldsymbol{\Delta}_{\mathrm{TX}}^{\mathrm{T}} \mathbf{k}(\theta_{\mathrm{BF},l})), \tag{6.3}$$

where $\mathbf{k}(\theta_{\mathrm{BF},l}) = \frac{2\pi}{\lambda}[\cos(\theta_{\mathrm{BF},l}), \sin(\theta_{\mathrm{BF},l})]^{\mathrm{T}}$ is the wavenumber vector, λ is the wavelength, $\boldsymbol{\Delta}_{\mathrm{TX}} \triangleq [\mathbf{u}_{\mathrm{TX},1}, \mathbf{u}_{\mathrm{TX},2}, ..., \mathbf{u}_{\mathrm{TX},N_{\mathrm{TX}}}]$ is a $2 \times N_{\mathrm{TX}}$ matrix which contains the positions of the transmit antenna elements in 2D Cartesian coordinates in its columns, i.e. the n^{th} column of $\boldsymbol{\Delta}_{\mathrm{TX}}$ is given by $\mathbf{u}_{\mathrm{TX},n} \triangleq [x_{\mathrm{TX},n}, y_{\mathrm{TX},n}]^{\mathrm{T}}$. To normalize the transmitted power, the following conventions are considered: $\mathrm{tr}\left(\mathbf{F}^{\mathrm{H}}\mathbf{F}\right) = 1$ and $\| \int_0^{N_{\mathrm{p}}T_{\mathrm{s}}} s_l(t)\mathrm{d}t\|^2 = 1, \forall l$.

[6] For notational convenience, the position $\mathbf{p}$ of the agent has no subscript, since only a single agent is considered. Moreover, only a single anchor is considered and in order to drop the anchor subscript, its position $\mathbf{q}$ is described by a different letter.

6.2.3 Channel Model

A radio channel with $J+1 \geq 1$ distinct paths between the anchor and the agent is considered. In mmWave mMIMO systems, the number of paths is generally small [PK11]. The LOS path - if it exists - is denoted by $j=0$, while $j>0$ corresponds to NLOS components. It is assumed that the presence of an LOS path is known. Due to the large path loss in the mmWave band and the high directionality of the antenna arrays of mMIMO systems, the received power of NLOS paths is significantly lower than that of the LOS path [RMSS15]. Hence it is assumed that the LOS path can be detected reliably. In addition, due to the high path loss and the high directionality of the transmitted beams, NLOS components are assumed to originate from single-bounce scattering or reflection only [SGD$^+$18a, HSZ$^+$16b, GGD18b, DS14]. Scatterers are objects that are much smaller than the wavelength of the signal, while reflectors are objects with specific reflection points that are much larger than the wavelength of the signal. The reflecting point and the location of the scatterer are represented by the *point of incidence* and they are denoted by $\mathbf{s}_j = [s_{\mathrm{x},j}, s_{\mathrm{y},j}]^{\mathrm{T}}$. Considering Fig. 6.2, it can be seen that each path is associated with three distinct channel parameters; AOA, AOD, and delay (TOA). The AOA, AOD, and delay of the j^{th} path are denoted by $\theta_{\mathrm{RX},j}$, $\theta_{\mathrm{TX},j}$, and τ_j, respectively. Note that the AOAs and TOAs are measured in the local reference frames of the agent. The local temporal reference frame of the agent is related to the global temporal reference frame of the anchor via the clock offset ϵ. Similarly, the orientation α of the agent connects the local and global reference frames of the agent and anchor, respectively. If the clock offset and orientation are unknown they have to be included in the estimation problem in order to be able to harness the AOA and TOA estimates.

Remark: Recall the principle of delay estimation presented in Section 2.2.1. A channel estimator cannot directly measure the delay of a path. Instead, it estimates the TOA of a path and concludes the delay and hence the distance from the estimated TOA $\hat{t}_{\mathrm{TOA},j}$ and the time of transmission t_{TX} (c.f. (2.2)). Suppose that the TOA of the j^{th} path can be estimated with standard deviation σ_{TOA_j}. If the clock offset between the transmitter and receiver is known, the standard deviations σ_{τ_j} and σ_{TOA_j} of delay estimation and TOA estimation are identical due to their linear relationship and the fact that the clock offset does not contain uncertainty, i.e. $\sigma_{\tau_j} = \sigma_{\mathrm{TOA}_j}$. Hence the terms range, delay, and TOA will be used interchangeably in the case, where the transmitter and receiver are perfectly synchronized.

Assuming a narrow-band array model[7], the channel impulse response is given by

$$\mathbf{H}(t) = \sum_{j=0}^{J} \underbrace{\sqrt{N_{\mathrm{RX}} N_{\mathrm{TX}}} h_j \mathbf{a}_{\mathrm{RX},j}(\theta_{\mathrm{RX},j}) \mathbf{a}_{\mathrm{TX},j}^H(\theta_{\mathrm{TX},j})}_{\mathbf{H}_j} \times \delta\left(t-\tau_j\right), \tag{6.4}$$

where $h_j = h_{\mathrm{R},j} + j h_{\mathrm{I},j}$ is the complex path gain while $\mathbf{a}_{\mathrm{TX},j}(\theta_{\mathrm{TX},j})$ and $\mathbf{a}_{\mathrm{RX},j}(\theta_{\mathrm{RX},j})$ denote the unit-norm array response vectors of the j^{th} path at the transmitter and receiver, respectively. It is worthwhile to emphasize that $\theta_{\mathrm{BF},l}$ determines the angle of the l^{th} transmitted beam, while $\theta_{\mathrm{TX},j}$ describes the angle between the anchor and the j^{th} point of incidence. For instance, if $\theta_{\mathrm{BF},l}$ differs greatly from $\theta_{\mathrm{TX},j}$, the l^{th} beam does not illuminate the j^{th} point of incidence. In other words, physical paths that are not illuminated by

[7] It is assumed that $A_{\max} << c/B$, where $A_{\max}$ is the maximum array aperture size, c is the speed of light, and B is the system bandwidth.

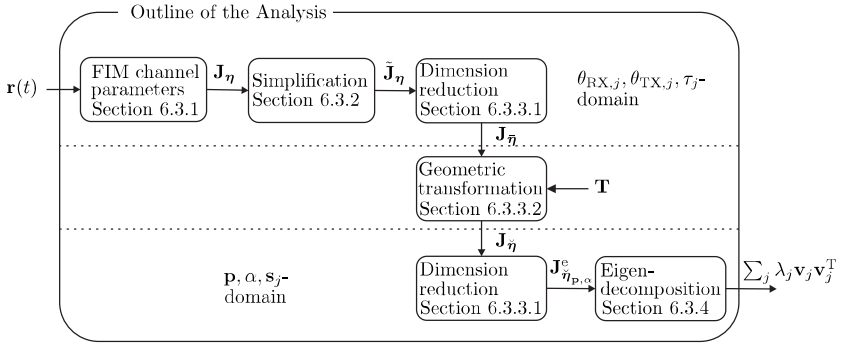

Figure 6.3: *Outline of the Fisher information analysis* - in order to analyze the role of NLOS components the FIM $\mathbf{J}_{\boldsymbol{\eta}}$ of the channel parameters is determined from the received signal and simplified subsequently. Parameters that are irrelevant for positioning are eliminated from the simplified FIM $\mathbf{J}_{\tilde{\boldsymbol{\eta}}}$ which results in a FIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ of reduced dimensions. This FIM is transformed from the AOA, AOD, and delay-domain to the position, orientation, and point of incidence-domain using the geometric transformation matrix $\mathbf{T}$. The EFIM $\mathbf{J}^{\mathrm{e}}_{\check{\boldsymbol{\eta}}_{\mathrm{p},\alpha}}$ of the position and orientation is determined based on $\mathbf{J}_{\check{\boldsymbol{\eta}}}$ and decomposed subsequently.

beams will not be seen by the agent. Note that $\mathbf{a}_{\mathrm{TX},j}(\theta_{\mathrm{TX},j})$ is defined via (6.3), while $\mathbf{a}_{\mathrm{RX},j}(\theta_{\mathrm{RX},j})$ can be defined analogously by (6.3) with appropriate subscripts.

6.2.4 Receiver Model

The noisy observed signal at the agent is given by

$$\mathbf{r}(t) \triangleq \sum_{j=0}^{J} \sqrt{E_{\mathrm{s}}} \mathbf{H}_j \mathbf{F} \mathbf{s}(t-\tau_j) + \mathbf{n}(t), \quad t \in [0, N_s T_s], \tag{6.5}$$

where $\mathbf{n}(t) = [n_1(t), n_2(t), ..., n_{N_{\mathrm{RX}}}(t)]^{\mathrm{T}}$ is zero-mean additive white Gaussian noise (AWGN) with power spectral density (PSD) $N_0/2$ per real dimension. Similar to [FNAB+10, VV10], it is assumed that a low-noise amplifier and a passband filter are attached to every receive antenna. This assumption might seem restrictive for the practical application, yet it simplifies the analysis of the EFIM. It can be regarded as the receiver architecture which results in the lowest PEB and OEB.

6.3 Positioning and Mapping with Known Clock Offset

This section derives the EFIM of the position, orientation, and map estimation problem assuming that the agent and anchor are perfectly synchronized. In order to reveal the nature of the estimation problem, the following steps will be performed (see Fig. 6.3). The FIM of the channel parameters will be derived based on the received signal in (6.5) and it will

be simplified subsequently. The irrelevant dimensions (from the positioning perspective) of this simplified FIM will be erased by means of the EFIM. The reduced-dimension FIM of the channel parameters will then be transformed in the position, orientation, and point of incidence domain. The dimensions of the resulting FIM will be reduced another time in order to derive the EFIM of the position and orientation which will be decomposed in the final step. It will be shown that the EFIM of the position and orientation can be expressed as the sum of rank one matrices.

6.3.1 Definition

Let $\boldsymbol{\eta}$ be the vector of all channel parameters, i.e.

$$\boldsymbol{\eta} \triangleq [\boldsymbol{\theta}_{\mathrm{RX}}^{\mathrm{T}}, \boldsymbol{\theta}_{\mathrm{TX}}^{\mathrm{T}}, \boldsymbol{\tau}^{\mathrm{T}}, \mathbf{h}_{\mathrm{R}}^{\mathrm{T}}, \mathbf{h}_{\mathrm{I}}^{\mathrm{T}}]^{\mathrm{T}}. \tag{6.6}$$

where the AOAs, AODs, delays, and channel gains are collected in the vectors $\boldsymbol{\theta}_{\mathrm{RX}} \triangleq [\theta_{\mathrm{RX},0}, \theta_{\mathrm{RX},1}, ..., \theta_{\mathrm{RX},J}]^{\mathrm{T}}$, $\boldsymbol{\theta}_{\mathrm{TX}} \triangleq [\theta_{\mathrm{TX},0}, \theta_{\mathrm{TX},1}, ..., \theta_{\mathrm{TX},J}]^{\mathrm{T}}$, $\boldsymbol{\tau} \triangleq [\tau_0, \tau_1, ..., \tau_J]^{\mathrm{T}}$, $\mathbf{h}_{\mathrm{R}} \triangleq [h_{\mathrm{R},0}, h_{\mathrm{R},1}, ..., h_{\mathrm{R},J}]^{\mathrm{T}}$, and $\mathbf{h}_{\mathrm{I}} \triangleq [h_{\mathrm{I},0}, h_{\mathrm{I},1}, ..., h_{\mathrm{I},J}]^{\mathrm{T}}$, respectively. Given the received signal $\mathbf{r}(t)$, the FIM of the channel parameters can be computed as

$$\mathbf{J}_{\boldsymbol{\eta}} \triangleq \begin{bmatrix} \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{RX}}\boldsymbol{\theta}_{\mathrm{RX}}} & \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{RX}}\boldsymbol{\theta}_{\mathrm{TX}}} & \cdots & \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{RX}}\mathbf{h}_{\mathrm{I}}} \\ \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{RX}}\boldsymbol{\theta}_{\mathrm{TX}}}^{\mathrm{T}} & \ddots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{RX}}\mathbf{h}_{\mathrm{I}}}^{\mathrm{T}} & \cdots & \cdots & \mathbf{J}_{\mathbf{h}_{\mathrm{I}}\mathbf{h}_{\mathrm{I}}} \end{bmatrix}, \tag{6.7}$$

where each entry is determined according to[8] [Kay93]

$$[\mathbf{J}_{\boldsymbol{\eta}}]_{u,v} \triangleq \frac{1}{N_0} \int_0^{N_{\mathrm{p}}T_{\mathrm{s}}} \mathbb{E}\left[\Re\left\{\frac{\partial \boldsymbol{\mu}_{\boldsymbol{\eta}}^{\mathrm{H}}(t)}{\partial [\boldsymbol{\eta}]_u} \frac{\partial \boldsymbol{\mu}_{\boldsymbol{\eta}}(t)}{\partial [\boldsymbol{\eta}]_v}\right\}\right] \mathrm{d}t \tag{6.8}$$

and $\mathbb{E}[\cdot]$ denotes the expectation with respect to the pilot symbols, $\Re\{\cdot\}$ is the real part operator, and $\boldsymbol{\mu}_{\boldsymbol{\eta}}(t)$ is defined as the noise-free observation which can be obtained from (6.5) as

$$\boldsymbol{\mu}_{\boldsymbol{\eta}}(t) = \sum_{j=0}^{J} \sqrt{E_{\mathrm{s}}} \mathbf{H}_j \mathbf{F} \mathbf{s}(t - \tau_j). \tag{6.9}$$

The FIM is related to the estimation error covariance matrix of any unbiased estimator via the information inequality [Abe93, TB07, TBT13]

$$\mathbb{E}\left[(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})^{\mathrm{T}}\right] \succeq \mathbf{J}_{\boldsymbol{\eta}}^{-1}, \tag{6.10}$$

where $\hat{\boldsymbol{\eta}}$ is the estimate of $\boldsymbol{\eta}$ and $\mathbf{A} \succeq \mathbf{B}$ is equivalent to $\mathbf{A} - \mathbf{B}$ being positive semi-definite. The inequality in (6.10) is the well-known Cramér-Rao lower bound (CRLB).

[8] This result holds whenever the signal $\mathbf{r}(t)$ is corrupted by AWGN.

6.3.2 Simplification

The blocks of the FIM in (6.7) obey certain scaling laws when the number of receive and transmit antennas as well as the bandwidth become sufficiently large[9]. In particular, it was shown in [ASZA$^+$18a, Section III-B] that some blocks can be well approximated by diagonal matrices, while others become zero matrices. The corresponding intuition is the following: If a block matrix on the diagonal of $\mathbf{J}_{\boldsymbol{\eta}}$ becomes diagonal. the involved parameters can be estimated independently of each other. Moreover, if an off-diagonal block matrix becomes zero, there is no coupling of the involved parameters. Examples will be provided in the following paragraph, where the results of [ASZA$^+$18a] that are relevant for this thesis will be summarized.

Let $\odot$ denote the Hadamard product and let $\mathbf{I}_{J+1}$ and $\mathbf{0}_{J+1}$ be the $J+1 \times J+1$ identity and all-zeros matrix, respectively. The following simplifications of the FIM $\mathbf{J}_{\boldsymbol{\eta}}$ in (6.7) can be made [ASZA$^+$18a]:

1. Since the AOAs of the different paths are assumed to be distinct, the steering vectors at the receiver do not interact considerably with each other, i.e. $\left\|\mathbf{a}_{\mathrm{RX},u}^{\mathrm{H}}\mathbf{a}_{\mathrm{RX},v}\right\| \ll \left\|\mathbf{a}_{\mathrm{RX},u}^{\mathrm{H}}\mathbf{a}_{\mathrm{RX},u}^{\mathrm{H}}\right\|$, $\forall u \neq v$. Hence AOAs can be estimated independently, i.e. $\tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{RX}}\boldsymbol{\theta}_{\mathrm{RX}}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{RX}}\boldsymbol{\theta}_{\mathrm{RX}}}$.

2. The array response $\mathbf{a}_{\mathrm{TX},u}$ and $\mathbf{a}_{\mathrm{TX},u}$ for different angles $\theta_{\mathrm{TX},u}$ and $\theta_{\mathrm{TX},v}, u \neq v$ do not interact considerably with each other, i.e. $\lim_{N_{\mathrm{TX}}\to\infty} \mathbf{a}_{\mathrm{TX},u}^{\mathrm{H}}\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{a}_{\mathrm{TX},v} \approx 0$. Hence AODs can be estimated independently, i.e. $\tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\boldsymbol{\theta}_{\mathrm{TX}}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{TX}}\boldsymbol{\theta}_{\mathrm{TX}}}$.

3. The temporal cross-correlations of different paths vanish as the bandwidth of the signal becomes large, since the paths can be resolved independently in time, i.e. $\tilde{\mathbf{J}}_{\boldsymbol{\tau}\boldsymbol{\tau}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\boldsymbol{\tau}\boldsymbol{\tau}}$.

4. As a consequence of the previous results, the channel gains can be estimated independently, i.e. $\tilde{\mathbf{J}}_{\mathbf{h}_{\mathrm{R}}\mathbf{h}_{\mathrm{R}}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\mathbf{h}_{\mathrm{R}}\mathbf{h}_{\mathrm{R}}}$ and $\tilde{\mathbf{J}}_{\mathbf{h}_{\mathrm{I}}\mathbf{h}_{\mathrm{I}}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\mathbf{h}_{\mathrm{I}}\mathbf{h}_{\mathrm{I}}}$.

5. All off-diagonal blocks, except for $\mathbf{J}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{R}}}$ and $\mathbf{J}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{I}}}$, in (6.7) become zero. It was shown in [ASZA$^+$18a] that the real and imaginary parts of the j^{th} channel gain couple only with the AOD of the j^{th} path, i.e. $\tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{R}}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{R}}}$ and $\tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{I}}} \approx \mathbf{I}_{J+1} \odot \mathbf{J}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{I}}}$.

Thus, when the bandwidth of the signal is large and the number of receive and transmit antennas is also large, $\mathbf{J}_{\boldsymbol{\eta}}$ can be well approximated by

$$\tilde{\mathbf{J}}_{\boldsymbol{\eta}} \triangleq \begin{bmatrix} \tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{RX}}\boldsymbol{\theta}_{\mathrm{RX}}} & \mathbf{0}_{J+1} & \mathbf{0}_{J+1} & \mathbf{0}_{J+1} & \mathbf{0}_{J+1} \\ \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\boldsymbol{\theta}_{\mathrm{TX}}} & \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{R}}} & \tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{I}}} \\ \mathbf{0}_{J+1} & \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\boldsymbol{\tau}\boldsymbol{\tau}} & \mathbf{0}_{J+1} & \mathbf{0}_{J+1} \\ \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{R}}}^{\mathrm{T}} & \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\mathbf{h}_{\mathrm{R}}\mathbf{h}_{\mathrm{R}}} & \mathbf{0}_{J+1} \\ \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{I}}}^{\mathrm{T}} & \mathbf{0}_{J+1} & \mathbf{0}_{J+1} & \tilde{\mathbf{J}}_{\mathbf{h}_{\mathrm{I}}\mathbf{h}_{\mathrm{I}}} \end{bmatrix}. \tag{6.11}$$

[9]It was shown numerically in [ASZA$^+$18a] that the approximation error of the PEB due to the simplification of the FIM is fairly small even under realistic assumptions on the bandwidth ($B = 125$ MHz) and the 2D array sizes ($N_{\mathrm{TX/RX}} = 12 \times 12$).

6.3.3 Derivation of the EFIM of the Position and Orientation

The parameter vector $\boldsymbol{\eta}$ is reordered for mathematical convenience such that parameters of the same path are grouped together, i.e.

$$\tilde{\boldsymbol{\eta}} \triangleq \left[\tilde{\boldsymbol{\eta}}_0^{\mathrm{T}}, ..., \tilde{\boldsymbol{\eta}}_J^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{6.12}$$

where $\tilde{\boldsymbol{\eta}}_j \triangleq [\tau_j, \theta_{\mathrm{TX},j}, h_{\mathrm{R},j}, h_{\mathrm{I},j}, \theta_{\mathrm{RX},j}]^{\mathrm{T}}$. Reordering the parameter vector results in a permutation of the entries of the FIM in (6.11). The reordered FIM is given by

$$\mathbf{J}_{\tilde{\boldsymbol{\eta}}} \triangleq \mathbf{P}_\pi \tilde{\mathbf{J}}_{\boldsymbol{\eta}} \mathbf{P}_\pi, \tag{6.13}$$

where $\mathbf{P}_\pi$ is a permutation matrix of size $5J \times 5J$ which is given by

$$\mathbf{P}_\pi \triangleq [\mathbf{P}_{\pi,1}, ..., \mathbf{P}_{\pi,J}]^{\mathrm{T}}, \tag{6.14}$$

Here, $\mathbf{P}_{\pi,i} = [\mathbf{e}_{2J+i}, \mathbf{e}_{J+i}, \mathbf{e}_{3J+i}, \mathbf{e}_{4J+i}, \mathbf{e}_i]$ and $\mathbf{e}_i$ denotes the i^{th} unit vector of the standard basis of appropriate length. Hence the FIM of the channel parameters has the following structure

$$\mathbf{J}_{\tilde{\boldsymbol{\eta}}} = \mathrm{blkdiag}\left(\mathbf{J}_{\tilde{\boldsymbol{\eta}}_0}, ..., \mathbf{J}_{\tilde{\boldsymbol{\eta}}_J}\right). \tag{6.15}$$

The FIM block $\mathbf{J}_{\tilde{\boldsymbol{\eta}}_j}$ of the j^{th} path is given by

$$\mathbf{J}_{\tilde{\boldsymbol{\eta}}_j} \triangleq \begin{bmatrix} \frac{1}{\sigma^2_{\tau_j}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \underbrace{\begin{bmatrix} \frac{1}{\tilde{\sigma}^2_{\theta_{\mathrm{TX},j}}} & b_{\mathrm{R},j} & b_{\mathrm{I},j} \\ b_{\mathrm{R},j} & \frac{1}{\sigma^2_{\mathrm{h}_{\mathrm{R},j}}} & 0 \\ b_{\mathrm{I},j} & 0 & \frac{1}{\sigma^2_{\mathrm{h}_{\mathrm{I},j}}} \end{bmatrix}}_{\mathbf{J}_{\theta_{\mathrm{TX},j}\mathbf{h}_j}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\sigma^2_{\theta_{\mathrm{RX},j}}} \end{bmatrix}, \tag{6.16}$$

where the diagonal entries of the reordered FIM matrix of the channel parameters are abbreviated by the respective $1/\sigma^2$-terms. The variables $b_{\mathrm{R},j}$ and $b_{\mathrm{I},j}$ denote the corresponding entries of $\tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{R}}}$ and $\tilde{\mathbf{J}}_{\boldsymbol{\theta}_{\mathrm{TX}}\mathbf{h}_{\mathrm{I}}}$, respectively.

6.3.3.1 Dimension Reduction of the FIM of the Channel Parameters

In contrast to the AOA, AOD, and delay of every path, the channel gains $h_{\mathrm{R,j}}$ and $h_{\mathrm{I,j}}$ have no geometric relationship with the parameters in the position, orientation, and point of incidence domain. The goal of this subsection is to absorb them into the parameters (AODs) with which they couple. In that way, the dimension of the FIM is reduced and all remaining channel parameters carry information about the position, orientation, and the points of incidence.

In particular, the uncertainty of $h_{\mathrm{R,j}}$ and $h_{\mathrm{I,j}}$ will be absorbed in the AOD-related term. For that, the notion of the EFIM will be employed. The EFIM is a measure of the information corresponding to certain parameters of interest, while accounting for the uncertainties of other irrelevant parameters.

Definition 1. *Given a parameter vector* $\boldsymbol{\xi} \triangleq [\boldsymbol{\xi}_1^{\mathrm{T}}, \boldsymbol{\xi}_2^{\mathrm{T}}]^{\mathrm{T}}$ *with corresponding FIM*

$$\mathbf{J}_{\boldsymbol{\xi}} \triangleq \begin{bmatrix} \mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_1} & \mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_2} \\ \mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_2}^{\mathrm{T}} & \mathbf{J}_{\boldsymbol{\xi}_2\boldsymbol{\xi}_2} \end{bmatrix}, \tag{6.17}$$

the EFIM of $\boldsymbol{\xi}_1$ *is obtained by [TMN98]*

$$\mathbf{J}_{\boldsymbol{\xi}_1}^{\mathrm{e}} \triangleq \mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_1} - \mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_2}\mathbf{J}_{\boldsymbol{\xi}_2\boldsymbol{\xi}_2}^{-1}\mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_2}^{\mathrm{T}}. \tag{6.18}$$

Intuitively, the fact that the parameters of $\boldsymbol{\xi}_2$ *are not perfectly known, leads to a loss in information which is quantified by* $\mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_2}\mathbf{J}_{\boldsymbol{\xi}_2\boldsymbol{\xi}_2}^{-1}\mathbf{J}_{\boldsymbol{\xi}_1\boldsymbol{\xi}_2}^{\mathrm{T}}$.

Using (6.16) and (6.18), the EFIM of $\mathbf{J}_{\theta_{\mathrm{TX},j}\mathbf{h}_j}$ with respect to $\theta_{\mathrm{TX},j}$ is given by

$$\mathbf{J}_{\theta_{\mathrm{TX},j}}^{\mathrm{e}} = \frac{1}{\tilde{\sigma}_{\theta_{\mathrm{TX},j}}^2} - \left(b_{\mathrm{R},j}^2\sigma_{h_{\mathrm{R},j}}^2 + b_{\mathrm{I},j}^2\sigma_{h_{\mathrm{I},j}}^2\right) \triangleq \frac{1}{\sigma_{\theta_{\mathrm{TX},j}}^2}. \tag{6.19}$$

Hence the (E)FIM[10] of the j^{th} path is given by $\mathbf{J}_{\bar{\boldsymbol{\eta}}_j} \triangleq \mathrm{diag}\left(1/\sigma_{\tau_j}^2, 1/\sigma_{\theta_{\mathrm{TX},j}}^2, 1/\sigma_{\theta_{\mathrm{RX},j}}^2\right)$, where $\bar{\boldsymbol{\eta}}_j = [\tau_j, \theta_{\mathrm{TX},j}, \theta_{\mathrm{RX},j}]^{\mathrm{T}}$. For notational convenience, the sub-FIMs $\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{LOS}}} \triangleq \mathbf{J}_{\bar{\boldsymbol{\eta}}_0}$ of the LOS path and of the NLOS paths $\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}} \triangleq \mathrm{blkdiag}\left(\mathbf{J}_{\bar{\boldsymbol{\eta}}_1}, ..., \mathbf{J}_{\bar{\boldsymbol{\eta}}_J}\right)$ are defined. Consequently, the FIM of all paths is given by

$$\mathbf{J}_{\bar{\boldsymbol{\eta}}} \triangleq \mathrm{blkdiag}\left(\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{LOS}}}, \mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\right). \tag{6.20}$$

Intuition: Each σ^2-term reflects how well the respective parameter can be estimated, e.g., large $\sigma_{\tau_0}^2$ means that the delay of the LOS path cannot be estimated accurately. Note that the σ^2-terms in (6.16) depend on many aspects, including the number of antennas, beamforming, bandwidth, and receiver location.

6.3.3.2 Geometric Transformation of the FIM of the Channel Parameters

In order to obtain the FIM $\mathbf{J}_{\breve{\boldsymbol{\eta}}}$ of the position, orientation, and points of incidence, a Jacobian transformation of the FIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ of the channel parameters in (6.20) is required. This transformation captures the geometric relationships between the channel parameters $\bar{\boldsymbol{\eta}} = [\bar{\boldsymbol{\eta}}_0, ..., \bar{\boldsymbol{\eta}}_J]$ and the position-related parameters $\breve{\boldsymbol{\eta}} \triangleq [\mathbf{p}^{\mathrm{T}}, \alpha, \mathbf{s}_1^{\mathrm{T}}, ..., \mathbf{s}_J^{\mathrm{T}}]^{\mathrm{T}}$, where $\mathbf{s}_j, j = 1, 2, \cdots J$ denotes the j^{th} point of incidence. For notational convenience and unlike the parameter vector used in the cooperative positioning problem in Chapter 5, the anchor $\mathbf{q}$ is not part of the parameter vector $\breve{\boldsymbol{\eta}}$. In particular, the FIM of the position-related parameters is given by [Kay93]

$$\mathbf{J}_{\breve{\boldsymbol{\eta}}} \triangleq \mathbf{T}\mathbf{J}_{\bar{\boldsymbol{\eta}}}\mathbf{T}^{\mathrm{T}}, \tag{6.21}$$

where $\mathbf{T} \triangleq \frac{\partial \bar{\boldsymbol{\eta}}^{\mathrm{T}}}{\partial \breve{\boldsymbol{\eta}}}$. In order to determine the transformation matrix $\mathbf{T}$, the geometric relationships between the channel and position-related parameters are depicted in the following

[10] There is a slight abuse of the notation here, since the EFIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}_j}^{\mathrm{e}}$ is denoted by $\mathbf{J}_{\bar{\boldsymbol{\eta}}_j}$ and it is called FIM. This inconsistency was purposely introduced to avoid confusion with another EFIM later on.

list:

$$\tau_0 = \frac{||\mathbf{q}-\mathbf{p}||}{c} + \epsilon \tag{6.22a}$$

$$\tau_j = \frac{||\mathbf{q}-\mathbf{s}_j|| + ||\mathbf{s}_j-\mathbf{p}||}{c} + \epsilon \tag{6.22b}$$

$$\theta_{\mathrm{TX},0} = \operatorname{atan2}\left(\frac{p_\mathrm{y}-q_\mathrm{y}}{p_\mathrm{x}-q_\mathrm{x}}\right) \tag{6.22c}$$

$$\theta_{\mathrm{TX},j} = \operatorname{atan2}\left(\frac{s_{\mathrm{y},j}-q_\mathrm{y}}{s_{\mathrm{x},j}-q_\mathrm{x}}\right), \quad j>0 \tag{6.22d}$$

$$\theta_{\mathrm{RX},0} = \operatorname{atan2}\left(\frac{q_\mathrm{y}-p_\mathrm{y}}{q_\mathrm{x}-p_\mathrm{x}}\right) - \alpha \tag{6.22e}$$

$$\theta_{\mathrm{RX},j} = \operatorname{atan2}\left(\frac{s_{\mathrm{y},j}-p_{\mathrm{y},j}}{s_{\mathrm{x},j}-p_{\mathrm{x},j}}\right) - \alpha, \quad j>0, \tag{6.22f}$$

where ϵ is the *known* clock offset between the anchor and agent. The partial derivatives of the channel parameters with respect to the position-related parameters will be required in later in this chapter. The non-zero partial derivatives are summarized in Appendix A.1. It is worth to emphasize that ϵ can be included in the estimation problem as a parameter in the case where the agent and anchor are not perfectly synchronized. This approach will be pursued in Section 6.4. For now, the clock offset is assumed to be known. Based on the geometric relationships in (6.22), the following transformation matrix results.

Lemma 1. *The transformation matrix* $\mathbf{T}$ *is an upper triangle block-matrix*

$$\mathbf{T} \triangleq \left[\begin{array}{c|ccc} \mathbf{T}_\mathbf{P}^{(0)} & \mathbf{T}_\mathbf{P}^{(1)} & \cdots & \mathbf{T}_\mathbf{P}^{(J)} \\ \hline \mathbf{0}_{2\times 3} & \mathbf{T}_{\mathbf{s}_1} & \cdots & \mathbf{0}_{2\times 3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2\times 3} & \mathbf{0}_{2\times 3} & \cdots & \mathbf{T}_{\mathbf{s}_J} \end{array}\right] \triangleq \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{0}_{J\times 3} & \mathbf{D} \end{array}\right], \tag{6.23}$$

where [SGD+18a]

$$\mathbf{T}_\mathbf{P}^{(j)} = \begin{bmatrix} \partial\tau_j/\partial\mathbf{p} & \partial\theta_{\mathrm{TX},j}/\partial\mathbf{p} & \partial\theta_{\mathrm{RX},j}/\partial\mathbf{p} \\ \partial\tau_j/\partial\alpha & \partial\theta_{\mathrm{TX},j}/\partial\alpha & \partial\theta_{\mathrm{RX},j}/\partial\alpha \end{bmatrix} \tag{6.24}$$

and

$$\mathbf{T}_{\mathbf{s}_j} = \begin{bmatrix} \partial\tau_j/\partial\mathbf{s}_j & \partial\theta_{\mathrm{TX},j}/\partial\mathbf{s}_j & \partial\theta_{\mathrm{RX},j}/\partial\mathbf{s}_j \end{bmatrix}. \tag{6.25}$$

Proof. The block-matrices in the top row of $\mathbf{T}$ are given by

$$\mathbf{T}_\mathbf{P}^{(0)} = \begin{bmatrix} \frac{1}{c}\cos(\theta_{\mathrm{TX},0}) & -\frac{1}{\|\mathbf{p}-\mathbf{q}\|}\sin(\theta_{\mathrm{TX},0}) & -\frac{1}{\|\mathbf{p}-\mathbf{q}\|}\sin(\theta_{\mathrm{TX},0}) \\ \frac{1}{c}\sin(\theta_{\mathrm{TX},0}) & \frac{1}{\|\mathbf{p}-\mathbf{q}\|}\cos(\theta_{\mathrm{TX},0}) & \frac{1}{\|\mathbf{p}-\mathbf{q}\|}\cos(\theta_{\mathrm{TX},0}) \\ 0 & 0 & -1 \end{bmatrix} \tag{6.26}$$

and, for $j > 0$,

$$\mathbf{T}_{\mathbf{P}}^{(j)} = \begin{bmatrix} -\frac{1}{c}\cos(\theta_{\mathrm{RX},j}) & 0 & +\frac{1}{\|\mathbf{p}-\mathbf{s}_j\|}\sin(\pi-\theta_{\mathrm{RX},j}) \\ -\frac{1}{c}\sin(\theta_{\mathrm{RX},j}) & 0 & -\frac{1}{\|\mathbf{p}-\mathbf{s}_j\|}\cos(\theta_{\mathrm{RX},j}) \\ 0 & 0 & -1 \end{bmatrix}, \tag{6.27}$$

respectively. The diagonal blocks corresponding to the points of incidence are given

$$\mathbf{T}_{\mathbf{s}_j} = \begin{bmatrix} \frac{1}{c}\left[\cos(\theta_{\mathrm{TX},j})+\cos(\theta_{\mathrm{RX},j})\right] & -\frac{1}{\|\mathbf{q}-\mathbf{s}_j\|}\sin(\theta_{\mathrm{TX},j}) & -\frac{1}{\|\mathbf{p}-\mathbf{s}_j\|}\sin(\theta_{\mathrm{RX},j}) \\ \frac{1}{c}\left[\sin(\theta_{\mathrm{TX},j})+\sin(\theta_{\mathrm{RX},j})\right] & \frac{1}{\|\mathbf{q}-\mathbf{s}_j\|}\cos(\theta_{\mathrm{TX},j}) & +\frac{1}{\|\mathbf{p}-\mathbf{s}_j\|}\cos(\theta_{\mathrm{RX},j}) \end{bmatrix}, \tag{6.28}$$

for $j > 0$. □

Remark: The units of the position-related and orientation-related entries in the transformation matrix are normalized by 1 m and 1 deg, respectively. By normalization, a mixed-units FIM $\mathbf{J}_{\breve{\boldsymbol{\eta}}}$ of the position, orientation, and points of incidence is avoided after the transformation. As a result, all entries in the transformed FIM $\mathbf{J}_{\breve{\boldsymbol{\eta}}}$ will be dimensionless and the standard inner product (dot product) can be employed to define the norm of a vector. For notational convenience, the normalization constants (1 m and 1 deg) are omitted.

6.3.3.3 Dimension Reduction of the FIM of Position-related Parameters

Recall that the goal of the analysis is to reveal the impact of different paths on the position and orientation estimation problem. In that regard, the points of incidence are just nuisance variables and the goal of this subsection is to remove them from the FIM via an appropriate EFIM formulation.

Lemma 2. *The FIM $\mathbf{J}_{\breve{\boldsymbol{\eta}}}$ of the position-related parameters is given by*

$$\mathbf{J}_{\breve{\boldsymbol{\eta}}} = \left[\begin{array}{c|c} \mathbf{A}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{LOS}}}\mathbf{A}^{\mathrm{T}} + \mathbf{B}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}} & \mathbf{B}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}} \\ \hline \mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}} & \mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}} \end{array}\right]. \tag{6.29}$$

Proof. Evaluate (6.21) using Lemma 1. □

Defining the parameter vector $\breve{\boldsymbol{\eta}}_{\mathbf{p},\alpha} \triangleq [\mathbf{p}^{\mathrm{T}}, \alpha]^{\mathrm{T}}$, the EFIM $\mathbf{J}^{\mathrm{e}}_{\breve{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}$ of the position and orientation is obtained using (6.18) and (6.29) from Lemma 2, i.e.

$$\begin{aligned} \mathbf{J}^{\mathrm{e}}_{\breve{\boldsymbol{\eta}}_{\mathbf{p},\alpha}} = & \underbrace{\mathbf{A}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{LOS}}}\mathbf{A}^{\mathrm{T}}}_{\triangleq \tilde{\mathbf{A}}^{(\mathrm{G})}\text{ - LOS info gain}} + \underbrace{\mathbf{B}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}}}_{\triangleq \tilde{\mathbf{B}}^{(\mathrm{G})}\text{ - NLOS info gain}} \\ & - \underbrace{\mathbf{B}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}}(\mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}})^{-1}\mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}}}_{\triangleq \tilde{\mathbf{B}}^{(\mathrm{L})}\text{ - NLOS info loss}}. \end{aligned} \tag{6.30}$$

It can be seen that the EFIM $\mathbf{J}^{\mathrm{e}}_{\breve{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}$ is composed of three terms. The first term $\tilde{\mathbf{A}}^{(\mathrm{G})}$ quantifies the information gain from the LOS path. The second term $\tilde{\mathbf{B}}^{(\mathrm{G})}$ quantifies the information gain from the NLOS components. Finally, the third term $\tilde{\mathbf{B}}^{(\mathrm{L})}$ specifies the

loss of information which accounts for the fact that the points of incident of NLOS paths are unknown.
Considering the structure of $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ and $\mathbf{T}$ the NLOS information gain and loss can be written as

$$\tilde{\mathbf{B}}^{(\mathrm{G})} = \sum_{j=1}^{J} \mathbf{T}_{\mathbf{P}}^{(j)} \mathbf{J}_{\bar{\boldsymbol{\eta}}_j} \left(\mathbf{T}_{\mathbf{P}}^{(j)} \right)^{\mathrm{T}}, \tag{6.31}$$

and

$$\tilde{\mathbf{B}}^{(\mathrm{L})} = \sum_{j=1}^{J} \mathbf{T}_{\mathbf{P}}^{(j)} \mathbf{J}_{\bar{\boldsymbol{\eta}}_j} \mathbf{T}_{\mathbf{s}_j}^{\mathrm{T}} \left(\mathbf{T}_{\mathbf{s}_j} \mathbf{J}_{\bar{\boldsymbol{\eta}}_j} \mathbf{T}_{\mathbf{s}_j}^{\mathrm{T}} \right)^{-1} \mathbf{T}_{\mathbf{s}_j} \mathbf{J}_{\bar{\boldsymbol{\eta}}_j} \left(\mathbf{T}_{\mathbf{P}}^{(j)} \right)^{\mathrm{T}}, \tag{6.32}$$

respectively.

Interpretation: The term in (6.31) reflects the gain of information from NLOS components if the points of incident $\mathbf{s}_j, j = 1, 2, ..., J$ were perfectly known. Since each point of incidence $\mathbf{s}_j$ has to be estimated as well, the estimation uncertainty leads to a loss of information which is quantified in (6.32).

6.3.4 Decomposition of the EFIM of the Position and Orientation

In the following, the terms $\mathbf{A}^{(\mathrm{G})}$ and $\mathbf{B}^{(\mathrm{N})} = \mathbf{B}^{(\mathrm{G})} - \mathbf{B}^{(\mathrm{L})}$ will be decomposed with the goal to express the EFIM $\mathbf{J}^{\mathrm{e}}_{\bar{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}$ in (6.30) as the sum of rank one matrices. Each rank one matrix is given by the outer-product of the unit-norm eigenvector $\mathbf{v}_j$ that corresponds to the only non-zero eigenvalue λ_j of the rank one matrix, i.e. $\mathbf{J}^{\mathrm{e}}_{\bar{\boldsymbol{\eta}}_{\mathbf{p},\alpha}} = \sum_j \lambda_j \mathbf{v}_j \mathbf{v}_j^{\mathrm{T}}$. It will be shown that each rank one matrix of an NLOS path is positive semi-definite, meaning that each NLOS path improves the position and orientation estimation accuracy.

The following matrix template will be defined for notational convenience

$$\boldsymbol{\Upsilon}_{n,m}(\theta, \phi, \rho) \triangleq \begin{bmatrix} \cos^2(\theta+\phi) & (-1)^n \sin(\theta)\cos(\theta) & (-1)^m \rho \sin(\theta) \\ (-1)^n \sin(\theta)\cos(\theta) & \sin^2(\theta+\phi) & (-1)^{m+1} \rho \cos(\theta) \\ (-1)^m \rho \sin(\theta) & (-1)^{m+1} \rho \cos(\theta) & \rho^2 \end{bmatrix}. \tag{6.33}$$

6.3.4.1 Decomposition of the LOS Information Gain Matrix

Using (6.23), (6.26), and by simple algebra, the LOS information gain matrix $\tilde{\mathbf{A}}^{(\mathrm{G})}$ in (6.30) can be easily shown to be

$$\begin{aligned}\tilde{\mathbf{A}}^{(\mathrm{G})} = & \underbrace{\frac{1}{\sigma_{\tau_0}^2 c^2}\boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{TX},0},0,0)}_{\tilde{\mathbf{A}}_R^{(\mathrm{G})}} \\ & + \underbrace{\frac{1}{\sigma_{\theta_{\mathrm{TX},0}}^2\left\|\mathbf{p}-\mathbf{q}\right\|^2}\boldsymbol{\Upsilon}_{1,0}(\theta_{\mathrm{TX},0},\pi/2,0)}_{\tilde{\mathbf{A}}_D^{(\mathrm{G})}} \\ & + \underbrace{\frac{1}{\sigma_{\theta_{\mathrm{RX},0}}^2\left\|\mathbf{p}-\mathbf{q}\right\|^2}\boldsymbol{\Upsilon}_{1,0}(\theta_{\mathrm{TX},0},\pi/2,\left\|\mathbf{p}-\mathbf{q}\right\|)}_{\tilde{\mathbf{A}}_A^{(\mathrm{G})}}.\end{aligned} \tag{6.34}$$

Proposition 1. *The eigenvalues of the matrices* $\tilde{\mathbf{A}}_R^{(\mathrm{G})}, \tilde{\mathbf{A}}_D^{(\mathrm{G})}$, *and* $\tilde{\mathbf{A}}_A^{(\mathrm{G})}$ *are given by*

$$\lambda_{R,0}^{(\mathrm{G})} = \frac{1}{\sigma_{\tau_0}^2 c^2}, \tag{6.35a}$$

$$\lambda_{D,0}^{(\mathrm{G})} = \frac{1}{\sigma_{\theta_{\mathrm{TX},0}}^2\left\|\mathbf{p}-\mathbf{q}\right\|^2}, \tag{6.35b}$$

$$\lambda_{A,0}^{(\mathrm{G})} = \frac{\left\|\mathbf{p}-\mathbf{q}\right\|^2+1}{\sigma_{\theta_{\mathrm{RX},0}}^2\left\|\mathbf{p}-\mathbf{q}\right\|^2}, \tag{6.35c}$$

respectively. The corresponding eigenvectors are given by

$$\mathbf{v}_{R,0}^{(\mathrm{G})} = [\cos(\theta_{\mathrm{TX},0}),\sin(\theta_{\mathrm{TX},0}),0]^{\mathrm{T}}, \tag{6.36a}$$

$$\mathbf{v}_{D,0}^{(\mathrm{G})} = [-\sin(\theta_{\mathrm{TX},0}),\cos(\theta_{\mathrm{TX},0}),0]^{\mathrm{T}}, \tag{6.36b}$$

$$\mathbf{v}_{A,0}^{(\mathrm{G})} = v_{A,0}^{(\mathrm{G})}\left[\sin(\theta_{\mathrm{TX},0}),-\cos(\theta_{\mathrm{TX},0}),\left\|\mathbf{p}-\mathbf{q}\right\|\right]^{\mathrm{T}}, \tag{6.36c}$$

where $v_{A,0}^{(\mathrm{G})} = \sqrt{1/(1+\left\|\mathbf{p}-\mathbf{q}\right\|^2)}$ *ensures that* $\mathbf{v}_{A,0}^{(\mathrm{G})}$ *is unit-norm. All eigenvalues are positive, since the variance term* σ^2, *the distance between two points* $\|\mathbf{p}-\mathbf{q}\|$, *and the speed of light* c *are positive by definition.*

Proof. See Appendix A.2. □

The following conclusions can be drawn from Proposition 1. Delay, AOD, and AOA of the LOS path contribute information to the EFIM. Each of these quantities provides information in *one* direction, where the direction is given by the eigenvector of the non-zero eigenvalue. When the variance of a channel parameter (e.g, $\sigma_{\theta_{\mathrm{TX},0}}^2$) is finite, information is contributed to the EFIM. For infinite variance, no information is contributed. This situation is equivalent to parameters being not observable. Considering (6.35b) and (6.35c), it can be seen that the information gain decreases as the separation between the anchor

and agent increases. For a large separation, AOA provides only orientation information, since the x-y components of the eigenvector in (6.36c) go to zero. In that case, the amount of orientation information is given by $1/\sigma^2_{\theta_{\mathrm{RX},0}}$, meaning that it only depends on the quality of the AOA estimate. Similarly, the information provided by the AOD decreases as the agent moves away from the anchor. These observations reflect the following intuition: As the distance between an agent and an anchor becomes large, small angular estimation errors translate into large position estimation errors and hence angular estimates become less useful for positioning.

6.3.4.2 NLOS Information Gain

In this subsection, the information gain matrix $\tilde{\mathbf{B}}^{(\mathrm{G})}$ of NLOS paths is analyzed and decomposed using (6.23) and (6.27). This decomposition will be done in a similar manner as compared to the decomposition in the previous subsection.

Lemma 3. *The NLOS information gain matrix $\tilde{\mathbf{B}}^{(\mathrm{G})}$ is given by*

$$\tilde{\mathbf{B}}^{(\mathrm{G})} = \sum_{j=1}^{J} \underbrace{\frac{1}{\sigma^2_{\tau_j} c^2} \boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{RX},j}, 0, 0)}_{\tilde{\mathbf{B}}^{(\mathrm{G})}_{R,j}} + \underbrace{\frac{1}{\sigma^2_{\theta_{\mathrm{RX},j}} \left\|\mathbf{p}-\mathbf{s}_j\right\|^2} \boldsymbol{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2, \left\|\mathbf{p}-\mathbf{s}_j\right\|)}_{\tilde{\mathbf{B}}^{(\mathrm{G})}_{A,j}}. \tag{6.37}$$

Proof. See Appendix A.3.1. □

Note that the matrices $\tilde{\mathbf{B}}^{(\mathrm{G})}_{R,j}$ and $\tilde{\mathbf{B}}^{(\mathrm{G})}_{A,j}$ have similar structures as $\tilde{\mathbf{A}}^{(\mathrm{G})}_{R}$ and $\tilde{\mathbf{A}}^{(\mathrm{G})}_{A}$, respectively. Hence the eigenvalues and eigenvectors have similar structures as well, and are given by

$$\lambda^{(\mathrm{G})}_{R,j} = \frac{1}{\sigma^2_{\tau_j} c^2}, \tag{6.38a}$$

$$\lambda^{(\mathrm{G})}_{A,j} = \frac{\left\|\mathbf{p}-\mathbf{s}_j\right\|^2 + 1}{\sigma^2_{\theta_{\mathrm{RX},j}} \left\|\mathbf{p}-\mathbf{s}_j\right\|^2}. \tag{6.38b}$$

and

$$\mathbf{v}^{(\mathrm{G})}_{R,j} = [\cos(\theta_{\mathrm{RX},j}), \sin(\theta_{\mathrm{RX},j}), 0]^{\mathrm{T}}, \tag{6.39a}$$

$$\mathbf{v}^{(\mathrm{G})}_{A,j} = v^{(\mathrm{G})}_{A,j} \left[-\sin(\theta_{\mathrm{RX},j}), \cos(\theta_{\mathrm{RX},j}), \left\|\mathbf{p}-\mathbf{s}_j\right\|\right]^{\mathrm{T}}, \tag{6.39b}$$

respectively. The term $v^{(\mathrm{G})}_{A,j} = \sqrt{1/(1+\left\|\mathbf{p}-\mathbf{s}_j\right\|^2)}$ normalizes the eigenvector in (6.39b).

Interpretation: AOA and delay of each NLOS component contribute to the EFIM. The eigenvalues can be interpreted as the achievable information gains if the points of incidence were perfectly known. From (6.39a), it can be seen that delay provides only position information. From (6.39b), one can infer that AOA provides both position and orientation information. The term in (6.38b) suggests that the information gain of NLOS components depends on the separation between the agent and point of incidence. For large distances, the information gain converges to the inverse of the AOA estimation accuracy $1/\sigma^2_{\mathrm{RX},j}$. Observe that, in this case, a NLOS component provides mostly orientation information,

since the x-y components of the eigenvector in (6.39b) go to zero. The interpretation is the same as in the previous subsection. Note that AOD estimates by themselves do provide any information. However, it will be shown in the next section that the quality of the AOD estimation influences the information loss strongly.

6.3.4.3 NLOS Information Loss

The NLOS information loss matrix $\tilde{\mathbf{B}}^{(\mathrm{L})}$ is not as obvious to decompose as the gain matrices $\tilde{\mathbf{A}}^{(\mathrm{G})}$ and $\tilde{\mathbf{B}}^{(\mathrm{G})}$, since it involves multiple matrix-matrix products including the inverse of a matrix-matrix product as can be seen from (6.32). The result of the decomposition is depicted in the following.

Lemma 4. *The NLOS information loss matrix is given by*

$$\tilde{\mathbf{B}}^{(\mathrm{L})} = \sum_{j=1}^{J} \underbrace{w_{R,j}^{(\mathrm{L})} \boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{TX},0}, 0, 0)}_{\tilde{\mathbf{B}}_{R,j}^{(\mathrm{L})}} + \underbrace{w_{A,j}^{(\mathrm{L})} \boldsymbol{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2, \|\mathbf{p}-\mathbf{s}_j\|)}_{\tilde{\mathbf{B}}_{A,j}^{(\mathrm{L})}} - \gamma_{\mathbf{s}_j} \mathbf{B}_j^{(\mathrm{L})}, \tag{6.40}$$

where the weights[11] $w_{R,j}^{(\mathrm{L})}, w_{A,j}^{(\mathrm{L})}$ *and* $\gamma_{\mathbf{s}_j}$*, and the matrix* $\mathbf{B}_j^{(\mathrm{L})}$ *are defined in Appendix A.3.2.*

Proof. See Appendix A.3.2. □

Interpretation: The information loss of each path is composed of three terms. Comparing (6.37) and (6.40), it can be observed that the first two terms of every path j have identical structure. In particular, the delay information gain $\tilde{\mathbf{B}}_{R,j}^{(\mathrm{G})}$ and loss $\tilde{\mathbf{B}}_{R,j}^{(\mathrm{L})}$ differ only in the weights. The same holds for the AOA information gain $\tilde{\mathbf{B}}_{A,j}^{(\mathrm{G})}$ and loss $\tilde{\mathbf{B}}_{A,j}^{(\mathrm{L})}$. Hence each information gain and loss pair has eigenvectors that are pointing in the same direction. Yet, the eigenvalues have opposite signs. For these terms, net information gains are immediately given by the differences of the respective weights. However, the third term has eigenvectors which are not aligned with other eigenvectors.

It will be shown in the following subsection that the combination of the gain $\tilde{\mathbf{B}}_j^{(G)}$ and loss $\tilde{\mathbf{B}}_j^{(L)}$ matrices of every NLOS component results in a net information gain matrix which has only a single non-zero eigenvalue.

6.3.4.4 Net NLOS Information Gain

Lemma 5. *The net NLOS gain matrix* $\tilde{\mathbf{B}}^{(\mathrm{N})} \triangleq \tilde{\mathbf{B}}^{(\mathrm{G})} - \tilde{\mathbf{B}}^{(\mathrm{L})}$ *is given by*

$$\tilde{\mathbf{B}}^{(\mathrm{N})} = \sum_{j=1}^{J} \epsilon_{\mathbf{s}_j} \boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{RX},0}, 0, 0) + \beta_{\mathbf{s}_j} \boldsymbol{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2, \|\mathbf{p}-\mathbf{q}\|) + \gamma_{\mathbf{s}_j} \mathbf{B}_j^{(\mathrm{L})}, \tag{6.41}$$

where $\epsilon_{\mathbf{s}_j} \triangleq \frac{1}{\sigma_{\tau_j}^2 c^2} - w_{R,j}^{(\mathrm{L})}$ *and* $\beta_{\mathbf{s}_j} \triangleq \frac{1}{\sigma_{\theta_{\mathrm{RX},j}}^2 \|\mathbf{p}-\mathbf{s}_j\|^2} - w_{A,j}^{(\mathrm{L})}$.

Proof. Applying Lemma 3 and 4, (6.41) is obtained immediately. □

[11] Each weight is a function of $(\theta_{\mathrm{TX},j}, \theta_{\mathrm{RX},j}, \tau_j, \sigma_{\theta_{\mathrm{TX},j}}^2, \sigma_{\theta_{\mathrm{RX},j}}^2, \sigma_{\tau_j}^2, \mathbf{p}, \mathbf{q}, \mathbf{s}_j)$. To simplify notation, the arguments are dropped.

For compactness of the notation, the following matrices $\Psi_{\mathbf{s}_j}, j = 1, 2, ..., J$ are defined $\tilde{\mathbf{B}}^{(N)} \triangleq \sum_{j=1}^{J} \mathbf{\Psi}_{\mathbf{s}_j}$, where

$$\mathbf{\Psi}_{\mathbf{s}_j} \triangleq \epsilon_{\mathbf{s}_j} \mathbf{\Upsilon}_{0,0}(\theta_{\mathrm{TX},0}, 0, 0) + \beta_{\mathbf{s}_j} \mathbf{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2, \|\mathbf{p} - \mathbf{q}\|) + \gamma_{\mathbf{s}_j} \mathbf{B}_j^{(\mathrm{L})}. \tag{6.42}$$

Theorem 1. *The net information gain matrix $\mathbf{\Psi}_{\mathbf{s}_j}$ of the j^{th} NLOS component is rank one. The only non-zero eigenvalue of $\mathbf{\Psi}_{\mathbf{s}_j}$ is always positive and given by*

$$\lambda_{\mathbf{s}_j} \triangleq \frac{2 + \|\mathbf{p} - \mathbf{s}_j\|^2 (1 + \cos(\Delta\theta_j))}{(1 - \cos(\Delta\theta_j)) c^2 \sigma_{\tau_j}^2 + (1 + \cos(\Delta\theta_j))(\|\mathbf{p} - \mathbf{s}_j\|^2 \sigma_{\theta_{\mathrm{RX},j}}^2 + \|\mathbf{q} - \mathbf{s}_j\|^2 \sigma_{\theta_{\mathrm{TX},j}}^2)}. \tag{6.43}$$

The corresponding unit-norm eigenvector is

$$\mathbf{v}_{\mathbf{s}_j} \triangleq v_{\mathbf{s}_j} \begin{bmatrix} -\frac{\epsilon_{\mathbf{s}_j}}{\gamma_{\mathbf{s}_j}} \cos(\theta_{\mathrm{RX},j}) - \sin(\theta_{\mathrm{RX},j}) \\ -\frac{\epsilon_{\mathbf{s}_j}}{\gamma_{\mathbf{s}_j}} \sin(\theta_{\mathrm{RX},j}) + \cos(\theta_{\mathrm{RX},j}) \\ \|\mathbf{p} - \mathbf{s}_j\| \end{bmatrix}, \tag{6.44}$$

where $\Delta\theta_j \triangleq \theta_{\mathrm{RX},j} - \theta_{\mathrm{TX},j}$ *and* $v_{\mathbf{s}_j} = \sqrt{(1 + \cos(\Delta\theta_j))/(2 + \|\mathbf{p} - \mathbf{s}_j\|^2 (1 + \cos(\Delta\theta_j)))}$.

Proof. The details of the proof can be found in Appendix A.3.3. The outline of the proof is as follows: Using Lemma 5, it can be shown that $\mathbf{\Psi}_{\mathbf{s}_j}$ is rank one. Rank one matrices have the property that their only non-zero eigenvalues are obtained by taking the trace of the matrix [Lau04], i.e. $\lambda_{\mathbf{s}_j} = \mathrm{tr}\left(\mathbf{\Psi}_{\mathbf{s}_j}\right)$. Positivity of $\lambda_{\mathbf{s}_j}$ is readily seen. since all terms in (6.43) are non-negative, i.e. $(1 + \cos(\theta_{\mathrm{RX},j} - \theta_{\mathrm{TX},j})) \geq 0$. $(1 - \cos(\theta_{\mathrm{RX},j} - \theta_{\mathrm{TX},j})) \geq 0$, and all other terms are positive by definition. In order to show that $\mathbf{v}_{\mathbf{s}_j}$ in (6.44) is the corresponding eigenvector, it will be shown that it is in the null space of $\left(\lambda_{\mathbf{s}_j} \mathbf{I}_3 - \mathbf{\Psi}_{\mathbf{s}_j}\right)$. □

Interpretation: Each NLOS component provides one dimensional Fisher information for all three parameters (p_x, p_y, and α). The information from all NLOS components is additive. Hence each path contributes to the EFIM which, in turn, increases the position and orientation estimation accuracy. In other words, NLOS components in 5G mmWave mMIMO systems can be employed in order to support position and orientation estimation. The amount of information contributed by each NLOS component depends strongly on the geometry (this can be seen from the $\cos(\Delta\theta_j)$-terms in (6.43)), i.e. points of incidence, which are located in certain areas, provide more information than points of incidence in other areas. The following essential insight can be gained from the denominator in (6.43) which is given by the sum of weighted channel parameter variances. In order to provide significant Fisher information. all terms in the sum must take small values. If at least one term is large, the overall sum of terms is also large. In that case, the denominator becomes large which erases the Fisher information. There are two reasons why a term in the denominator can become large: (i) The variance σ^2 of a channel parameter (AOA, AOD, and delay) is large and/or (ii) the separation between the anchor or agent and a point of incidence is large. Hence the following conclusion can be drawn:

- If and only if AOA, AOD, and delay can be estimated with sufficient accuracy, NLOS components provide considerable Fisher information.

- NLOS paths, whose points of incidence are far away from the agent or anchor do not increase the Fisher information notably.

While the first statement might have been intuitively clear to a positioning expert, the detailed relationship presented in (6.43) and (6.44) are certainty not obvious. The second finding has the following interpretation: NLOS paths whose points of incidence are far away from the agent or anchor do not provide significant Fisher information, since their incidence points contain a lot of position uncertainty. Suppose that the AOD can be estimated with a certain accuracy, then the point of incidence is somewhere on the line which starts at the anchor and corresponds to the measured AOD. The further an incidence point moves away from the anchor, the larger is the positioning error of the incidence point due to an AOD measurement error. In other words, small angular errors translate into large positioning errors if the incidence points are far from the anchor. Identical considerations can be made regarding the incidence point, the agent, and the AOA.

The two observations above have substantial impact on the design of position and orientation estimators that want to employ NLOS paths. In particular, these findings advocate to ignore NLOS paths which have (i) long path lengths compared to the LOS path and (ii) whose parameters (AOA, AOD, and delay) cannot be estimated accurately[12]. In both cases (i) and (ii), the expected position and orientation estimation gain would be only marginal.

Corollary 1.1. *The EFIM of the position and orientation is given by the sum of the outer-products of the eigenvectors weighted by the corresponding eigenvalues, i.e.,*

$$\mathbf{J}^{\mathrm{c}}_{\check{\boldsymbol{\eta}}_{\mathbf{p},\alpha}} = \sum_{j\in R,D,A} \lambda^{(\mathrm{G})}_{j,0} \mathbf{v}^{(\mathrm{G})}_{j,0} \left(\mathbf{v}^{(\mathrm{G})}_{j,0}\right)^{\mathrm{T}} + \sum_{j=1}^{J} \lambda_{\mathbf{s}_j} \mathbf{v}_{\mathbf{s}_j} \mathbf{v}^{\mathrm{T}}_{\mathbf{s}_j}. \tag{6.45}$$

Proof. From Lemma 1 and Theorem 1, it is known that each matrix in (6.30) is rank one. Any rank one matrix can be written as the weighted outer-product of the unit-norm eigenvector which corresponds to the only non-zero eigenvalue [Lau04]. Hence (6.45) follows. □

6.3.5 Discussion and Implications

Note that the position and orientation of the receiver can be determined accurately if and only if $\mathbf{J}^{\mathrm{c}}_{\check{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}$ is non-singular. The sum of three rank one matrices with distinct eigenvectors provides a full rank (non-singular) matrix [Lau04]. Hence there are three ways of obtaining a non-singular Fisher information matrix $\mathbf{J}^{\mathrm{c}}_{\check{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}$:

1. The LOS path exists and its AOA, AOD, and delay are observable.
2. The LOS path is absent, but at least $J \geq 3$ NLOS paths with observable AOAs, AODs, and delays exist.
3. Any combination of the LOS and NLOS paths such that the sum in (6.45) contains three terms with distinct eigenvectors.

[12] One criterion to assess whether the parameters of a path can be estimated accurately is the received signal-to-noise-power ratio (SNR) of the respective path. As the SNR decreases, the estimation accuracy of the channel parameters generally decreases as well.

Another key observation that can be inferred from the fact that NLOS paths provide position and orientation information in the Fisher sense is the following: On top of the position and orientation of the agent, the points of incidence must be observable if NLOS paths provide significant Fisher information. This interesting byproduct of harnessing NLOS paths gives rise to an unforeseen application: simultaneous localization and mapping (SLAM). In particular, an agent can estimate its own location and orientation together with a map of the points of incidence in its proximity. Even more interestingly, the agent can estimate all of the aforementioned quantities based on a single transmission burst from the anchor. In contrast to classical radio-based SLAM [LMR+15, WML+16, GJW+16], neither mobility of the agent nor long convergence times are necessary.

6.4 Positioning and Mapping with Unknown Clock Offset

The previous section provided fundamental insights into the problem of estimating the position and orientation of an agent in a mmWave mMIMO system. It was revealed that NLOS components can be harnessed for position, orientation, and map estimation. Yet, these results rely on the assumption that the anchor and agent are perfectly synchronized in time. Even though, time synchronization is generally a prerequisite for communication systems, small clock offsets (in the order of nanoseconds) are inevitable. Recall from Section 2.2.1 that precise time synchronization is critical for the achievable ranging accuracy. Every nanosecond of synchronization error translates into a systematic ranging error of approximately 30 cm.

The common way of mitigating systematic ranging errors due to clock offsets is to employ two-way ranging protocols [LS02, ZLK07, ASWASG18]. For instance, if the agent attempts to determine the distance to the anchor, it has to send a message and wait for the reply of the base station. Then, the agent can calculate the range from the measured round-trip time which is independent of the clock offset. The issues with two-way ranging protocols are that they incur delays and two transmissions per range estimate are necessary which lead to congestion of the wireless network.

An alternative approach to two-way protocols is to incorporate the clock offset into the estimation problem. Obviously, this increases the difficulty of the estimation problem, since one more parameter (compared to the synchronized case) has to be taken into consideration. It is also not obvious whether the resulting estimation problem is solvable in an unambiguous way. In this section, an extended estimation problem will be formulated and the corresponding FIM will be derived in a closed form. In contrast to the previous section, the EFIM of the agent states (position, orientation, and clock offset) will not be derived in a closed form, since the additional parameter complicates the analysis significantly. However, necessary conditions regarding the observability[13] will be presented and it will be shown numerically in Section 6.5 that the extended estimation problem (which includes the clock offset) can be solved unambiguously if the LOS path and at least one NLOS component exist.

This section builds upon the results of the previous section. In particular, the definition of the channel parameter vector $\boldsymbol{\eta}$ and its FIM $\mathbf{J}_{\boldsymbol{\eta}}$ from Section 6.3.1 will be reused. The simplification step described in Section 6.3.2 and the dimension reduction procedure from

[13] An estimation problem is called observable if all parameters can be estimated unambiguously.

Section 6.3.3.1 are also considered. In other words, the analysis starts at the output of the third block in Fig. 6.3 where the FIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ of the reduced channel parameters is considered. The new parameter vector $\check{\boldsymbol{\eta}}$ and the modified geometric transformation matrix $\check{\mathbf{T}}$ will be presented in the following.

The new parameter vector $\check{\boldsymbol{\eta}}$ contains all parameters of the original problem and, in addition, it includes the clock offset ϵ. It is defined as

$$\check{\boldsymbol{\eta}} \triangleq [\mathbf{p}^{\mathrm{T}}, \alpha, \epsilon, \mathbf{s}_1^{\mathrm{T}}, ..., \mathbf{s}_J^{\mathrm{T}}]^{\mathrm{T}}. \tag{6.46}$$

As the clock offset ϵ is now part of the estimation problem, the transformation matrix $\mathbf{T}$, which transforms the FIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ of the reduced channel parameters into the domain of the position-related parameters, must be modified. Since the transformation of the FIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ of the reduced channel parameters requires determining the partial derivatives of all channel parameters with respect to all position-related parameters, the coupling of the channel parameters in (6.22) and the clock offset ϵ must be studied. It can be seen from (6.22a) and (6.22b) that the clock bias ϵ is only coupled with two channel parameters; the delay τ_0 of the LOS path and the delays $\tau_j, j = 1, 2, ..., J$ of the NLOS components. Hence the modified transformation matrix $\check{\mathbf{T}}$ can be obtained by inserting the following row into the original transformation matrix $\mathbf{T}$:

$$\mathbf{t}_{\epsilon} \triangleq \begin{bmatrix} \mathbf{t}_{\epsilon,0}^{\mathrm{T}} & \cdots & \mathbf{t}_{\epsilon,J}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \tag{6.47}$$

where

$$\mathbf{t}_{\epsilon,j} \triangleq \begin{bmatrix} \partial\tau_j/\partial\epsilon & \partial\theta_{\mathrm{TX},j}/\partial\epsilon & \partial\theta_{\mathrm{RX},j}/\partial\epsilon \end{bmatrix}^{\mathrm{T}}. \tag{6.48}$$

It is easy to verify that $\partial\tau_j/\partial\epsilon, \forall j$ are the only non-zero entries, i.e.

$$\mathbf{t}_{\epsilon,j} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\mathrm{T}}. \tag{6.49}$$

Thus, the modified transformation matrix $\check{\mathbf{T}}$ is given by

$$\check{\mathbf{T}} \triangleq \left[\begin{array}{c|ccc} \mathbf{T}_{\mathbf{P}}^{(0)} & \mathbf{T}_{\mathbf{P}}^{(1)} & \cdots & \mathbf{T}_{\mathbf{P}}^{(J)} \\ \mathbf{t}_{\epsilon,0}^{\mathrm{T}} & \mathbf{t}_{\epsilon,1}^{\mathrm{T}} & \cdots & \mathbf{t}_{\epsilon,J}^{\mathrm{T}} \\ \hline \mathbf{0}_{2\times 3} & \mathbf{T}_{\mathbf{s}_1} & \cdots & \mathbf{0}_{2\times 3} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{2\times 3} & \mathbf{0}_{2\times 3} & \cdots & \mathbf{T}_{\mathbf{s}_J} \end{array}\right]. \tag{6.50}$$

The FIM $\mathbf{J}_{\check{\boldsymbol{\eta}}}$ of the position, orientation, clock offset, and points of incidence can be determined based on the modified transformation matrix $\check{\mathbf{T}}$ and the FIM $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ of the reduced channel parameters. It is given by

$$\mathbf{J}_{\check{\boldsymbol{\eta}}} \triangleq \check{\mathbf{T}}\mathbf{J}_{\bar{\boldsymbol{\eta}}}\check{\mathbf{T}}^{\mathrm{T}}. \tag{6.51}$$

Remark: It is hard to draw conclusions from the FIM $\mathbf{J}_{\check{\boldsymbol{\eta}}}$, since it does not disclose any obvious structure. However, it is worth reviewing an important property of any FIM: A necessary condition for an observable estimation problem is that the corresponding FIM is full rank. The following necessary conditions for non-observability can be derived from this property.

Lemma 6. *The position, orientation, clock offset, and map estimation problem is not observable if*

1. *the LOS path exists, but no NLOS components are present ($J = 0$) or*
2. *the LOS path is absent and the number of NLOS paths is less or equal to three ($J \leq 3$).*

Proof. The estimation problem contains $4 + 2J$ parameters and the rank of the FIM $\mathbf{J}_{\check{\boldsymbol{\eta}}}$ is given by $\mathrm{Rank}(\mathbf{J}_{\check{\boldsymbol{\eta}}}) = \mathrm{Rank}(\check{\mathbf{T}}\mathbf{J}_{\bar{\boldsymbol{\eta}}}\check{\mathbf{T}}^{\mathrm{T}}) \leq \min(\mathrm{Rank}(\check{\mathbf{T}}), \mathrm{Rank}(\mathbf{J}_{\bar{\boldsymbol{\eta}}}))$. If the LOS is present, $\mathrm{Rank}(\mathbf{J}_{\bar{\boldsymbol{\eta}}}) = 3(J+1)$ and if the LOS is absent, $\mathrm{Rank}(\mathbf{J}_{\bar{\boldsymbol{\eta}}}) = 3J$. The previous statement holds, since $\mathbf{J}_{\bar{\boldsymbol{\eta}}}$ is a diagonal matrix and the elements on the diagonal are different from zero. Therefore, $\mathrm{Rank}(\mathbf{J}_{\check{\boldsymbol{\eta}}}) < 4 + 2J$ if the conditions 1. or 2. are true. □

In the following section, it will be shown numerically that the estimation problem is observable if the converse of condition 1. from Lemma 6 holds. In other words, the FIM $\mathbf{J}_{\check{\boldsymbol{\eta}}}$ in (6.51) is full rank and the estimation problem is observable if the LOS path and at least one NLOS component are present.

6.5 Numerical Results

This section contains numerical examples regarding the theoretical results derived in the previous sections of this chapter. The Cramér-Rao lower bounds are investigated in different scenarios. In particular, the effects of varying geometries are studied and, in addition, the behavior of the bounds with respect to the channel parameter variances is examined. The following bounds will be considered. If the clock offset is known (anchor and agent are synchronized), the PEB, OEB, and the scatterer error bound (SEB) are considered:

$$\mathrm{PEB}_{\mathrm{sync}} \triangleq \sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}^{\mathrm{e}}_{\check{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}\right)^{-1}\right]_{1:2,1:2}\right)}, \tag{6.52a}$$

$$\mathrm{OEB}_{\mathrm{sync}} \triangleq \sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}^{\mathrm{e}}_{\check{\boldsymbol{\eta}}_{\mathbf{p},\alpha}}\right)^{-1}\right]_{3,3}\right)}, \tag{6.52b}$$

$$\mathrm{SEB}_{\mathrm{sync}} \triangleq \sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}_{\check{\boldsymbol{\eta}}}\right)^{-1}\right]_{4:4+2J,4:4+2J}\right)}. \tag{6.52c}$$

The bounds presented in (6.52) embody fundamental lower limits on the achievable standard deviations of the respective parameter estimates obtained by an unbiased estimator. For instance, the PEB in (6.52a) is the lowest possible position standard deviation that an unbiased estimator can achieve. Similarly, the SEB in (6.52c) represents the lowest possible standard deviation of the estimate of all incidence points that is achievable by an unbiased estimator. The mathematical relationship is the following: Let $\mathbf{s} = [\mathbf{s}_1^{\mathrm{T}}, \cdots, \mathbf{s}_J^{\mathrm{T}}]$ be the vector of all incidence points and $\hat{\mathbf{s}}$ its estimate achieved by an unbiased estimator. Then, $\mathrm{SEB}_{\mathrm{sync}} \leq \sqrt{\mathrm{tr}(\mathbb{E}[(\mathbf{s}-\hat{\mathbf{s}})(\mathbf{s}-\hat{\mathbf{s}})^{\mathrm{T}}])}$.

If the clock bias ϵ is unknown, the bias error bound (BEB) will be considered in addition to the PEB, OEB, and SEB. Note that the definitions of the bounds in (6.53) differ from those in (6.52), since they are based on different Fisher information matrices.

$$\mathrm{PEB}_{\mathrm{nonsync}} \triangleq \sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}_{\check{\boldsymbol{\eta}}}\right)^{-1}\right]_{1:2,1:2}\right)}, \tag{6.53a}$$

$$\mathrm{OEB}_{\mathrm{nonsync}} \triangleq \sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}_{\check{\boldsymbol{\eta}}}\right)^{-1}\right]_{3,3}\right)}, \tag{6.53b}$$

$$\mathrm{BEB}_{\mathrm{nonsync}} \triangleq c\sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}_{\check{\boldsymbol{\eta}}}\right)^{-1}\right]_{4,4}\right)}, \tag{6.53c}$$

$$\mathrm{SEB}_{\mathrm{nonsync}} \triangleq \sqrt{\mathrm{tr}\left(\left[\left(\mathbf{J}_{\check{\boldsymbol{\eta}}}\right)^{-1}\right]_{5:5+2J,5:5+2J}\right)}. \tag{6.53d}$$

6.5.1 Impact of the Geometry

For now, it is assumed that the anchor and agent are synchronized, i.e. the clock offset ϵ is known. The main result from Section 6.3.4.4 (Theorem 1) revealed that the gain and direction of information in the Fisher sense depend exclusively on variances of the channel parameters and the geometry of the scenario. However, translating the Fisher information gain provided by an NLOS path into a gain in terms of positioning accuracy is not possible in a closed form. The reason is that this translation requires determining the inverse of an arbitrary 3×3 matrix which is not possible in closed-form. Therefore, a numerical analysis will be conducted. For the simplicity of the analysis, the variances of the channel parameters will be fixed and different geometries will be considered.

6.5.1.1 Simulation Setup

The considered geometries are depicted in Fig. 6.4. All position-related and angle-related quantities are considered in meters and degrees, respectively. They contain an anchor and an agent as well as up to three distinct points of incidence. The anchor is located at $\mathbf{q} = [0, 0]^{\mathrm{T}}$, while the agent is positioned at $\mathbf{p} = [60, 60]^{\mathrm{T}}$. The points of incidence are located at $\mathbf{s}_1 = [40, 0]^{\mathrm{T}}$, $\mathbf{s}_2 = [10, 50]^{\mathrm{T}}$, and $\mathbf{s}_3 = [70, 20]^{\mathrm{T}}$, respectively. The orientation of the agent is assumed to be $\alpha = 0$. The errors on the observed channel parameters (TOAs[14], AODs, and AOAs) are zero-mean Gaussian[15] distributed with the following

[14]Recall that a channel estimator does not directly measure the delay of a path but rather its TOA (see Section 2.2.1). From the measured TOA, the delay is inferred via the time t_{TX} of transmission. In the absence of precise synchronization, the time t_{TX} of transmission is not known at the receiver, leading to highly inaccurate delay estimates. However, the estimated TOA of a path may still be measured precisely and hence the standard deviation σ_{TOA_j} of TOA estimation takes small values. Thus, precise ranging is possible if the clock offset ϵ can be resolved *and* if the TOAs can be estimated accurately. Also, note that, in the synchronized case, the standard deviations σ_{τ_j} and σ_{TOA_j} of a delay and the corresponding TOA estimate are identical since no uncertainty regarding the clock offset exists and the error of the delay estimate originates solely from the TOA estimation error.

[15]Note that the measurement error statistics differ from those assumed in Chapter 5. The reason is that, in contrast to the assumptions in Chapter 5, paths are resolvable due to large bandwidth and antenna apertures.

standard deviations: $c\sigma_{\mathrm{TOA}_j} = 0.2$ and $\sigma_{\mathrm{TX},j} = \sigma_{\mathrm{RX},j} = 1,\ \forall j$. Similar error statistics have been observed in studies on channel estimators [SGD+18a].

Remark: It is worth noting that the errors on the observations (TOAs and angles) are not necessarily Gaussian in practice. So far, the error statistics are not sufficiently studied. Some studies [TVDW17, SGD+18b] suggest that the distribution of the errors is zero-mean Gaussian for the considered channel estimators. The standard deviations of the errors depend on many different aspects, including the number of transmit and receive antennas, the bandwidth, the transmit power, the surface of the scattering or reflecting object, etc. Taking all of these effects into consideration makes the analysis unnecessarily obscure and veils the effects that should be shown. Therefore, the standard deviations are considered as fictional values that are achievable by practical estimators [SGD+18b].

6.5.1.2 Results and Discussion

The following list contains the scenarios that will be studied and discussed in this subsection.

1. The LOS link exists, but no NLOS links are available.
2. The LOS link is obstructed and all points of incidence $\mathbf{s}_1$, $\mathbf{s}_2$. and $\mathbf{s}_3$ generate NLOS paths.
3. The LOS link exists and the point of incidence $\mathbf{s}_1$ generates an NLOS path.
4. The LOS link exists and the point of incidence $\mathbf{s}_3$ generates an NLOS path.

Fig. 6.4 depicts the geometries of the four scenarios mentioned above. The PEB of each scenario is depicted as the title of the respective sub-figure. The normalized eigenvectors of the corresponding FIMs in the x-y plane are also shown via colored arrows. The top left figure depicts the results of scenario 1. The eigenvector $\mathbf{v}_{\mathrm{R},0}^{(G)}$ of the LOS range estimate (c.f. (6.36a)) points in the direction of the line which connects the anchor and agent. However, the eigenvectors $\mathbf{v}_{\mathrm{D},0}^{(G)}$ and $\mathbf{v}_{\mathrm{A},0}^{(G)}$ (c.f. (6.36b) and (6.36c)) of the angular estimates are orthogonal to that of the range estimates. Since the corresponding FIM is invertible, a finite PEB of $\mathrm{PEB}_{\mathrm{sync},1} = 0.767$ is achieved. In other words, the LOS path is sufficient to determine position of an agent if the anchor and agent are synchronized.

The top right sub-figure demonstrates that LOS-free positioning is possible if the number of NLOS paths equals three. It is worth noting that the achievable positioning accuracy is generally lower compared to the LOS case, i.e. $\mathrm{PEB}_{\mathrm{sync},2} = 1.631 > \mathrm{PEB}_{\mathrm{sync},1}$. Intuitively, the positioning accuracy has to degrade compared to the LOS case, since the estimate on the position of the agent is based on the estimated points of incidence. In contrast to the position of the anchor, the estimates on the points of incidence inherently contain some estimation errors and uncertainties. Thus, the positioning accuracy of LOS-free positioning is generally lower compared to that of LOS-based positioning. How much the achievable estimation accuracy degrades, depends on the geometry of the scenario and how well the channel parameters of the NLOS paths can be estimated. For instance, if the geometry of the scenario is unfavorable and the estimation accuracy of the channel parameters are low, LOS-free positioning is still theoretically possible, but the achievable accuracy can be poor.

The left and right sub-figure at the bottom of Fig. 6.4 reflect the scenarios 3 and 4, respectively. In both cases, the LOS path and one NLOS path are available. The difference

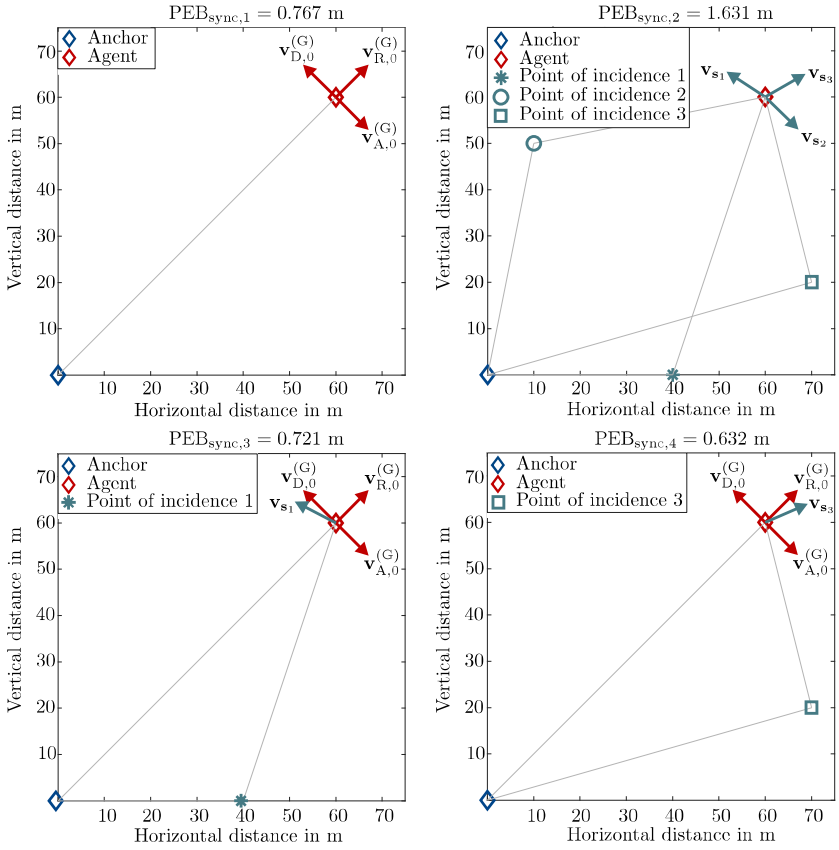

Figure 6.4: *Position error bound and direction of information* - the PEB and the directions of the eigenvectors of the FIMs from the scenarios 1-4 are depicted. Finite PEBs are achieved in all considered scenarios.

between the scenarios is the location of the incidence point. As expected, the presence of the additional NLOS path decreases the PEB in both cases, i.e. $\text{PEB}_{\text{sync},3} = 0.721 < \text{PEB}_{\text{sync},1}$ and $\text{PEB}_{\text{sync},4} = 0.632 < \text{PEB}_{\text{sync},1}$. It is worthwhile to mention, however, that the point of incidence $\mathbf{s}_3$ leads to a more pronounced decrease of the PEB compared to that of the point of incidence $\mathbf{s}_1$, i.e. $\text{PEB}_3 > \text{PEB}_4$. In other words, the position of the point of incidence matters and certain points result in larger reductions of the PEB compared to others. The specific relationship of the position of a point of incidence and the PEB is yet unknown. The important message – as shown theoretically in Theorem 1 and numerically in Fig. 6.4 – is that NLOS paths can be harnessed for position estimation as they reduce the PEB. It should be stressed, however, that the LOS link already provides high positioning accuracy and the reduction of the PEB due to the NLOS path is only moderate (here, $< 20\%$). In other words, it may not always be very rewarding to include the information from NLOS paths into the position estimation problem if the LOS path exists.

In summary, it was shown numerically that single-anchor positioning is possible if the LOS path exists. Moreover, it was presented that NLOS paths increase the positioning accuracy whenever they complement the LOS path. Though, the increase of the positioning accuracy is relatively moderate, since the LOS path already provides considerable position information. Finally, it was shown that unambiguous positioning is possible even in the absence of the LOS path if at least three distinct NLOS paths exist. Compared to the case where the LOS path exists, however, the positioning accuracy is degraded.

6.5.2 Impact of the Observation Variances

From the previous subsection, it was observed that the geometry of the scenario has a profound impact on the achievable estimation accuracy. The other important aspect concerns the accuracy of the channel parameter estimates. One may ask the question: *What is the effect of poor angle or delay estimation on the performance bounds?* Or: *Which angular estimate (AOD or AOA) is more important for achieving low performance bounds?* Questions like these will be answered in this subsection. As the results in the synchronized case differ considerably from the results in the non-synchronized case, both cases will be analyzed here.

6.5.2.1 Simulation Setup

The geometry from scenario 3 is considered (depicted in the bottom left sub-figure of Fig. 6.4), where $\mathbf{q} = [0, 0]^{\mathrm{T}}$and $\mathbf{p} = [60, 60]^{\mathrm{T}}$. All position-related and angle-related quantities are considered in meters and degrees, respectively. The point of incidence is located at $\mathbf{s}_1 = [40, 0]^{\mathrm{T}}$. For the simplicity of the analysis, only one standard deviation will be varied at a time, while the others remain unchanged. The values of fixed standard deviations are $c\sigma_{\mathrm{TOA}_j} = 0.2$ and $\sigma_{\mathrm{TX},j} = \sigma_{\mathrm{RX},j} = 1$, $\forall j$. The angular standard deviations are varied in the range from $\sigma_{\mathrm{TX},j}/\sigma_{\mathrm{RX},j} = 0.1$ to $\sigma_{\mathrm{TX},j}/\sigma_{\mathrm{RX},j} = 10, \forall j$, while the scaled TOA standard deviations from $c\sigma_{\mathrm{TOA}_j} = 0.05$ to $c\sigma_{\mathrm{TOA}_j} = 2, \forall j$ are considered.

6.5.2.2 Results and Discussion

Fig. 6.5 depicts the PEB, SEB, and BEB versus the standard deviation σ_{RX} of the AOAs. The left sub-figure depicts the synchronized case, while the right sub-figure shows the non-synchronized case. Note that the BEB is only defined in the non-synchronized case. If the anchor and agent are synchronized, two important observations can be made.

1. Precise AOA estimates ($\sigma_{\mathrm{RX}} \ll 1$) increase the position and point of incidence estimation accuracy greatly.

2. Imprecise AOA estimates ($\sigma_{\mathrm{RX}} \gg 2$) do not affect the position and point of incidence estimation accuracy considerably.

Observation 1 has the following implication: The positioning and mapping accuracy benefits significantly from highly accurate AOA estimates. Since the AOA estimation accuracy scales cubic in the number of receive antennas [ASZA+18a], an agent with a large antenna array can experience highly accurate position and point of incidence estimates. Observation 2 provides also interesting insights into the estimation problem. Since the PEB and SEB go into saturation for large[16] σ_{RX}, the AOA estimates can be regarded as non-essential

[16]Recall that $\sigma \to \infty$ corresponds to a non-available estimate on a certain channel parameter.

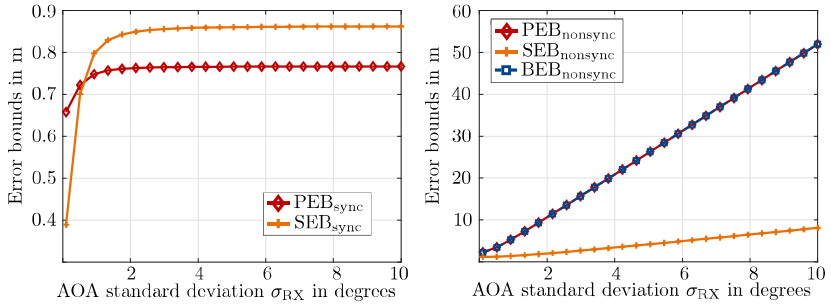

Figure 6.5: *Error bounds versus AOA standard deviation* σ_{RX} - synchronized anchor and agent (left) and non-synchronized anchor and agent (right).

for the estimation problem. The reason is that the AOD and delay of the LOS path resolve the position of the agent and, when this position is determined, the AOD and delay of multipath components are sufficient to resolve the positions of the incidence points. In other words, positioning and mapping can be performed even in the absence of the AOA estimates. Moreover, the AOA estimates are only helpful if their standard deviations are relatively low.
A completely different situation can be observed if the anchor and agent are not synchronized (right sub-figure in Fig. 6.5). Two significant effects can be concluded:

1. Accurate AOA estimation is a prerequisite in order to resolve the clock offset. Only if the AOAs can be estimated accurately, low PEB, SEB, and BEB can be achieved.

2. Mapping can still be performed relatively accurately if the standard deviation of the AOA is large, but positioning and clock offset estimation is poor.

Observation 1. implies that high-accuracy joint positioning, mapping, and synchronization can only be performed if the AOA can be estimated accurately. Otherwise, the clock offset cannot be resolved well and the positioning accuracy suffers dramatically. From the second observation, it can be deduced that the AOA estimates do not seem to play a critical role during the estimation of the points of incidence. In other words, AOD and TOA are the important measurements regarding the point of incidence estimation if the anchor and agent are not synchronized.

Fig. 6.6 depicts the PEB, SEB, and BEB versus the standard deviation σ_{TX} of the AODs. The left sub-figure depicts the synchronized case, while the right sub-figure shows the non-synchronized case. Compared to the previous analysis of the AOA standard deviation σ_{RX}, it can be seen that the AODs have more profound impact on the achievable estimation accuracy. Poor AOD estimation immediately leads to poor positioning and mapping accuracy. In contrast to the AOAs in the synchronized case, the AODs play a key role in the estimation of the positions of the agent and the points of incidence. The reason is that – in the synchronized case – the AOD and TOA estimate of the LOS path determine the position of the agent unambiguously. However, if the position is known, the AOD and TOA estimates of the NLOS paths directly provide estimates on the points of incidence. Hence poor AOD estimates result in poor position and point of incidence estimates.

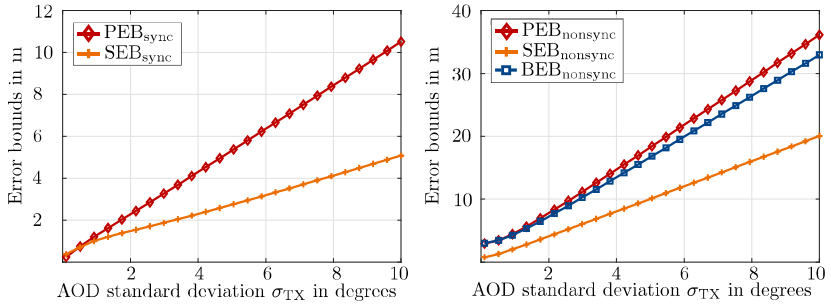

Figure 6.6: *Error bounds versus AOD standard deviation* σ_{TX} - synchronized anchor and agent (left) and non-synchronized anchor and agent (right).

If the anchor and agent are not synchronized, a similar trend compared to the AOA can be observed, i.e. accurate AOD estimation is necessary to obtain low PEB, SEB, and BEB. The reason is that the clock offset cannot be resolved well if the AOD estimation accuracy is poor.

Fig. 6.7 depicts the PEB, SEB, and BEB versus the scaled TOA standard deviation $c\sigma_{\text{TOA}}$. The left sub-figure depicts the synchronized case, while the right sub-figure shows the non-synchronized case. The same observations as compared to the Fig. 6.6 can be made. In particular, a degradation of the TOA estimation accuracy leads to a persistent increase of the performance bounds in both the synchronized and non-synchronized case. It is worth mentioning that, in the synchronized case, relatively large TOA estimation errors still result in moderate positioning errors. For instance, if the scaled TOA standard deviation $c\sigma_{\text{TOA}} = 2$ m, the achievable positioning accuracy is approximately 1.5 m. The reason is that accurate angular measurements can rectify large TOA estimation errors to some extend.
In the non-synchronized case, the increase of the error bounds is more pronounced as the TOA standard deviation $c\sigma_{\text{TOA}}$increases. It can be concluded that accurate TOA measurements are also a prerequisite for resolving the clock offset. Note that, in the absence of anchor-agent synchronization, the error bounds in Fig. 6.7 flatten out for small TOA measurement errors. These behaviors of the bounds mean that even large improvements of the TOA measurement quality will not improve the estimation accuracy considerably, unless the angular estimates become more accurate as well.

In summary, it can be concluded that the achievable estimation accuracy is relatively sensitive to large measurement errors. However, there is a discrepancy depending on the synchronization state of the anchor and agent. If the anchor and agent are synchronized, a small increase of measurement errors does not affect the estimation accuracy notably. Yet, agents that are not synchronized to anchors suffer drastically from an increase in the standard deviation of the measurement errors. While, in the synchronized case, precise AOD and TOA measurements are sufficient to obtain high position and map estimation accuracies, all measurements (TOA, AOD, and AOA) need to be precise in the absence of anchor-agent synchronization. In other words, a single poor measurement can deteriorate the estimation accuracy considerably. This observation puts a heavy burden on the joint position, map, and clock offset estimation problem in practice.

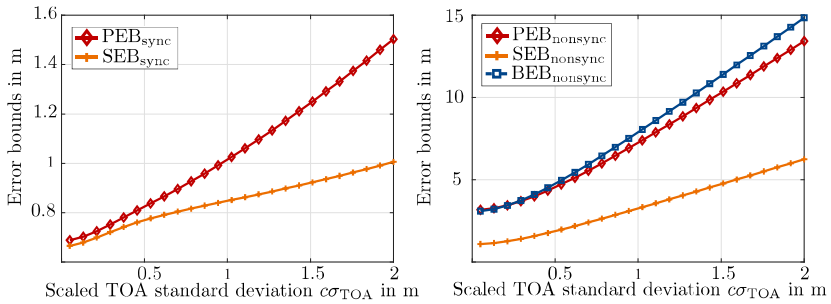

Figure 6.7: *Error bounds versus scaled TOA standard deviation* $c\sigma_{\mathrm{TOA}}$ - synchronized anchor and agent (left) and non-synchronized anchor and agent (right).

6.6 Conclusion

The findings in this chapter provided insights into the problem of position, orientation, map, and clock offset estimation in 5G mmWave mMIMO. It was shown both theoretically and numerically that single-anchor positioning is possible in mmWave mMIMO. Beyond the sole purpose of localization, other states of the agent such as its orientation and clock offset to the anchor can also be estimated. On top of these agent states, mmWave mMIMO systems have the capabilities to provide estimates on the points of incidence of the reflected or scattered NLOS paths which allows for a mapping of the radio environment. Compared to conventional radio-based positioning systems, mmWave mMIMO systems bear the potential to provide an unprecedented degree of location and situation awareness.

During the course of this chapter, the following insights into the capabilities of mmWave mMIMO systems have been gained. It was shown that single-anchor localization can be rendered possible if the LOS path between a synchronized anchor-agent pair exists. Any single-bounce NLOS path was proven to increase the achievable localization accuracy compared to LOS-only positioning. This marks a paradigm shift in the domain of radio-based positioning where NLOS propagation was regarded as a source of perturbation for a long time. Under certain circumstances (if at least three NLOS paths exist), even LOS-free positioning becomes possible.

Whenever NLOS paths are harnessed for positioning, estimates on the points of incidence of the NLOS paths are immediate byproducts. Using these estimates, a map with the relevant features in the transmission environment can be generated. Thus, simultaneous localization and mapping can be performed based on mmWave mMIMO systems that are used for communication purposes anyways. Another interesting aspect which mmWave mMIMO signals bring to the table is the fact that precise anchor-agent synchronization is not necessary if at least one NLOS path complements the LOS path. The clock offset between the anchor and agent can be resolved by including it in the position, orientation, and map estimation problem. Clearly, this leads to a complication of the estimation problem, but (sub-)nanosecond synchronization required for accurate ranging can be mitigated by estimating the clock offset.

It is worthwhile to mention that all of the above findings have been derived based on simplified models of the reality which enabled tractable analyses. The findings summarized above display deep insights into the potential of mmWave mMIMO systems. How this

potential can be exploited in practice, though, remains to be seen. It was already mentioned earlier that the number of NLOS paths with significant energy is usually small. While it was shown theoretically that LOS-free localization is possible, it is not obvious how often three or more strong NLOS paths will be available in actual mmWave mMIMO systems. In other words, it is not yet clear whether LOS-free localization will be feasible in practice. Moreover, even if strong NLOS paths exist, it is not guaranteed that estimators will be able to capitalize on those paths. As a first step towards practicability, feasible estimators are required that attain the fundamental limits embodied by the performance bounds shown in this chapter. This step will be taken in the next chapter where a message passing-based estimator will be presented.

Chapter 7

Single-Anchor Positioning Part II: Algorithmic Advances

The theoretical limits of millimeter wave (mmWave) massive multiple-input multiple-output (mMIMO) systems were explored in the last chapter, where it was shown that a number of agent states (position, orientation, and clock offset) can be estimated along with a map of reflecting or scattering objects in the transmission environment. These insights have the potential to push the boundaries of radio-based positioning to unprecedented heights. Yet, tractable estimators will be required in order to fully harness the potential of mmWave mMIMO systems. In this chapter, an estimation framework will be presented that comes very close to the ultimate performance limits as it attains the Cramér-Rao lower bounds (CRLBs) in certain scenarios. The presented estimation method is capable of jointly estimating the position, orientation, and clock offset of an agent along with the incidence points of NLOS paths in the radio environment.

The chapter is structured as follows: First, the state of the art will be reviewed and the novelty of the contribution will be highlighted in Section 7.1. The estimation problem will be formulated mathematically in Section 7.2. Section 7.3 contains the main contribution of this chapter where the joint posterior distribution will be presented, the corresponding factor graph will be developed, and the mechanics of the algorithm will be elaborated. A numerical analysis of the proposed estimator will be conducted in Section 7.4 and Section 7.5 concludes this chapter.

7.1 Introduction

The previous chapter presented deep insights into the positioning capabilities of mmWave mMIMO systems. Theoretical analyses showed that many position-related parameters can be estimated in such systems and theoretical limits regarding the achievable estimation accuracy were presented. Yet, much like Shannon's channel capacity observations in 1948, knowing the fundamental limits does not imply that it is obvious how they can be achieved. In fact, designing methods that achieve these ultimate performance limits are not obvious at all. In this section, state-of-the-art methods will be discussed and subsequently, the contributions of this chapter will be summarized.

7.1.1 State of the Art

The domain of high-dimensional mmWave mMIMO positioning problems has evolved very recently. On a broader scale, concurrent positioning of an agent and the extraction of features in its radio environment have been investigated in the last few years [LMR+15, WML+16, GJW+16] under the umbrella of multipath-assisted positioning. Preceding studies have focused on ultrawideband (UWB) signals [LMR+15, WML+16] and signals of opportunity [GJW+16] which embody a class of signals that includes, e.g., Long Term Evolution (LTE) and wireless fidelity (WiFi) systems. These approaches were not derived in the context of mmWave mMIMO systems. Thus, the considered measurements are restrict to TOA and angle-of-arrival (AOA) estimates. As it was shown in the previous section, the angles-of-departure (AODs) are required to obtain unambiguous estimates in a static scenario. In mobile scenarios, the ambiguity of the estimates of these approaches can be resolved over time. However, long convergence times are usually required and the computational complexity of the estimators can be high.

The availability of precise AOD, AOA, and delay estimates in fifth generation (5G) mmWave mMIMO systems has the potential to mark a turning point in the field of multipath-assisted positioning as shown theoretically in Chapter 6. Yet, positioning and mapping techniques in that domain are scarce and those that exist are subject to various restrictions and limiting assumptions. Current efforts [SGD+18b, RG18, TVDW17, MWB18, WGK+18] consider different estimation problems and take different approaches to tackle those problems. A nonlinear least squares (NLS) estimator[1] that determines the position and orientation of an agent was presented in [SGD+18b]. Another estimator that tackles the previous problem is based on Gibbs sampling and was proposed in [TVDW17]. In [RG18], a gradient-assisted particle filter was introduced which estimates the points of incidence of non-line-of-sight (NLOS) paths in addition to the position of an agent. However, the orientation of the agent must be known. While the estimators in [SGD+18b, RG18, TVDW17] explicitly rely on the presence of the line-of-sight (LOS) link. [MWB18] puts forth a message passing-based estimator that can perform position and orientation estimation along with mapping of the incidence points entirely based on NLOS paths. [WGK+18] presents a belief propagation-based approach that allows for the joint estimation of the position, orientation, and clock offset of an agent along with a map of the features in its radio environment. However, the presented approach does not attain the CRLBs. Except for [WGK+18], none of the discussed estimators considers the non-synchronized case, where the anchor and agent have a clock offset. However, this problem is particularly interesting, since including the clock offset into the estimation problem relaxes the synchronization requirements between an anchor and an agent. Recall that (sub)nanosecond synchronization is required in order to employ delay estimates. Achieving and maintaining such a high synchronization accuracy between anchors and agents is generally a heavy burden for any positioning system. Thus, developing estimators for joint positioning and synchronization can provide various benefits from the positioning perspective.

[1] It should be noted that the NLS approach from [SGD+18b] was shown to achieve the CRLB if the anchor and agent are synchronized. Therefore, this problem will not be studied in this chapter.

7.1.2 Contribution

Even though the field of position estimation in mmWave mMIMO systems is in its early stages, [SGD+18b] presented an estimator which attains the CRLB in the case where the anchor and agent are synchronized. Therefore, the research presented in this chapter focuses on the extended and more complicated estimation problem in which the clock offset between anchor and agent is unknown. This chapter presents an estimation framework that is capable of tackling the extended estimation problem. The proposed estimator performs a variant of sample-based belief propagation. In that way, the marginal distributions of the agent states (position, orientation, and clock offset) and the points of incidence are approximately obtained by passing messages along the edges of the underlying factor graph of the inference problem. Messages are approximated by particle representations as described in Chapter 3. MMSE estimates are derived from these marginals in order to conclude the estimation procedure. The following list contains the main contributions of this chapter.

1. A unified and modular estimation framework will be presented which can jointly estimate the position, orientation, and clock offset of an agent together with the points of incidence in its radio environment.

2. Modularity of the framework will be presented in the sense that it can also target simplified estimation problems by solely deleting nodes and the corresponding edges from the underlying factor graph.

3. It will be shown numerically that the proposed estimator attains the CRLBs in scenarios where the LOS path and a single NLOS path are present.

7.2 Definition of the Estimation Problem

The estimation of the agent states along with the map of the radio environment is based on measurements of the channel parameters as outlined in the previous chapter. These parameters include the ranges $d_j = c\tau_j$, AODs $\theta_{\mathrm{TX},j}$, and AOAs $\theta_{\mathrm{RX},j}$ of all received paths j. Recall that the range estimate of a path can be concluded from the respective time-of-arrival (TOA) estimate. In the absence of anchor-agent synchronization, the range estimate can be inaccurate, even though the TOA estimate itself is very accurate. The reason is the unknown clock offset ϵ. Channel estimators for mmWave mMIMO systems which provide TOA, AOD, and AOA estimates of every path have been recently presented in [XGJ16, SGD+18b, CW18, YZZ18]. Note that AOD estimation requires knowledge regarding the beamforming matrix $\mathbf{F}$ that was chosen at the anchor, i.e. AOD estimation requires the incorporation of overhead. Such information could be included, e.g., in the header of a data package or in a dedicated signal such as a positioning reference signal (PRS). The measurements $\hat{d}_j$, $\hat{\theta}_{\mathrm{TX},j}$, and $\hat{\theta}_{\mathrm{RX},j}$ of those parameters are collected path-wise in the observation vector

$$\mathbf{z}_j = [\hat{d}_j, \hat{\theta}_{\mathrm{TX},j}, \hat{\theta}_{\mathrm{RX},j}]^{\mathrm{T}} \tag{7.1}$$

where $j = 0$ indexes the LOS path and $j = 1, \ldots, J$ summarize the NLOS paths.

Recall the definition of the LOS, NLOS, and obstructed-line-of-sight (OLOS) channel states used in this thesis. LOS refers to the channel condition, where only the direct (LOS) path exists and no indirect (NLOS) paths are present. The NLOS condition includes the

direct (LOS) path and at least one indirect (NLOS) path. Finnally, OLOS describes the channel condition, where the direct (LOS) path is obstructed and at least one indirect (NLOS) path exists.

The estimator to be presented in Section 7.3 makes it possible to tackle a number of different estimation problems. The estimation problem that will be studied in detail in this chapter is the following. Let

$$\check{\boldsymbol{\eta}} \triangleq [\mathbf{p}^{\mathrm{T}}, \alpha, \epsilon, \mathbf{s}_1^{\mathrm{T}}, ..., \mathbf{s}_J^{\mathrm{T}}]^{\mathrm{T}}$$

be the parameter vector of the joint position, orientation, synchronization and mapping problem. Given the LOS measurements $\mathbf{z}_0$ and a number of NLOS observations $\mathbf{z}_j, j = 1, \ldots, J$ as well as their statistical models, the estimation problem can be formulated in the maximum a posteriori (MAP) sense like in (2.1). However, it is an issue that the parameter vector $\check{\boldsymbol{\eta}}$ is high-dimensional ($\dim(\check{\boldsymbol{\eta}}) = 4 + 2J$) and the relationships between the measurements (AOAs, AODs, and delays) and the estimation parameters (position, orientation, clock offset, and points of incidence) are highly nonlinear. These two issues render the numerical optimization required for MAP estimation impractical. However, it is generally tractable to obtain estimates based on the marginal distributions. Prominent examples are marginal MAP or minimum mean square error (MMSE) estimates. In this context, the focus is on marginal MMSE estimates as they are easy to determine if the marginal distributions are given as particle representations. Hence the estimation problem is defined as follows:

$$\hat{\mathbf{p}}_{\mathrm{MMSE}} = \int \mathbf{p} f(\mathbf{p}|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\mathbf{p}, \tag{7.2a}$$

$$\hat{\alpha}_{\mathrm{MMSE}} = \int \alpha f(\alpha|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\alpha, \tag{7.2b}$$

$$\hat{\epsilon_{\mathrm{MMSE}}} = \int \epsilon f(\epsilon|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\epsilon, \tag{7.2c}$$

$$\hat{\mathbf{s}}_{j,\mathrm{MMSE}} = \int \mathbf{s}_j f(\mathbf{s}_j|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\mathbf{s}_j \quad j = 1, ..., J, \tag{7.2d}$$

where $\mathbf{z}_{\mathrm{NLOS}} = [\mathbf{z}_0^{\mathrm{T}}, \ldots, \mathbf{z}_J^{\mathrm{T}}]^{\mathrm{T}}$. The marginal distributions in (7.2) are approximated by the beliefs that are obtained using belief propagation. For instance, $f(\mathbf{p}|\mathbf{z}_{\mathrm{NLOS}}) \approx b_{\mathbf{p}}^{(l)}(\mathbf{p})$, where $b_{\mathbf{p}}^{(l)}(\mathbf{p})$is the belief of $\mathbf{p}$ in the l^{th} iteration. Note that the necessary conditions for the estimation problem to have an unambiguous solution are that (i) the LOS path exists and (ii) at least one NLOS path is present [WGK+18]. Otherwise, the problem is ill-posed.

The simplified version of this problem, where the anchor and agent are synchronized. has the following parameter $\check{\boldsymbol{\eta}} \triangleq [\mathbf{p}^{\mathrm{T}}, \alpha, \mathbf{s}_1^{\mathrm{T}}, ..., \mathbf{s}_J^{\mathrm{T}}]^{\mathrm{T}}$ and is defined as:

$$\hat{\mathbf{p}}_{\mathrm{MMSE}} = \int \mathbf{p} f(\mathbf{p}|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\mathbf{p}, \tag{7.3a}$$

$$\hat{\alpha}_{\mathrm{MMSE}} = \int \alpha f(\alpha|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\alpha, \tag{7.3b}$$

$$\hat{\mathbf{s}}_{j,\mathrm{MMSE}} = \int \mathbf{s}_j f(\mathbf{s}_j|\mathbf{z}_{\mathrm{NLOS}}) \mathrm{d}\mathbf{s}_j \quad j = 1, ..., J. \tag{7.3c}$$

This estimation problem can be shown to have an unambiguous solution as long as the LOS path exists [WGK+18]. Another interesting problem is the OLOS problem, where the LOS path is absent:

$$\hat{\mathbf{p}}_{\mathrm{MMSE}} = \int \mathbf{p} f(\mathbf{p}|\mathbf{z}_{\mathrm{OLOS}}) \mathrm{d}\mathbf{p}, \tag{7.4a}$$

$$\hat{\alpha}_{\mathrm{MMSE}} = \int \alpha f(\alpha|\mathbf{z}_{\mathrm{OLOS}}) \mathrm{d}\alpha, \tag{7.4b}$$

$$\hat{\mathbf{s}}_{j,\mathrm{MMSE}} = \int \mathbf{s}_j f(\mathbf{s}_j|\mathbf{z}_{\mathrm{OLOS}}) \mathrm{d}\mathbf{s}_j \quad j = 1, ..., J, \tag{7.4c}$$

where $\mathbf{z}_{\mathrm{OLOS}} = [\mathbf{z}_1^{\mathrm{T}}, \ldots, \mathbf{z}_J^{\mathrm{T}}]^{\mathrm{T}}$. As was shown in the last chapter, this problem has an unambiguous solution if at least three NLOS paths exist. However, it was shown in Chapter 6 that the channel parameter measurements must be accurate in order to allow for position estimates with reasonable accuracy.

Note that many other problem formulations are possible. Yet, they all have one thing in common. Their solutions require to determine the marginal distributions. A unified and modular framework based on belief propagation will be presented in the following section. The problem which will be studied in depth in this chapter is the one defined in (7.2). Though, it will be emphasized what adjustments to the belief propagation procedure have to be made in order to transform the procedure for the estimation problem in (7.2) into approaches for the related estimation problems in (7.3) and (7.4).

7.3 COMPAS: Concurrent Mapping, Positioning, and Synchronization

The main contribution of this chapter is contained in this section. First, the joint posterior distribution of the estimation problem will be developed and the corresponding factor graph will be concluded from it. Subsequently, a belief propagation-based algorithm will be presented that exchanges messages along the edges of the underlying factor graph in order to determine the approximate marginal distributions. The position-related parameters will be estimated based on these approximated marginals. Finally, a particle-based implementation will be presented.

7.3.1 Joint Posterior Distribution

The cornerstone to develop a belief propagation algorithm is to find a suitable factorization of the joint posterior distribution as discussed in Chapter 4. In order to factorize the joint posterior distribution, independencies between the measurements and parameters as well as inter-parameter independencies must be identified. Using Bayes' rule, the joint posterior distribution can be rewritten as being proportional to the joint likelihood times the joint prior, i.e.

$$f(\check{\boldsymbol{\eta}}|\mathbf{z}_{\mathrm{NLOS}}) \propto f(\mathbf{z}_{\mathrm{NLOS}}|\check{\boldsymbol{\eta}}) f(\check{\boldsymbol{\eta}}). \tag{7.5}$$

Based on this representation, the goal is to further break up the two factors $f(\mathbf{z}_{\mathrm{NLOS}}|\check{\boldsymbol{\eta}})$ and $f(\check{\boldsymbol{\eta}})$ in (7.5). Starting with the joint prior distribution $f(\check{\boldsymbol{\eta}})$, it is reasonable to assume that the priors of all individual parameters are independent. For instance, there

is no reason to believe that the clock offset ϵ and the position $\mathbf{p}$ of the agent share a dependency. Moreover, if the points $\mathbf{s}_j, j = 1, \cdots, J$ of incidence are simply treated as points with no spatial extend. they can be assumed independent of the position $\mathbf{p}$ of the agent. Hence the joint prior distribution can be written as

$$f(\check{\boldsymbol{\eta}}) = f_{\mathbf{p}}(\mathbf{p}) f_{\alpha}(\alpha) f_{\epsilon}(\epsilon) \prod_{j=1}^{J} f_{\mathbf{s}_j}(\mathbf{s}_j). \tag{7.6}$$

Now, it is time to open the factor $f(\mathbf{z}_{\mathrm{NLOS}}|\check{\boldsymbol{\eta}})$ of the joint likelihood function and reveal its structure in the form of an appropriate factorization. Results from the previous chapter turn out to be helpful here. As discussed in Section 6.3.2, different paths can be estimated independently which is reflected by the block-diagonal Fisher information matrix (FIM) of the channel parameters in (6.15). Recall that this block-diagonal structure of the FIM is no assumption but an inherent property as the bandwidth and antenna array sizes approach infinity. Hence the joint likelihood function can be written as

$$f(\mathbf{z}_{\mathrm{NLOS}}|\check{\boldsymbol{\eta}}) = \prod_{j=0}^{J} f(\mathbf{z}_j|\check{\boldsymbol{\eta}}). \tag{7.7}$$

Moreover, the delay (range), AOD, and AOA of every path j can be estimated independently, since the (E)FIM-block defined by (6.16) and (6.19) is diagonal. Based on these insights and the relationships in (6.22), the joint likelihood function $f(\mathbf{z}_j|\check{\boldsymbol{\eta}})$ of the j^{th} path can be further factorized as

$$f(\mathbf{z}_0|\check{\boldsymbol{\eta}}) = f_{d_0}(\hat{d}_0|\mathbf{p}, \epsilon) f_{\mathrm{TX}_0}(\hat{\theta}_{\mathrm{TX}_0}|\mathbf{p}) f_{\mathrm{RX}_0}(\hat{\theta}_{\mathrm{RX}_0}|\mathbf{p}, \alpha), \tag{7.8a}$$

$$f(\mathbf{z}_j|\check{\boldsymbol{\eta}}) = f_{d_j}(\hat{d}_j|\mathbf{p}, \epsilon, \mathbf{s}_j) f_{\mathrm{TX}_j}(\hat{\theta}_{\mathrm{TX},j}|\mathbf{s}_j) f_{\mathrm{RX}_j}(\hat{\theta}_{\mathrm{RX},j}|\mathbf{p}, \alpha, \mathbf{s}_j), \quad j = 1, \ldots, J. \tag{7.8b}$$

For the compactness of the notation and unlike the factors in cooperative localization, the anchor position $\mathbf{q}$ is not explicitly visible in the factors in (7.8), but it is implicitly contained. Combining (7.5)-(7.8) yields the following factorization of the joint posterior distribution:

$$\begin{aligned} f(\check{\boldsymbol{\eta}}|\mathbf{z}_{\mathrm{NLOS}}) \propto & \underbrace{f_{d_0}(\hat{d}_0|\mathbf{p}, \epsilon) f_{\mathrm{TX}_0}(\hat{\theta}_{\mathrm{TX},0}|\mathbf{p}) f_{\mathrm{RX}_0}(\hat{\theta}_{\mathrm{RX},0}|\mathbf{p}, \alpha)}_{\text{LOS information}} \\ & \times \underbrace{\prod_{j=1}^{J} f_{\mathrm{TX}_j}(\hat{\theta}_{\mathrm{TX},j}|\mathbf{s}_j) \overbrace{f_{d_j}(\hat{d}_j|\mathbf{p}, \epsilon, \mathbf{s}_j) f_{\mathrm{RX}_j}(\hat{\theta}_{\mathrm{RX},j}|\mathbf{p}, \alpha, \mathbf{s}_j)}^{\triangleq v_j(\hat{d}_j, \hat{\theta}_{\mathrm{RX},j}|\mathbf{p}, \alpha, \epsilon, \mathbf{s}_j)}}_{\text{NLOS information}} \\ & \times \underbrace{f_{\mathbf{p}}(\mathbf{p}) f_{\alpha}(\alpha) f_{\epsilon}(\epsilon) \prod_{j=1}^{J} f_{\mathbf{s}_j}(\mathbf{s}_j)}_{\text{prior information}}. \end{aligned} \tag{7.9}$$

Note that the factorization of the synchronized estimation problems in (7.3) and (7.4) can be obtained by simply substituting the unknown clock offset ϵ with the true clock offset

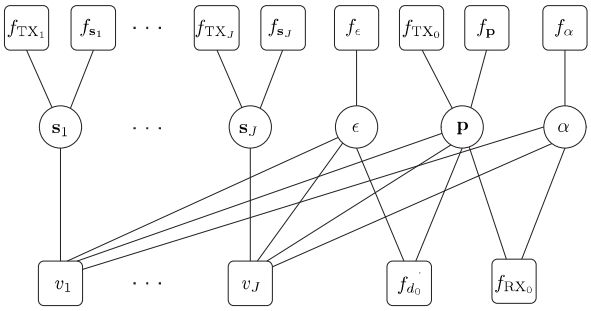

Figure 7.1: *Factor graph of the joint posterior distribution in* (7.9) - the factor graph contains numerous cycles. Thus, several belief propagation iterations are necessary to approximate the marginal posterior distributions of the variable nodes.

ϵ^* in all terms in (7.9). This is equivalent to imposing a Dirac-shaped prior distribution on ϵ. i.e. $f_\epsilon(\epsilon) = \delta(\epsilon - \epsilon^*)$. In order to obtain the estimation problem displayed in (7.4). the LOS factors have to be removed as well. This can be achieved by setting all LOS-related factors to one, i.e.
$f_{d_0}(\hat{d}_0|\mathbf{p}, \epsilon) f_{\mathrm{TX}_0}(\hat{\theta}_{\mathrm{TX},0}|\mathbf{p}) f_{\mathrm{RX}_0}(\hat{\theta}_{\mathrm{RX},0}|\mathbf{p}, \alpha) = 1$.

7.3.2 Factor Graph

From the factorization in (7.9), the factor graph can be drawn. It will help to decipher (7.9), since the complicated relationships between measurements and parameters are depicted in a graphical rather than a mathematical manner. Moreover. the graph will also help to improve the understanding of the belief propagation algorithm that will be proposed in the next section. The full graph is depicted in Fig. 7.1.

The factor graph in Fig. 7.1 is depicted in the style of a Tanner graph. The leaf nodes at the top of the graph can be regarded as the equivalent of the log-likelihood ratios from the channel output in coding applications. Variable nodes are the equivalent of bit nodes and the factor nodes at the bottom of the graph can be interpreted as the check nodes in a Tanner graph. Iterative message passing will be performed between the variable nodes and the factor nodes at the bottom of the graph.

Observe that the factors $f_{d_j}(\hat{d}_j|\mathbf{p}, \epsilon, \mathbf{s}_j)$ and $f_{\mathrm{RX}_j}(\hat{\theta}_{\mathrm{RX},j}|\mathbf{p}, \alpha, \mathbf{s}_j)$ of every NLOS path j are lumped together to the factor $v_j(\hat{d}_j, \hat{\theta}_{\mathrm{RX},j}|\mathbf{p}, \alpha, \epsilon, \mathbf{s}_j)$. This action has been taken in order to reduce the number of cycles of the graph. Note that even with this counter-measure. the factor graph in Fig. 7.1 still contains numerous short cycles (e.g., $\alpha \to f_{\mathrm{RX}_0} \to \mathbf{p} \to v_j \to \alpha$). While combining the factors $f_{d_j}(\hat{d}_j|\mathbf{p}, \epsilon, \mathbf{s}_j)$ and $f_{\mathrm{RX}_j}(\hat{\theta}_{\mathrm{RX},j}|\mathbf{p}, \alpha, \mathbf{s}_j)$ reduces the number of cycles. it also leads to more complicated message update rules, since each factor node v_j has four incoming edges, i.e. it depends on four variables. The implications of these factor mergers will be discussed in Section 7.3.3.2.

Observe that the factor graphs of the estimation problems in (7.3) and (7.4) are directly given by removing the clock offset variable node ϵ and all of its connected edges from the graph. To obtain the factor graph corresponding to the estimation problem in (7.4), the LOS-factor nodes f_{TX_0}, f_{RX_0} and f_{d_0} and their edges must be removed as well.

7.3.3 Algorithm Blueprint

The proposed message passing algorithm will be presented in this section. The message passing strategy will be discussed first. It consists of two phases; the initialization phase and the iterative message update phase. Finally, the particle-based implementation proposed in this work will be elaborated. The algorithm will be described by elaborating the essential steps based on exemplary messages. Once those steps are understood, it is straightforward to derive the computation rules of all other messages.

7.3.3.1 Methodology

Initialization:
As discussed in Chapter 4, belief propagation is initialized at the leaf nodes of the graph, i.e. nodes which have a single edge. In the factor graph in Fig. 7.1, these nodes have been arranged at the top. Tab. 7.1 summarizes the messages in the initialization phase. Note that these messages are only computed once and do not change in subsequent iterations. Fig. 7.2 depicts these messages.

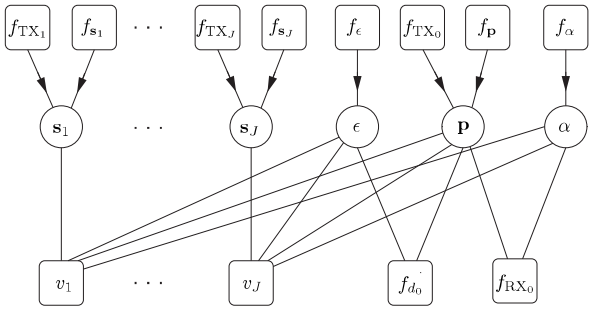

Figure 7.2: *Factor graph of the joint posterior distribution in* (7.9) - during the initialization phase, messages are sent from the leaf factor nodes at the top of the graph to the variable nodes in the middle.

Iterative Message Update:
In this phase of the algorithm, the downward messages from the variable nodes to the factor nodes at the bottom of the graph are computed (see Fig. 7.3). This requires multiplying the incoming messages from the factor nodes. The computation rules obey the standard belief propagation conventions explained in Chapter 4. For instance,

$$\mu^{(l)}_{\mathbf{p}\to v_j}(\mathbf{p}) = \mu^{(0)}_{f_{\mathrm{TX}_0}\to\mathbf{p}}(\mathbf{p})\mu^{(0)}_{f_{\mathbf{p}}\to\mathbf{p}}(\mathbf{p})\mu^{(l-1)}_{f_{\mathrm{RX}_0}\to\mathbf{p}}(\mathbf{p})\mu^{(l-1)}_{f_{d_0}\to\mathbf{p}}(\mathbf{p})\prod_{j'\neq j}\mu^{(l-1)}_{v_{j'}\to\mathbf{p}}(\mathbf{p}). \tag{7.10}$$

For $j = 1$, this message is highlighted in red in Fig. 7.3 The beliefs (approximated marginals) are determined as the product of all incoming messages to the variable nodes.

Table 7.1: Messages transmitted during the initialization phase.

Leaf Node	Transmitted Message
f_{TX_1}	$\mu^{(0)}_{f_{\mathrm{TX}_1}\to\mathbf{s}_1} = f_{\mathrm{TX}_1}(\hat{\theta}_{\mathrm{TX},1}\vert\mathbf{s}_1)$
$f_{\mathbf{s}_1}$	$\mu^{(0)}_{f_{\mathbf{s}_1}\to\mathbf{s}_1} = f_{\mathbf{s}_1}(\mathbf{s}_1)$
$\vdots$	$\vdots$
f_{TX_J}	$\mu^{(0)}_{f_{\mathrm{TX}_J}\to\mathbf{s}_J} = f_{\mathrm{TX}_J}(\hat{\theta}_{\mathrm{TX},J}\vert\mathbf{s}_J)$
$f_{\mathbf{s}_J}$	$\mu^{(0)}_{f_{\mathbf{s}_J}\to\mathbf{s}_J} = f_{\mathbf{s}_J}(\mathbf{s}_J)$
f_ϵ	$\mu^{(0)}_{f_\epsilon\to\epsilon} = f_\epsilon(\epsilon)$
f_{TX_0}	$\mu^{(0)}_{f_{\mathrm{TX}_0}\to\mathbf{p}} = f_{\mathrm{TX}_0}(\hat{\theta}_{\mathrm{TX},0}\vert\mathbf{p})$
$f_\mathbf{p}$	$\mu^{(0)}_{f_\mathbf{p}\to\mathbf{p}} = f_\mathbf{p}(\mathbf{p})$
f_α	$\mu^{(0)}_{f_\alpha\to\alpha} = f_\alpha(\alpha)$

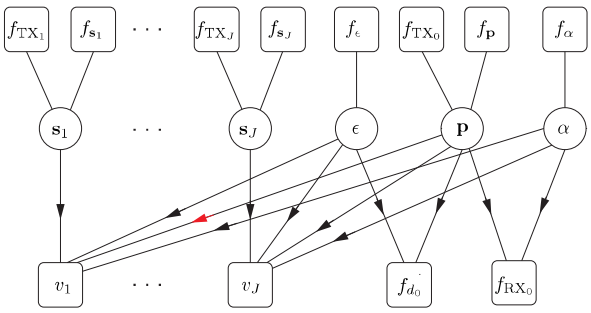

Figure 7.3: *Factor graph of the joint posterior distribution in* (7.9) - during the iterative update phase, downward messages are sent from variable nodes to factor nodes.

For instance,

$$
\begin{aligned}
b^{(1)}_\mathbf{p}(\mathbf{p}) &= \mu^{(0)}_{f_{\mathrm{TX}_0}\to\mathbf{p}}(\mathbf{p})\mu^{(0)}_{f_\mathbf{p}\to\mathbf{p}} \\
b^{(l)}_\mathbf{p}(\mathbf{p}) &= \mu^{(0)}_{f_{\mathrm{TX}_0}\to\mathbf{p}}(\mathbf{p})\mu^{(0)}_{f_\mathbf{p}\to\mathbf{p}}(\mathbf{p})\mu^{(l-1)}_{f_{\mathrm{RX}_0}\to\mathbf{p}}(\mathbf{p})\mu^{(l-1)}_{f_{d_0}\to\mathbf{p}}(\mathbf{p})\prod_j \mu^{(l-1)}_{v_j\to\mathbf{p}}(\mathbf{p}), l>1.
\end{aligned}
\tag{7.11}
$$

When all outgoing messages from the variable nodes are determined (downward messages), the outgoing messages from the factor nodes at the bottom of the graph to the variable nodes in the middle can be computed (upward messages). These messages are also obtained according to the standard belief propagation conventions explained in Chapter 4. Since two distinct types of factor nodes exist (v_j versus f_{d_0} and f_{RX_0}), messages of two different types have to be determined. Note of course that more than two different upward messages have to be computed. Yet, the structure of the computations of all other messages can be reduced to two distinct types. Message related to the node v_j are the

most sophisticated ones as they contain triple integrals. For instance,

$$\mu^{(l)}_{v_j\to\mathbf{p}}(\mathbf{p}) = \int\int\int v_j(\mathbf{p},\alpha,\epsilon,\mathbf{s}_j;\hat{d}_j,\hat{\theta}_{\mathrm{RX},j})\mu^{(l)}_{\epsilon\to v_j}(\epsilon)\mu^{(l)}_{\mathbf{s}_j\to v_j}(\mathbf{p})\mu^{(l)}_{\alpha\to v_j}(\alpha)\mathrm{d}\epsilon\mathrm{d}\mathbf{s}_j\mathrm{d}\alpha \quad (7.12)$$

All other messages related to the factor node v_j are determined similarly with appropriate incoming messages inside of the integral in (7.12). Compared to the message in (7.12), the messages related to f_{d_0} and f_{RX_0} are much simpler, since only a single integral needs to be solved. For instance,

$$\mu^{(l)}_{f_{d_0}\to\mathbf{p}}(\epsilon) = \int f_{d_0}(\epsilon,\mathbf{p})\mu^{(l)}_{\epsilon\to f_{d_0}}(\mathbf{p})\mathrm{d}\epsilon. \quad (7.13)$$

All upward messages are depicted in Fig. 7.4. The messages described in (7.12) and (7.13) are highlighted in magenta (for $j=1$) and red, respectively.

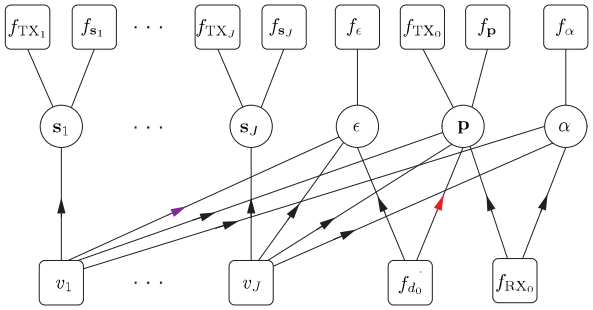

Figure 7.4: *Factor graph of the joint posterior distribution in* (7.9) - during the iterative update phase, upward messages are sent from the factor nodes at the bottom of the graph to the variable nodes.

An iteration is completed after the downward and upward messages as well as the beliefs are determined. This procedure is repeated N_{BP} times until the beliefs on the marginals required for solving the estimation problem in (7.2) are approximated sufficiently well. Note that in the first iteration ($l=1$), only the messages from the initialization phase are available to the variable nodes. Hence the downward messages and the beliefs are calculated only based on the available messages. For instance, the variable nodes ϵ and α receive the messages $\mu^{(0)}_{f_\alpha\to\alpha}(\alpha)$ and $\mu^{(0)}_{f_\epsilon\to\epsilon}(\epsilon)$, respectively. Hence the incoming message is simply forwarded to all other edges, e.g., $\mu^{(1)}_{\alpha\to v_j}(\alpha) = \mu^{(0)}_{f_\alpha\to\alpha}(\alpha) = f_\alpha(\alpha)$. The points of incidence $\mathbf{s}_j, \forall j$ and the position of the agent have two incoming messages. Thus, their outgoing messages are given as the product of these two messages, e.g., $\mu^{(1)}_{\mathbf{p}\to v_j} = \mu^{(0)}_{f_{\mathrm{TX}_0}\to\mathbf{p}_0}(\mathbf{p})\mu^{(0)}_{f_\mathbf{p}\to\mathbf{p}}(\mathbf{p}) = f_{\mathrm{TX}_0}(\hat{\theta}_{\mathrm{TX},0}|\mathbf{p})f_\mathbf{p}(\mathbf{p})$.

7.3.3.2 Particle-based Implementation

Considering the computations required to determine the messages discussed in the previous subsection, it becomes evident that neither the filtered nor the multiplied messages can be determined in closed forms. Hence a variant of sample-based belief propagation is considered in this section, where all messages are represented by particles. In the previous

chapters, filtered messages have been approximated by direct sampling and these messages have been multiplied using their kernel density estimates. Due to the complexity of the integrals such as in (7.12), direct sampling becomes relatively difficult. This gives rise to an alternative way of approximating these messages with the help of particles. Using Monte Carlo integration instead of direct sampling is one way of solving the problem mentioned above. This approach will be elaborated using (7.12) as an example. Once the approach is understood, it is straightforward to generalize it for other messages.

Observe that (7.12) contains three integrals. These integrals can be solved for an arbitrary position $\mathbf{p}_{\mu_{v_j\to\mathbf{p}}}^{(k,l)}$ by applying three concatenated Monte Carlo integrations. Suppose that the incoming messages $\mu_{\epsilon\to v_j}^{(l)}$, $\mu_{\mathbf{s}_j\to v_j}^{(l)}$, and $\mu_{\alpha\to v_j}^{(l)}$ are given as particle representations, i.e.

$$\mathcal{R}_{Ns}\left(\mu_{\epsilon\to v_j}^{(l)}(\epsilon)\right) = \left\{w_{\mu_{\epsilon\to v_j}}^{(k,l)}, \epsilon_{\mu_{\epsilon\to v_j}}^{(k,l)}\right\}_{k=1}^{N_s}, \tag{7.14a}$$

$$\mathcal{R}_{Ns}\left(\mu_{\mathbf{s}_j\to v_j}^{(l)}(\mathbf{s}_j\right) = \left\{w_{\mu_{\mathbf{s}_j\to v_j}}^{(k,l)}, \mathbf{s}_{j,\mu_{\mathbf{s}_j\to v_j}}^{(k,l)}\right\}_{k=1}^{N_s}, \tag{7.14b}$$

$$\mathcal{R}_{Ns}\left(\mu_{\alpha\to v_j}^{(l)}(\alpha)\right) = \left\{w_{\mu_{\alpha\to v_j}}^{(k,l)}, \alpha_{\mu_{\alpha\to v_j}}^{(k,l)}\right\}_{k=1}^{N_s}. \tag{7.14c}$$

Using (7.14), the integral in (7.12) can be approximated at the position $\mathbf{p}_{\mu_{v_j\to\mathbf{p}}}^{(k,l)}$ as follows:

$$\begin{aligned}\mu_{v_j\to\mathbf{p}}^{(l)}(\mathbf{p}_{\mu_{v_j\to\mathbf{p}}}^{(k,l)}) \approx \sum_{\tilde{k}} w_{\mu_{\mathbf{s}_j\to v_j}}^{(\tilde{k},l)} \sum_{\bar{k}} w_{\mu_{\alpha\to v_j}}^{(\bar{k},l)} \sum_{\check{k}} w_{\mu_{\epsilon\to v_j}}^{(\check{k},l)} \\ \times v_j(\mathbf{p}_{\mu_{\mathbf{p}\to v_j}}^{(k,l)}, \alpha_{\mu_{\alpha\to v_j}}^{(\bar{k},l)}, \epsilon_{\mu_{\epsilon\to v_j}}^{(\check{k},l)}, \mathbf{s}_{j,\mu_{\mathbf{s}_j\to v_j}}^{(\tilde{k},l)}; \hat{d}_j, \hat{\theta}_{\mathrm{RX},j}).\end{aligned} \tag{7.15}$$

While (7.15) approximates (7.12) arbitrarily closely if the number N_s of particles goes to infinity, the computational complexity of the operation described in (7.15) is $\mathcal{O}(N_s^4)$, rendering it intractable for any practical sample size. One way of circumventing this complexity trap is to resort to an important property of particle representations. Particularly, a particle representation of multiple *independent* messages can be obtained by stacking the individual particle representations [MHW+16].

Stacking of Particle Representations [MHW+16]

Given the particle representations $\mathcal{R}_{Ns}(\mu_1(\eta_1)) = \{w_1^{(k)}, \eta_1^{(k)}\}_{k=1}^{N_s}$ and $\mathcal{R}_{Ns}(\mu_2(\eta_2)) = \{w_2^{(k)}, \eta_2^{(k)}\}_{k=1}^{N_s}$ of the messages $\mu_1(\eta_1)$ and $\mu_2(\eta_2)$, a particle representation of the product $\mu_1(\eta_1)\mu_2(\eta_2)$ of the messages can be obtained by combining the particles and weights, i.e.

$$\mathcal{R}_{Ns}\left(\mu_1(\eta_1)\mu_2(\eta_2)\right) = \left\{w^{(k)}, \eta_1^{(k)}, \eta_2^{(k)}\right\}_{k=1}^{N_s}, \tag{7.16}$$

where $w^{(k)} = w_1^{(k)} w_2^{(k)} / \sum_k w_1^{(k)} w_2^{(k)}$. In other words, the particles $\eta_1^{(k)}$ and $\eta_2^{(k)}$ are simply stacked on top of each other and their weights $w_1^{(k)}$ and $w_2^{(k)}$ are combined and normalized.

Using the insights from (7.16), a particle representation $\mathcal{R}_{Ns}(\mu^{(l)}_{\epsilon\to v_j}(\epsilon)\mu^{(l)}_{\mathbf{s}_j\to v_j}(\mathbf{s}_j)\mu^{(l)}_{\alpha\to v_j}(\alpha))$ of the three incoming messages can be obtained as follows:

$$\mathcal{R}_{Ns}\left(\mu^{(l)}_{\epsilon\to v_j}(\epsilon)\mu^{(l)}_{\mathbf{s}_j\to v_j}(\mathbf{s}_j)\mu^{(l)}_{\alpha\to v_j}(\alpha)\right)=\left\{w^{(k,l)}_{\to v_j},\epsilon^{(k,l)}_{\mu_{\epsilon\to v_j}},\mathbf{s}^{(k,l)}_{j,\mu_{\mathbf{s}_j\to v_j}},\alpha^{(k,l)}_{\mu_{\alpha\to v_j}}\right\}_{k=1}^{N_s}, \tag{7.17}$$

where $w^{(k,l)}_{\to v_j}=w^{(k,l)}_{\mu_{\epsilon\to v_j}}w^{(k,l)}_{\mu_{\mathbf{s}_j\to v_j}}w^{(k,l)}_{\mu_{\alpha\to v_j}}/\sum w^{(k,l)}_{\mu_{\epsilon\to v_j}}w^{(k,l)}_{\mu_{\mathbf{s}_j\to v_j}}w^{(k,l)}_{\mu_{\alpha\to v_j}}$. Using (7.17), the integral in (7.12) can be approximated as follows

$$\mu^{(l)}_{v_j\to\mathbf{p}}(\mathbf{p}^{(k,l)}_{\mu_{v_j\to\mathbf{p}}})\approx\sum_{\tilde{k}}w^{(\tilde{k},l)}_{\to v_j}v_j(\mathbf{p}^{(k,l)}_{\mu_{\mathbf{p}\to v_j}},\alpha^{(\tilde{k},l)}_{\mu_{\alpha\to v_j}},\epsilon^{(\tilde{k},l)}_{\mu_{\epsilon\to v_j}},\mathbf{s}^{(\tilde{k},l)}_{j,\mu_{\mathbf{s}_j\to v_j}};\hat{d}_j,\hat{\theta}_{\mathrm{RX},j}). \tag{7.18}$$

Hence a particle representation of the message $\mu^{(l)}_{v_j\to\mathbf{p}}(\mathbf{p})$ can be obtained by approximating the integral in (7.12) by (7.18) and evaluating it in the positions $\mathbf{p}^{(k,l)}_{\mu_{v_j\to\mathbf{p}}}, k=1,...,N_s$, i.e.

$$\mathcal{R}_{Ns}\left(\mu^{(l)}_{v_j\to\mathbf{p}}(\mathbf{p})\right)=\left\{w^{(k,l)}_{\mu_{v_j\to\mathbf{p}}},\mathbf{p}^{(k,l)}_{\mu_{v_j\to\mathbf{p}}}\right\}_{k=1}^{N_s}, \tag{7.19}$$

where $w^{(k,l)}_{\mu_{v_j\to\mathbf{p}}}=\mu^{(l)}_{v_j\to\mathbf{p}}(\mathbf{p}^{(k,l)}_{\mu_{v_j\to\mathbf{p}}})/\sum'_k\mu^{(l)}_{v_j\to\mathbf{p}}(\mathbf{p}^{(k',l)}_{\mu_{v_j\to\mathbf{p}}})$ is the normalized weight according to (7.18). The samples $\mathbf{p}^{(k,l)}_{\mu_{v_j\to\mathbf{p}}}$ are directly drawn from the prior distribution, i.e. $\mathbf{p}^{(k,l)}_{\mu_{v_j\to\mathbf{p}}}\sim f_{\mathbf{p}}(\mathbf{p})$. Comparing (7.15) and (7.18), it can be observed that the complexity of message filtering has been reduced by two orders of magnitude, i.e. $\mathcal{O}(N_s^2)$. Thus, a tractable way of solving the integrals required during message filtering has been identified.

In order to facilitate easy multiplication at the variable nodes, the samples in which the messages are evaluated should be identical. For instance, the sample sets $\{\mathbf{p}^{(k,l)}_{\mu_{v_1\to\mathbf{p}}}\}_{k=1}^{N_s}$,...,$\{\mathbf{p}^{(k,l)}_{\mu_{v_J\to\mathbf{p}}}\}_{k=1}^{N_s}$, $\{\mathbf{p}^{(k,l)}_{\mu_{f_{d_0}\to\mathbf{p}}}\}_{k=1}^{N_s}$, and $\{\mathbf{p}^{(k,l)}_{\mu_{f_{\mathrm{RX}_0}\to\mathbf{p}}}\}_{k=1}^{N_s}$ should be the same. Then, the belief and the outgoing messages at the variable node $\mathbf{p}$ can be determined by directly multiplying the weights of the incoming messages with computational complexity $\mathcal{O}(N_s)$. For instance, a particle representation of the belief on the position $\mathbf{p}$ of the agent can be obtained as

$$\mathcal{R}_{Ns}\left(b^{(l)}_{\mathbf{p}}(\mathbf{p})\right)=\left\{w^{(k,l)}_{b_{\mathbf{p}}},\mathbf{p}^{(k,l)}_{b_{\mathbf{p}}}\right\}, \tag{7.20}$$

where $\mathbf{p}^{(k,l)}_{b_{\mathbf{p}}}=\mathbf{p}^{(k,l)}_{\mu_{v_1\to\mathbf{p}}}=\ldots=\mathbf{p}^{(k,l)}_{\mu_{v_J\to\mathbf{p}}}=\mathbf{p}^{(k,l)}_{\mu_{f_{d_0}\to\mathbf{p}}}=\mathbf{p}^{(k,l)}_{\mu_{f_{\mathrm{RX}_0}\to\mathbf{p}}}$. $w^{(k,l)}_{b_{\mathbf{p}}}=\tilde{w}^{(k,l)}_{b_{\mathbf{p}}}/\sum_k\tilde{w}^{(k,l)}_{b_{\mathbf{p}}}$, and $\tilde{w}^{(k,l)}_{b_{\mathbf{p}}}=w^{(k,l)}_{\mu_{f_{d_0}\to\mathbf{p}}}w^{(k,l)}_{\mu_{f_{\mathrm{RX}_0}\to\mathbf{p}}}\prod_j w^{(k,l)}_{\mu_{v_j\to\mathbf{p}}}$ for all samples k. Often, after multiplication of the weights, numerous samples $\mathbf{p}^{(k,l)}_{b_{\mathbf{p}}}$ have insignificant weights $w^{(k,l)}_{b_{\mathbf{p}}}$. Resampling as described in Section 3.2.1.2 is performed in order to eliminate these low-weight samples.

Finally, approximate MMSE estimates on the parameters of interest can be obtained by computing the centroid of the particle representations of the beliefs. For instance,

$$\hat{\mathbf{p}}^{(l)}_{\mathrm{MMSE}}\approx\sum_{k=1}^{N_s}w^{(k,l)}_{b_{\mathbf{p}}}\mathbf{p}^{(k,l)}_{b_{\mathbf{p}}}. \tag{7.21}$$

The entire procedure is summarized as pseudo-code in algorithm 5.

Remark: It is worth emphasizing that stacking of particle representations comes at a

Algorithm 5 COMPAS: Concurrent Mapping, Positioning, and Synchronization

given $\mathbf{z}_{\mathrm{NLOS}}$
for $l = 1$ to N_{BP} **do**
  if $l = 1$ **then** (*initialization*)
    receive messages from prior and AOD nodes
    - see Fig. 7.2 and Table 7.1 -
  else (*subsequent iterations*)
    receive messages on all edges
  end if
  compute the downward messages from $\mathbf{p}$, ϵ, α, $\mathbf{s}_1$,....,$\mathbf{s}_J$
  - see Fig. 7.3 and e.g., (7.10) -
  compute the upward messages from f_{d_0}, f_{RX_0}, v_1,...,v_J
  - see Fig. 7.4 and e.g., (7.18) -
  compute the beliefs of $\mathbf{p}$, ϵ, α, $\mathbf{s}_1$,....,$\mathbf{s}_J$
  - see e.g., (7.20) -
  determine estimates on $\mathbf{p}$, ϵ, α, $\mathbf{s}_1$,....,$\mathbf{s}_J$
  - see e.g., (7.21) -
end for

certain expense. Stacking is only asymptotically ($N_s \to \infty$) exact. In other words, for a finite number of particles, stacking of particle representations will incur an approximation error into the Monte Carlo integration. This error is relatively small if the uncertainty (spread of the particles to be stacked) is moderate. If the uncertainty is large, stacking can have a detrimental effect which results in poor approximation accuracy of the Monte Carlo integral. Hence stacking of particle representations needs to be used carefully and may require a large number of particles.

7.4 Numerical Results

Algorithm 5 will be analyzed numerically in this section. As observed in Section 6.5.1, the geometry influences the achievable accuracy. Hence different geometric constellations will be considered and the estimation accuracies in these scenarios will be studied. First, the simulation setup will be described. Then, the results will be presented and discussed subsequently.

7.4.1 Simulation Setup

For the numerical evaluation, the geometries described in Section (6.5.1.1) will be revisited. Unless stated differently, all position-related and angle-related quantities are considered in meters and degrees, respectively. In particular, the anchor is located at $\mathbf{q} = [0, 0]^{\mathrm{T}}$ and the agent is positioned at $\mathbf{p} = [60, 60]^{\mathrm{T}}$. The points of incidence are located at $\mathbf{s}_1 = [40, 0]^{\mathrm{T}}$, $\mathbf{s}_2 = [10, 50]^{\mathrm{T}}$, and $\mathbf{s}_3 = [70, 20]^{\mathrm{T}}$, respectively. The orientation and clock offset of the agent are assumed to be $\alpha = 0$ and $c\epsilon = 35$, respectively. As was shown in 6.5.2, reasonable estimation accuracy is only achieved if the standard deviations on the observed channel parameters are low. Thus, small standard deviations σ are considered. Particularly, the errors on the observed channel parameters (delays, AODs, and AOAs)

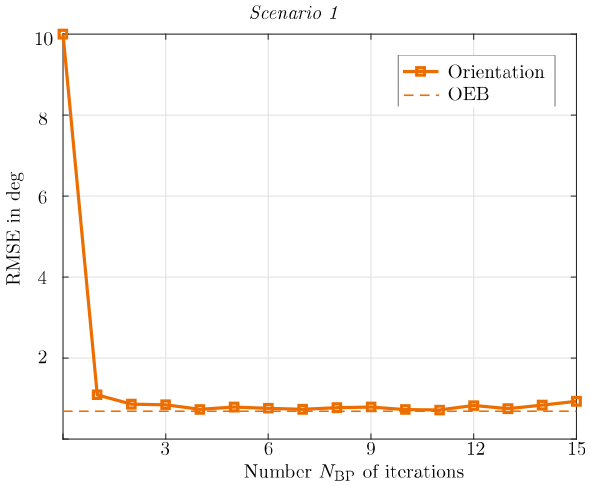

Figure 7.5: *Scenario 1: RMSE of the orientation of the agent* - the root-mean-square error (RMSE) almost attains the orientation error bound (OEB) after approximately 4 iterations.

are zero-mean Gaussian distributed with the following standard deviations: $c\sigma_{\mathrm{TOA}_j} = 0.2$ and $\sigma_{\mathrm{TX},j} = \sigma_{\mathrm{RX},j} = 0.5$. $\forall j$ [WGK+18]. Relatively broad prior information[2] is assumed on all agent states. The prior information on each agent state consists of a mean vector[3] $\mathbf{m}$ and a covariance matrix $\mathbf{\Sigma}$. In every realization, the mean vector is randomly chosen according to the true agent state plus a zero-mean Gaussian offset with covariance matrix $\mathbf{\Sigma}$. For instance, the mean vector $\mathbf{m}_{\mathbf{p}}$ of the agent's position is given by a random sample of $\mathbf{m}_{\mathbf{p}} = \mathcal{N}(\mathbf{m}_{\mathbf{p}}; \mathbf{p}, \mathbf{\Sigma}_{\mathbf{p}})$. Then, the prior information on the position of the agent is given by $f_{\mathbf{p}}(\mathbf{p}) = \mathcal{N}(\mathbf{p}; \mathbf{m}_{\mathbf{p}}, \mathbf{\Sigma}_{\mathbf{p}})$. Numerically, the following values were chosen: $\mathbf{\Sigma}_{\mathbf{p}} = \mathrm{diag}([9, 9])$, $\Sigma_\alpha = 100$ and $\Sigma_\epsilon = 100$. The units of the (co)variances are m^2, deg^2, and m^2, respectively. It should be stressed that no prior information on the points of incidence is assumed, i.e. $\mathbf{\Sigma}_{\mathbf{s}_j} = \mathrm{diag}([\infty, \infty]), \forall j$. For statistical significance, $N_{\mathrm{MC}} = 100$ realization were considered. The number of particles was set to $N_s = 2000$ unless stated differently.

Remark: It is worth mentioning that prior information increases the amount of available Fisher information and thus, the performance bounds are reduced. In particular, the FIM matrix $\mathbf{J}_{\tilde{\boldsymbol{\eta}}}$ of all parameters is increased by $\mathbf{J}_{\mathrm{prior}} = \mathrm{blkdiag}(\mathbf{\Sigma}_{\mathbf{p}}^{-1}, 1/\Sigma_\alpha, 1/\Sigma_\epsilon, \mathbf{\Sigma}_{\mathbf{s}_1}^{-1}, \ldots, \mathbf{\Sigma}_{\mathbf{s}_J}^{-1})$. The performance bounds are still calculated according (6.53) but with $\tilde{\mathbf{J}}_{\tilde{\boldsymbol{\eta}}} = \mathbf{J}_{\tilde{\boldsymbol{\eta}}} + \mathbf{J}_{\mathrm{prior}}$ instead of $\mathbf{J}_{\tilde{\boldsymbol{\eta}}}$.

[2]Prior information can be obtained from past time steps via sequential estimation or from additional hardware. In the former case, past estimates are usually transferred into prior knowledge with the help of tracking filters [JU04, Cos94, Wan06]. In the latter case, devices like Global Positioning System (GPS)-receivers and gyroscopes can provide coarse prior information on the position and orientation. Rough clock offset prior information is inherently available since communication systems require time synchronization in general.

[3]Note that, in case of the clock offset and orientation, the entries of the mean vector $\mathbf{m}$ and the covariance matrix $\mathbf{\Sigma}$ are only scalars.

7.4.2 Results and Discussion

The estimation accuracies of the position, orientation, and clock offset of the agent as well as the mapping accuracy will be studied in this subsection. Four scenarios will be considered. The LOS link is present in all scenarios. The scenarios 1-3 consider a single point of incidence $\mathbf{s}_j$ at a time. In the fourth scenario, all points of incidence are present at the same time. The results will be compared to the performance bounds (position error bound (PEB), OEB, bias error bound (BEB), and scatterer error bound (SEB)) that embody the ultimate limits regarding the estimation accuracies. Figs. 7.5-7.8 depict the results where a single point of incidence $\mathbf{s}_j$ is present. The RMSEs in the iteration $N_{\mathrm{BP}} = 0$ reflect the standard deviations of the prior distributions. Since no prior information on the points of incidence is assumed, the corresponding standard deviation is infinite. Thus, the corresponding (red) curves have no marker at $N_{\mathrm{BP}} = 0$.

The orientation estimation accuracy of scenario 1 is depicted in Fig. 7.5. Since the orientation of the agent is not considered as relevant for positioning as the agent's position and clock offset, numerical results regarding the orientation will be presented only for scenario 1. The following observations can be made from Fig. 7.5. The orientation is almost completely resolved after one iteration. Subsequent iterations do not increase the orientation accuracy notably. The orientation accuracy in the presence of the LOS link is very high (sub-1°) and the OEB is almost attained. It should be noted that the results for the scenarios 2 and 3 are almost identical to that of scenario 1 (not shown).

It can be inferred from the Figs. 7.6-7.8 that the RMSEs converge to the CRLBs after $N_{\mathrm{BP}} > 10$ iterations. Though, the largest reductions of the RMSEs are achieved during the first two iterations. After two iterations, the RMSEs reduce only gradually until the CRLBs are met. Hence it will often be sufficient to perform only a small number of iterations ($N_{\mathrm{BP}} < 5$) in order to achieve reasonable estimation accuracy. Considering that every additional iteration increases the computational complexity linearly and that the estimation accuracy does not improve significantly beyond $N_{\mathrm{BP}} > 2$, a few iterations may be sufficient in practice.

Fig. 7.9 depicts the fourth scenario, where the incidence points $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$ are present. Certain similarities to the previous scenarios can be observed. First, the RMSEs converge within a few iterations and the RMSEs of the clock offset and agent position are approximately identical. However, it can be seen that none of the performance bounds is met by Algorithm 5. There are two potential explanations why the proposed estimator cannot fully harness the available information and reach the CRLBs. Possible reasons are as follows: There exists a numerical problem due to the stacking of particles or there exists a systematic problem due to the cyclic[4] structure of the graph which prevents the beliefs from converging to the true marginals. The results depicted in Fig. 7.10 suggest that the reasons for not attaining the CRLBs is an insufficient number of particles since increasing the number of particles reduces the gap to the performance bounds.

Numerically, the estimation problem becomes more challenging as the number of incidence points increases. From that perspective, it is possible that the number of particles is insufficient to represent the uncertainties of all three incidence points at the same time. This can be troublesome especially during message multiplication. Note that the number of messages to be multiplied at all agent variable nodes (position, orientation and clock offset) increases with the points of incidence. Also, recall that a sample will have a significant weight after multiplication if and only if the weights of all incoming messages of that

[4]Every incidence point increase the number of cycles in the graph.

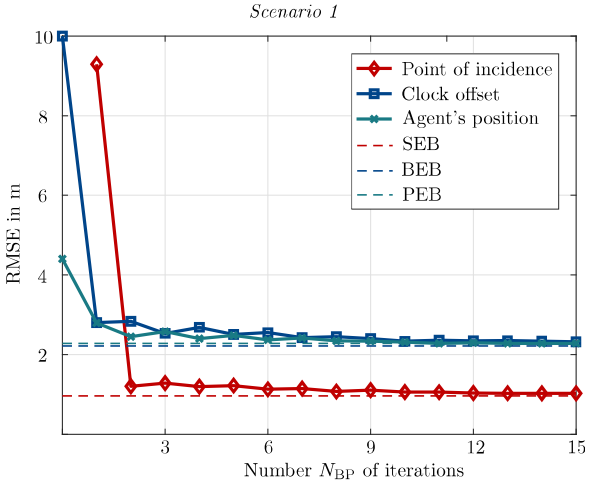

Figure 7.6: *Scenario 1: RMSEs of the position and clock offset of the agent as well as of the point of incidence* $\mathbf{s}_1$ - the RMSEs on all considered parameters almost attain the performance bounds after approximately ten iterations.

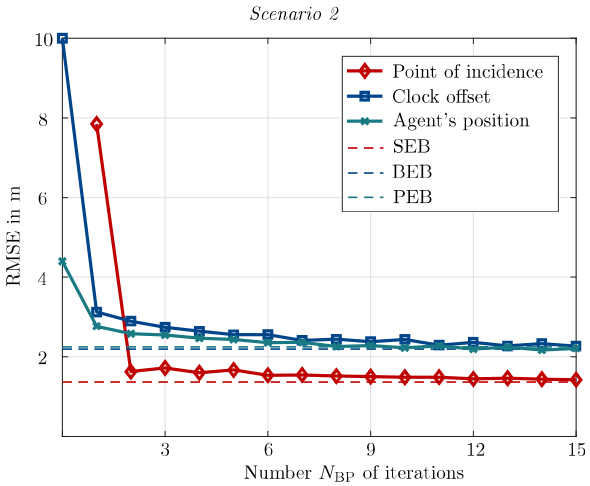

Figure 7.7: *Scenario 2: RMSE of the position and clock offset of the agent as well as of the point of incidence* $\mathbf{s}_2$ - the RMSEs on all considered parameters almost attain the performance bounds after approximately ten iterations.

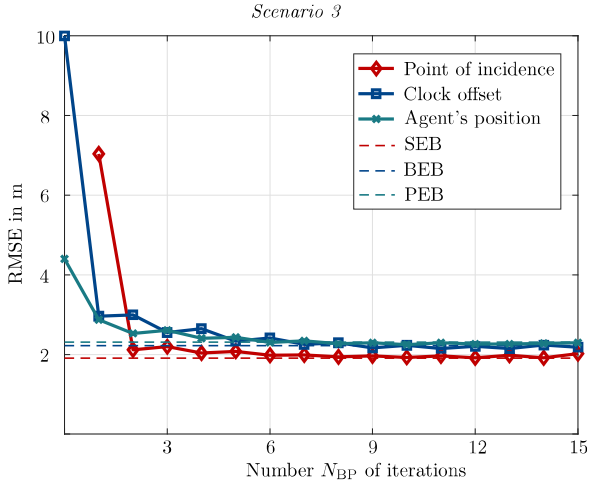

Figure 7.8: *Scenario 3: RMSEs of the position and clock offset of the agent as well as of the point of incidence* $\mathbf{s}_3$ - the RMSEs on all considered parameters almost attain the performance bounds after approximately ten iterations.

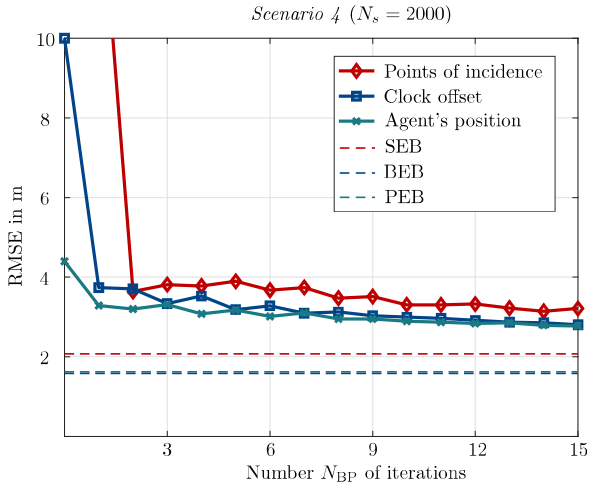

Figure 7.9: *Scenario 4: RMSEs of the position and clock offset of the agent as well as all the points of incidence* $\mathbf{s}_j, j = 1, 2, 3$ - premature convergence is observed, i.e. the RMSEs converge to values that are above the CRLBs.

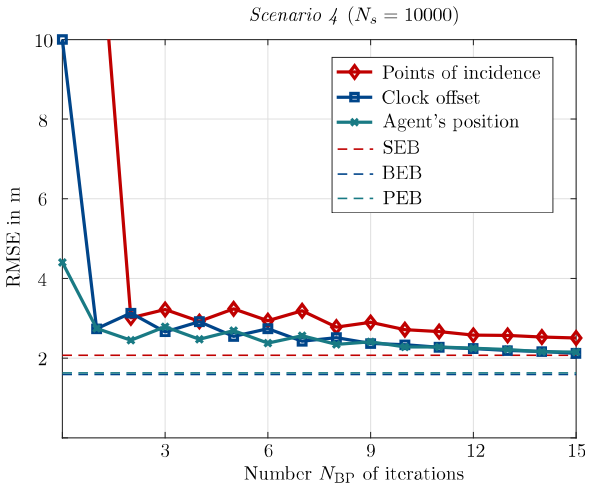

Figure 7.10: *Scenario 4: RMSEs of the position and clock offset of the agent as well as all the points of incidence* $\mathbf{s}_j, j = 1, 2, 3$ - compared to Fig. 7.9, the gap between the RMSEs and the performance bounds reduces if $N_s = 10000$ particles are used.

sample are considerable. The more incidence points exist, the lower is the probability that a particular sample will have a significant weight after multiplication. Loosely speaking, all points of incidence must agree that a certain sample is relevant. As long as one point of incidence has the opinion that a sample is irrelevant, a low weight will be assigned to the sample after multiplication. This behavior can lead to *sample impoverishment* which degrades the estimation accuracy.

Another possible explanation for the degradation of the estimation accuracy is that the corresponding factor graph has a poor structure. It contains many short cycles of length four. Every additional point of incidence adds multiple cycles of length four. This may explain the premature convergence in scenario 4, where the number of received paths is large.

In order to clarify the reason why the CRLBs are not reached, scenario 4 was re-evaluated with a large number of particles. In particular, $N_s = 10000$ was considered. Such a large number of particles is clearly prohibitive for real-time operation on most state-of-the-art devices. However, the results will help to understand the deficiencies of the proposed algorithm better. The results are depicted in Fig. 7.10. It can be inferred from that figure that the gap between the RMSEs and the CRLBs decreased compared to the results with $N_s = 2000$ particles. This observations indicates that the reason for premature convergence is a numerical issue. In order to achieve high estimation accuracies with a reasonable number N_s of particles, samples must be drawn in a smart manner. Similar to the approach presented in Chapter 5, the goal is to use samples efficiently. Especially from the practical perspective, extensions of the polygon outer-approximation (POA) algorithm to the single-anchor positioning problem are an interesting field of future research.

7.5 Conclusion and Outlook

This chapter presented a message passing-based estimation framework that jointly determines the position, orientation, and clock offset of an agent along with a map of the incidence points in its transmission environment. First, a variant of sample-based belief propagation is performed which approximately determines the marginal posterior distributions of the parameters of interest. Then, MMSE estimates are derived from the approximated marginal distributions. It was shown numerically that the proposed estimator attains the ultimate performance limits as it attains to the CRLBs. However, achieving the CRLBs requires some degree of prior information on the states of the agent. In addition, with the current sample-based implementation, the CRLBs are not reached if more than one NLOS path is present and the CRLBs are not reached.

The presented estimator embodies a unified and modular framework for different positioning problems. It can be adjusted to many different estimation problems by simple adding or deleting nodes in the presented factor graph. Various simpler estimation problems than the one studied in this chapter can be tackled. For instance, if the agent and anchor are synchronized, only the variable node of the clock offset and all its edges have to be erased from the presented factor graph. The message passing strategy, however, remains unchanged.

From a practical perspective, the proposed estimator provides a way of obtaining accurate position information in the absence of (sub-)nanosecond anchor-agent synchronization. The estimator performs well with the considered amount of prior information on the states (position, orientation, and clock offset) of the agent. None of the requirements on the prior information is extremely restrictive for a practical application, since the necessary priors are relatively broad. In the context of 5G mmWave mMIMO systems, it can be expected that agents (cellphones, vehicles, etc.) are equipped with a GPS receiver and a gyroscope which provide rough priors on the position and orientation. Even in GPS-challenged environments, position information with several meters of uncertainty can be obtained. Coarse-grained information on the clock offset is usually available, since symbol-level synchronization with the anchor is required for communications.

Nonetheless, the proposed estimator still carries a considerable burden in terms of its computational complexity. The complexity is quadratic in the number of particles. Since the uncertainties on the parameters are large, the number N_s of samples has to be large as well, i.e. $\mathcal{O}(N_s^2)$. Smart sampling techniques similar to those described in Section 5.4 would be required to reduces the algorithm's complexity. Alternatively, parametrized belief propagation or constrained optimization algorithms could be interesting ways to identify accurate estimators with lower complexity.

Chapter 8

Summary and Outlook

This chapter summarizes the contents presented in this thesis and discusses future research directions. During the summary, the main contributions of this work will be highlighted. Interesting research directions that are outside the scope of this thesis will be outlined to conclude this dissertation.

8.1 Summary

The theme of this thesis is statistical signal processing for radio-based positioning in the context of sparse infrastructure. Radio-based positioning refers to the challenges of determining the unknown positions of network devices (agents) based on the received radio signals from infrastructure nodes whose location is perfectly known (anchors). In conventional approaches, multiple anchors are required in order to determine unambiguous position estimates. If anchors are deployed in a sparse fashion, conventional position methods (e.g., trilateration or triangulation) fail and alternative concepts are required. In this thesis, the paradigms of cooperative positioning and single-anchor positioning have been studied. Cooperative approaches attempt to compensate the lack of anchors by cooperation among neighboring agents. Single-anchor positioning takes a different route in overcoming the lack of infrastructure. Particularly, the measurements (angles and delays) from all propagation paths of the received radio signal are extracted; not only those of the line-of-sight (LOS) path. Both paradigms theoretically increase the amount of available position information to an extend that enables unambiguous positioning even in the context of highly sparse infrastructure. However, it is not immediately obvious how precise position estimates can be achieved by practical position estimators. At this point, the methodology of statistical signal processing plays an important role. As it was shown throughout this thesis, the integral part of cooperative and single-anchor positioning is to account for the uncertainty of all parameters in the estimation problem. In the context of cooperative positioning, this uncertainty refers to the positions of neighboring agents. However, the positions of reflecting or scattering points of non-line-of-sight (NLOS) paths and other agent parameters (e.g., orientation and clock offset) are uncertain in single-anchor positioning. Statistical signal processing inherently addresses these uncertainties. Particularly, the belief propagation framework played an essential role throughout this thesis.

Before the research challenges were discussed, a tutorial-style review of radio-based positioning, statistical signal processing, and the belief propagation framework was presented. This review built the foundation for the studies presented in this thesis. In the context of

cooperative localization, the novel paradigm of constrained stochastic inference was proposed. The goal of this approach is to determine geometric constraints on the positions of agents. Particularly, polygons have been introduced to embody the constraints. A novel algorithm called *polygon outer-approximation* was presented that determines these polygons in a tractable and distributed way. It was shown that this algorithm has negligible computational costs, converges quickly, and provides tight constraints on the locations of agents. Based on these constraints, it was proposed to formulate the position estimation problem as a constrained inference problem. In particular, the belief propagation algorithm was combined with polygon outer-approximation. This formulation helped to relax the notoriously high computational burden of belief propagation. It was shown that the constrained inference problem can be solved faster and more accurately than the unconstrained problem. Nonetheless, the remaining computational burden is still relatively high and may be intractable for devices with low-cost hardware. This issue will be further discussed in the next section, where open research problems are presented.

The other big item that was studied in this thesis is the concept of single-anchor positioning. In contrast to cooperative positioning, single-anchor positioning is an emerging field of study and theoretical aspects are not yet fully understood. Hence the research that was presented in this thesis focused on analytic and algorithmic studies. First, the role of NLOS paths was reevaluated in the context of millimeter wave (mmWave) massive multiple-input multiple-output (mMIMO) systems. It was shown that each NLOS path increases the positioning accuracy of an agent. Moreover, it was proven that positioning is possible in the absence of the LOS path. From the presented analytical results, it was concluded that the positions of reflectors or scatterers in the radio environment can be estimated together with the position of the agent. In other words, important features in the radio environment can be inferred as a byproduct of positioning. It was also shown that the orientation of an agent and its clock offset to the anchor can be included in the estimation problem. These findings led to the conclusion that mmWave mMIMO systems bear an enormous positioning potential. Motivated by those findings, a belief propagation-based estimation framework was derived that can jointly estimate the position, orientation, and clock offset of an agent together with the positions of reflecting or scattering objects in the radio environment. It was pointed out that this estimation framework can tackle a number of similar estimation problems by simply adjusting its graphical representation. In that way, a modular and unified framework was introduced that can adjust to the presence or absence of the LOS path as well as known or unknown orientations and clock offsets.

In summary, it can be concluded that both cooperative and single-anchor positioning enable unambiguous positioning in the context of sparse infrastructure. Cooperative approaches overcome the lack of anchors by cooperation among agents. The concept of single-anchor positioning is that multiple measurements, which carry information regarding the position of an agent, are derived from the received signal of a single anchor. The results of both approaches are high-dimensional, nonlinear, non-convex estimation problems that are tough to solve. In order to successfully infer the parameters of interest, it is essential to account for the uncertainties of the parameters to be estimated. statistical signal processing – particularly, belief propagation – inherently considers these uncertainties. While the estimation accuracy of stochastic estimators is unprecedented, their computational burden is usually high. The main takeaway from this observation is that sparsity of infrastructure can be compensated via cooperative or single-anchor positioning, but the price that needs to be paid is high computational complexity.

8.2 Outlook

This thesis has illuminated certain aspects of radio-based positioning in the context of sparse infrastructure. Yet, there are still many interesting problems to be studied. A selection of future research directions will be provided in this section. While the field of cooperative positioning is already fairly mature, single-anchor positioning with mmWave mMIMO systems still has numerous unexplored spaces. Hence most of the research challenges that will be outlined are from the domain of single-anchor positioning.

Unarguably, cooperative positioning has received a lot of attention in the past decade, but systems that deploy cooperative approaches in practice are still rare. One of the main reasons is that cooperation among agents still requires relatively complicated algorithms to provide considerable gains in terms of accuracy. Even though the combination of geometric constraints and belief propagation presented in this thesis provides remarkable benefits, the overall algorithmic burden is still heavy. An important research direction includes studies of cooperative positioning methods whose computational requirements are by orders of magnitude lower than those of belief propagation. Certainly, it would be naive to assume that similar estimation accuracy can be expected with such methods. However, if these methods can provide reasonable positioning accuracy, the reduction of the computational burden can outweigh the degradation in terms of accuracy compared to high-accuracy approaches such as belief propagation. One way of achieving reasonable accuracy with low computational complexity could be to formulate the cooperative positioning problem as a constrained nonlinear least squares problem. Such problems can be solved very efficiently using interior point methods. Yet, the problem is that the unconstrained cooperative positioning problem is non-convex and has many local extrema. Constraints such as those delivered by the presented polygon outer-approximation algorithm can assist the interior point methods to converge to the global minimum or at least, to local minima that are close to the global minimum. Since polygon outer-approximation and interior point methods are computationally tractable, this combination could be very intriguing if the achievable accuracy turns out to be sufficiently high.

On the side of single-anchor positioning in mmWave mMIMO systems, there are numerous open questions. The most obvious research challenge from the context of this thesis is cooperation among agents. Cooperation is even more challenging, since the position of an agent is not the only uncertain parameter. Particularly, the uncertainty regarding the positions, orientations, and clock offsets of neighboring agents must be encoded properly. With such a high degree of uncertainty, cooperative positioning algorithms are likely to carry a significant computational burden, rendering them impractical for low-complexity devices. For more powerful devices, however, they may enhance the positioning accuracy notably. Another interesting aspect to study is the positioning and mapping capability of mmWave mMIMO systems in a dynamic setup. When mobility of agents is considered, the positions of reflecting or scattering objects can be estimated sequentially. This makes it possible to build a spatially consistent map of the radio environment, where large reflectors such as walls of buildings can be inferred. The step from a static to a dynamic scenario brings along another important challenge: *How are NLOS paths associated with previously observed reflectors or scatters?* This problem is usually referred to as data association and it can be performed in a deterministic or stochastic way. Hence data association must be studied in the context of single-anchor positioning and mapping in mmWave mMIMO systems. While there are certainly many other interesting research directions, one problem to study is the synergy of communications and positioning in mmWave mMIMO systems.

It was discussed earlier in this thesis that in such systems, the communication channel resembles the geometry of the environment closely. Identifying certain time-invariant features in the radio environment, such as reflectors, can probably improve certain aspects of communication systems. Knowledge regarding the transmission environment can simplify channel estimation and increase the reliable and/or throughput of communication systems. For instance, redundant information can be transmitted via NLOS paths if the locations of strong reflectors are known. Alternatively, the data rate could be increased by transmitting independent information via these NLOS side links. Finally, it is worthwhile to study an extension of the presented polygon outer-approximation algorithm in the context of mmWave mMIMO systems. Deriving polygonal constraints on the parameters to be estimated may benefit the proposed belief propagation framework. This list of future research problems is by no means complete and many other problems exist that deserve attention. Its sole purpose is to outline a subset of relevant problems.

Appendix A

Eigenvalues and Eigenvectors of the LOS and NLOS Information Matrices

A.1 Partial Derivatives of the Channel Parameters

The following list of equations summarizes all non-zero partial derivates of the channel parameters.

$$\frac{\partial \tau_0}{\partial \mathbf{p}} = \frac{1}{c}[\cos(\theta_{\mathrm{TX},0}), \sin(\theta_{\mathrm{TX},0})]^{\mathrm{T}}, \tag{A.1a}$$

$$\frac{\partial \theta_{\mathrm{TX},0}}{\partial \mathbf{p}} = \frac{1}{\|\mathbf{p}-\mathbf{q}\|_2}[-\sin(\theta_{\mathrm{TX},0}), \cos(\theta_{\mathrm{TX},0})]^{\mathrm{T}}, \tag{A.1b}$$

$$\frac{\partial \theta_{\mathrm{RX},0}}{\partial \mathbf{p}} = \frac{1}{\|\mathbf{p}-\mathbf{q}\|_2}[-\sin(\theta_{\mathrm{TX},0}), \cos(\theta_{\mathrm{TX},0})]^{\mathrm{T}}, \tag{A.1c}$$

$$\frac{\partial \tau_j}{\partial \mathbf{p}} = \frac{1}{c}[\cos(\theta_{\mathrm{RX},j}), -\sin(\theta_{\mathrm{RX},j})]^{\mathrm{T}}, j \neq 0, \tag{A.1d}$$

$$\frac{\partial \theta_{\mathrm{RX},j}}{\partial \mathbf{p}} = \frac{1}{\|\mathbf{p}-\mathbf{s}_j\|_2}[\sin(\theta_{\mathrm{RX},j}), \cos(\theta_{\mathrm{RX},j})]^{\mathrm{T}}, j \neq 0. \tag{A.1e}$$

$$\frac{\partial \tau_j}{\partial \mathbf{s}_j} = \frac{1}{c}[\cos(\theta_{\mathrm{TX},j}) + \cos(\theta_{\mathrm{RX},j}), \sin(\theta_{\mathrm{TX},j}) + \sin(\theta_{\mathrm{RX},j})]^{\mathrm{T}}, j \neq 0, \tag{A.1f}$$

$$\frac{\partial \theta_{\mathrm{TX},j}}{\partial \mathbf{s}_j} = \frac{1}{\|\mathbf{q}-\mathbf{s}_j\|_2}[-\sin(\theta_{\mathrm{TX},j}), \cos(\theta_{\mathrm{TX},j})]^{\mathrm{T}}, j \neq 0, \tag{A.1g}$$

$$\frac{\partial \theta_{\mathrm{RX},j}}{\partial \mathbf{s}_j} = -\frac{1}{\|\mathbf{p}-\mathbf{s}_j\|_2}[\sin(\theta_{\mathrm{RX},j}), \cos(\theta_{\mathrm{RX},j})]^{\mathrm{T}}, j \neq 0, \tag{A.1h}$$

$$\frac{\partial \theta_{\mathrm{RX},j}}{\partial \alpha} = -1, \tag{A.1i}$$

$$\frac{\partial \tau_j}{\partial \epsilon} = 1. \tag{A.1j}$$

A.2 Decomposition of the LOS Information Gain Matrix

Since the matrices in (6.34) are rank one, the eigenvalues are given by the traces of the respective matrices [Lau04]. Hence (6.35) follow. Having obtained the eigenvalues of the matrices $\tilde{\mathbf{A}}_R^{(\mathrm{G})}$,$\tilde{\mathbf{A}}_D^{(\mathrm{G})}$, and $\tilde{\mathbf{A}}_A^{(\mathrm{G})}$, it can be seen by straightforward algebra that the vectors $\mathbf{v}_{R,0}^{(\mathrm{G})}$, $\mathbf{v}_{D,0}^{(\mathrm{G})}$, and $\mathbf{v}_{A,0}^{(\mathrm{G})}$ in (6.36) are in the null space of $(\lambda_R^{(\mathrm{G})}\mathbf{I}_3 - \tilde{\mathbf{A}}_R^{(\mathrm{G})})$, $(\lambda_D^{(\mathrm{G})}\mathbf{I}_3 - \tilde{\mathbf{A}}_D^{(\mathrm{G})})$, and $(\lambda_A^{(\mathrm{G})}\mathbf{I}_3 - \tilde{\mathbf{A}}_A^{(\mathrm{G})})$. respectively, i.e.

$$\left(\lambda_R^{(\mathrm{G})}\mathbf{I}_3 - \tilde{\mathbf{A}}_R^{(\mathrm{G})}\right)\mathbf{v}_{R,0}^{(\mathrm{G})} = \mathbf{0}_{3\times 1}. \tag{A.2a}$$

$$\left(\lambda_D^{(\mathrm{G})}\mathbf{I}_3 - \tilde{\mathbf{A}}_D^{(\mathrm{G})}\right)\mathbf{v}_{D,0}^{(\mathrm{G})} = \mathbf{0}_{3\times 1}. \tag{A.2b}$$

$$\left(\lambda_A^{(\mathrm{G})}\mathbf{I}_3 - \tilde{\mathbf{A}}_A^{(\mathrm{G})}\right)\mathbf{v}_{A,0}^{(\mathrm{G})} = \mathbf{0}_{3\times 1}. \tag{A.2c}$$

A.3 Decomposition of the NLOS Information Gain Matrices

Terms $\tilde{\mathbf{B}}^{(\mathrm{G})}$, $\tilde{\mathbf{B}}^{(\mathrm{L})}$. and $\tilde{\mathbf{B}}^{(\mathrm{N})} = \tilde{\mathbf{B}}^{(\mathrm{G})} - \tilde{\mathbf{B}}^{(\mathrm{L})}$ which are associated with NLOS paths will be decomposed in this section.

A.3.1 NLOS Information Gain

The NLOS information gain is given by

$$\begin{aligned}\tilde{\mathbf{B}}^{(\mathrm{G})} &\triangleq \mathbf{B}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}} \\ &= \left[\mathbf{T}_{\mathbf{P}}^{(1)}, \mathbf{T}_{\mathbf{P}}^{(2)}, ..., \mathbf{T}_{\mathbf{P}}^{(J)}\right]\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\left[\mathbf{T}_{\mathbf{P}}^{(1)}, \mathbf{T}_{\mathbf{P}}^{(2)}, ..., \mathbf{T}_{\mathbf{P}}^{(K)}\right]^{\mathrm{T}}.\end{aligned} \tag{A.3}$$

Using matrix-matrix multiplication. the NLOS information gain matrix $\tilde{\mathbf{B}}^{(\mathrm{G})}$ can be decomposed as follows

$$\begin{aligned}\tilde{\mathbf{B}}^{(\mathrm{G})} &= \sum_{j=1}^{J}\mathbf{T}_{\mathbf{P}}^{(j)}\mathbf{J}_{\bar{\boldsymbol{\eta}}_j}\left(\mathbf{T}_{\mathbf{P}}^{(j)}\right)^{\mathrm{T}} \\ &= \sum_{j=1}^{J}\frac{1}{\sigma_{\tau_j}^2 c^2}\boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{RX},j}, 0, 0) + \frac{1}{\sigma_{\theta_{\mathrm{RX},j}}^2\left\|\mathbf{p}-\mathbf{s}_j\right\|^2}\boldsymbol{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2, \|\mathbf{p}-\mathbf{s}_j\|).\end{aligned} \tag{A.4}$$

A.3.2 NLOS Information Loss

It is easy to verify that $\mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}}$ is block diagonal since $\mathbf{D}$ and $\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}$ are block diagonal. i.e.

$$\mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}} = \begin{bmatrix}\mathbf{T}_{\mathbf{s}_1}\mathbf{J}_{\bar{\boldsymbol{\eta}}_1}\mathbf{T}_{\mathbf{s}_1}^{\mathrm{T}} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{T}_{\mathbf{s}_J}\mathbf{J}_{\bar{\boldsymbol{\eta}}_J}\mathbf{T}_{\mathbf{s}_J}^{\mathrm{T}}\end{bmatrix}. \tag{A.5}$$

For the compactness of notation, the following shorthand $\tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{s}_j} \triangleq \mathbf{T}_{\mathbf{s}_j}\mathbf{J}_{\bar{\boldsymbol{\eta}}_j}\mathbf{T}_{\mathbf{s}_j}^{\mathrm{T}}$ will be used. The inverse of a block diagonal matrix is determined by inverting the blocks on the diagonal. The j^{th} block is given by

$$\tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{s}_j} \triangleq \begin{bmatrix} a_j & b_j \\ b_j & d_j \end{bmatrix}, \tag{A.6}$$

where

$$a_j \triangleq \frac{(\cos(\theta_{\mathrm{TX},j}) + \cos(\theta_{\mathrm{RX},j}))^2}{\sigma_{\tau_j}^2 c^2} + \frac{\sin^2(\theta_{\mathrm{TX},j})}{\sigma_{\theta_{\mathrm{TX},j}}^2 \left\|\mathbf{q} - \mathbf{s_j}\right\|^2} + \tag{A.7}$$

$$\frac{\sin^2(\theta_{\mathrm{RX},j})}{\sigma_{\theta_{\mathrm{RX},j}}^2 \left\|\mathbf{p} - \mathbf{s_j}\right\|^2} \tag{A.8}$$

$$b_j \triangleq \frac{(\cos(\theta_{\mathrm{TX},j}) + \cos(\theta_{\mathrm{RX},j}))(\sin(\theta_{\mathrm{TX},j}) + \sin(\theta_{\mathrm{RX},j}))}{\sigma_{\tau_j}^2 c^2} \tag{A.9}$$

$$- \frac{\sin(\theta_{\mathrm{TX},j})\cos(\theta_{\mathrm{TX},j})}{\sigma_{\theta_{\mathrm{TX},j}}^2 \left\|\mathbf{q} - \mathbf{s_j}\right\|^2} - \frac{\sin(\theta_{\mathrm{RX},j})\cos(\theta_{\mathrm{RX},j})}{\sigma_{\theta_{\mathrm{RX},j}}^2 \left\|\mathbf{p} - \mathbf{s_j}\right\|^2} \tag{A.10}$$

$$d_j \triangleq \frac{(\sin(\theta_{\mathrm{TX},j}) + \sin(\theta_{\mathrm{RX},j}))^2}{\sigma_{\tau_j}^2 c^2} + \frac{\cos^2(\theta_{\mathrm{TX},j})}{\sigma_{\theta_{\mathrm{TX},j}}^2 \left\|\mathbf{q} - \mathbf{s_j}\right\|^2} + \tag{A.11}$$

$$\frac{\cos^2(\theta_{\mathrm{RX},j})}{\sigma_{\theta_{\mathrm{RX},j}}^2 \left\|\mathbf{p} - \mathbf{s_j}\right\|^2}. \tag{A.12}$$

Thus, the inverse of the j^{th} block is given by

$$\tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{s}_j}^{-1} = \frac{1}{\left|\tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{s}_j}\right|} \begin{bmatrix} d_j & -b_j \\ -b_j & a_j \end{bmatrix}, \tag{A.13}$$

where $\left|\tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{s}_j}\right| = a_j d_j - b_j^2$.

The left term $\mathbf{B}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{D}^{\mathrm{T}}$ and the right term $\mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}}$ in (6.32) can be evaluated by straightforward matrix-matrix multiplications. Observe that the left term can be obtained by taking the transpose of the right term. Due to the block diagonal structure of $\mathbf{D}$, the following matrix results

$$\mathbf{D}\mathbf{J}_{\bar{\boldsymbol{\eta}}_{\mathrm{NLOS}}}\mathbf{B}^{\mathrm{T}} = \begin{bmatrix} \mathbf{T}_{\mathbf{s}_1}\mathbf{J}_{\bar{\boldsymbol{\eta}}_1}\left(\mathbf{T}_{\mathbf{P}}^{(1)}\right)^{\mathrm{T}} \\ \vdots \\ \mathbf{T}_{\mathbf{s}_J}\mathbf{J}_{\bar{\boldsymbol{\eta}}_J}\left(\mathbf{T}_{\mathbf{P}}^{(J)}\right)^{\mathrm{T}} \end{bmatrix}. \tag{A.14}$$

The following shorthand will be used to abbreviate each block in (A.14): $\tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{P}} \triangleq \mathbf{T}_{\mathbf{s}_j}\mathbf{J}_{\bar{\boldsymbol{\eta}}_j}\left(\mathbf{T}_{\mathbf{P}}^{(j)}\right)^{\mathrm{T}}$. Considering (A.13) and (A.14), the NLOS information loss matrix $\tilde{\mathbf{B}}^{(\mathrm{L})}$ can be determined

by straightforward matrix-matrix multiplication and is given by

$$\begin{aligned}\tilde{\mathbf{B}}^{(\mathrm{L})} &= \sum_{j=1}^{J} \tilde{\mathbf{T}}^{\mathrm{T}}_{\mathbf{s}_j\mathbf{p}} \tilde{\mathbf{T}}^{-1}_{\mathbf{s}_j\mathbf{s}_j} \tilde{\mathbf{T}}_{\mathbf{s}_j\mathbf{p}} \\ &= \sum_{j=1}^{J} \underbrace{w^{(\mathrm{L})}_{R,j} \boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{TX}.0}, 0, 0)}_{\tilde{\mathbf{B}}^{(\mathrm{L})}_{R,j}} + \underbrace{w^{(\mathrm{L})}_{A,j} \boldsymbol{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2. \, \|\mathbf{p}-\mathbf{s}_j\|)}_{\tilde{\mathbf{B}}^{(\mathrm{L})}_{A,j}} - \gamma_{\mathbf{s}_j} \mathbf{B}^{(\mathrm{L})}_j,\end{aligned} \tag{A.15}$$

where the weights are defined as

$$w^{(\mathrm{L})}_{R,j} \triangleq \frac{1}{a_j d_j - b_j^2} \left(\frac{1+\cos(\Delta\theta_j)}{\sigma^2_{\tau_j} c^2} \right)^2 \left(\frac{1}{\sigma^2_{\theta_{\mathrm{RX},j}} \left\|\mathbf{p}-\mathbf{s_j}\right\|^2} + \frac{1}{\sigma^2_{\theta_{\mathrm{TX},j}} \left\|\mathbf{q}-\mathbf{s_j}\right\|^2} \right) \tag{A.16a}$$

$$w^{(\mathrm{L})}_{A,j} \triangleq \left(\frac{(1+\cos(\Delta\theta_j))^2}{\sigma^2_{\tau_j} c^2 \sigma^4_{\theta_{\mathrm{RX},j}} \left\|\mathbf{p}-\mathbf{s_j}\right\|^4} + \frac{\sin^2(\Delta\theta_j)}{\sigma^2_{\theta_{\mathrm{TX},j}} \left\|\mathbf{q}-\mathbf{s_j}\right\|^2 \sigma^4_{\theta_{\mathrm{RX},j}} \left\|\mathbf{p}-\mathbf{s_j}\right\|^4} \right) \frac{1}{a_j d_j - b_j^2} \tag{A.16b}$$

$$\gamma_{\mathbf{s}_j} \triangleq \frac{(1+\cos(\Delta\theta_j))\sin(\Delta\theta_j)}{\sigma^2_{\theta_{\mathrm{TX},j}} \left\|\mathbf{q}-\mathbf{s_j}\right\|^2 \sigma^2_{\theta_{\mathrm{RX},j}} \left\|\mathbf{p}-\mathbf{s_j}\right\|^2 \sigma^2_{\tau_j} c^2} \frac{1}{a_j d_j - b_j^2} \tag{A.16c}$$

and the matrix $\mathbf{B}^{(\mathrm{L})}_j$ is defined as

$$\mathbf{B}^{(\mathrm{L})}_j \triangleq \begin{bmatrix} -2\sin(\theta_{\mathrm{RX},j})\cos(\theta_{\mathrm{RX},j}) & \cos^2(\theta_{\mathrm{RX},j}) - \sin^2(\theta_{\mathrm{RX},j}) & \cos(\theta_{\mathrm{RX},j})\|\mathbf{p}-\mathbf{s}_j\| \\ \cos^2(\theta_{\mathrm{RX},j}) - \sin^2(\theta_{\mathrm{RX},j}) & 2\sin(\theta_{\mathrm{RX},j})\cos(\theta_{\mathrm{RX},j}) & \sin(\theta_{\mathrm{RX},j})\|\mathbf{p}-\mathbf{s}_j\| \\ \cos(\theta_{\mathrm{RX},j})\|\mathbf{p}-\mathbf{s}_j\| & \sin(\theta_{\mathrm{RX},j})\|\mathbf{p}-\mathbf{s}_j\| & 0 \end{bmatrix} \tag{A.17}$$

A.3.3 Net NLOS Gain

The net information gain matrix of the j^{th} NLOS component is collected in the matrix

$$\begin{aligned}\boldsymbol{\Psi}_{\mathbf{s}_j} &\triangleq \epsilon_{\mathbf{s}_j} \boldsymbol{\Upsilon}_{0,0}(\theta_{\mathrm{TX},0}, 0, 0) \\ &\quad + \beta_{\mathbf{s}_j} \boldsymbol{\Upsilon}_{1,1}(\theta_{\mathrm{RX},j}, \pi/2. \, \|\mathbf{p}-\mathbf{q}\|) + \gamma_{\mathbf{s}_j} \mathbf{B}^{(\mathrm{L})}_j \\ &\triangleq [\boldsymbol{\psi}_{\mathbf{s}_j,1}, \boldsymbol{\psi}_{\mathbf{s}_j,2}. \boldsymbol{\psi}_{\mathbf{s}_j,3}].\end{aligned} \tag{A.18}$$

It will be shown in the following that $\boldsymbol{\Psi}_{\mathbf{s}_j}$ has only one non-zero eigenvalue by showing that $\boldsymbol{\Psi}_{\mathbf{s}_j}$ is rank one, i.e. $\boldsymbol{\Psi}_{\mathbf{s}_j}$ has only one linearly independent column. By inspection of the columns of $\boldsymbol{\Psi}_{\mathbf{s}_j}$, it is straightforward to show that

$$a\boldsymbol{\psi}_{\mathbf{s}_j,1} - b\boldsymbol{\psi}_{\mathbf{s}_j,2} = \begin{bmatrix} \cos(\theta_{\mathrm{RX},j})\|\mathbf{p}-\mathbf{s}_j\|(\epsilon_{\mathbf{s}_j}\beta_{\mathbf{s}_j} - \gamma^2_{\mathbf{s}_j}) \\ \sin(\theta_{\mathrm{RX},j})\|\mathbf{p}-\mathbf{s}_j\|(\epsilon_{\mathbf{s}_j}\beta_{\mathbf{s}_j} - \gamma^2_{\mathbf{s}_j}) \\ 0 \end{bmatrix}, \tag{A.19}$$

where $a \triangleq [\mathbf{\Psi}_{\mathbf{s}_j}]_{2,3} = (\beta_{\mathbf{s}_j} \cos(\theta_{\mathrm{RX},j}) \|\mathbf{p} - \mathbf{s}_j\| + \gamma_{\mathbf{s}_j} \sin(\theta_{\mathrm{RX},j}) \|\mathbf{p} - \mathbf{s}_j\|)$ and $b \triangleq [\mathbf{\Psi}_{\mathbf{s}_j}]_{1,3} = (-\beta_{\mathbf{s}_j} \sin(\theta_{\mathrm{RX},j}) \|\mathbf{p} - \mathbf{s}_j\| + \gamma_{\mathbf{s}_j} \cos(\theta_{\mathrm{RX},j}) \|\mathbf{p} - \mathbf{s}_j\|)$. Using simple algebra it can be seen that

$$(\epsilon_{\mathbf{s}_j} \beta_{\mathbf{s}_j} - \gamma_{\mathbf{s}_j}^2) = 0. \tag{A.20}$$

Hence $\boldsymbol{\psi}_{\mathbf{s}_j,1}$ and $\boldsymbol{\psi}_{\mathbf{s}_j,2}$ are linearly dependent. All other combinations of the columns of $\mathbf{\Psi}_{\mathbf{s}_j}$ can be shown to linearly dependent in the same fashion. Hence $\mathbf{\Psi}_{\mathbf{s}_j}$ is rank one and has one non-zero eigenvalue. The only non-zero eigenvalue is given by the trace of $\mathbf{\Psi}_{\mathbf{s}_j}$ [Lau04] which can be shown to be

$$\lambda_{\mathbf{s}_j} = \epsilon_{\mathbf{s}_j} + \beta_{\mathbf{s}_j} (1 + \|\mathbf{p} - \mathbf{s}_j\|). \tag{A.21}$$

Note that the terms $\epsilon_{\mathbf{s}_j}$ and $\beta_{\mathbf{s}_j}$ in (A.21) are functions of $\theta_{\mathrm{RX},j}$. $\theta_{\mathrm{TX},j}$, $\|\mathbf{q} - \mathbf{s}_j\|$, $\|\mathbf{s}_j - \mathbf{p}\|$, $\sigma_{\mathrm{RX},j}$, $\sigma_{\mathrm{TX},j}$, and σ_{τ_j}, i.e. the eigenvalue in (A.21) is a function of the geometry of the scenario and the quality of the observations (AOA, AOD, and TOA). This results is presented in a more obvious format in (6.43).

Using the results for $\epsilon_{\mathbf{s}_j}$ and $\beta_{\mathbf{s}_j}$, as well as straightforward algebraic manipulations, the main result of this paper in (6.43) is readily obtained. That being said, it can be shown that

$$\mathbf{v}_{\mathbf{s}_j} = \begin{bmatrix} -\frac{\epsilon_{\mathbf{s}_j}}{\gamma_{\mathbf{s}_j}} \cos(\theta_{\mathrm{RX},j}) - \sin(\theta_{\mathrm{RX},j}) \\ -\frac{\epsilon_{\mathbf{s}_j}}{\gamma_{\mathbf{s}_j}} \sin(\theta_{\mathrm{RX},j}) + \cos(\theta_{\mathrm{RX},j}) \\ \|\mathbf{p} - \mathbf{s}_j\| \end{bmatrix}. \tag{A.22}$$

is in the null space of $\left(\epsilon_{\mathbf{s}_j} + \beta_{\mathbf{s}_j} (1 + \|\mathbf{p} - \mathbf{s}_j\|)\right) \mathbf{I}_3 - \mathbf{\Psi}_{\mathbf{s}_j}$, i.e.

$$\left(\left(\epsilon_{\mathbf{s}_j} + \beta_{\mathbf{s}_j} (1 + \|\mathbf{p} - \mathbf{s}_j\|)\right) \mathbf{I}_3 - \mathbf{\Psi}_{\mathbf{s}_j}\right) \mathbf{v}_{\mathbf{s}_j} = \mathbf{0}_{3\times 1}. \tag{A.23}$$

Hence $\mathbf{v}_{\mathbf{s}_j}$ is the eigenvector corresponding to the eigenvalue in (A.21).

Bibliography

[Abe93] Abel, J. S.: A bound on mean-square-estimate error. In: *IEEE Transactions on Information Theory* 39 (1993). Sep. Nr. 5, S. 1675–1680. http://dx.doi.org/10.1109/18.259655. – DOI 10.1109/18.259655. – ISSN 0018–9448

[AP04] Ash, J.N. ; Potter, L.C.: Sensor Network Localization via Received Signal Strength Measurements with Directional Antennas. In: *Proc. 2004 Allerton Conf. Communication, Control, and Computing* (2004), S. 1861–1870

[ASWASG18] Abu-Shaban, Z. ; Wymeersch, H. ; Abhayapala, T. ; Seco-Granados, G.: *Single-Anchor Two-Way Localization Bounds for 5G mmWave Systems: Two Protocols*. 2018

[ASZA+18a] Abu-Shaban, Z. ; Zhou, X. ; Abhayapala, T. ; Seco-Granados, G. ; Wymeersch, H.: Error Bounds for Uplink and Downlink 3D Localization in 5G Millimeter Wave Systems. In: *IEEE Transactions on Wireless Communications* 17 (2018), Aug, Nr. 8, S. 4939–4954

[ASZA+18b] Abu-Shaban, Z. ; Zhou, X. ; Abhayapala, T. ; Seco-Granados, G. ; Wymeersch. H.: Performance of location and orientation estimation in 5G mmWave systems: Uplink vs downlink. In: *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, 2018. – ISSN 1558–2612, S. 1–6

[BHE00] Bulusu, N. ; Heidemann, J. ; Estrin, D.: GPS-less low-cost outdoor localization for very small devices. In: *IEEE Personal Communications* 7 (2000), Oct. Nr. 5, S. 28–34. http://dx.doi.org/10.1109/98.878533. – DOI 10.1109/98.878533. – ISSN 1070–9916

[BHM98] Borras, J. ; Hatrack, P. ; Mandayam, N. B.: Decision theoretic framework for NLOS identification. In: *VTC '98. 48th IEEE Vehicular Technology Conference. Pathway to Global Wireless Revolution (Cat. No.98CH36151)* Bd. 2, 1998. – ISSN 1090–3038. S. 1583–1587 vol.2

[Bot07] Botev, Z.: *Nonparametric Density Estimation via Diffusion Mixing*. https://espace.library.uq.edu.au/view/UQ:120006, 2007. – Online Accessed : 2016-07-08

[BV04] Boyd, S. ; Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, 2004

[BWV18] Buehrer, R. M. ; Wymeersch, H. ; Vaghefi, R. M.: Collaborative Sensor Network Localization: Algorithms and Practical Issues. In: *Proceedings of the IEEE* 106 (2018), June, Nr. 6, S. 1089–1114. http://dx.doi.org/10.1109/JPROC.2018.2829439. – DOI 10.1109/JPROC.2018.2829439. – ISSN 0018–9219

[Cos94] Costa, P. J.: Adaptive model architecture and extended Kalman-Bucy filters. In: *IEEE Transactions on Aerospace and Electronic Systems* 30 (1994), April, Nr. 2, S. 525–533. http://dx.doi.org/10.1109/7.272275. – DOI 10.1109/7.272275. – ISSN 0018–9251

[CW18] Chen, C. ; Wu, W.: Joint AoD, AoA, and Channel Estimation for MIMO-OFDM Systems. In: *IEEE Transactions on Vehicular Technology* 67 (2018), July, Nr. 7, S. 5806–5820. http://dx.doi.org/10.1109/TVT.2018.2798360. – DOI 10.1109/TVT.2018.2798360. – ISSN 0018–9545

[CWW07] Chong, C. ; Watanabe, F. ; Win, M. Z.: Effect of Bandwidth on UWB Ranging Error. In: *2007 IEEE Wireless Communications and Networking Conference*, 2007. – ISSN 1525–3511, S. 1559–1564

[DCF+09] Dardari, D. ; Conti, A. ; Ferner, U. ; Giorgetti, A. ; Win, M. Z.: Ranging With Ultrawide Bandwidth Signals in Multipath Environments. In: *Proceedings of the IEEE* 97 (2009), Feb, Nr. 2, S. 404–426. http://dx.doi.org/10.1109/JPROC.2008.2008846. – DOI 10.1109/JPROC.2008.2008846. – ISSN 0018–9219

[DGA00] Doucet, Arnaud ; Godsill, Simon ; Andrieu, Christophe: On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. In: *Statistics and Computing* 10 (2000), Juli, Nr. 3, 197–208. http://dx.doi.org/10.1023/A:1008935410038. – DOI 10.1023/A:1008935410038. – ISSN 0960–3174

[DR07] Davis, P. ; Rabinowitz, P.: *Introduction to Numerical Analysis: Second Edition*. Dover Publications, 2007

[DS14] Deng, H. ; Sayeed, A.: Mm-wave MIMO channel modeling and user localization using sparse beamspace signatures. In: *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2014. – ISSN 1948–3244, S. 130–134

[DSW15] Dai, W. ; Shen, Y. ; Win, M. Z.: Distributed Power Allocation for Cooperative Wireless Network Localization. In: *IEEE Journal on Selected Areas in Communications* 33 (2015), Jan, Nr. 1, S. 28–40. http://dx.doi.org/10.1109/JSAC.2014.2369631. – DOI 10.1109/JSAC.2014.2369631. – ISSN 0733–8716

[DSW18] Dai, W. ; Shen, Y. ; Win, M. Z.: A Computational Geometry Framework for Efficient Network Localization. In: *IEEE Transactions on Information Theory* 64 (2018), Feb, Nr. 2, S. 1317–1339. http://dx.doi.org/10.1109/TIT.2017.2674679. – DOI 10.1109/TIT.2017.2674679. – ISSN 0018–9448

[DW12] Das, K. ; Wymeersch, H.: Censoring for Bayesian Cooperative Positioning in Dense Wireless Networks. In: *IEEE Journal on Selected Areas in Communications* 30 (2012), October, Nr. 9, S. 1835–1842. http://dx.doi.org/10.1109/JSAC.2012.121029. – DOI 10.1109/JSAC.2012.121029. – ISSN 0733–8716

[EBHS13] Etzlinger, B. ; Bartel, D. ; Haselmayr, W. ; Springer, A.: Mean field message passing for cooperative simultaneous ranging and synchronization. In: *2013 IEEE Global Conference on Signal and Information Processing*. 2013, S. 583–586

[EMH+17] Etzlinger, B. ; Meyer, F. ; Hlawatsch. F. ; Springer, A. ; Wymeersch, H.: Cooperative Simultaneous Localization and Synchronization in Mobile Agent Networks. In: *IEEE Transactions on Signal Processing* 65 (2017), July, Nr. 14, S. 3587–3602. http://dx.doi.org/10.1109/TSP.2017.2691665. – DOI 10.1109/TSP.2017.2691665. – ISSN 1053–587X

[FDMW06] Falsi, Chiara ; Dardari, Davide ; Mucchi, Lorenzo ; Win, Moe Z.: Time of Arrival Estimation for UWB Localizers in Realistic Environments. In: *EURASIP J. Appl. Signal Process.* 2006 (2006), Januar, 152–152. http://dx.doi.org/10.1155/ASP/2006/32082. – DOI 10.1155/ASP/2006/32082. – ISSN 1110–8657

[FNAB+10] Fakharzadeh, M. ; Nezhad-Ahmadi, M. R. ; Biglarbegian, B. ; Ahmadi-Shokouh, J. ; Safavi-Naeini, S.: CMOS Phased Array Transceiver Technology for 60 GHz Wireless Applications. In: *IEEE Transactions on Antennas and Propagation* 58 (2010), April, Nr. 4, S. 1093–1104. http://dx.doi.org/10.1109/TAP.2010.2041140. – DOI 10.1109/TAP.2010.2041140. – ISSN 0018–926X

[GC09] Guvenc, I. ; Chong, Chia-Chin: A Survey on TOA Based Wireless Localization and NLOS Mitigation Techniques. In: *Commun. Surveys Tuts.* 11 (2009), Juli, Nr. 3, 107–124. http://dx.doi.org/10.1109/SURV.2009.090308. – DOI 10.1109/SURV.2009.090308. – ISSN 1553–877X

[GCW07] Guvenc, I. ; Chong, C.-C. ; Watanabe, F.: NLOS Identification and Mitigation for UWB Localization Systems. In: *IEEE Wireless Communications and Networking Conference* (2007)

[GCWI07] Güvenç. Ismail ; Chong, Chia-Chin ; Watanabe, Fujio ; Inamura, Hiroshi: NLOS Identification and Weighted Least-Squares Localization for UWB Systems Using Multipath Channel Statistics. In: *EURASIP Journal on Advances in Signal Processing* 2008 (2007), Aug, Nr. 1, 271984. http://dx.doi.org/10.1155/2008/271984. – DOI 10.1155/2008/271984. – ISSN 1687–6180

[GGD18a] Guerra. A. ; Guidi, F. ; Dardari, D.: Single-Anchor Localization and Orientation Performance Limits Using Massive Arrays: MIMO vs.Beamforming. In: *IEEE Transactions on Wireless Communications* 17 (2018), Aug, Nr. 8, S. 5241–5255. http://dx.doi.org/10.1109/TWC.2018.2840136. – DOI 10.1109/TWC.2018.2840136. – ISSN 1536–1276

[GGD18b] Guerra, A. ; Guidi, F. ; Dardari, D.: Single-Anchor Localization and Orientation Performance Limits Using Massive Arrays: MIMO vs.Beamforming. In: *IEEE Transactions on Wireless Communications* 17 (2018), Aug, Nr. 8, S. 5241–5255. http://dx.doi.org/10.1109/TWC.2018.2840136. – DOI 10.1109/TWC.2018.2840136. – ISSN 1536–1276

[GGM11] Giorgetti, G. ; Gupta. S. K. S. ; Manes, G.: Understanding the Limits of RF-Based Collaborative Localization. In: *IEEE/ACM Transactions on Networking* 19 (2011), Dec, Nr. 6. S. 1638–1651. http://dx.doi.org/10.1109/TNET.2011.2141681. – DOI 10.1109/TNET.2011.2141681. – ISSN 1063–6692

[Gho11] Gholami. M.: Positioning Algorithms for Wireless Sensor Networks. In: *PhD Thesis, Chalmers University of Technology* (2011)

[GJW+16] Gentner. C. ; Jost, T. ; Wang, W. ; Zhang, S. ; Dammann. A. ; Fiebig, U.: Multipath Assisted Positioning with Simultaneous Localization and Mapping. In: *IEEE Transactions on Wireless Communications* 15 (2016). Sept, Nr. 9, S. 6104–6117. http://dx.doi.org/10.1109/TWC.2016.2578336. – DOI 10.1109/TWC.2016.2578336. – ISSN 1536–1276

[GP09] Gezici, S. ; Poor, H. V.: Position Estimation via Ultra-Wide-Band Signals. In: *Proccedings of the IEEE* 97 (2009), Feb, Nr. 2, S. 386–403. http://dx.doi.org/10.1109/JPROC.2008.2008840. – DOI 10.1109/JPROC.2008.2008840. – ISSN 0018–9219

[GWG13] Gholami. M. ; Wymeersch, H. ; Gezici. S.: Distributed Bounding of Feasible Sets in Cooperative Wireless Network Positioning. In: *IEEE Communication Letters* 17 (2013)

[GWL+17] Garcia, N. ; Wymeersch, H. ; Larsson, E. G. ; Haimovich, A. M. ; Coulon, M.: Direct Localization for Massive MIMO. In: *IEEE Transactions on Signal Processing* 65 (2017), May, Nr. 10. S. 2475–2487. http://dx.doi.org/10.1109/TSP.2017.2666779. – DOI 10.1109/TSP.2017.2666779. – ISSN 1053–587X

[GWSR11a] Gholami. M. R. ; Wymeersch, H. ; Ström, E. G. ; Rydström, M.: Robust distributed positioning algorithms for cooperative networks. In: *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications*, 2011. – ISSN 1948–3244, S. 156–160

[GWSR11b] Gholami. Mohammad R. ; Wymeersch, Henk ; Ström, Erik G. ; Rydström, Mats: Wireless network positioning as a convex feasibility problem. In: *EURASIP Journal on Wireless Communications and Networking* 2011 (2011). Nov, Nr. 1, 161. http://dx.doi.org/10.1186/1687-1499-2011-161. – DOI 10.1186/1687–1499–2011–161. – ISSN 1687–1499

[HSZ+16a] Han. Y. ; Shen. Y. ; Zhang, X. ; Win, M. Z. ; Meng, H.: Performance Limits and Geometric Properties of Array Localization. In: *IEEE Transactions on Information Theory* 62 (2016), Feb, Nr. 2, S.

1054–1075. http://dx.doi.org/10.1109/TIT.2015.2511778. – DOI 10.1109/TIT.2015.2511778. – ISSN 0018–9448

[HSZ+16b] Han, Y. ; Shen, Y. ; Zhang, X. P. ; Win, M. Z. ; Meng, H.: Performance Limits and Geometric Properties of Array Localization. In: *IEEE Transactions on Information Theory* 62 (2016), Feb, Nr. 2, S. 1054–1075. http://dx.doi.org/10.1109/TIT.2015.2511778. – DOI 10.1109/TIT.2015.2511778. – ISSN 0018–9448

[IFMW05] Ihler, A. ; Fisher, J. ; Moses, R. ; Willsky, A.: Nonparametric belief propagation for self-localization of sensor networks. In: *IEEE Journal on Selected Areas in Communications* 23 (2005)

[JDW08] Jourdan, D. B. ; Dardari, D. ; Win, M. Z.: Position error bound for UWB localization in dense cluttered environments. In: *IEEE Transactions on Aerospace and Electronic Systems* 44 (2008), April, Nr. 2, S. 613–628. http://dx.doi.org/10.1109/TAES.2008.4560210. – DOI 10.1109/TAES.2008.4560210. – ISSN 0018–9251

[JR06] Jourdan, D. B. ; Roy, N.: Optimal Sensor Placement for Agent Localization. In: *2006 IEEE/ION Position, Location, And Navigation Symposium*, 2006. – ISSN 2153–358X, S. 128–139

[JU04] Julier, S. J. ; Uhlmann, J. K.: Unscented filtering and nonlinear estimation. In: *Proceedings of the IEEE* 92 (2004), March, Nr. 3, S. 401–422. http://dx.doi.org/10.1109/JPROC.2003.823141. – DOI 10.1109/JPROC.2003.823141. – ISSN 0018–9219

[JW03] Jaakkola, M. Wainwright T. ; Willsky, A.: Tree-based Reparameterization Analysis of Sum-Product and its Generalizations. In: *IEEE Transactions on Information Theory* 49 (2003)

[Kay93] Kay: *Fundamentals of Statistical Signal Processing: Estimation Theory*. NJ, USA: Prentice-Hall, Inc.,, 1993

[KF09] Koller, D. ; Friedman. N.: *Probabilistic Graphical Models: Principles and Techniques (Adaptive Computation and Machine Learning)*. The MIT Press, 2009

[KHC+17] Koivisto. M. ; Hakkarainen, A. ; Costa, M. ; Kela, P. ; Leppanen, K. ; Valkama, M.: High-Efficiency Device Positioning and Location-Aware Communications in Dense 5G Networks. In: *IEEE Communications Magazine* 55 (2017), Aug, Nr. 8, S. 188–195. http://dx.doi.org/10.1109/MCOM.2017.1600655. – DOI 10.1109/MCOM.2017.1600655. – ISSN 0163–6804

[KTB11] Kroese, D. ; Taimre, T. ; Botev, Z.: *Handbook of Monte Carlo Methods*. Wiley, 2011

[Lau04] Laub, A. J.: *Matrix Analysis for Scientists and Engineers*. Society for Industrial and Applied Mathematics, 2004

[LCP05] Li, Ji ; Conan, J. ; Pierre, S.: Joint Estimation of Channel Parameters for MIMO Communication Systems. In: *2005 2nd International Symposium on Wireless Communication Systems*, 2005. – ISSN 2154–0217, S. 22–26

[LFS+12] Lien, J. ; Ferner, U. ; Srichavengsup, W. ; Wymeersch, H. ; Win, M.: A Comparison of Parametric and Sample-Based Message Representation in Cooperative Localization. In: *Hindawi International Journal of Navigation and Observation* (2012)

[LMR+15] Leitinger, E. ; Meissner, P. ; Rüdisser, C. ; Dumphart, G. ; Witrisal, K.: Evaluation of Position-Related Information in Multipath Components for Indoor Positioning. In: *IEEE Journal on Selected Areas in Communications* 33 (2015), Nov, Nr. 11, S. 2313–2328. http://dx.doi.org/10.1109/JSAC.2015.2430520. – DOI 10.1109/JSAC.2015.2430520. – ISSN 0733–8716

[LS02] Lee, Joon-Yong ; Scholtz, R. A.: Ranging in a Dense Multipath Environment Using an UWB Radio Link. In: *IEEE Journal on Selected Areas in Communications* 20 (2002), Dec, Nr. 9, S. 1677–1683. http://dx.doi.org/10.1109/JSAC.2002.805060. – DOI 10.1109/JSAC.2002.805060. – ISSN 0733–8716

[LSS09] Larsen, M. D. ; Swindlehurst, A. L. ; Svantesson, T.: Performance Bounds for MIMO-OFDM Channel Estimation. In: *IEEE Transactions on Signal Processing* 57 (2009), May, Nr. 5, S. 1901–1916. http://dx.doi.org/10.1109/TSP.2009.2014269. – DOI 10.1109/TSP.2009.2014269. – ISSN 1053–587X

[LWG+18] Luo, F. ; Wang, S. ; Gong, Y. ; Jing, X. ; Zhang, L.: Geographical Information Enhanced Cooperative Localization in Vehicular Ad-Hoc Networks. In: *IEEE Signal Processing Letters* 25 (2018), April, Nr. 4, S. 556–560. http://dx.doi.org/10.1109/LSP.2018.2806374. – DOI 10.1109/LSP.2018.2806374. – ISSN 1070–9908

[LWWK14] Li, B. ; Wu, N. ; Wang, H. ; Kuang, J.: Nodes localization with inaccurate anchors via EM algorithm in wireless sensor networks. In: *2014 IEEE International Conference on Communications Workshops (ICC)*, 2014. – ISSN 2164–7038, S. 121–126

[Mar63] Marquardt, D.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters. In: *SIAM Journal Journal on Numerical Analysis* 11 (1963)

[MB17] Mendrzik, R. ; Bauch, G.: Smart Sampling for Indoor Ultrawide-Band Localization using Nonparametric Belief Propagation. In: *11th ITG Conference on Signals, Systems, and Coding* (2017)

[MGWW10] Marano, S. ; Gifford, W. M. ; Wymeersch, H. ; Win, M. Z.: NLOS identification and mitigation for localization based on UWB experimental data. In: *IEEE Journal on Selected Areas in Communications* 28 (2010), September, Nr. 7, S. 1026–1035. http://dx.doi.org/10.1109/JSAC.2010.100907. – DOI 10.1109/JSAC.2010.100907. – ISSN 0733–8716

[MHW13] Meyer, F. ; Hlawatsch, F. ; Wymeersch, H.: Cooperative simultaneous localization and tracking (coslat) with reduced complexity and communication. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013. – ISSN 1520–6149, S. 4484–4488

[MHW+16] Meyer, Florian ; Hlinka, Ondrej ; Wymeersch, Henk ; Riegler, Erwin ; Hlawatsch, Franz: Distributed localization and tracking of mobile networks including noncooperative objects. In: *IEEE Trans. Signal Inf. Process. Netw.* 2 (2016), März, Nr. 1, S. 57–71

[Min01] Minka, T.: Expectation Propagation for Approximate Bayesian Inference. In: *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*. San Francisco, CA, USA : Morgan Kaufmann Publishers Inc., 2001 (UAI '01). – ISBN 1–55860–800–1, 362–369

[Min05] Minka, T.: Divergence Measures and Message Passing. In: *Technical Report MSR-TR-2005-173*. Cambridge, UK : Microsoft Research Ltd, 2005, S. 1–17

[MPG15] Malajner, M. ; Planinsic, P. ; Gleich, D.: UWB ranging accuracy. In: *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2015. – ISSN 2157–8672, S. 61–64

[MSW13] Mazuelas, S. ; Shen, Y. ; Win, M. Z.: Belief Condensation Filtering. In: *IEEE Transactions on Signal Processing* 61 (2013), Sept, Nr. 18, S. 4403–4415. http://dx.doi.org/10.1109/TSP.2013.2261991. – DOI 10.1109/TSP.2013.2261991. – ISSN 1053–587X

[MWB18] Mendrzik, R. ; Wymeersch, H. ; Bauch, G.: Joint Localization and Mapping through Millimeter Wave MIMO in 5G Systems. In: *2018 IEEE Globecom*, 2018, S. 1–6

[MWBAS18] Mendrzik, R. ; Wymeersch, H. ; Bauch. G. ; Abu-Shaban, Z.: Harnessing NLOS Components for Position and Orientation Estimation in 5G Millimeter Wave MIMO. In: *IEEE Transactions on Wireless Communications* (2018), S. 1–1. http://dx.doi.org/10.1109/TWC.2018.2877615. – DOI 10.1109/TWC.2018.2877615. – ISSN 1536–1276

[N. 03] N. Patwari and A. O. Hero and M. Perkins and N. S. Correal and R. J. O'Dea: Relative location estimation in wireless sensor networks. In: *IEEE Transactions on Signal Processing* 51 (2003). Aug, Nr. 8, S. 2137–2148. http://dx.doi.org/10.1109/TSP.2003.814469. – DOI 10.1109/TSP.2003.814469. – ISSN 1053–587X

[Nic04] Niculescu, D.: Positioning in ad hoc sensor networks. In: *IEEE Network* 18 (2004), July. Nr. 4. S. 24–29. http://dx.doi.org/10.1109/MNET.2004.1316758. – DOI 10.1109/MNET.2004.1316758. – ISSN 0890–8044

[NJSW15] Nguyen, T. V. ; Jeong, Y. ; Shin, H. ; Win, M. Z.: Least Square Cooperative Localization. In: *IEEE Transactions on Vehicular Technology* 64 (2015), April, Nr. 4, S. 1318–1330. http://dx.doi.org/10.1109/TVT.2015.2398874. – DOI 10.1109/TVT.2015.2398874. – ISSN 0018–9545

[NN01] Niculescu, D. ; Nath, B.: Ad hoc positioning system (APS). In: *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE* Bd. 5, 2001, S. 2926–2931 vol.5

[NN03a] Niculescu, D. ; Nath, B.: Ad hoc positioning system (APS) using AOA. In: *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Socictics (IEEE Cat. No.03CH37428)* Bd. 3, 2003. – ISSN 0743–166X, S. 1734–1743 vol.3

[NN03b] Niculescu, Dragoş ; Nath, Badri: DV Based Positioning in Ad Hoc Networks. In: *Telecommunication Systems* 22 (2003), Jan, Nr. 1, 267–280. http://dx.doi.org/10.1023/A:1023403323460. – DOI 10.1023/A:1023403323460. – ISSN 1572–9451

[PAK+05] Patwari, N. ; Ash, J. ; Kyperoutas, S. ; A. Hero III, R. M. ; Correal, S.: Locating the Nodes: Cooperative Localization in Wireless Sensor Networks. In: *IEEE Signal Processing Magazine* 22 (2005)

[PK11] Pi, Z. ; Khan, F.: An Introduction to Millimeter-wave Mobile Broadband Systems. In: *IEEE Communications Magazine* 49 (2011), June, Nr. 6, S. 101–107. http://dx.doi.org/10.1109/MCOM.2011.5783993. – DOI 10.1109/MCOM.2011.5783993. – ISSN 0163–6804

[RG18] Ruble, M. ; Guvenc, I.: Wireless Localization for mmWave Networks in Urban Environments. (2018). http://dx.doi.org/10.1186/s13634-018-0556-6. – DOI 10.1186/s13634–018–0556–6

[RMSS15] Rappaport, T. S. ; MacCartney, G. R. ; Samimi, M. K. ; Sun, S.: Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design. In: *IEEE Transactions on Communications* 63 (2015), Sept, Nr. 9, S. 3029–3056. http://dx.doi.org/10.1109/TCOMM.2015.2434384. – DOI 10.1109/TCOMM.2015.2434384. – ISSN 0090–6778

[SGD+18a] Shahmansoori, A. ; Garcia, G. E. ; Destino, G. ; Seco-Granados, G. ; Wymeersch, H.: Position and orientation estimation through millimeter-wave MIMO in 5G systems. In: *IEEE Transactions on Wireless Communications* 17 (2018), March, Nr. 3, S. 1822–1835. http://dx.doi.org/10.1109/TWC.2017.2785788. – DOI 10.1109/TWC.2017.2785788. – ISSN 1536–1276

[SGD+18b] Shahmansoori, A. ; Garcia, G.-E. ; Destino, G. ; Seco-Granados, G. ; Wymeersch, H.: Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems. In: *IEEE Transactions on Wireless Communications* 17 (2018), March, Nr. 3, S. 1822–1835. http://dx.doi.org/10.1109/TWC.2017.2785788. – DOI 10.1109/TWC.2017.2785788. – ISSN 1536–1276

[SH74] Sutherland, I. ; Hodgman, G.: Reentrant Polygon Clipping. In: *Communications of the ACM* 17 (1974)

[SIFW03] Sudderth, E. ; Ihler, A. ; Freeman, W. ; Willsky, A.: Nonparametric belief propagation. In: *IEEE Conference on Computer Vision and Pattern Recognition* (2003)

[SII+10] Sudderth, E. ; Ihler, A. ; Isard, M. ; Freeman, W. ; Willsky, A.: Nonparametric Belief Propagation. In: *Communications of the ACM* 53 (2010)

[Sil86] Silverman, B. W.: *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986

[Sor82] Sorensen, D. C.: Newton's Method with a Model Trust Region Modification. In: *SIAM Journal Journal on Numerical Analysis* 19 (1982)

[SRL02] Savarese, Chris ; Rabaey, Jan M. ; Langendoen, Koen: Robust Positioning Algorithms for Distributed Ad-Hoc Wireless Sensor Networks. In: *Proceedings of the General Track of the Annual Conference on USENIX Annual Technical Conference*. Berkeley, CA, USA : USENIX Association. 2002 (ATEC '02). – ISBN 1–880446–00–6, 317–327

[STK05] Sayed, A. ; Tarighat, A. ; Khajehnouri, N.: Network-based Wireless Location: Challenges Faced in Developing Techniques for Accurate Wireless Location Information. In: *IEEE Signal Processing Magazine* 22 (2005), Juli, S. 24–40

[SW10a] Shen, Y. ; Win, M. Z.: Fundamental Limits of Wideband Localization; Part I: A General Framework. In: *IEEE Transactions on Information Theory* 56 (2010), Oct, Nr. 10, S. 4956–4980. http://dx.doi.org/10.1109/TIT.2010.2060110. – DOI 10.1109/TIT.2010.2060110. – ISSN 0018–9448

[SW10b] Shen, Y. ; Win, M. Z.: On the accuracy of localization systems using wideband antenna arrays. In: *IEEE Transactions on Communications* 58 (2010), January, Nr. 1, S. 270–280. http://dx.doi.org/10.1109/TCOMM.2010.01.080141. – DOI 10.1109/TCOMM.2010.01.080141. – ISSN 0090–6778

[SWW10] Shen, Y. ; Wymeersch, H. ; Win, M. Z.: Fundamental Limits of Wideband Localization; Part II: Cooperative Networks. In: *IEEE Transactions on Information Theory* 56 (2010), Oct, Nr. 10, S. 4981–5000. http://dx.doi.org/10.1109/TIT.2010.2059720. – DOI 10.1109/TIT.2010.2059720. – ISSN 0018–9448

[SZ09] Savic, V. ; Zazo, S.: Nonparametric Boxed Belief Propagation for Localization in Wireless Sensor Networks. In: *2009 Third International Conference on Sensor Technologies and Applications*, 2009, S. 520–525

[SZ13] Savic, Vladimir ; Zazo, Santiago: Cooperative Localization in Mobile Networks Using Nonparametric Variants of Belief Propagation. In: *Ad Hoc Netw.* 11 (2013), Januar, Nr. 1, 138–150. http://dx.doi.org/10.1016/j.adhoc.2012.04.012. – DOI 10.1016/j.adhoc.2012.04.012. – ISSN 1570–8705

[TB07] Trees, H. V. ; Bell, K. L.: *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley, 2007

[TBT13] Trees, H. V. ; Bell, K. L. ; Tian, Z.: *Detection Estimation and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory, 2nd Edition*. Wiley. 2013

[TMN98] Tichavsky, P. ; Muravchik, C. H. ; Nehorai, A.: Posterior Cramer-Rao Bounds for Discrete-time Nonlinear Filtering. In: *IEEE Transactions on Signal Processing* 46 (1998), May, Nr. 5, S. 1386–1396. http://dx.doi.org/10.1109/78.668800. – DOI 10.1109/78.668800. – ISSN 1053–587X

[Tre02] Trees, H. V.: *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Wiley, 2002

[TVDW17] Talvitie, J. ; Valkama, M. ; Destino, G. ; Wymeersch, H.: Novel Algorithms for High-Accuracy Joint Position and Orientation Estimation in 5G mmWave Systems. In: *2017 IEEE Globecom Workshops (GC Wkshps)*, 2017, S. 1–7

[VB07] Venkatesh, S. ; Buehrer, R. M.: Non-line-of-sight identification in ultra-wideband systems based on received signal statistics. In: *IET Microwaves, Antennas Propagation* 1 (2007), Dec, Nr. 6, S. 1120–1130. http://dx.doi.org/10.1049/iet-map:20060273. – DOI 10.1049/iet–map:20060273. – ISSN 1751–8725

[VV10] Venkateswaran, V. ; Veen, A. J. d.: Analog Beamforming in MIMO Communications With Phase Shift Networks and Online Channel Estimation. In: *IEEE Transactions on Signal Processing* 58 (2010), Aug, Nr. 8, S. 4131–4143. http://dx.doi.org/10.1109/TSP.2010.2048321. – DOI 10.1109/TSP.2010.2048321. – ISSN 1053–587X

[VWS12] Velde, S. V. ; Wymeersch, H. ; Steendam, H.: Comparison of message passing algorithms for cooperative localization under NLOS conditions. In: *2012 9th Workshop on Positioning, Navigation and Communication*, 2012, S. 1–6

[Wan06] Wan, E.: Sigma-Point Filters: An Overview with Applications to Integrated Navigation and Vision Assisted Control. In: *2006 IEEE Nonlinear Statistical Signal Processing Workshop*, 2006, S. 201–202

[WCM+11] Win, M. Z. ; Conti, A. ; Mazuelas, S. ; Shen, Y. ; Gifford, W. M. ; Dardari, D. ; Chiani, M.: Network localization and navigation via cooperation. In: *IEEE Communications Magazine* 49 (2011), May, Nr. 5, S. 56–62. http://dx.doi.org/10.1109/MCOM.2011.5762798. – DOI 10.1109/MCOM.2011.5762798. – ISSN 0163–6804

[WF01] Weiss, Y. ; Freeman, W. T.: Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology. In: *Neural Computation* 13 (2001)

[WGK+18] Wymeersch, H. ; Garcia, N. ; Kim, H. ; Seco-Granados, G. ; Kim, S. ; Wen, F. ; Fröhle, M.: 5G mmWave Downlink Vehicular Positioning. In: *2018 IEEE Globecom*, 2018, S. 1–7

[WLW09] Wymeersch, H. ; Lien, J. ; Win, M.: Cooperative Localization in Wireless Networks. In: *Proceedings of the IEEE* 97 (2009)

[WMGW12] Wymeersch, H. ; Marano, S. ; Gifford, W. ; Win, M.: A Machine Learning Approach to Ranging Error Mitigation for UWB localization. In: *IEEE Transactions on Communications* 60 (2012)

[WML+16] Witrisal, Klaus ; Meissner, Paul ; Leitinger, Erik ; Shen, Yuan ; Gustafson, Carl ; Tufvesson, Fredrik ; Haneda, Katsuyuki ; Dardari, Davide ; Molisch, Andreas F. ; Conti, Andrea ; Win, Moe Z.: High-Accuracy Localization for Assisted Living. In: *IEEE Signal Processing Magazine* 33 (2016), März, Nr. 2, S. 59–70

[WSD18] Win, Moe Z. ; Shen, Yuan ; Dai. Wenhan: A Theoretical Foundation of Network Localization and Navigation. In: *Proceedings of the IEEE* 106 (2018), Juli, Nr. 7, S. 1136–1165. `http://dx.doi.org/10.1109/JPROC.2018.2844553`. – DOI 10.1109/JPROC.2018.2844553. – special issue on *Foundations and Trends in Localization Technologies*

[WSGD+17] Wymeersch, H. ; Seco-Granados, G. ; Destino, G. ; Dardari. D. ; Tufvesson, F.: 5G mmWave Positioning for Vehicular Networks. In: *IEEE Wireless Communications* 24 (2017), Dec, Nr. 6, S. 80–86. `http://dx.doi.org/10.1109/MWC.2017.1600374`. – DOI 10.1109/MWC.2017.1600374. – ISSN 1536–1284

[Wym07] Wymeersch, H.: *Iterative Receiver Design*. Cambridge University Press, 2007

[XGJ16] Xie, H. ; Gao, F. ; Jin, S.: An overview of low-rank channel estimation for massive MIMO systems. In: *IEEE J. Sel. Topics Signal Process.* 4 (2016), S. 7313–7321

[YFGZ13] Yin, F. ; Fritsche, C. ; Gustafsson, F. ; Zoubir, A. M.: TOA-Based Robust Wireless Geolocation and Cramer-Rao Lower Bound Analysis in Harsh LOS/NLOS Environments. In: *IEEE Transactions on Signal Processing* 61 (2013), May, Nr. 9, S. 2243–2255. `http://dx.doi.org/10.1109/TSP.2013.2251341`. – DOI 10.1109/TSP.2013.2251341. – ISSN 1053–587X

[YFW05] Yedidia, J. ; Freeman, W. ; Weiss, Y.: Constructing Free-Energy Approximations and Generalized Belief Propagation Algorithms. In: *IEEE Transactions on Information Theory* 51 (2005)

[YWE+16] Yuan, W. ; Wu, N. ; Etzlinger, B. ; Wang, H. ; Kuang, J.: Cooperative Joint Localization and Clock Synchronization Based on Gaussian Message Passing in Asynchronous Wireless Networks. In: *IEEE Transactions on Vehicular Technology* 65 (2016), Sept, Nr. 9, S. 7258–7273. `http://dx.doi.org/10.1109/TVT.2016.2518185`. – DOI 10.1109/TVT.2016.2518185. – ISSN 0018–9545

[YWI08] Yan, K. ; Wu, H. C. ; Iyengar, S. S.: Robustness Analysis of Source Localization Using Gaussianity Measure. In: *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, 2008. – ISSN 1930–529X, S. 1–5

[YZZ18] Yang, L. ; Zeng, Y. ; Zhang, R.: Channel Estimation for Millimeter-Wave MIMO Communications With Lens Antenna Arrays. In: *IEEE Transactions on Vehicular Technology* 67 (2018), April, Nr. 4, S. 3239–3251. `http://dx.doi.org/10.1109/TVT.2017.2779828`. – DOI 10.1109/TVT.2017.2779828. – ISSN 0018–9545

[ZB11] Zekavat, R. ; Buehrer, M.: *Handbook of Position Location: Theory, Practice and Advances*. Wiley-IEEE Press, 2011

[ZCW+17] Zhou, B. ; Chen, Q. ; Wymeersch, H. ; Xiao, P. ; Zhao, L.: Variational Inference-Based Positioning with Nondeterministic Measurement Accuracies and Reference Location Errors. In: *IEEE Transactions on Mobile Computing* 16 (2017), Oct, Nr. 10, S. 2955–2969. `http://dx.doi.org/10.1109/TMC.2016.2640294`. – DOI 10.1109/TMC.2016.2640294. – ISSN 1536–1233

[ZLK07] Zhen, B. ; Li, H. B. ; Kohno, R.: Clock management in Ultra-wideband Ranging. In: *2007 16th IST Mobile and Wireless Communications Summit*, 2007. – ISSN 2167–1753, S. 1–5

List of Figures