th·

TECHNISCHE UNIVERSITÄT ILMENAU

Parameter Estimation for Broadband mm-Wave FMCW MIMO Radars – A Model-Based System Identification Perspective

Dissertation

to obtain a doctoral degree in electrical engineering (Dr.–Ing.)

by

M. Sc. Stephan Häfner

submitted to

Faculty of Electrical Engineering and Information Technology

1. Reviewer: Univ.–Prof. Dr.–Ing. habil. Reiner Thomä
2. Reviewer: Prof. Dr.–Ing. Christian Waldschmidt
3. Reviewer: Prof. Dr.–Ing. Dr.–Ing. habil. Robert Weigel

Date of submission: 4/12/2020
Date of defence: 11/11/2021

Zusammenfassung

Radaranwendungen für das tägliche Leben haben in den letzten Jahrzehnten immer mehr an Bedeutung gewonnen. Radarsensoren werden in verschiedensten Bereichen eingesetzt, z.B. in Autos, in der industriellen Automatisierung, in medizinischen Anwendungen usw. FMCW-Radare sind hierbei populär. Ihre architektonische Einfachheit und die jüngsten Entwicklungen im Bereich integrierter Schaltkreise erleichtern die System-on-Chip-Realisierung von FMCW-Radaren. Darüber hinaus sind MIMO-Radare ein aufkommender Technologietrend zukünftiger Radarsysteme. Solche Radare bieten neben der Entfernungs- oder Dopplerbestimmung eine Vielzahl neuartiger Möglichkeiten. Kombiniert mit der Einfachheit der Hardware von FMCW-Radaren werden somit neue Anwendungsfelder adressierbar.

Neben der Radar-Hardware ist die Signalverarbeitung ein elementarer Bestandteil einer jeden Radaranwendung. Ein wesentlicher Teil der Radarsignalverarbeitung ist die Zielerfassung und die Schätzung der zugehörigen Parameter aus den Radarbeobachtungen. Die geschätzten Parameter dienen als Eingangsdaten für alle nachfolgenden Verarbeitungsschritte. Die Radarsignalverarbeitung ist hierbei untrennbar mit der Hardware und dem Anwendungszweck verbunden. Breitbandige Radare, d.h. Radare mit großer Signalbandbreite und räumlicher Apertur, bieten eine hohe räumliche und zeitliche Auflösung. Klassische Methoden der Schätzung von Zielparametern beruhen auf Annahmen bezüglich der Radararchitektur oder der Bandbreite. Diese Annahmen sind bei breitbandigen Radaren nicht erfüllt. Daher sind neuartige Signalverarbeitungsmethoden notwendig um einerseits die Möglichkeiten des Radarsystems voll auszuschöpfen, sowie Verarbeitungsproblemen, die durch das Radar selbst verursacht werden, zu bewältigen.

In dieser Arbeit wird die Signalverarbeitung für ein breitbandiges FMCW-MIMO-Radar dargestellt, welches im mm-Wellen-Frequenzbereich arbeitet. Die Entwicklung geeigneter Signalverarbeitungsmethoden erfordert hierbei einen Paradigmenwechsel in der Betrachtung der Radaranwendung. Im Gegensatz zu der klassischen Betrachtung von Radar, bei der Radar hauptsächlich vor dem Hintergrund der hardwarebedingten Zielerkennung betrachtet wird, soll hier ein anderer Gesichtspunkt herangezogen werden. Radar wird als eine Anwendung der Systemidentifikation im Bereich der drahtlosen Wellenausbreitung betrachtet. Die Parameter eines Modells beschreiben die Ausbreitung elektromagnetischer Wellen in Raum-, Frequenz- und Zeitbereich. Diese Parameter können Radarzielen zugeordnet, und deren Auswertung zur Zielortung ausgenutzt werden. Daher wird das Radar in Verbindung mit der Schätzung von Modellparametern als ein Problem der modellbasierten Systemidentifikation betrachtet. Die modellbasierte Systemidentifikation aus verrauschten Messungen ist ein stochastisch inverses Problem. Aus der Literatur ist eine Vielzahl von Verfahren zur Lösung solcher Probleme bekannt, welche hinsichtlich des Modells sehr flexibel sind. Diese Flexibilität ermöglicht die Entwicklung von dedizierten Methoden der Signalverarbeitung, welche auf die Radar-Hardware zugeschnitten, und nicht auf dedizierte Annahmen über z.B. die Signalbandbreite beschränkt sind. Die Betrachtung von Radarsignalverarbeitung als modellbasierte Systemidentifikation ermöglicht es somit, geeignete Signalverarbeitungsmethoden für das betrachtete Radar zu entwickeln.

Abstract

Radar has gained an increased popularity in daily life applications over the past decades. Radar sensors are deployed for multiple sensing purposes in e.g. cars, industrial automation, medical applications etc. Quite often, FMCW radars are utilised. Their architectural simplicity and recent developments in microwave-integrated circuits makes FMCW radars easily realisable as system-on-chip solution. Furthermore, MIMO radars are an emerging technology trend of future radar systems. Such radars offer a variety of novel applications apart from pure distance or Doppler determination. Combined with the hardware simplicity of FMCW radars, new application fields become addressable.

Apart from sophisticated radar hardware, signal processing is an elementary component of every radar application. An essential goal of radar signal processing is target detection and estimation of related parameters from the radar observations. The gained parameters are the input for all subsequent processing chains. The radar signal processing is inseparably linked to the radar hardware and the application purpose. Broadband radars, i.e. radars with a large bandwidth and huge spatial aperture, offer tremendous resolution capabilities, such that classical methods to estimate target parameters are not applicable. Classical methods are constrained by assumptions regarding e.g. the radar architecture or the bandwidth. However, these assumptions are not fulfilled by broadband radars. Hence, novel signal processing methods have to be developed, in order to exploit the full set of abilities provided by the radar system, as well as to tackle processing issues due to the radar itself.

In this thesis, signal processing for a broadband FMCW MIMO radar operating in the mm-Wave frequency range will be outlined. The development of appropriate signal processing methods necessitates a paradigm change in treating radar. Apart from the classical perspective to radar, where the radar application of e.g. target recognition is the driving point, a different viewpoint will be considered here. Radar will be considered as a system identification application in the area of wireless propagation. A parametric model describes the propagation of electromagnetic waves in spatial, frequency and time domain. The respective parameters can be assigned to radar targets, and their evaluation can be used for target localisation. Hence, radar in conjunction with the estimation of target parameters will be treated as a model-based system identification problem. Model-based system identification from noisy measurements is a stochastic inverse problem. A variety of methods, which are highly flexible regarding the model, are known from literature. This flexibility enables the development of dedicated methods for radar signal processing, which are in accordance with the radar hardware and not inherently limited to a certain radar architecture or assumptions on e.g. the signal bandwidth. Therefore, the treatment of radar signal processing as model-based system identification is more suited to develop appropriate signal processing methods for the considered broadband FMCW MIMO radar.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A

ACF	auto-correlation function
ADC	analog-to-digital converter
AoA	azimuth of arrival
AoD	azimuth of departure

B

BCRB	Bayesian Cramér-Rao bound
BIM	Bayesian information matrix
BLUE	best-linear-unbiased estimator

C

CCF	cross-correlation function
CDF	cummulative distribution function
CRLB	Cramér-Rao lower bound
CRM	Cramér-Rao matrix

D

DFT	discrete Fourier transform
DoA	direction of arrival
DoD	direction of departure

E

EADF	effective aperture distribution function
EMW	electromagnetic wave
EoA	elevation of arrival
EoD	elevation of departure
ESPRIT	estimation of signal parameters via rotational invariance technique

ETADF   effective time-aperture distribution function

F

FDR	the false discovery rate
FFT	fast Fourier transform
FIM	Fisher information matrix
FMCW	frequency modulated continuous wave
FWER	family-wise error rate

I

i.i.d.	independent and identical distributed
Ix	incident

K

KKT	Karush–Kuhn–Tucker

L

LFM	linear frequency modulated
LO	local oscillator
LOS	line of sight
LTI	linear time-invariant
LTV	linear time-variant

M

MAP	maximum a posteriori
MCMC	Markov chain Monte Carlo
MDL	minimum description length
MIMO	multiple input multiple output
MISO	multiple input single output
ML	maximum-likelihood

MMIC	monolithic microwave integrated circuit	**SINR**	signal-to-interference-plus-noise ratio
MMSE	minimum mean-square error	**SISO**	single input single output
MUSIC	multiple signal classification	**SME**	Slepian mode expansion
		SNR	signal-to-noise ratio
O		**SSH**	scalar spherical harmonics
		STAP	space-time adaptive processing
OTA	over-the-air	**SVD**	singular value decomposition
OU	Ornstein-Uhlenbeck	**Sx**	scattering
P		**T**	
PARAFAC	parallel factors	**TDM**	time-division multiplexing
PDF	probability density function	**ToA**	time of arrival
PLL	phase locked loop	**Tx**	transmitter
PSD	power spectral density	**U**	
R		**UCA**	uniform circular array
RCS	radar cross section	**ULA**	uniform linear array
RMSE	root-mean-square error	**URA**	uniform rectangular array
RO	ramp oscillator	**V**	
Rx	receiver		
		VCO	voltage controlled oscillator
S		**VSH**	vector spherical harmonics
SAGE	space-alternating generalized expectation-maximization	**W**	
SEM	singularity expansion method	**w.r.t.**	with respect to
SIMO	single input multiple output	**WSS**	wide-sense stationary

LIST OF SYMBOLS

Coordinates

x	x coordinate
y	y coordinate
z	z coordinate
φ	azimuth angle
ϑ	elevation angle
R	range
φ_{Rx}	azimuth of arrival
ϑ_{Rx}	elevation of arrival
φ_{Tx}	azimuth of departure
ϑ_{Tx}	elevation of departure
φ_{Ix}	azimuth of incidence
ϑ_{Ix}	elevation of incidence
R_{Ix}	range to incidence
φ_{Sx}	azimuth of scattering
ϑ_{Sx}	elevation of scattering
R_{Sx}	range to scattering

Unit Vectors

$\boldsymbol{\kappa}_x$	unit vector of x-coordinate
$\boldsymbol{\kappa}_y$	unit vector of y-coordinate
$\boldsymbol{\kappa}_z$	unit vector of z-coordinate
$\boldsymbol{\kappa}_\varphi$	unit vector of φ-coordinate
$\boldsymbol{\kappa}_\vartheta$	unit vector of ϑ-coordinate
$\boldsymbol{\kappa}_R$	unit vector of R-coordinate
$\boldsymbol{\kappa}_{\varphi,Rx}$	unit vector of φ_{Rx}-coordinate
$\boldsymbol{\kappa}_{\vartheta,Rx}$	unit vector of ϑ_{Rx}-coordinate
$\boldsymbol{\kappa}_{\varphi,Tx}$	unit vector of φ_{Tx}-coordinate
$\boldsymbol{\kappa}_{\vartheta,Tx}$	unit vector of ϑ_{Tx}-coordinate
$\boldsymbol{\kappa}_{\varphi,Ix}$	unit vector of φ_{Ix}-coordinate
$\boldsymbol{\kappa}_{\vartheta,Ix}$	unit vector of ϑ_{Ix}-coordinate
$\boldsymbol{\kappa}_{\varphi,Sx}$	unit vector of φ_{Sx}-coordinate
$\boldsymbol{\kappa}_{\vartheta,Sx}$	unit vector of ϑ_{Sx}-coordinate

Spatial Vectors

$\mathbf{d}$	spatial vector
$\mathbf{d}_{Rx}$	spatial vector from Rx
$\mathbf{d}_{Tx}$	spatial vector from Tx

Sets of Numbers

$\mathbb{N}$	set of natural numbers
$\mathbb{Z}$	set of integer numbers
$\mathbb{R}$	set of real numbers
$\mathbb{R}_+$	set of positive real numbers
$\mathbb{C}$	set of complex numbers

Sets

T	set of the truth
O	set of observations
D	scenario set
E	model error set
N	noise set
T'	set of predictions
T^*	set of hypothesis on truth T
$\mathsf{T}^{*'}$	set of predictions on hypothesis T^*

Models

$\mathcal{M}^P$	model of the physical system
$\mathcal{M}^N$	model of the noise
$\mathcal{M}^E$	model of model errors
$\mathcal{M}^S$	model of the system
$\mathcal{M}^D$	model of the sensing device

Parameters

ι	hyperparameter
θ	parameter of the data model
σ	parameter of the noise model

ϖ	parameter of the physical model	μ_0	permeability in vacuum
ϵ	parameter of the error model	ε_0	permittivity in vacuum
ζ	parameter of the device model	c_0	propagation velocity

Parameter Spaces

$\mathbf{1}_M$ unity vector of length M

$\mathbf{I}_M$ identity matrix of size $M \times M$

Θ	space of data model parameters
Σ	space of noise model parameters
Π	space of physical model parameters
E	space of error model parameters
Z	space of device model parameters
Ω	sampling space

Electromagnetic Theory

$\vec{\mathrm{E}}$	electric field vector
$\vec{\mathrm{H}}$	magnetic field vector
$\vec{\mathrm{S}}$	Poynting vector
κ	wavenumber
$\boldsymbol{\kappa}$	normalised wave vector
$\mathbf{k}_0$	wave vector
$\boldsymbol{\kappa}_{\mathrm{Px}}$	propagation direction
$\boldsymbol{\kappa}_{\mathrm{Tx}}$	normalised wave vector at Tx
$\boldsymbol{\kappa}_{\mathrm{Rx}}$	normalised wave vector at Rx
f	frequency
t	time
q	polarisation
λ	wavelength
λ_{c}	carrier wavelength
f_{c}	carrier frequency

Propagation Model Parameters

τ	time of arrival or delay
ν	Doppler shift
υ	velocity
$\gamma(t)$	path weight function
γ	path weight
ξ	radar cross section

Noise Model Parameters

η	noise power
ς	leakage noise power
ε	leakage noise coherence time
ϱ	leakage noise frequency shift
α	phase noise power
β	phase noise coherence time

Statistics

$\mathcal{H}_0$	null hypothesis
$\mathcal{H}_1$	alternative hypothesis
t_{val}	test statistic or T-value
t_{W}	test statistic of Wald test
t_{RS}	test statistic of Rao-Score test
t_{LR}	test statistic of likelihood-ratio test
t_{S}	test statistic of Student t-test
p_{val}	probability of the test results
c_{val}	critical value
α_{s}	significance level
E	expectation
ρ	root-mean-square error
r	remaining signal

FMCW Signal Parameters

W	bandwidth
f_0	start frequency of chirp
T_{W}	waiting or settling time
T_{R}	recovery time
T_{M}	modulation time
T_{P}	signal period time
f_{b}	beat frequency

ADC Parameters

T_{S}	sampling time
f_{S}	sampling frequency

Array Parameters

M_{Tx}	no. antennas or channels at Tx
M_{Rx}	no. antennas or channels at Rx

Constants

$\jmath$	imaginary unit
$\exp$	Euler's number
π	Pi

CHAPTER 1

INTRODUCTION

RADAR sensors are an emerging technology to ease live of the human society in the 21st century. For example, radar sensors are considered as a key enabling technology for autonomous mobility, advanced robotics in industry 4.0, automation in agriculture, medical imaging, assistant robots or smart cities (traffic monitoring and accident avoidance or pedestrian recognition), see Fig. 1.1. Further development trends are towards the combination of radar and communication system (RadCom [WS14]) or radar abilities piggyback to a communication system, e.g. for the upcoming 5G mobile radio network [Tho+19].

Utilisation of radar, i.e. the exploitation of electromagnetic waves for surveillance or sensing of objects, has a history of more than 100 years. First investigations date back to the beginning of the 20th century. In 1904 the German inventor Christian Hülsmeyer applied for a patent on a "Telemobiloskop", an apparatus for the detection of moving metallic bodies like ships by using electromagnetic waves; and practically demonstrated its abilities [Goo19]. However, the realised Telemobiloskop was capable to do detection only [Gua10]. In the mid 30s of the 20th century the Scottish physicist Robert Alexander Watson-Watt received a patent for the usage of the reflection of electromagnetic waves at objects like aircraft, in order to detect them and determine their range. At this time, the term "RADAR" (**Ra**dio **D**etection **a**nd **R**anging) was born. During the second World War, radar gained an increased popularity for air and sea surveillance, and reconnaissance. Up to this point, radar technologies has been pushed mainly due to military considerations and needs. In the recent years, civilian applications of radar as e.g. automotive radar [WHM21] come more and more into play. It is notable, that from the earliest beginning of radar by Hülsmeyer, to radar by Watson-Watt, the first bistatic passive radar "Klein Heidelberg" [GW10] to the recently idea of quantum radar [Bal18] it took only a century! This indicates, that radar technology gains a tremendous research attention. Accordingly, there is a huge variety of radar systems in terms of e.g. used wave forms, receiver concepts, or mono- and bistatic radars [CB93, Sko90, Wil05]. Further distinction has to be drawn between active and passive radar [GB17], i.e. is the transmitter operated by the radar system or is ambient radiation as e.g. from radio networks exploited for target illumination. Recently, the idea of MIMO radar gained much attention, as these radars offer superior performance compared to standard phased-array radars in terms of resolution, sensitivity and parameter identifiability [BF03, Fis+04a, Fis+04b, Fis+06, LS08]. In MIMO radars it has to be distinguished between distributed [HBC08] or co-located [LS07] transmit and receive arrays.

In order to properly detect targets in crowded environments as e.g. industrial halls, high
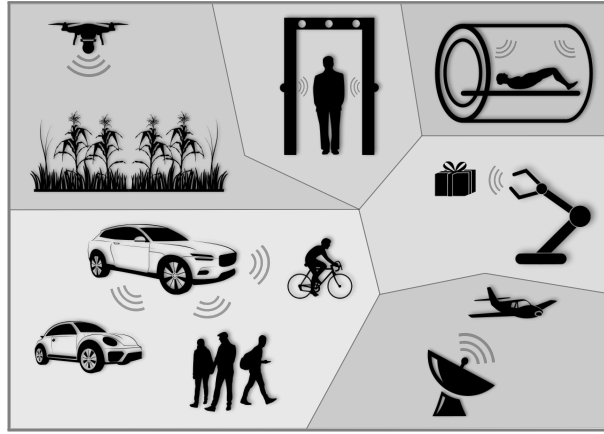
Figure 1.1: *Exemplary areas for radar applications. Pictures taken from [Sil20]*

bandwidths and large antenna apertures are required to ensure a good delay and angular resolution. Due to the dense occupation of low frequency bands and the respective large antenna size, the usage of higher frequency bands such as the millimetre wave spectrum (mm-Wave) is anticipated. Utilisation of these frequencies is furthermore pushed by the progress in semiconductor technology, which allows the low-cost realisation of complete radar frontends as monolithic microwave integrated circuit (MMIC) with on-chip antennas [2, Hit+18, Hit+17].

Radar applications are driven by their signal processing capabilities. A major aim of radar signal processing is the localisation of observed targets from the radar measurements. The target location is uniquely determined by *target parameters* like the delay of the echo signal (related to the target range) and the azimuth and elevation angle (observation angles of the radar with respect to (w.r.t.) to the target) [Wil05]. MIMO radars do observations in space, slow-time (time) and fast-time (delay) domain. These observations are gathered into the *radar data cube*. For localisation purpose, the target parameters have to be estimated from the radar data cube.

The title of the thesis implies two research fields of electrical engineering, whose interrelation will be outlined and exploited: *system identification* and *radar*. System identification, i.e. the process of gaining knowledge of a physical system of interest from measurements by an observer, arises in many scientific fields as e.g. control engineering, signal processing, biology or econometrics [SS89]. In this thesis, its application to signal processing for mm-Wave broadband MIMO FMCW radar will be considered. A broadband radar, i.e., a radar with a large array aperture and a huge signal bandwidth, is considered. The radar is the observer system to gain measurements from the system under consideration, which is the wireless propagation channel. A model of the physical system of interest will be exploited for signal processing, i.e. estimation of the target parameters. However, the consideration of model-based system identification for processing of radar observations is not limited to radars following the FMCW concept. In fact, it is a quite general signal processing concept, applicable to a variety of radar systems.

1.1 Motivation and Topic Under Consideration

1.1.1 Model-Based System Identification

System identification can be defined as the determination of the mathematical relation between the input and the output of a system using measurements [SS89]. In practise, the system is stimulated with a known signal (input) and the response of the system (output) to this stimulus is captured. The instrument, which stimulates the system and captures the response, is denoted as *observer*. From the known input and output conclusions on the *system under identification* can be drawn. Depending on the available information priorly to the identification step, one may distinguish between a *black box* and *grey box* system identification. In black box identification no prior information is considered, whereas grey box assumes a model structure regarding the system under identification priorly to the identification step. This mathematical model $\mathcal{M}(\boldsymbol{\theta})$ describes the input-output relation by modelling the underlying physics of the system under identification [SS89]. For a linear system the input-output relation is given by the *impulse response*, hence, $\mathcal{M}(\boldsymbol{\theta})$ is a model of that. The model is characterised by parameter vector $\boldsymbol{\theta}$, which has to be estimated from the observations (the known input and output). The parameter estimation step is therefore the system identification step. Hence, grey box modelling can be considered as *model-based system identification*. Because the model is of parametric nature, model-based system identification is also denoted as *parametric system identification*.

Why Model-Based System Identification?

Estimation of the parameters of the model of the system under identification is a problem of (statistical) signal processing, whereas the observations are the signal to be processed. Hence, model-based system identification belongs to the class of *model-based* or *parametric signal processing*. Model-based signal processing has superior performance over standard signal processing techniques [Can05]. Models improve the detection and estimation performance, as they enable the signal processor to account for the underlying physical phenomena of the problem under consideration, effects of the instrumentation and occurring noise sources. Two main advantages of model-based signal processing will be pointed out in this thesis:

- De-Embedding of instrumentation disturbances

- Enhanced resolution.

Of course, model-based signal processing encounters drawbacks as well. The biggest drawback is the applied model itself, because model-based processing encounters its superiority only for a properly chosen model. If the applied model poorly describes the system under identification, model errors are very severe and the quality of the estimated model parameters breaks down, e.g. estimates have a huge variance or become even biased. Hence, one can say that, if the model inputted to the signal processor is bad, the outputted parameter values from the signal processor are bad too ('garbage in, garbage out'). In order to circumvent improper models one can go for complicated models. However, complicated models can become a problem too under the view of processing effort. Roughly speaking, the parameter estimation effort scales with the model accuracy. Therefore, if processing speed matters, one goes for more simple models, which have a possibly higher inaccuracy. Summarised, a dualism between estimation
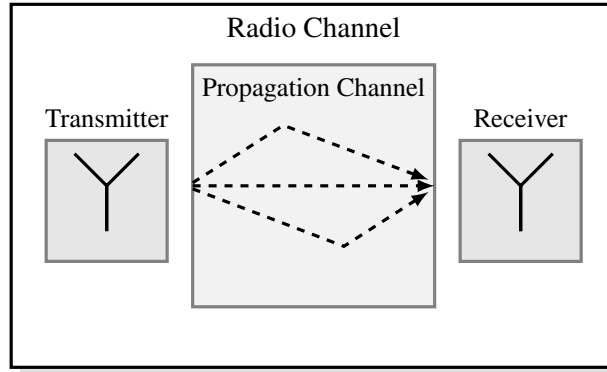
Figure 1.2: *Distinction of radio channel and propagation channel according to [SMB01]. The radio channel is the propagation channel embedded in the observer system. One wants to observe the response of the propagation channel (i.e. the system under identification) but actually the response of the radio channel has been observed.*

effort and model accuracy is present, a problem every practitioner encounters. A proper trade-off is necessary to tackle this dualism, however, a general trade-off is not given as it severely depends on the observer system, the application scenario or the operator's needs.

The De-Embedding and Mitigation Aspect

First it has to be clarified, what *should has been observed* and what actually *has been observed*. As an example for clarification, lets consider system identification for the *wireless propagation channel*, which will be termed as *channel sounding* [Mol12]. Determination of the input-output relation of the wireless propagation channel requires the utilisation of appropriate observer systems, so called *channel sounders* [TLS+04]. In general, channel sounders consist of a transmitter (Tx) and a receiver (Rx), each featuring transmitting/receiving chains and being equipped with antennas. A response of the propagation channel is measured by exciting the channel with an appropriate sounding signal at the Tx, and gathering the respective response at the Rx. Due to the characteristics of e.g. the utilised antennas or transceiver chains, the observations are influenced by the channel sounder. To separate the influence of the observer and the propagation channel itself, it is commonly distinguished between the *radio channel* and the *propagation channel* [SMB01], see Fig. 1.2. The radio channel can be assumed as the multidimensional embedding of the propagation channel (the system under identification) in the response of the observer system. As the operator is interested in the response of the propagation channel only, it is of interest to de-embed effects of the observer system from the observations. By developing a model of the gathered data, called *data model*, the multidimensional embedding of the propagation channel in the response of the observer system is described. To do so, a model of the propagation channel (i.e. the system under identification) and a model of the response of the observer system is required. Exploitation of the data model for the estimation of the model parameters of the system under identification inherently accomplishes the mitigation or de-embedding of the observer system.

Summarised, model-based signal processing allows for de-embedding and mitigation of instrumentation related effects on the observations. A proper model of the instrumentation system is required for this.

Enhanced Resolution – The High-Resolution Aspect

Sometimes, the term *resolution* is mixed up with the term *accuracy* regarding the assignment of capabilities to a radar system. Resolution will defined as the capability of separating events like e.g. radar targets or propagation paths in a certain observation dimension. Accuracy will be defined as how well a single event like a target or propagation path can be estimated.

A lower bound for resolution Δ is given by the inverse aperture size A of the observation system in the respective dimension [DEH19].

$$\Delta \propto \frac{1}{A} \tag{1.1}$$

This limit is denoted as *Rayleigh resolution* limit. For example, the delay resolution of an observer is determined by the bandwidth ($A = W$), the Doppler resolution by the observation time ($A = T$), and the angular resolution by the array aperture size an impinging wave is facing ($A = M$ for ULA in broadside direction). According to relation (1.1), one may want to enhance the aperture, if the resolution in the respective dimension is not appropriate for the operator. Enhancing the aperture requires hardware changes or increases the hardware effort, which can become quite fast very expensive or complicated. By taking into account high-resolution signal processing the resolution can be enhanced without additional hardware. Hence, "hardware effort" to gain a higher resolution can be swapped to "software effort", which is typically less costly.

In the "IEEE Standard Radar Definitions 686-2017" [Pan17] *high- or super-resolution* is defined as:

> An algorithmic technique for achieving higher resolution than system parameters would imply, by exploiting prior knowledge.

Hence, high-resolution is the improvement of the Rayleigh resolution limit, using model-based (prior knowledge-based) processing. High-resolution has been already considered for channel sounding [HTR03, Ric05] and radar [DEH19, HMS15] applications. From a modelling point of view, high-resolution deals with the recovery of a sequence of superimposing spikes or Dirac pulses (i.e. signals of infinite resolution) from aperture-limited observations [BE19, CF13, CF14, Fer16]. The aperture-limitation of the observer does a pulse-shaping, such that the resolution of the spikes inside of the measurements is decreased. Figure 1.3 shows a signal of four superimposed spikes, which are gathered by a aperture-limited observer. Due to his aperture limitation the observer acts like a low-pass filter. The filtering causes the observed signal to become a blurred version of the spike signal. As a result, the observations feature a much lower resolution. In Fig. 1.3 only three peaks can be recognised in the observed signal, because the presence of the spike at approx. 0 is hard to detect. High-resolution attempts to recover the four spikes from the observation, such that their original resolution is regained and even spikes close to each other become detectable. However, high-resolution also encounters its limits, which are determined by the signal-to-noise ratio (SNR) of the observations and the model accuracy.

Summarised, high-resolution is a promising approach to enhance the resolution limits of an observer, without increasing the hardware effort. In order to enable high-resolution in all sensing dimensions, the model of the system under identification must be represented as
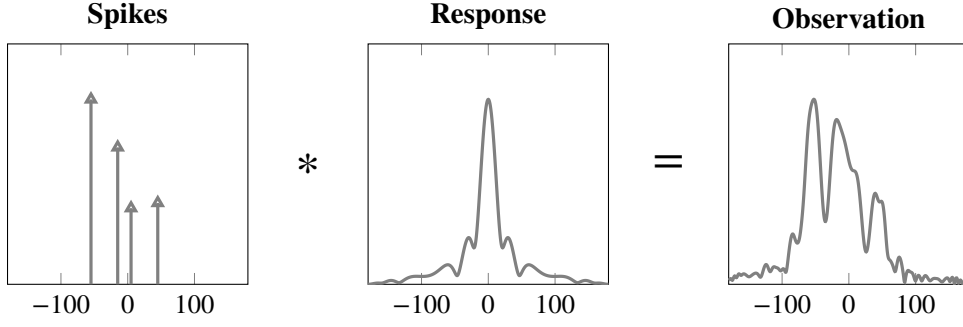
Figure 1.3: *Spike signal of four superimposing Dirac pulses, which are observed by a aperture-limited observer. The observed signal is a blurred version of the spike signal, having a reduced resolution. High-resolution attempts to recover the four spikes from the observation.*

a *sequence of superimposed (multidimensional) Dirac pulses*, which is known as *spreading function* [DEH19].

1.1.2 Radar Feats System Identification – The Close Relationship to Channel Sounding

A radar will be considered as a sensor, which radiates electromagnetic energy and receives the echo return from an scattering object, termed as *target* [Sko90]. From the received echo signal information on the target (location, velocity, signature) can be gained. According to this (quite rough) definition of a radar functionality it becomes already clear, that radar is a kind of a system identification application. In radar, the system under identification is termed as the *radar channel*, which is the wireless propagation channel from a top-level viewpoint.

Previously, channel sounding has been introduced as system identification application, where the wireless propagation channel is the system under identification. Channel sounding is an important field in wireless communication engineering [Mol12], as it is required to develop proper channel models, which are used to specify and test wireless communication systems. The propagation channel is sometimes denoted as wireless *communication channel* in this field of research.

In channel sounding as well as radar, a observer is employed to sense the system under identification. Besides this, model parameters are inferred in radar or channel sounding to describe the radar or communication channel, respectively. Hence, the idea of model-based system identification can be assigned to both applications. As the radar and communication channel undergo similar propagation phenomena, they are describable by the same mathematical model $\mathcal{M}(\boldsymbol{\theta})$ [Tre01]. Hence, from a pure system identification viewpoint, radar and channel sounding can be considered as similar applications, both facing the model-based identification of the wireless propagation channel. However, in radar, the estimated parameters are later on used for e.g. localisation or classification of targets, whereas in channel sounding the estimated parameters are used for e.g. statistical modelling of the propagation. Figure 1.4 summarises the information capturing and processing flow in radar and channel sounding, their commons in terms of model-based system identification, and their differences in terms of interpretation and post-processing of the identification results.
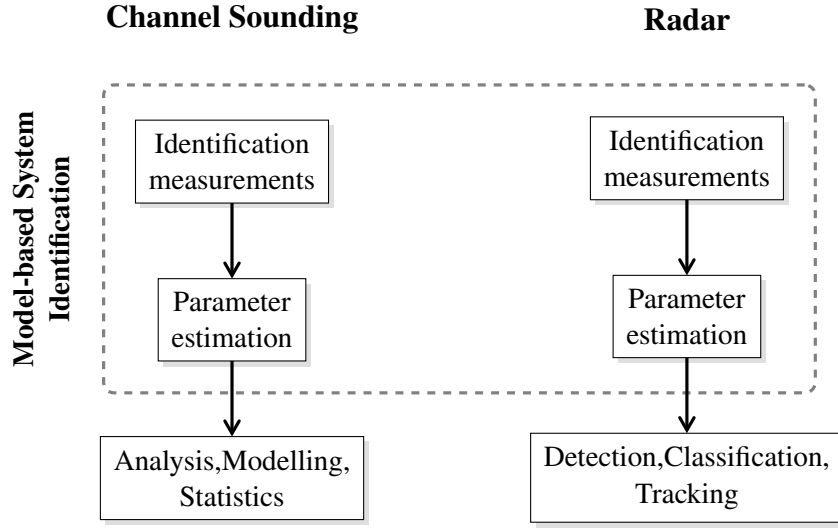
Figure 1.4: *Flow of information capturing and processing in radar and channel sounding application. Both applications have identification measurements and parameter estimation in common, i.e., feature model-based system identification.*

Furthermore, in radar applications, propagation due to reflections/scattering at targets is of interest only. Other echoes, which are contributions from the surrounding environment or which cannot be assigned to a target, are classified as disturbances or *clutter*. In channel sounding, it is not distinguished between valuable contributions and disturbances or clutter, because each energy transfer from Tx to Rx is considered as valuable. Roughly speaking, a specific kind of transmission (i.e. scattering at a target) is of interest in radar, whereas the transmission itself is of interest in channel sounding. Summarised, radar has a very specific interpretation of the propagation, whereas channel sounding is more general to this.

As a conclusion, radar applications can be considered as system identification applications, especially as a subset of channel sounding from the viewpoint of how the results of the system identification step are later on interpreted. By treating radar as system identification or model-based signal processing problem, the advantages of model-based system identification can be transferred to radar applications.

1.2 Major Contributions and Thesis Outline

1.2.1 Major Contributions

The major contributions of this thesis are in the field of modelling a broadband MIMO radar device, modelling of noise sources in FMCW radars and statistical signal processing of noisy radar observations.

- Noisy measurements require the application of statistical signal processing methods to estimate the intended parameters. The thesis gives an introduction for the non-familiar reader to the theory of inverse problems and statistical signal processing in system identification. The required information to solve the inverse problem and its solution by a Bayesian inversion framework will be explained. The mathematical background of

Bayesian inversion will be reviewed, and the maximum a posteriori (MAP) and minimum mean-square error (MMSE) parameter estimator will be presented. Own contributions in the area of statistical signal processing for radar and localisation purposes are [22, 27, 28].

- State-of-the-art signal processing methods for MIMO radars assume narrowband conditions, especially for array signal processing. Broadband radars require new approaches, because the narrowband assumption is not ensured. Therefore, a wideband array model will be developed for direction estimation, and compared to a classical narrowband array model. Own contributions in the area of modelling and calibration of antenna arrays for direction of arrival (DoA) estimation are [23, 7, 24, 26]. Own publications in the are of array signal processing under wideband conditions are [7, 25, 8].

- A novel device response model for an FMCW radar is developed, in order to mitigate disturbances of the receive signal due to the hardware impairments. As this model is non-linear, the parameter estimation and the disturbance mitigation have to be jointly accomplished by the model-based signal processing. The contributions in the field of modelling the response of FMCW radars are [21, 5, 9].

- Radars with closely located Tx and Rx channel typically suffer from leakage, the leak of signal energy from the Tx channel into the Rx channel. A novel method to cope with leakage will be presented, which employs background subtraction and a noise model for remaining signal parts. This has been published in [6].

- MIMO systems with time-division multiplexing (TDM) cause disturbances of angle and Doppler estimates (causes angle-Doppler coupling). A method to cope with angle-Doppler coupling in the parameter estimation framework will be presented in this thesis. It has been published in [10].

- A novel model to take phase noise of the radar hardware into account in processing of observations of an FMCW radar. To the best of the authors knowledge, this is done for the first time in radar signal processing. It is shown, that consideration of phase noise improves the detection performance of weak paths/targets. Furthermore, uncertainties of the data model are suppressed by the phase noise model.

- Model-based signal processing requires the proper selection of the model. Here, selection of the model refers to the determination of the model order, i.e. the number of point targets or propagation paths. A novel framework to determine the model order using stepwise regression search in conjunction with improved hypothesis tests is proposed.

1.2.2 Outline of the Thesis

The outline of the thesis is as follows:

Chapter 2: Model-Based System Identification – A Stochastic Inverse Problem

This chapter deals with the general description of the problem of identifying model parameters from noisy observations and under model misspecification. Treating noise and model errors as random disturbances, the parameter estimation problem becomes a stochastic inverse problem, whose solution by Bayesian inversion will be pointed out.

Chapter 3: Physical Model – Modelling of the Propagation

A physical model of the propagation channel, which is the considered system under identification, will be derived in this chapter. Description of the propagation of electromagnetic wave (EMW) will be briefly reviewed, and modelling by geometric ray optics will be described. The parameters of the geometric optical model and their relation to radar targets will be clarified. Last, the multidimensional path function and the multipath propagation channel model is introduced. The developed model is suitable to exploit high-resolution capability of the model-based signal processing.

Chapter 4: Device Model – Modelling of the Radar System

This chapter comprises the development of the device model of the broadband FMCW MIMO radar, which will be considered in this thesis. First, the basic theory of FMCW radars will be briefly reviewed, including signal theory and the stretch processing or homodyning receiver architecture. Afterwards, a model of the response of the transceiver chain is developed, which is suitable for de-embedding and mitigation of signal disturbances by the radar hardware. Because a MIMO radar is considered, the radar model also confounds a model of the antenna array. First, the description of antennas and modelling of the antenna output in terms of the antenna response is reviewed. Afterwards, the wideband and narrowband model of the direction- and polarisation dependent response of an antenna array is introduced, and respective conditions on the displacement of array elements are discussed. The joint calibration of the transceiver model and the narrowband or wideband array model from calibration measurements is discussed thereafter. Afterwards, the calibrated narrowband and wideband array model are compared regarding their usability to describe the array response for direction estimation purpose. Sampling of the baseband signal by the ADC of the radar, and interference free gathering of all MIMO channels by multiplexing will be discussed. Due to the time-division multiplexing of the considered radar, angle-Doppler coupling is present, whose compensation will be discussed too. In the last section, the coping with leakage, and the considered noise sources and their modelling as random processes is presented.

Chapter 5: Parameter Estimation from Observations by an FMCW MIMO Radar

The developed parameter estimator will be presented in this chapter. First, the statistical model of the observed data will be derived. This model bases on the derived propagation and device model from the previous two chapters. Afterwards, a parameter estimator is derived from Bayes' principle. This estimator is a ML estimator with bound constraints. Theoretical properties

of this estimator will be briefly summarised. As the estimator requires the optimisation of a non-convex objective function, a two-stage iterative and gradient-based optimisation method is presented.

Chapter 6: Model Order Estimation – Determination of the Number of Propagation Paths

In this chapter, the decision of the model order by parametric hypothesis testing will be outlined, and the developed methods to estimate the model order will be presented. The performance of the proposed model order estimators are verified and compared by simulations.

Chapter 7: Experimental Verification

The performance of the developed parameter estimator will be investigated in this chapter. Monte Carlo simulations will be used to do so. The achievable root-mean-square error (RMSE) of parameter estimates will be compared to theoretical findings. Also, the influence of model errors on the estimation performance will be investigated. Furthermore, the estimator is applied to measurements and selected results will be outlined and discussed. Last, the concept of a binocular MIMO radar (BinoMimo radar) is presented. Estimation results from applying the developed estimator to observations with the BinoMimo radar are presented, and the benefits of a binocular radar are outlined.

Chapter 8: Summary and Outlook

This chapter concludes the thesis and briefly summarises the major findings. Also, an outlook for further research and open issues will be given.

CHAPTER 2

MODEL-BASED SYSTEM IDENTIFICATION – A STOCHASTIC INVERSE PROBLEM

AS stated in the introduction, radar in conjunction with target parameter estimation will be treated as a model-based approach to identify a physical system. In this chapter, model-based system identification will be introduced as an *inverse stochastic problem*. Inverse, because the model parameters are mapped to the observations and shall be estimated from them. Stochastic, because the measurements are subject to probabilistic errors like noise and uncertainties. The chapter is intended to give an introduction to the topics of model-based or parametric system identification from measured data, the occurring inverse stochastic problem, and its solution by Bayes' theorem and point estimators. In order to understand the necessity of the subsequent chapters for the solution of the system identification problem, a profound knowledge of these topics is necessary.

Related Work and Own Contributions

In radar, the physical system under identification is the propagation channel of an electromagnetic wave. Empirical studies are conducted by special measurement systems, called radar observer, in order to physically access the system under identification. The physical system under identification is described by a mathematical model, which is driven by means of a set of parameters. These parameters are mapped to the measurement by the observation or sensing process, and will be subsequently estimated from the measurements using an estimator rule. This chain of relationships is depicted in Fig. 2.1.

The utilised observer is far away from being ideal. Therefore, the observations are influenced by the observer itself, such that properties of the physical system cannot be directly accessed from the measurements. In order to describe the influence of the observer on the observations, a model of the observer is necessary. Hence, a model of the sensing process (physical model and observer model) has to be developed, rather than a model for the system of interest only. The whole procedure of modelling, measuring and estimation will be wrapped up as model-based system identification [SS89]. According to the relationship between observations and the physical system of interest, there exists an inherent relationship between the parameters of the model of the physical system and the observations. The parameters are inferred from

---
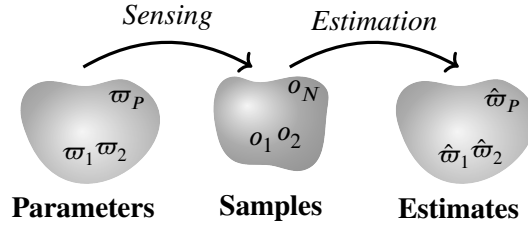
Figure 2.1: *Relationship of parameters ϖ, observations o and estimates $\hat{\varpi}$ in model-based system identification [TB07, p. 97]. The model being calibrated/identified is driven by a set of parameters from the parameter space $\Pi \subset \mathbb{R}^P$. The sensing process maps these parameters to the measurement or sampling space $\Omega \subset \mathbb{R}^N$. In the model calibration or parameter estimation step, these parameters are inferred from the observations by an estimation rule.*

the measurements to calibrate the model. Model calibration is the reverse of the observation process and therefore represents an inverse problem.

Inverse problems are well known in practical applications [ABT13, Han17, KS05, MS12] as e.g. communication engineering, control engineering, mathematics, econometrics, geophysics [Tar04, Tar05], imaging [SW12], or sensing. Also, inverse problems occur in many radar related applications [And03]. A bunch of methods and concepts to solve inverse problems are available, which are either specific for the problem under consideration, e.g. only applicable for linear problems, or which can be applied to a wide class of inverse problems. A very general and hence widely applicable concept is the Bayesian approach [DS17], which will be considered here.

First, the inverse problem and the parameter estimation procedure will be discussed more generally in Section 2.1. The estimation of the model parameters is influenced by the underlying uncertainties. Uncertainties are due to noise in the measurements and because the used mathematical model is subject to modelling errors. The quantification and classification of the occurring uncertainties as well as their modelling is described in Section 2.2. Due to the random nature of the measurements and uncertainties, the inverse problem cannot be tackled by deterministic inversion methods. Consequently, statistical inversion methods have to be employed, where uncertainties on the parameter estimates are assigned by probability density functions (PDFs) and point estimates are gained from these PDFs. Many methods for statistical inversion base on the Bayesian inversion approach, which exploits the Bayes theorem. The Bayesian inversion approach and derivation of parameter point estimates will be outlined in Section 2.3. Section 2.4 concludes the chapter.

The content of this chapter bases on own contributions in the area of statistical signal processing for radar and localisation purposes, e.g. [22, 7, 42, 8, 27, 28]. Especially the publication [28] deals with many aspects of this chapter, including the application of Bayesian inversion and utilisation of MAP and MMSE point estimators for noisy radar observations.

2.1 Relation of Observation and Truth

Estimating the parameters of a model describing a physical system has been introduced as model-based system identification. In order to properly introduce the system identification task, a general view point will be given, whereas the relation to the problem under consideration will be given thoroughly. Subsequently, the case of noiseless observations/measurements will be assumed.

Consider a set of observations O as a collection of physical events, which are observed by a sensing device. Also, the corresponding set of the truth T is introduced [Sha17]. The truth represents the reality, i.e. what really happened. The set of observations O can be considered as the set of the truth T, superimposed by uncertainties due to the operator's modelling errors and measurement errors due to the sensing device. In an ideal and perfect world, i.e. no errors or uncertainties are present, both sets are equivalent: $\mathsf{O} \equiv \mathsf{T}$. Vice versa, the truth can never be observed in practise and must be deduced from the observations. The set of the observation and truth may emerge from K physical events, with the k-th event denoted by vector $\mathbf{o}_k \in \mathsf{O}$ and $\mathbf{t}_k \in \mathsf{T}$.

$$\mathsf{O} = \left\{ \mathbf{o}_1, \dots, \mathbf{o}_K \right\} \tag{2.1a}$$

$$\mathsf{T} = \left\{ \mathbf{t}_1, \dots, \mathbf{t}_K \right\} \tag{2.1b}$$

Each event comprises N quantities, which describe characteristics of the physical system under identification in the N-dimensional sampling space $\Omega \subset \mathbb{R}^N$. Thus, the event vectors represent a single point in the sampling space, and the sets of observations and truth are subsets: $\mathsf{O}, \mathsf{T} \subseteq \Omega^K$. The sampling space is spanned by the measurement dimensions, which are time, frequency and spatial domain in MIMO radar.

2.1.1 Forward Problem − Prediction of the Truth

Suppose a mathematical model or a hypothesis w.r.t. T, which describes the physical quantities and events in T perfectly. This model will be called *system model* $\mathcal{M}^{\mathrm{S}}$. For the considered problem, this model describes the radio channel. As stated in Section 1.1.1, the radio channel is made up by the propagation channel and the observer, which are both described by their respective model.

The model for the propagation channel is the model of the physical system under identification. Hence, this model will be denoted as *physical model* $\mathcal{M}^{\mathrm{P}}(\mathsf{T}, \boldsymbol{\varpi})$. The physical model may be deterministic and generally non-linear, and is driven by *physical model parameters* $\boldsymbol{\varpi} \in \Pi \subset \mathbb{R}^P$ and the truth set. Here, the set of physical parameters are the parameters of the propagation channel model, see Chapter 3. The output of the physical model is the set of predictions T'' of the truth, excluding the observer influence.

$$\mathsf{T}'' = \mathcal{M}^{\mathrm{P}}(\mathsf{T}, \boldsymbol{\varpi}) \tag{2.2}$$

The set of observations O is gathered by a sensing device or observer as e.g. a radar. The observer is described by a set of features, conditions and properties: the *physical scenario* [Sha17]. The physical scenario denotes the known experimental setup (device, measurement setup etc.) and is represented by the set $\mathsf{D} \subset \mathbb{R}^{N \times K}$. This set is generated by
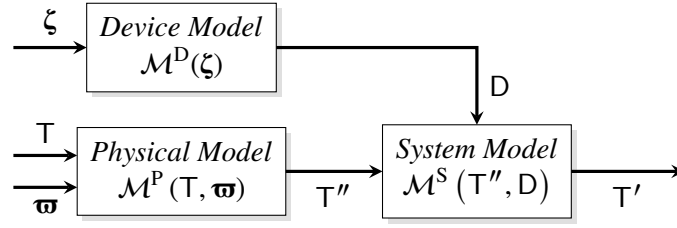
Figure 2.2: *Forward problem, which describes the mapping of the truth set T and the physical scenario set D on the set of predictions T'.*

the *device model* $\mathcal{M}^{\mathrm{D}}$, which is driven by the known parameters $\zeta \in Z \subset \mathbb{R}^D$. Chapter 4 deals with the derivation of a device model for a FMCW MIMO radar.

$$\mathsf{D} = \mathcal{M}^{\mathrm{D}}\left(\zeta\right) \tag{2.3}$$

In the simplest case the device model $\mathcal{M}^{\mathrm{D}}$ is independent of the truth. Also, it is not related to the physical model $\mathcal{M}^{\mathrm{P}}$, and the model parameters ζ and ϖ are disjunct. For example a linear time-invariant (LTI) model of the transceiver response of a radar fulfils this. However, if the device model is non-linear or the physical scenario depends on the truth, the device model and physical model become interrelated regarding their parameters (see e.g. Section 4.2.2 or Section 4.3). For example, the response of the antenna array depends on the direction of impingement, which is a parameter of ϖ.

Knowing the physical model and the device model, the system model $\mathcal{M}^{\mathrm{S}}$ can be defined. The system model describes a *set of predictions* $\mathsf{T}' = \left\{\mathbf{t}'_1, \ldots, \mathbf{t}'_K\right\} \subseteq \Omega^K$ about the truth.

$$\mathsf{T}' = \mathcal{M}^{\mathrm{S}}\left(\mathsf{T}'', \mathsf{D}\right) = \mathcal{M}^{\mathrm{S}}\left(\mathcal{M}^{\mathrm{P}}\left(\mathsf{T}, \varpi\right), \mathcal{M}^{\mathrm{D}}\left(\zeta\right)\right) \tag{2.4}$$

Model $\mathcal{M}^{\mathrm{S}}$ is also called the *forward model or prediction model*. Figure 2.2 sketches the relationship in equation (2.4). In an ideal and perfect world, i.e. no uncertainties like errors or noise are present, the set of predictions coincides with the sets of truth and observations $\mathsf{T} = \mathsf{T}' = \mathsf{O}$.

2.1.2 The Inverse Problem – Inference of Parameters

The forward problem makes a prediction T' of the truth T knowing the physical parameters ϖ according to the truth. These parameters are unknown (and the truth too) and have to be deduced from the set of observations. Assume, that equation (2.4) holds for the forward problem. Than, there exists a set of physical parameters $\hat{\varpi} \in \Pi$, which fulfils in the noiseless case ($\mathsf{T} = \mathsf{O}$)

$$\mathsf{T} - \mathcal{M}^{\mathrm{S}}\left(\mathcal{M}^{\mathrm{P}}\left(\mathsf{T}, \hat{\varpi}\right), \mathsf{D}\right) = \mathsf{T} - \mathsf{T}' = \mathbf{0} \quad . \tag{2.5}$$

The relationship described by equation (2.5) is pictured in Fig. 2.3. Solution of equation (2.5) is an *inverse problem* [Sul15, p. 92]. The term *inverse* is stated, because problem (2.5) is the inverse description of the forward problem (2.4). Inverse problems are always present, if one wants to get information about a physical process, which is not directly observable. Instead, another process is observed, which incorporates the physical process of interest. The
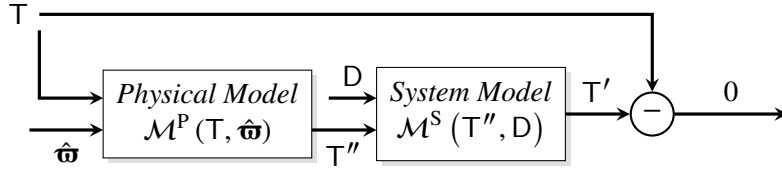
Figure 2.3: *Inverse problem, which generates the zero set according to the difference of the truth set and the model prediction using parameters* $\hat{\varpi}$.

physical process of interest is inferred from the observed process, which is the inversion of the incorporation of the process of interest in the observed process.

The question arises, whether the inverse problem is solvable, i.e. the physical parameters can be inferred. Basically, inverse problems are solvable, if they are *well-posed*. According to the definition suggested by Jacques Hadamard [Had02], a well-posed problem fulfils the following conditions:

1. **existence**: at least one solution exist

2. **uniqueness**: the solution is unique

3. **stability**: the solution is non-sensitive to the input data .

The third criterion states, that the inverse of the forward model is a continuous function. Hence, a small change in the input data results in small changes on the solution. Especially in case of erroneous input data this condition has a tremendous meaning, because even "simple" inverse problems can become quickly ill-posed. Typically, forward problems are well-posed, whereas inverse problems are ill-posed [DS17, MS12, Tar05].

The inverse problem in equation (2.5) is a continuum model on an infinite dimensional spaces. However, their practical solution faces some limitations as e.g. numerical instability due to finite computational precision, finite number of measurements or solution w.r.t. a finite number of unknowns. Hence, even if the original problem is well-posed, the problem may become ill-conditioned in the practise. Ill-conditioned is equivalent to the third Hadamard criterion, meaning that a small error in the input data result in large errors in the solution.

2.2 Uncertainties – Model Errors and Measurement Noise

Assuming an ideal and perfect world, the physical model fully describes the observations without any imperfections or uncertainties. However, this assumption does not hold in reality, because the presumed perfect world does not exist and the observations are always confounded by uncertainties. Understanding the uncertainties, accounting for their sources and quantifying the contributions of each individual uncertainty source is stated as *uncertainty quantification* [Cou+12, Owh+13]. Sources of uncertainties are e.g. limited accuracy of the physical model or device model, modelling assumptions, measurement errors, calibration accuracy, numerical limitations or approximations [Sul15, p. 2]. Uncertainties are divided into the category of *epistemic or systematic uncertainties*, and *aleatory or statistical uncertainties* [KD09].

Aleatory Uncertainties are a property of the physical system. They are presumed to originate from unpredictable randomness in the observations or not-reproducible errors [KD09, Sul15]. This uncertainties cannot be reduced in the estimation or observation process.

Epistemic Uncertainties are a property of the observer itself. They state any lack of knowledge or data about the problem under study or the experimental setup. This uncertainties are assumed to be reducible by e.g. gathering more data, using better observers or refining the model [KD09, Sul15]. They can be further divided in model and parametric uncertainties, describing doubts on the model itself or on the model parameters, respectively [Sul15, p. 4].

In model-based system identification both types of uncertainties are present, whereas it is difficult to determine from which category a certain uncertainty originates. The categorisation is up to the operator during the modelling phase. The two major sources of uncertainties considered here are model errors and measurement noise (or model inadequacy and measurement uncertainties according to [Sha17]). Measurement noise is assumed as epistemic uncertainties [Sha17]), because the influence of measurement noise can be reduced by e.g. more accurate measurement devices or averaging. Model errors cannot be easily assigned to one of these groups [Sha17]. First, the sources of model errors have to be clarified in more detail. Quite often, models are proposed, which base on measurements itself. However, the knowledge of the measurement setup as well as the experiment's input is always subject to incompleteness (epistemic uncertainties), which leads to the development of inaccurate or incomplete models of the underlying physics. Furthermore, physical models are often subject to mathematical idealisation due to a trade-off between accuracy and effort. Such models are inherently inadequate. On the other hand, model errors are quite often assumed as aleatory uncertainties, because the improvement of the knowledge encounters its limits in reality and uncertainties always remain. Summarised, model errors are assumed to contribute to both categories of uncertainties. Depending on the type of uncertainty, being measurement noise or model errors, different modelling approaches have to be chosen. Measurement noise and its modelling for a FMCW radar will be discussed in Section 4.6. Model errors as epistemic uncertainties and their reduction by developing better or improvement of existing models will be discussed in Section 4.2, Section 4.5 and Section 4.4 for an broadband FMCW radar.

Developing an advanced model to account for these uncertainties will change the nature of the model for the predictions. So far, the system model $\mathcal{M}^S\left(\mathsf{T}'', \mathsf{D}\right)$ as a physical, deterministic model has been considered for prediction. By taking into account the uncertainties the model becomes a stochastic model, because of the stochastic modelling of the uncertainties. However, this advanced model no longer describes the physics of the system, rather it describes the processes of observation and modelling. Hence, the advanced model is commonly denoted as *observation model* or *data model*.

2.2.1 Modelling the Presence of Measurement Noise

Experimental observations are affected by measurement noise like e.g. thermal noise, quantisation noise or flicker noise [SSR06]. Measurement noise is a non-reproducible error, hence, despite the same experimental settings different values are observed. The unavoidable presence of noise implies that the truth about an event can never be gathered by an observer

and a subsequent compensation is impossible. Furthermore, the influence of the noise on the observation is not known in a deterministic fashion. Consequently, the effects of noise are treated as being of stochastic or probabilistic nature. Therefore, the only valuable statement about the noisy observation in terms of the truth is, that the observation is the output of a stochastic process, driven by the truth as input [Sha17].

Assume the set of observations O is superimposed by noise. Under a well-known experimental setup and procedure, the noise-caused effects $\mathbf{n}_k$ in the k-th observed event $\mathbf{o}_k$ may be quantised by the model $\mathcal{M}^{\mathrm{N}}{}_k \left(\mathbf{t}_k, \boldsymbol{\sigma}_k \right)$, which approximates the influence of all noise sources on the observations [Sha17].

$$\mathbf{o}_k \sim \mathcal{M}^{\mathrm{N}}{}_k \left(\mathbf{t}_k, \boldsymbol{\sigma}_k \right) \tag{2.6}$$

In Section 4.6 noise processes and their models for an FMCW MIMO radar will be discussed. The set of known parameters $\boldsymbol{\sigma}_k \in \Sigma \subset \mathbb{R}^W$ and the k-th event of the truth $\mathbf{t}_k \in \mathsf{T}$ are the inputs. The model relates to the k-th event only, because the model itself or its parameters may change from event to event, e.g. the parameters are time-variant. Here, the noise model and the respective parameters are assumed as time-invariant: $\mathcal{M}^{\mathrm{N}}{}_k \Rightarrow \mathcal{M}^{\mathrm{N}}$ and $\boldsymbol{\sigma}_k \Rightarrow \boldsymbol{\sigma}$. Using model (2.6) the probability of observing $\mathbf{o}_k$ under the truth $\mathbf{t}_k$ can be assigned. Because the truth is unknown, this cannot be easily accomplished under practical considerations, see Section 2.3.

Due to the noise the observations become a random process too. Thus, model (2.6) is a stochastic model, rather than a deterministic one-to-one mapping between truth and observation. Hence, the model does not predict the observations, the model predicts the statistics of the observations. Commonly, such models are given in terms of a PDF. For a brief recap of the theory of the description of random processes and PDFs see APPENDIX B.2. A model, which describes the noise process itself and not the noisy observations, will be introduced. Define the set of K events emerging from the noise process $\mathbf{n}_k \in \mathsf{N}$.

$$\mathsf{N} = \left\{ \mathbf{n}_1, \ldots, \mathbf{n}_K \right\} \tag{2.7}$$

The model, precisely the distribution, for the random noise process is

$$\mathbf{n}_k \sim \mathcal{M}^{\mathrm{N}} \left(\mathbf{0}, \boldsymbol{\sigma} \right) \quad . \tag{2.8}$$

Again, the model does not predict the noise process, rather than describes its statistics by a PDF. Assume the noise as additive [Sul15] the following model for the observations can be stated.

$$\mathbf{o}_k = \mathbf{t}_k + \mathbf{n}_k \quad . \tag{2.9}$$

This model for the noisy observations is depicted in Fig. 2.4.

The noise model $\mathcal{M}^{\mathrm{N}}$ and its parameters $\boldsymbol{\sigma}$ have to be known for the inference procedure. Commonly, the structure of the noise model is known only, whereas the parameters are unknown in advance. Hence, they are subject to the inference process, but they are not of immediate interest. Parameters, which are not of interest but have to be account for in the estimation process are denoted as *nuisance parameters*.
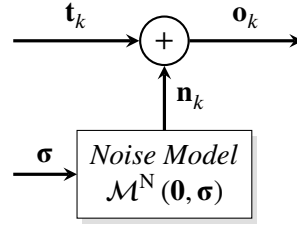
Figure 2.4: *Model of measurement noise and its incorporation in the noisy observation.*

2.2.2 Modelling the Presence of Model Errors

There is a quote from George E. P. Box in his publication [Box76], which slightly describes the subsequent problem under consideration: 'all models are wrong'. According to this phrase, there seems to be no physical model to properly predict the truth in a general sense, aren't they? However, this phrase states that no *reasonable model*, reasonable in the sense of a finite parameter set of the model, provides a *perfect* prediction of the truth. Hence, the model is subject to imperfections or incompleteness, which are commonly stated as *model discrepancy or model inadequacy* [Sha17]. Quantification of this inadequateness by modelling approaches is discussed in the following. Dealing with this uncertainties allows to apply models, which 'are false but which we believe may be useful nonetheless' [Box76] for the parameter estimation problem.

The set E will be introduced, which represents any *unpredicted or unknown* variability, prediction errors and uncertainties of $\mathcal{M}^{\mathrm{S}}$ in predicting T.

$$\mathsf{E} = \left\{ \mathbf{e}_1, \ldots, \mathbf{e}_K \right\} \tag{2.10}$$

Each event $\mathbf{e}_k \in \mathsf{E}$ corresponds to one event $\mathbf{t}_k$ of the truth. Consider observations without noise. If the physical model does not match the observations, the forward problem (2.4) is subject to uncertainties. Hence, the inverse problem (2.5) cannot be equal to the zero set. Considering the set of model errors E, the inverse problem can be stated as [Sha17]

$$\mathsf{T} - \mathcal{M}^{\mathrm{S}} \left(\mathcal{M}^{\mathrm{P}} \left(\mathsf{T}, \boldsymbol{\varpi} \right), \mathsf{D} \right) = \mathsf{T} - \mathsf{T}' = \mathsf{E} \quad . \tag{2.11}$$

The model errors have to be described under statistical as well as physical considerations, because the errors are generally assumed of mixed deterministic and stochastic nature. For example, the deterministic part introduces e.g. a bias and the stochastic part may result in increased variance of the solution of the inverse problem. Introducing the model $\mathcal{M}^{\mathrm{E}}$, which features the above stated dependencies and reflects the set of model errors, the set E is given by this model as

$$\mathsf{E} \simeq \mathcal{M}^{\mathrm{E}} \left(\mathsf{T}, \mathsf{T}', \boldsymbol{\epsilon} \right) \quad , \tag{2.12}$$

where $\boldsymbol{\epsilon} \in E \subset \mathbb{R}^E$ denotes the set of parameters of the error model. Because the errors are assumed as deterministic and stochastic, the symbol $\simeq$ is used in equation (2.12). Just as with measurement noise, the model describes the statistics of the errors rather than predicting them. Hence, the error model is a statistical model. On the other hand, deterministic errors, e.g. errors due to simplifications, are predicted by the model as well. Hence, the error model
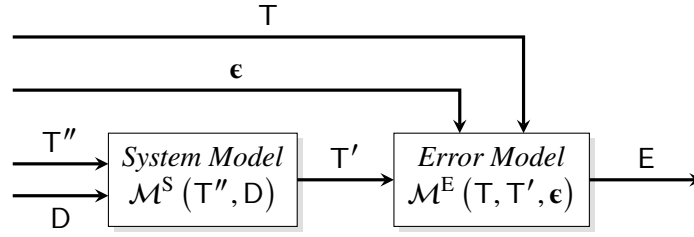
Figure 2.5: *Model for the errors of the predictions* T' *w.r.t. the truth* T

is a deterministic model too. The error generation and the relations of the error model w.r.t. the other sets is depicted in Fig. 2.5. Note, that the error model depends on the truth, which is unknown in the inference process. Hence, the model for the prediction errors has to be re-stated under practical considerations, see Section 2.3.

Choosing a certain structure for the model of the errors is quite complicated. The prediction errors depend on the truth, such that a model cannot be easily found in advance. This is a fundamental difference to measurement noise, because noise and truth are independent. Hence, a model for the noise can be found easily in advance, e.g. approximation as a Gaussian random process. Design of a proper error model is subject to empirical studies. There, observations of sets of known truths and their description by the system model are employed. By analysing the discrepancy of the known and predicted truth an "average" model of prediction errors can be derived, averaged over the set of considered truths. An example is the estimation of a global calibration matrix for direction estimation with arrays [7]. The error model and its parameters have to be known in order to solve the inverse problem. If the model is given only and the parameters are not known in advance, the parameters become subject to the parameter estimation process as nuisance parameters.

Quite often and for reasons of simplification, model errors are assumed as small, such that they are hidden by the measurement noise [7, 8, 26]. Hence, a specific error model is not taken into account by the estimator rule. Also, model errors are sometimes assumed as part of the noise process, which is not completely accurate, because both processes emerge from different and independent sources. Also, model errors are not purely random in their nature [26], because they are reproducible under stationary measurement circumstances contrarily to measurement noise. Here it will be assumed, that the model errors are *hidden by the noise process*, see Section 4.6 for a discussion on that. As a consequence, the system model has to be very accurate, in order to presume this assumption.

2.2.3 Forward Problem – Forward Uncertainty Propagation

The joint influence of all sources of uncertainties is considered now. An overall description is derived, quantifying the influence of uncertainties on the predictions. This is called *forward uncertainty quantification*. On the contrary to most forward uncertainty quantification procedures, uncertainties of the model parameters $\boldsymbol{\varpi}$ are not considered here.

The previously described uncertainties account for the deviation between predicted and observed set. Assume the model errors as negligible or hidden by the noise process, the

following model for the observations can be stated.

$$O = T + N = T' + E + N \approx T' + N \tag{2.13}$$

Prediction of observations becomes a prediction in a statistical sense, i.e. the distribution of the observations is predicted.

$$O \sim \mathcal{M}^N \left(T', \boldsymbol{\sigma} \right) = \mathcal{M}^N \left(\mathcal{M}^S \left(\mathcal{M}^P \left(T, \boldsymbol{\varpi} \right), D \right), \boldsymbol{\sigma} \right) \tag{2.14}$$

For short hand notation the operator $\mathcal{G}$: $\boldsymbol{\varpi} \mapsto O$ is introduced, which denotes the above a probabilistic mapping. The truth T is unknown under practical considerations, such that the above mapping cannot be easily given. Instead, it has to be described according to a hypothesis of the truth, see Section 2.3.

2.2.4 Inverse Problem – Inverse Uncertainty Quantification

So far, uncertainties w.r.t. the measurements and the mathematical model of the system were stated. In the inference process, which is considered in this thesis, the parameters of the model have to be deduced from the observations. Hence, the *stochastic inverse problem*

$$\mathcal{G}^{-1} : O \mapsto \hat{\boldsymbol{\varpi}} \tag{2.15}$$

has to be solved. Because $\mathcal{G}$ is a probabilistic mapping to account for uncertainties, the estimates $\hat{\boldsymbol{\varpi}}$ are subject to uncertainties too. Hence, a simple inversion of $\mathcal{G}$ on the observations O is not meaningful [DS17]. The question arises: "How to propagate uncertainties on observations and model to uncertainties on the parameter estimates?". Uncertainties on the estimates $\hat{\boldsymbol{\varpi}}$ can be assigned by a respective distribution of them. This is *inverse uncertainty quantification*. The uncertainties in the sampling space (model errors and measurement noise) are back propagated to the parameter space by solving the inverse stochastic problem. Contrarily to the forward uncertainty quantification, uncertainties on the model parameters are not propagated through the model, rather than deduced by analysing the measurements. A method for inverse uncertainty quantification is the Bayesian approach, where the distribution of the parameter estimates (the so called posterior distribution) is assigned, see Section 2.3.

Quantification of the uncertainties on the estimates is divided into *bias correction* and *model calibration* [Sul15]. Bias correction is the attempt to describe (and also minimise) the discrepancy between observation and the model. This part of the quantification can be seen as choosing the correct model or improving the model according to the bias. Model calibration denotes the estimation of the parameters of the model, see Chapter 5. Basically, bias correction and model calibration are both present in model-based system identification. In many cases, the correct or best suited model to describe the system under identification is unknown and has to be chosen from a family of candidate models. A famous example is the deduction of the number of propagation paths P, which sufficiently describes the observed propagation channel. This number is commonly not known in advance, and is therefore subject to the estimation task too, see Chapter 6. However, the model selection problem is not covered by the discussion on the inverse problem, because the model is assumed as fully specified their.

2.3 Bayesian Inversion – Statistical Inversion by Bayes Theorem

Due to the presence of random uncertainties the inverse problem, i.e. estimating the unobserved model parameters (the unknowns) from the observations, cannot be solved deterministically. Instead, methods of *statistical inversion* have to be applied. A common method therefore is *Bayesian inversion*, where the statistical inverse problem is recast as a problem of *Bayesian inference*. There, the *inverse probability* is assigned using *Bayes' theorem*. The inverse probability is in fact the inverse uncertainty quantification, i.e. the assignment of the uncertainties on the unknown parameters after the observation. On the contrary to frequentist procedures, where the model parameters are assumed as fix and deterministic, the observations as well as the unknowns are assumed as random in the Bayesian setting. Hence, uncertainties on the system parameters can be taken into account as well.

However, the goal of parameter estimation is to find the "best estimate" $\hat{\varpi} \in \Pi$ of the unknown parameters in a statistical sense, rather than assigning uncertainties. This casts the statistical inverse problem into a *point estimation problem* [BS94].

2.3.1 Bayes' Theorem – Assigning the Posterior Distribution

Solution of the statistical inverse problem in Bayesian sense is the assignment of the inverse probability or the probability of the parameters after the observation. Three key components are used therefore

- the observations

- a model of the observation process

- a-priori information on e.g. the unknowns.

According to Bayes theorem [ABT13] the inverse probability or *posterior distribution* can be calculated from the *marginal distribution*, the *likelihood function* and the *prior distribution*.

- $p(\varpi | O)$ posterior distribution of the parameters

- $p(\varpi | \iota)$: prior distribution of the parameters

- $p(O | \varpi)$: observation distribution or likelihood

- $p(O | \iota)$: marginal distribution or marginal likelihood.

The likelihood is the conditional distribution of the observations given the parameters. This is the model of the measurement process itself, i.e. the distribution of gathering the observations under a certain realisation of the model parameters. The prior distribution states a-priori information on the unknowns as e.g. their uncertainties or constraints like parameter bounds, and is parametrised by $\iota \in \mathbb{R}^H$. These parameters are termed as *hyperparameters*. Sometimes, these hyperparameters are unknown in advance, and hence become additional unknowns in

Bayesian inference. The posterior distribution is the conditional probability distribution of the unknowns *after* the observation, i.e. propagates the uncertainties on the observations to uncertainties on the estimated unknowns. The posterior distribution is given by Bayes' theorem [BS94].

$$p(\boldsymbol{\varpi}|\, \mathsf{O}) = \frac{p(\mathsf{O}|\,\boldsymbol{\varpi}) \cdot p(\boldsymbol{\varpi}|\boldsymbol{\iota})}{p(\mathsf{O}|\boldsymbol{\iota})} \tag{2.16}$$

For sake of notational convenience, hyperparameters will be neglected in the following.

$$p(\boldsymbol{\varpi}|\, \mathsf{O}) = \frac{p(\mathsf{O}|\,\boldsymbol{\varpi}) \cdot p(\boldsymbol{\varpi})}{p(\mathsf{O})} \tag{2.17}$$

The marginal likelihood can be calculated by integrating the product of prior and observation distribution over the parameter space.

$$p(\mathsf{O}) = \int_{\Pi} p(\mathsf{O}|\,\boldsymbol{\varpi}) \cdot p(\boldsymbol{\varpi}) \; d\boldsymbol{\varpi} \tag{2.18}$$

The marginal likelihood is independent on the unknown parameters, and just a normalisation constant in order to have a proper PDF. Thus, the marginal likelihood becomes negligible in the parameter estimation step [28].

2.3.2 Probability Distribution Under Measurement Noise

The experimentally collected observations O are subject to measurement noise and the physical model is inadequate, turning the observations into a random process. Hence, the noisy observations cannot be predicted in a deterministic fashion, only their statistics can be predicted. This is accomplished by assigning the distribution of the observations. As stated in Section 2.2, errors due to model inadequacy will be neglected or assumed as being hidden by the noise. Hence, the distribution of the observations under measurement noise will be taken into account only. Model (2.6) describes the distribution of the noisy observations. According to model (2.6), the k-th event $\mathbf{t}_k$ of the unknown truth is a model input. However, the truth is never known and subject to the inference process itself. Hence, the problem has to be re-stated.

Lets state the hypothesis, that the observations O are gathered under the truth $\mathsf{T}^* \subseteq \Omega^{*^K}$. This hypothesis on the truth is drawn by the operator, and it is a candidate solution of the inverse problem. May the superset Ω^{*^K} contain all candidate events or hypothetical truths $\mathbf{t}_k^* \in \mathsf{T}^*$, which itself fulfils the inverse problem (2.5).

$$\mathsf{T}^* - \mathcal{M}^{\mathrm{S}}\left(\mathcal{M}^{\mathrm{P}}\left(\mathsf{T}^*, \hat{\boldsymbol{\varpi}}\right), \mathsf{D}\right) = \mathbf{0} \tag{2.19}$$

Hence, the truly predictable statistic of the observation is always given under a certain hypothesis on the truth.

$$\mathsf{O} \sim \mathcal{M}^{\mathrm{N}}\left(\mathsf{T}^{*'}, \boldsymbol{\sigma}\right) = \mathcal{M}^{\mathrm{N}}\left(\mathcal{M}^{\mathrm{S}}\left(\mathcal{M}^{\mathrm{P}}\left(\mathsf{T}^*, \boldsymbol{\varpi}\right), \mathsf{D}\right), \boldsymbol{\sigma}\right) \tag{2.20}$$

The probability density $p\left(\mathbf{t}_k^* \middle|\, \mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right)$ of each candidate solution or hypothesis $\mathbf{t}_k^*$ being the

truth while observing $\mathbf{o}_k \in \mathsf{O}$ can be assigned using Bayes' theorem [Sha17].

$$p\left(\mathbf{t}_k^*\middle|\mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right) = \frac{p\left(\mathbf{o}_k\middle|\mathbf{t}_k^*, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right) \cdot p\left(\mathbf{t}_k^*\right)}{p\left(\mathbf{o}_k\middle|\boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right)} \tag{2.21}$$

Density $p\left(\mathbf{t}_k^*\right)$ assigns the prior probability of stating hypothesis $\mathbf{t}_k^*$ of the truth.

Next, consider equation (2.4) and equation (2.19) and take into account, that model errors are neglected. Thus, the hypothesis $\mathbf{t}_k^*$ is equal to the output of the system model $\mathbf{t}_k^{*'} \in \mathsf{T}^{*'}$ under this hypothesis.

$$\mathbf{t}_k^* = \mathbf{t}_k^{*'} = \mathcal{M}^{\mathrm{S}}\left(\mathcal{M}^{\mathrm{P}}\left(\mathbf{t}_k^*, \boldsymbol{\varpi}\right), \mathsf{D}\right) \tag{2.22}$$

The likelihood of the model parameters being the true parameters can be assigned by integrating over the entire set of hypothetical truths Ω^*.

$$p\left(\mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\middle|\boldsymbol{\varpi}, \mathcal{M}^{\mathrm{S}}\right) = \int_{\Omega^*} p\left(\mathbf{t}_k^*\middle|\mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right) \cdot p\left(\mathbf{t}_k^*\middle|\boldsymbol{\varpi}, \mathcal{M}^{\mathrm{S}}\right) \, d\mathbf{t}_k^* \tag{2.23}$$

Now the distribution of the physical model parameters, while observing $\mathbf{o}_k$, can be given [Sha17].

$$p\left(\boldsymbol{\varpi}\middle|\mathbf{o}_k, \mathcal{M}^{\mathrm{S}}, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right) = \frac{p\left(\mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\middle|\boldsymbol{\varpi}, \mathcal{M}^{\mathrm{S}}\right) \cdot p\left(\boldsymbol{\varpi}\middle|\mathcal{M}^{\mathrm{S}}\right)}{p\left(\mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\middle|\mathcal{M}^{\mathrm{S}}\right)} \tag{2.24}$$

Note, that relation (2.24) describes the solution of the statistical inverse problem. Hence, having defined a statistical model of the observations $p\left(\mathbf{o}_k, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\middle|\boldsymbol{\varpi}, \mathcal{M}^{\mathrm{S}}\right)$ and a priori information about the unknowns $p\left(\boldsymbol{\varpi}\middle|\mathcal{M}^{\mathrm{S}}\right)$, the inverse distribution $p\left(\boldsymbol{\varpi}\middle|\mathbf{o}_k, \mathcal{M}^{\mathrm{S}}, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right)$ can be assigned. Taking into account all events K and assuming them as independent, the complete solution of the statistical inverse problem is

$$p\left(\boldsymbol{\varpi}\middle|\mathsf{O}, \mathcal{M}^{\mathrm{S}}, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right) = \prod_{k=1}^{K} p\left(\boldsymbol{\varpi}\middle|\mathbf{o}_k, \mathcal{M}^{\mathrm{S}}, \boldsymbol{\sigma}, \mathcal{M}^{\mathrm{N}}\right) \quad . \tag{2.25}$$

2.3.3 Prior Distribution

The prior distribution or shortly prior represents any information regarding the unknown parameters, which is available *before* the observation. The prior assigns additional information to the estimation problem, which can be helpful to narrow the uncertainty on the parameters after the observation or to account for additional uncertainties before the observation. In case of an ill-posed problem, priors may assign some constraints, which result in regularisations and the problem becomes well-posed. Prior information has to be assigned by the operator, but there is no practical guidance to do so. It can be chosen by deductive reasoning as e.g. experiences of the operator, inspection of recent experimental results, exploitation of postulates on the physical system under consideration, or using physical axioms excluding certain model parameter values. Selection of the prior is purely subjective, such that Bayesian inversion gets a strong subjective nature. The somewhat arbitrariness in the selection of the prior is a major criticism of Bayesian analysis [KS05].
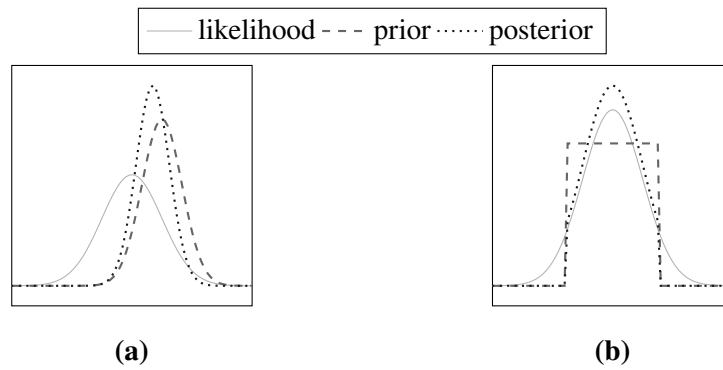
Figure 2.6: *Resulting posterior for (a) informative and (b) non-informative prior and equal likelihood.*

The prior must be properly selected, because he influences the posterior and hence the estimation result. In Fig. 2.6 a posterior distribution is shown for two different priors but the same likelihood. Obviously, the choice of the prior highly influences the posterior and hence the inference result.

Priors are divided into two classes: informative and non-informative priors. A prior is assigned to one of this classes according to his influence on the posterior. A prior is denoted as informative, if he has a significant impact on the posterior. Such priors are also denoted as *subjective priors*, because they are chosen from e.g. other data or the operator's experiences. Figure 2.6(a) shows an example for an informative prior. Priors are non-informative, if they minimally influence the posterior and are dominated by the likelihood. Such priors are denoted as *objective priors*, because they are constructed from some formal rules [KW96]. Figure 2.6(b) shows an example for a non-informative prior. A discussion on objectivist versus subjectivist Bayesianism is given in [Ber06, Gol06]. Because a proper prior is quite difficult to select the question arises, why priors should be taken into account. For example, a-priori information is not taken into account in frequentist inference, because this information was not observed and therefore cannot be verified in the inference step. However, as stated, a prior can be useful to improve the estimation quality or to make the estimation problem tractable.

Summarised, the prior is a major concern in Bayesian inference. In order to lower this concern, non-informative or objective priors are common choices. Such prior distributions play a minimal role in the posterior distribution and hence on the inversion result.

Uninformative Priors

Considering an uninformative prior means, that the "data speaks for itself". The prior shall have a minor influence on the posterior, such that the likelihood (i.e. the data) dominates. Here, two uninformative priors will be briefly outlined.

Principle of Indifference   Under the *principle of indifference* no certain realisation of parameter values is preferred over the other [ABT13, p. 255], because no prior information is available to do so. Hence, the parameters are assumed as uniformly distributed $\mathcal{U}$ in their

space Π.

$$p\left(\boldsymbol{\varpi}\middle|\mathcal{M}^{\mathrm{S}}\right) = \mathcal{U}(\boldsymbol{\varpi}) \tag{2.26}$$

Jeffrey's Principle   Another method to derive an uninformative prior is given by the so called *Jeffrey's principle of re-parametrisation* [Jay68]. Jeffrey argued, that an uninformative prior should be invariant w.r.t. any re-parametrisation of the model, i.e., remains uninformative under the re-parametrisation. The respective prior is

$$p\left(\boldsymbol{\varpi}\middle|\mathcal{M}^{\mathrm{S}}\right) \propto \sqrt{|\mathbf{F}(\boldsymbol{\varpi})|} \quad , \tag{2.27}$$

with $\mathbf{F}$ the Fisher information matrix (FIM). The FIM is a measure of the amount of information that an observation O carries about the model parameters [BS94, p. 361].

2.3.4 Point Estimates from the Posterior Distribution

So far, the stochastic inverse problem has been solved by assigning the posterior. However, parameter estimates $\hat{\boldsymbol{\varpi}}$, hence values of the parameters itself, shall be inferred; rather than just assigning their distribution after observation. Parameter estimates $\hat{\boldsymbol{\varpi}}$ are given as *point estimates* from the posterior [28]. Point estimates from the posterior can be calculated by a *Bayes risk estimator*, where an *average loss function* is used to justify the goodness of the point estimates [SD91]. Point estimates are inferred from the posterior by minimisation of this average loss function.

Loss Function and Risk

Defining a general loss function $C(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}) : \Pi \times \Pi \mapsto \mathbb{R}$, which is a general margin to denote an estimation error or estimate goodness. Despite knowing the true parameters $\boldsymbol{\varpi}$ some randomness in the loss function still remains, described by the observation distribution $p(\mathsf{O}|\boldsymbol{\varpi})$. In order to overcome this remaining randomness the expected loss function, which is the averaged loss function over the observation's distribution, is considered as margin [BS94, SD91]

$$\mathcal{R}(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}) = \int_{\Omega} C(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}) \cdot p(\mathsf{O}|\boldsymbol{\varpi}) \; d\mathsf{O} \quad . \tag{2.28}$$

This average loss function is denoted as the *risk*, which coincides with the frequentist definition of the average loss function [BS94]. Having the risk function defined the point estimate is given by minimising the risk subject to the parameters.

$$\hat{\boldsymbol{\varpi}} = \arg\min_{\hat{\boldsymbol{\varpi}}} \mathcal{R}(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}) \tag{2.29}$$

Bayes Rule – Minimising the Bayes Risk

The Bayes risk function is given by averaging the risk function (2.28) over the prior distribution.

$$\mathcal{R}\left(\hat{\boldsymbol{\varpi}}\right) = \int_{\Pi} \mathcal{R}\left(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}\right) p\left(\boldsymbol{\varpi}\right) d\boldsymbol{\varpi} = \int_{\Pi} p\left(\boldsymbol{\varpi}\right) \cdot \int_{\Omega} C\left(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}\right) \cdot p\left(\mathsf{O} \mid \boldsymbol{\varpi}\right) \, d\mathsf{O} \, d\boldsymbol{\varpi} \qquad (2.30)$$

The product of prior and observation distribution is the joint distribution. According to Bayes theorem, this joint distribution is equal to

$$p\left(\mathsf{O} \mid \boldsymbol{\varpi}\right) \cdot p\left(\boldsymbol{\varpi}\right) = p\left(\boldsymbol{\varpi} \mid \mathsf{O}\right) \cdot p\left(\mathsf{O}\right) \quad . \qquad (2.31)$$

Plugging into the Bayes risk function and having in mind, that the marginal likelihood $p(\mathsf{O})$ is a constant w.r.t. estimates and hence can be neglected under the minimisation, one gets the Bayesian conditional risk estimator [SD91].

$$\hat{\boldsymbol{\varpi}} = \arg \min_{\hat{\boldsymbol{\varpi}}} \int_{\Pi} C\left(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}\right) \cdot p\left(\boldsymbol{\varpi} \mid \mathsf{O}\right) \, d\boldsymbol{\varpi} \qquad (2.32)$$

So far, a general loss function has been considered for the derivations, without any specification. Considering a uniform and a quadratic loss function the Bayes risk estimator becomes the well known maximum a posteriori (MAP) and minimum mean-square error (MMSE) estimator, respectively [28, SD91].

Minimum-Mean-Square Error Estimator   A quadratic loss function is the Mahalanobis distance between the vectors of true parameters and the estimated ones [SD91].

$$C\left(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}\right) = \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right)^{\mathrm{T}} \cdot \mathbf{H} \cdot \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right) \qquad (2.33)$$

Plugging in to the conditional risk estimator (2.32) yields

$$\hat{\boldsymbol{\varpi}}_{\mathrm{MMSE}} = \arg \min_{\hat{\boldsymbol{\varpi}}} \int_{\Pi} \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right)^{\mathrm{T}} \cdot \mathbf{H} \cdot \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right) \cdot p(\boldsymbol{\varpi} \mid \mathsf{O}) \, d\boldsymbol{\varpi} \qquad (2.34)$$

Calculating the risk function's gradient w.r.t. the estimator $\hat{\boldsymbol{\varpi}}$

$$\frac{\partial}{\partial \hat{\boldsymbol{\varpi}}} \int_{\Pi} \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right)^{\mathrm{T}} \cdot \mathbf{H} \cdot \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right) \cdot p(\boldsymbol{\varpi} \mid \mathsf{O}) \, d\boldsymbol{\varpi} = -2 \cdot \mathbf{H} \cdot \int_{\Pi} \left(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}\right) \cdot p(\boldsymbol{\varpi} \mid \mathsf{O}) \, d\boldsymbol{\varpi}$$
$$(2.35)$$

and setting the gradient to zero yields

$$\mathbf{H} \cdot \hat{\boldsymbol{\varpi}} = \mathbf{H} \cdot \int_{\Pi} \boldsymbol{\varpi} \cdot p\left(\boldsymbol{\varpi} \mid \mathsf{O}\right) \, d\boldsymbol{\varpi} \quad . \qquad (2.36)$$
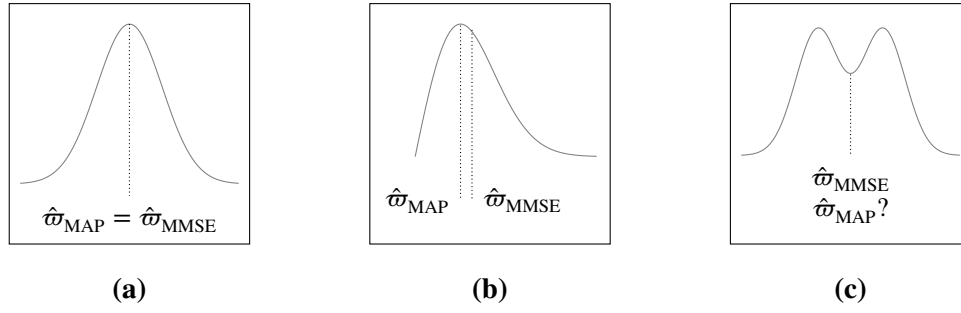
Figure 2.7: *Exemplary posterior distributions and the MAP and MMSE estimate $\hat{\varpi}_{MAP}$ and $\hat{\varpi}_{MMSE}$, respectively. The MAP and MMSE estimate are the mode and the first central moment of the distribution, respectively, and therefore depend on the shape of the distribution.*

If the matrix inverse $\mathbf{H}^{-1}$ exists, the Bayes risk estimator under quadratic loss function is given by

$$\hat{\boldsymbol{\varpi}}_{\text{MMSE}} = \int_{\Pi} \boldsymbol{\varpi} \cdot p(\boldsymbol{\varpi}|O) \, d\boldsymbol{\varpi} \quad . \tag{2.37}$$

Summarised, the Bayesian risk estimator under quadratic loss is given by the first central moment of the posterior distribution.

Maximum A Posteriori Estimator A uniform loss function treats the errors equally over the parameter space Π, except at the true parameter location [SD91].

$$C(\boldsymbol{\varpi}, \hat{\boldsymbol{\varpi}}) = 1 - \delta(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}) \tag{2.38}$$

Plugging in the uniform loss function into the conditional risk estimator (2.32) yields

$$\hat{\boldsymbol{\varpi}}_{\text{MAP}} = \arg\min_{\hat{\boldsymbol{\varpi}}} \int_{\Pi} [1 - \delta(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}})] \cdot p(\boldsymbol{\varpi}|O) \, d\boldsymbol{\varpi} \quad . \tag{2.39}$$

Due to the negative sign, the minimisation becomes a maximisation, and the Bayes risk estimator under uniform loss function is given by

$$\hat{\boldsymbol{\varpi}}_{\text{MAP}} = \arg\max_{\hat{\boldsymbol{\varpi}}} \int_{\Pi} \delta(\boldsymbol{\varpi} - \hat{\boldsymbol{\varpi}}) \cdot p(\boldsymbol{\varpi}|O) \, d\boldsymbol{\varpi} = \arg\max_{\boldsymbol{\varpi}} p(\boldsymbol{\varpi}|O) \quad . \tag{2.40}$$

Summarised, the Bayesian conditional risk estimator under uniform loss is given by the mode of the posterior distribution, i.e., maximising the probability of the parameters to be included in the observation.

Examples of MAP and MMSE Estimates The shape of the distribution, being e.g. symmetric and/or unimodal, determines the MAP and MMSE estimate. Figure 2.7 shows three exemplary distributions and the MAP and MMSE estimate $\hat{\varpi}_{MAP}$ and $\hat{\varpi}_{MMSE}$, respectively. In Fig. 2.7(a) the distribution is symmetric and unimodal, such that the MAP and MMSE estimate equals. In Fig. 2.7(b) the distribution is nonsymmetric and unimodal, such that the MAP and

MMSE estimate differs. Last, in Fig. 2.7(c), the distribution is symmetric but bimodal. The MMSE estimator exist but the MAP estimator does not. The distribution has two modes, such that the MAP is not unambiguously given. As a conclusion, existence of the MAP cannot be ensured in general, because it depends on the shape of the distribution.

Calculation of MAP and MMSE Estimator   The MAP estimator is the solution of an optimisation problem [28, Lue+20]. If the parameters ϖ are non-linearly included in the model, the optimisation problem may become a non-convex optimisation problem [25, 28]. Non-convex optimisation is a NP-hard problem [BV04] and no general algorithm for its solution exist. Only some heuristic methods [Wei11, Yan17] or problem specific solutions are given in literature, see [25, Ric05] and Section 5.3.

The MMSE estimator is the solution of a high-dimensional integral [28, Lue+20]. This can be non-trivial, especially if parameters are non-linearly included in the model. Therefore, a numerical solution by Markov chain Monte Carlo (MCMC) or importance sampling [28, Lue+20] like the Gibbs sampler, Metropolis-Hastings sampler or the slice sampler [GL06] is commonly employed. In MCMC sampling random samples are drawn from a distribution from which direct sampling is difficult. For calculation of the MMSE estimates the posterior is selected as this distribution. Only the prior and the likelihood distribution is required, whereas the marginal distribution is obsolete [28], because the marginal distribution is independent of the parameters and only a normalisation to assure a proper posterior distribution. Hence, drawn samples from an improper posterior properly represent the distribution, because the marginal distribution does not alter the shape of the posterior, which is the requirement to get the MMSE estimator. An estimate of the MMSE estimator is given by averaging the drawn samples [28].

2.4 Summary

In this chapter, model-based system identification from observed data has been discussed. There, the parameters of a model, which approximates the system under identification, have to be estimated from noisy observations.

In case of noiseless data, parameter estimation is an inverse problem, which can be described by its respective forward model. The forward model describes the observations and the mapping of the parameters of the system model into them. As it has been pointed out, the forward model cannot be the model of the system under identification only, rather than has to be a combined model of the system under identification and the observer. This stems from the fact, that a real observer is far from being ideal. Hence, the observer influences the observations, which has to be account for in the forward model. The existence of a proper solution of the inverse problem has been given by he known Hadamard criteria. Especially the uniqueness of a solution can be critical, as it requires a proper setting of the observer system.

As pointed out, real measurements are subject to uncertainties as e.g. measurement noise and model errors. Noise is treated as a random process, such that noisy observations are random too. Therefore, the estimates will become random and deterministic inversion is not applicable. As a consequence, the inverse problem turns into a stochastic inverse problem, which attempts to account for the randomness of the estimates. A solution of the inverse problem has been proposed by Bayes' rule. There, the distribution of the parameter estimates after the observation

can be addressed. Using a point estimator as e.g. the MAP estimates for the model parameters can be derived from this distribution.

Three main components to solve the stochastic inverse problem are stated in this chapter.

- System model $\mathcal{M}^{\mathrm{S}}$, composed of
    - Physical model $\mathcal{M}^{\mathrm{P}}$ (see Chapter 3) and
    - Device model $\mathcal{M}^{\mathrm{D}}$ (see Chapter 4)

- Model of the distribution of the noise sources (see Chapter 4)

- Parameter estimation rule (see Chapter 5)

These components will be addressed in the subsequent chapters for a considered broadband FMCW MIMO radar.

PHYSICAL MODEL – MODELLING OF THE PROPAGATION

MODEL-BASED identification of the physical system of interest requires the mathematical model $\mathcal{M}^{\mathrm{P}}$ of the physical system itself. Here, the wireless propagation channel is the system under identification. Goal of this chapter is to derive a mathematical model $\mathcal{M}^{\mathrm{P}}$ for the wireless propagation channel.

Related Work and Own Contributions

The physical system under identification is the *wireless propagation channel*. Therefore, model $\mathcal{M}^{\mathrm{P}}$ will be termed as *channel model*. The wireless propagation channel can be interpreted for wireless communication or radar applications. In wireless communication every component of the channel, which contributes to the energy transmission from Tx to Rx, is taken as valuable. In radar applications, scattering from the target are considered as valuable, and unwanted scattering from e.g. the surrounding is considered as distortion (termed as *clutter*). Hence, the *radar scattering channel* is, in the sense of interpretation of valuable contributions, a subset of the *wireless communication channel*. In the subsequent descriptions it will not be distinguished between the wireless communication or radar scattering channel. The term propagation channel is used to address both channel types.

The term channel model is possibly ambiguous, because the term may refer to

(i) Model as a transmission system

(ii) Model for simulation purpose, like the 3GPP [3GP] or WINNER [Kyö+08] model.

In this thesis, the first definition will be considered. The channel model describes the transfer function of the propagation channel in terms of a parametric mathematical model. The set of model parameters is unknown and subject to the estimation problem. In radar applications, these parameters have to be related to the targets for e.g. localisation and identification purposes.

The content presented in this chapter is well known in literature, see for example [Bal97, Hay94, Mau05, Mol12, Ric05, SMB01, TLS+04]. For sake of completeness and a holistic presentation of the system identification problem, a brief review of the modelling of the propagation of EMWs and derivation of a respective model of the propagation channel will be given. First, modelling of the propagation of EMW by ray optics will be described in Section 3.1.

An explanation, why a Dirac model is a sufficient modelling approach for wave propagation is intended to be given. Afterwards, a transmission model for the propagation channel is derived. The propagation channel will be described as a linear time-variant (LTV) system [Hay94, Mol12], whose impulse response is made up of the weighted superposition of multiple propagation paths. Section 3.3 concludes the chapter.

The described propagation model has been used in many own publications as e.g. [21, 22, 25, 42, 8, 27, 28].

3.1 Ray Optical Modelling of Electromagnetic Waves

The propagation channel is a *continuum*, featuring propagation mechanisms like reflection, diffraction or scattering [YI15]. It is widely accepted, that this continuum can be adequately approximated by the superposition of a *finite* number of EMWs. Each wave emerges from a transmitter and propagates to a receiver by the stated propagation mechanisms. The solution of Maxwell's equations form the basis to describe the propagation of . However, their solution is practically intractable even for moderately complicated scenarios. Hence, asymptotic and approximative approaches are employed for simplifications. Here, the ray-based modelling of EMW propagation will be considered. An EMW can be adequately described by a ray, because electromagnetic theory can be asymptotically described by geometric optics [Kli62, Mau05].

3.1.1 Theoretical Foundations of Propagating Electromagnetic Waves

Coordinate Systems

Describing the propagation of an EMW necessitates a coordinate system. Figure 3.1 shows two types of coordinate systems with a common origin: a *Cartesian* and a *spherical* coordinate system. Both coordinate systems will be considered subsequently.

Homogeneous Helmholtz Equations

An electromagnetic wave is described in time domain by the electric and magnetic field vectors $\vec{E}(\mathbf{d}, t)$ and $\vec{H}(\mathbf{d}, t)$, respectively, which depend on the location vector $\mathbf{d}$ and time t [Bal97, Mau05]. Consider a source-free, isotropic, homogeneous, linear and lossless medium as e.g. the vacuum [Orf16]. Then, the *wave equations* according to the Maxwell's equations are [Orf16]

$$\triangle \vec{E}(\mathbf{d}, t) = \varepsilon_0 \mu_0 \frac{\partial^2}{\partial t^2} \vec{E}(\mathbf{d}, t) \tag{3.1a}$$

$$\triangle \vec{H}(\mathbf{d}, t) = \varepsilon_0 \mu_0 \frac{\partial^2}{\partial t^2} \vec{H}(\mathbf{d}, t) \quad , \tag{3.1b}$$
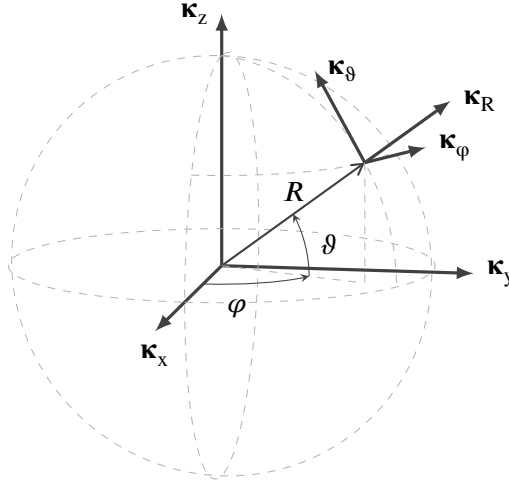
Figure 3.1: *Spherical and Cartesian coordinate system featuring a common origin. Their respective basis vectors are* κ_x,κ_y,κ_z *and* κ_φ,κ_ϑ,κ_R. *(Picture taken from [Hei09b].)*

with *electric permittivity in vacuum* ε_0 and *magnetic permeability in vacuum* μ_0, and $\triangle$ the Laplace operator. By separation of variables, the time-space dependent vectors $\vec{E}(\mathbf{d}, t)$ and $\vec{H}(\mathbf{d}, t)$ can be separated into vectors $\vec{E}(\mathbf{d})$ and $\vec{H}(\mathbf{d})$, which account for the spatial dependence, and scalars $E(t)$ and $H(t)$, which account for the time dependence.

$$\vec{E}(\mathbf{d}, t) = \vec{E}(\mathbf{d}) \cdot E(t) \tag{3.2a}$$

$$\vec{H}(\mathbf{d}, t) = \vec{H}(\mathbf{d}) \cdot H(t) \tag{3.2b}$$

Harmonic signals with frequency f_0 and initial phase ϕ_0 will be assumed, with E_0 and H_0 the respective magnitudes.

$$E(t) = E(t, f_0) = E_0 \cdot \Re\left\{\exp\left\{j2\pi f_0 t + \phi_0\right\}\right\} \tag{3.3a}$$

$$H(t) = H(t, f_0) = H_0 \cdot \Re\left\{\exp\left\{j2\pi f_0 t + \phi_0\right\}\right\} \tag{3.3b}$$

Now, the wave equations simplify to the *homogeneous Helmholtz equations* [Mau05].

$$\left[\triangle + \kappa^2(f_0)\right] \cdot \vec{E}(\mathbf{d}) \cdot E(t, f_0) = 0 \tag{3.4a}$$

$$\left[\triangle + \kappa^2(f_0)\right] \cdot \vec{H}(\mathbf{d}) \cdot H(t, f_0) = 0 \tag{3.4b}$$

The variable κ is the *wave number*, which depends on the propagation medium. For an isotropic and lossless medium, the wave number is [Mau05]

$$\kappa(f) = 2\pi f/c_0 = 2\pi/\lambda \quad . \tag{3.5}$$

Summarised, an EMW can be described by a time and location varying harmonic, whose propagation can be described by the homogeneous Helmholtz equations. The electric and magnetic field are perpendicular to each other and to the propagation direction $\kappa_{Px} = \kappa_E \times \kappa_H$ of the wave, see Fig. 3.2.
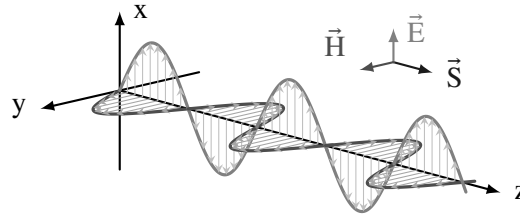
Figure 3.2: *Electromagnetic wave given by the orthogonal electric $\vec{E}$ and magnetic $\vec{H}$ field, and Poynting vector $\vec{S}$. (Picture taken from [Neu18].)*

Poynting Vector

The Poynting vector $\vec{S}(\mathbf{d}, t)$ describes the direction of propagation of the power flux. The Poynting vector is perpendicular to the electric and magnetic field vector ($\vec{S} \perp \vec{E} \perp \vec{H}$). Hence, it can be calculated by (see also Fig. 3.2)

$$\vec{S}(\mathbf{d}, t) = \vec{E}(\mathbf{d}, t) \times \vec{H}(\mathbf{d}, t) = S_0(\mathbf{d}, t)\,\boldsymbol{\kappa}_{\mathrm{Px}} \quad . \tag{3.6}$$

Solution of the Helmholtz Equations

Solution in Cartesian Coordinates – Plane Waves Solution of the Helmholtz equations in Cartesian coordinates $\mathbf{d} = [x, y, z]^{\mathrm{T}}$ results in *plane waves* (see Fig. 3.3). A plane wave is an EMW, where each wave front (surface of constant phase) is an infinite plane orthogonal to the direction of propagation [Orf16]. The fields may be represented in time domain as [Bal97, Mau05]

$$\vec{E}(\mathbf{d}, t) = E_0 \cdot \Re\left\{\exp\left\{-J\mathbf{d}^{\mathrm{T}}\mathbf{k}_0\right\} \cdot \exp\left\{J2\pi f_0 t + \phi_0\right\}\right\}\boldsymbol{\kappa}_{\mathrm{E}} \tag{3.7a}$$

$$\vec{H}(\mathbf{d}, t) = H_0 \cdot \Re\left\{\exp\left\{-J\mathbf{d}^{\mathrm{T}}\mathbf{k}_0\right\} \cdot \exp\left\{J2\pi f_0 t + \phi_0\right\}\right\}\boldsymbol{\kappa}_{\mathrm{H}} \quad , \tag{3.7b}$$

with $\boldsymbol{\kappa}_{\mathrm{E}}$ and $\boldsymbol{\kappa}_{\mathrm{H}}$ the normalized field vectors. The *wave vector* $\mathbf{k}_0 = \kappa_0 \boldsymbol{\kappa}$ is perpendicular to the wave fronts and depends on the wave number $\kappa_0 = \kappa(f_0) = 2\pi f_0/c_0$. For a lossless and isotropic medium, the direction of wave propagation $\boldsymbol{\kappa}_{\mathrm{Px}}$ and the direction of the normalised wave vector $\boldsymbol{\kappa}$ are equal. The direction of propagation of an EMW in terms of a Cartesian coordinate system is

$$\boldsymbol{\kappa}_{\mathrm{Px}} = \frac{1}{\sqrt{x^2 + y^2 + z^2}} \left[\boldsymbol{\kappa}_{\mathrm{x}}, \boldsymbol{\kappa}_{\mathrm{y}}, \boldsymbol{\kappa}_{\mathrm{z}}\right] \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad . \tag{3.8}$$

Plane waves are not physically, because the wave front is infinite large and thus requires infinite energy. Therefore, the spherical wave will be introduced next.

Solution in Spherical Coordinates – Spherical Waves Solution of the Helmholtz equations in spherical coordinates φ, ϑ, R results in *spherical waves* (see Fig. 3.4), i.e. the wave fronts become spheres. Assume a point source in the origin of the coordinate system. In a spherical coordinate system the direction of propagation corresponds to the radial direction
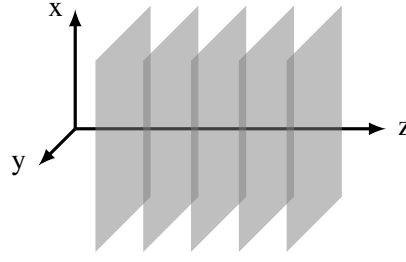
Figure 3.3: *Equi phase surfaces or wave fronts of a plane EMW, propagating in z-direction.*
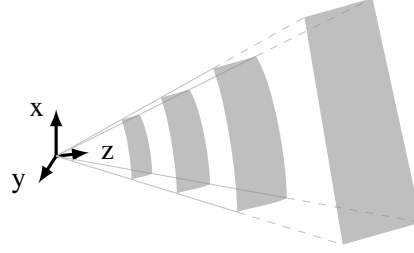
Figure 3.4: *Equi phase surfaces or wave fronts of a spherical EMW, propagating in radial direction. The spherical wave front becomes a locally plane wave front in the far field. Picture adapted from [Sta15a].*

of the basis vector $\boldsymbol{\kappa}_R$. Therefore, the direction of propagation is uniquely described by the elevation angle ϑ and azimuth angle φ.

$$\boldsymbol{\kappa}_{Px} = \boldsymbol{\kappa}_R \cdot 1 = \begin{bmatrix} \boldsymbol{\kappa}_x, \boldsymbol{\kappa}_y, \boldsymbol{\kappa}_z \end{bmatrix} \cdot \begin{bmatrix} \cos(\varphi)\cos(\vartheta) \\ \sin(\varphi)\cos(\vartheta) \\ \sin(\vartheta) \end{bmatrix} \qquad (3.9)$$

The solution of the Helmholtz equations in spherical coordinates yields [Orf16]

$$\vec{E}(\mathbf{d}, t) = E_0 \cdot \Re \left\{ \frac{J}{\kappa_0 R} \exp\left\{ -JR\kappa_0 \right\} \cdot \exp\left\{ J2\pi f_0 t + \phi_0 \right\} \right\} \boldsymbol{\kappa}_E \qquad (3.10a)$$

$$\vec{H}(\mathbf{d}, t) = H_0 \cdot \Re \left\{ \frac{J}{\kappa_0 R} \exp\left\{ -JR\kappa_0 \right\} \cdot \exp\left\{ J2\pi f_0 t + \phi_0 \right\} \right\} \boldsymbol{\kappa}_H \quad . \qquad (3.10b)$$

Unlike a plane wave, a spherical wave decreases in amplitude as it propagates outwards, which is indicated by the coefficient $1/R$.

Because the spherical wave model is hard to be considered in a mathematical model of the wave propagation (especially in ray optics, see next section), spherical waves are typically approximated as locally plane waves. Therefore, the plane wave expansion is considered [AS72], where a plane wave is described as the linear combination of spherical waves.

$$\exp\left\{ J\mathbf{d}^T\mathbf{k}_0 \right\} = \sum_{n=0}^{\infty} J^n \cdot j_n \left(\|\mathbf{d}\| \, \|\mathbf{k}_0\| \right) \cdot (2n+1) \cdot P_n \left(\mathbf{d}^T\mathbf{k}_0 \big/ \|\mathbf{d}\| \, \|\mathbf{k}_0\| \right) \qquad (3.11)$$

Function j_n is the spherical Bessel function of the first kind and *n*-th order, and P_n is the
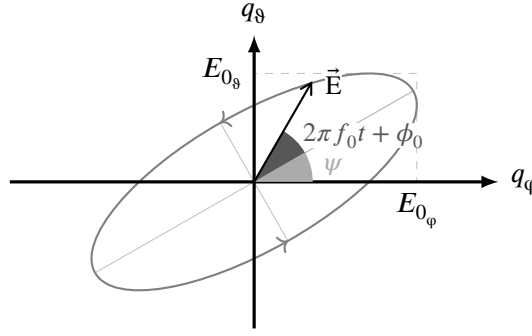
Figure 3.5: *Polarisation ellipse of a plane electromagnetic wave. The ellipse is the trace of the electric field vector $\vec{\text{E}}$ in the φ-ϑ-plane over time [Bal97]. Angle ψ is the slanting angle of the ellipse. (Picture adapted from [Hei09a].)*

Legendre polynomial of n-th degree. Under far field conditions ($R = \|\mathbf{d}\| \to \infty$) and in a local region ($\mathbf{d}^{\text{T}}\mathbf{k}_0 \approx \|\mathbf{d}\| \, \|\mathbf{k}_0\|$) only the contribution for $n = 0$ remains important. Hence, the plane wave expansion can be approximated.

$$\exp\left\{J\mathbf{d}^{\text{T}}\mathbf{k}_0\right\} \approx \frac{J}{\|\mathbf{k}_0\| \, \|\mathbf{d}\|} \exp\left\{J \, \|\mathbf{k}_0\| \, \|\mathbf{d}\|\right\} = \frac{J}{\kappa_0 R} \exp\left\{J\kappa_0 R\right\} \qquad (3.12)$$

Accordingly, a spherical wave can be approximated in a local region of the far field as a plane wave. Hence, a plane wave may be described in the far field by spherical coordinates.

A widely accepted criterion to ensure far field conditions at a distance R from the radiator and for a region of size D^2 is [SJ17]

$$R = \frac{\kappa_0}{8} \frac{D^2}{\phi} = \frac{\pi}{4} \frac{D^2}{\lambda\phi} \qquad . \qquad (3.13)$$

Variable ϕ denotes the phase variation of the spherical wave over the far field region. For example, a phase variation of $\pi/8$ or less requires a distance of $R \gg 2D^2/\lambda$. The above criterion is also important for array modelling, where the assumption of a plane wave across the whole array is of importance, see Section 4.3.

Polarisation

For a (locally) plane EMW, polarisation is defined as the shape, orientation, and sense of the ellipse, which is traced by the electric field vector $\vec{\text{E}}$ over time [Bal97, Com14], see Fig. 3.5. The electrical field vector is decomposed according to the φ-ϑ-plane, which is spanned by the basis vectors $\boldsymbol{\kappa}_\varphi$ and $\boldsymbol{\kappa}_\vartheta$ of the spherical coordinate system. Accordingly, the electrical field vector of a plane wave can be expressed in terms of the polarisation components q_φ and q_ϑ as

$$\vec{\text{E}} = E_0 \cdot \Re\left\{ \left[\boldsymbol{\kappa}_\varphi, \boldsymbol{\kappa}_\vartheta\right] \cdot \begin{bmatrix} q_\varphi \\ q_\vartheta \end{bmatrix} \cdot \exp\left\{-J\mathbf{d}^{\text{T}}\mathbf{k}_0\right\} \cdot \exp\left\{J\left(2\pi f_0 t + \phi_0\right)\right\} \right\} \qquad , \qquad (3.14)$$
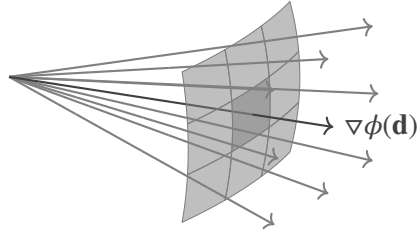
Figure 3.6: *Spherical wave front, approximated by multiple locally plane wave fronts. The propagation direction of a plane wave is described by a ray, orthogonal to the wave front. The propagation direction is given by the gradient of the phase or eikonal function $\phi(\mathbf{d})$, which is collinear to the Poynting vector. (Picture adapted from [Sta15b].)*

with polarisation components given by

$$q_{\varphi} = \cos(\psi) \cdot \exp\{J\beta_{\varphi}\} \tag{3.15a}$$

$$q_{\vartheta} = \sin(\psi) \cdot \exp\{J\beta_{\vartheta}\} \quad . \tag{3.15b}$$

Angle ψ is the slanting angle of the ellipse, and angle β_{φ} and β_{ϑ} is the phase of the field components $E_{0_{\varphi}}$ and $E_{0_{\vartheta}}$, respectively.

$$E_{0_{\varphi}} = E_0 \cdot \cos(\psi) \tag{3.16a}$$

$$E_{0_{\vartheta}} = E_0 \cdot \sin(\psi) \tag{3.16b}$$

3.1.2 Description of Wave Propagation by Ray Optics

Description of the propagation of EMWs is widely accomplished by geometrical optics, where a wave is treated like an optical ray and wave effects are ignored [Arn86, Kel78, Kli62]. The ray optical approach provides an asymptotic and approximative solution of Maxwell's wave equation for EMW at high frequencies [Arn86, Kel78, Mau05]. At higher frequencies, the object size becomes comparable to the wavelength, such that objects act like mirrors and do not cause additional sources when interacting with the wave. Also, the properties of the propagation medium stay constant over multiple wave lengths [Mau05].

A Brief Derivation of Geometrical Optics from Electromagnetic Theory

Subsequently, (locally) plane wave fronts are assumed. Spherical wave fronts are approximated by multiple plane wave fronts, see Fig. 3.6. The assumption of plane waves is a key prerequisite for the ray optical modelling [Kli62]. Consider the field components $\vec{E}(\mathbf{d}, f)$ and $\vec{H}(\mathbf{d}, f)$ of an EMW in frequency domain. By expansion in a Laurent series and asymptotic consideration of $f \to \infty$, the field components become the *geometrical optics fields* [Kli62, Mau05], which are

the leading orders of the expansion.

$$\vec{E}(\mathbf{d}) \propto \vec{E}_0(\mathbf{d}) \exp\{-\jmath\kappa\phi(\mathbf{d})\} \tag{3.17a}$$

$$\vec{H}(\mathbf{d}) \propto \vec{H}_0(\mathbf{d}) \exp\{-\jmath\kappa\phi(\mathbf{d})\} \tag{3.17b}$$

Argument $\phi(\mathbf{d})$ is the *phase or eikonal function* of the field, which describes the phase change of the field w.r.t. $\mathbf{d}$. The eikonal equation is well known in geometric optics, and hereby describes the link between wave propagation and ray optics. Plugging the geometrical optics fields in the Helmholtz equation (3.4) one can define the *eikonal equation* (∇ is the Nabla operator) [Arn86, Kli62, Mau05].

$$\|\nabla\phi(\mathbf{d})\| = 1 \tag{3.18}$$

A solution of the eikonal equation is given by $\phi(\mathbf{d}) = const$ [Kli62]. Thus, the eikonal equation describes a surface on which the phase changes are zero (equi phase surface). Hence, the eikonal equation describes the phase fronts of the waves. The geometrical optics fields $\vec{E}(\mathbf{d})$ and $\vec{H}(\mathbf{d})$ are parallel to this phase fronts [Arn86, Mau05]. The gradient $\nabla\phi(\mathbf{d})$ of the eikonal function is perpendicular to the wave front, and describes the direction of propagation, see Fig. 3.6. For a lossless medium the direction coincides with the propagation direction of the power flux $\vec{S}$. The direction of propagation can be described by a *ray*, such that the eikonal equation becomes the *ray equation* [Arn86]. In general, rays propagate on curved lines for arbitrary propagation medium. If the propagation medium is homogeneous, the ray path simplifies to a straight line [Arn86, Mau05].

Summarised, a ray can be used to model the propagation of a (locally) plane EMW. The propagation direction of the ray coincides with the Poynting vector of the plane EMW. A ray can be described by means of geometric rules [Arn86], and geometrical parameters can be assigned. A ray is treated as an infinite thin quantity in space. Hence, its propagation in spatial domain can be modelled in terms of Dirac impulses δ. Remember, that the Dirac impulse model enables high-resolution estimation in the respective domain (see Section 1.1.1).

3.2 Specular Propagation Path Model

From a system theoretic point of view the propagation channel is a system, which "transmits" electromagnetic waves from a source (transmitter) to a sink (receiver). This point-to-point transmission is accomplished by various propagation mechanisms as e.g. scattering [Mol12]. These mechanisms are deterministic and assumed as linear. If the propagation channel varies over time, due to movement of either the scatterers, the Tx or the Rx or combinations of them, the channel appears as a *linear time-variant* (LTV) system. Description of the propagation channel from a system theoretic viewpoint as a LTV system is a widely accepted modelling approach for the wireless communication channel [Mol12] as well as the radar scattering channel [JC17]. Putting together the concept of ray optical modelling from Section 3.1 and the LTV system model results in a model of the propagation channel. The geometric- and ray-based modelling of EMW propagation will be utilised to describes the LTV impulse response of the propagation channel in terms of geometrical parameters. An elementary quantity/event

Figure 3.7: *Extended radar target, which consists of multiple individual scattering surfaces. Observer O_1 is close to the target and resolves the scattering surfaces as individual point targets. Observer O_2 is far apart from the target and cannot resolve the scattering surfaces. Therefore, O_1 views the target as extended target and O_2 as point target.*

of the propagation channel model is introduced therefore: the *specular propagation path*.

3.2.1 Propagation Path – Parameters and Modelling

Specular Propagation Path and Point Target

The ray optical approach describes the propagation channel in terms of rays, which feature an infinite resolution in spatial domain. In order to properly resolve all rays, an infinite resolution would be necessary. However, an infinite resolution is not attainable in practise, because the resolution of the measurement system is limited by the system bandwidth and the resolution enhancement by model-based system identification is limited by the available SNR and model accuracy. Consider a bunch of multiple superimposed rays, which are so close in the spatial and temporal domain, such that they cannot be resolved by neither the measurement system nor the parameter estimation algorithm. Hence, the superimposed rays can be considered as a single "average" ray. This average ray is denoted as a specular propagation path [Ric05]. Summarised, a specular propagation path is defined as the smallest entity of a propagation channel, which transfers from the Tx to the Rx and which can be resolved.

In order to model the propagation channel for a radar application (i.e. the radar scattering channel), specular propagation paths have to be assigned to target scattering. A radar target is an extended physical structure, which comprises of several scattering surfaces with different reflectivity [HBC08, Tre01]. Electing each scattering surface by the Tx causes a contribution to the propagation channel as e.g. a path, which can be resolved. If the distance of the target from the radar is large, such that individual contributions are closer than the radar's resolution capability, the contributions of each scattering surface are not resolvable, see Fig. 3.7. As a consequence, an *average scatterer with average reflectivity* is observed. Thus, the target acts like a point in space from the viewpoint of the radar and is therefore termed as *point target* [HBC08, Sko90, Wil05]. Roughly speaking, a point target is a target which scatterers only a single wave to the Rx, if it is excited with a single wave by the Tx. However, if the target is close to the radar, such that several contributions become resolvable, the target is no longer viewed as a point, see Fig. 3.7. Indeed, the target becomes a spatially enlarged objected or *extended target* w.r.t. the radar. A common modelling approach for an extended target is the representation as a "cloud" of point sources [Gil+05, Sko90]. This cloud of point sources/targets can be assumed as a cluster, known from channel modelling [Czi07, Koc08, SC10]. Summarised, a single propagation path is assumed to emerge from scattering at a point target. A single propagation path (or point target, respectively) is assumed to coincide with the smallest resolvable scattering contribution of an extended target.
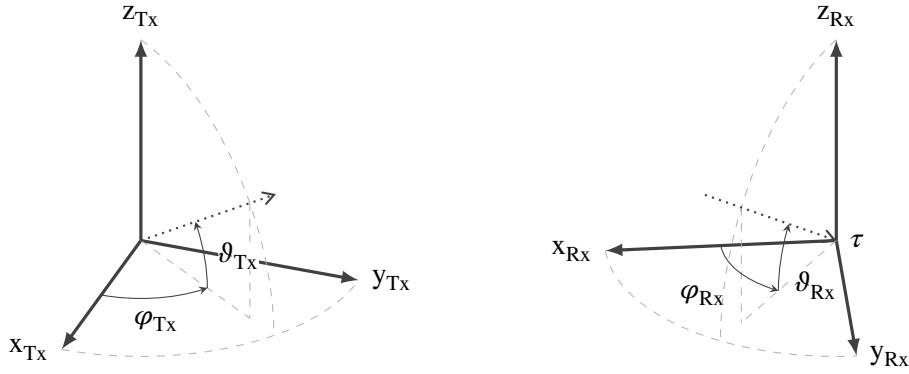
Figure 3.8: *Geometrical parameters of a specular propagation path in the double-directional channel model [Ric05, SMB01, TLS+04]. Picture adapted from [43].*

Table 3.1: *Geometrical parameters of a single propagation path, emerging from e.g. a target reflection or scattering, according to the double-directional channel model [Ric05, SMB01, TLS+04].*

Parameter	Symbol
Time of arrival or delay	τ
Azimuth of arrival	φ_{Rx}
Azimuth of departure	φ_{Tx}
Elevation of arrival	ϑ_{Rx}
Elevation of departure	ϑ_{Tx}

Geometrical Parameters

According to the ray optical approach, a specular propagation path is described in terms of geometrical parameters at Tx and Rx side. The respective model is quite often stated as double-directional channel model [Ric05, SMB01, TLS+04]. The geometrical parameters describe the propagation path in the spatial domain in terms of the angles at Tx and Rx side, i.e. the angles of arrival and departure, see in Fig. 3.8. The propagation duration in the temporal domain is assigned by a delay. Table 3.1 summarises the geometrical parameters per propagation path. Assuming a stationary propagation channel, the parameters of a propagation path are time-invariant.

Direction of Arrival and Direction of Departure The spatial domain is described by the azimuth and elevation angle of a single wave impinging at Rx or being radiated by Tx, as depicted in Fig. 3.8. Because angles are assigned, the spatial domain is often stated as angular domain. Here, the direction of departure (DoD) comprises azimuth of departure (AoD) and the elevation of departure (EoD), and DoA comprises azimuth of arrival (AoA) and elevation of arrival (EoA).

Time of Arrival The time of arrival (ToA) is defined as the waves travelling time from the coordinate system origin at Tx to the coordinate system origin at the Rx, see Fig. 3.8.

Table 3.2: *Mappings of polarisation at Tx and Rx, and corresponding entries of the scattering matrix* $\mathbf{\Gamma}$ *of a propagation path [Ric05, TLS+04].*

Polarisation Mapping	Matrix Entry
$q_{\varphi,\mathrm{Tx}} \rightarrow q_{\varphi,\mathrm{Rx}}$	$\gamma_{\varphi\varphi}(\tau)$
$q_{\vartheta,\mathrm{Tx}} \rightarrow q_{\varphi,\mathrm{Rx}}$	$\gamma_{\varphi\vartheta}(\tau)$
$q_{\varphi,\mathrm{Tx}} \rightarrow q_{\vartheta,\mathrm{Rx}}$	$\gamma_{\vartheta\varphi}(\tau)$
$q_{\vartheta,\mathrm{Tx}} \rightarrow q_{\vartheta,\mathrm{Rx}}$	$\gamma_{\vartheta\vartheta}(\tau)$

Polarimetric Scattering Matrix

A polarimetric scattering matrix $\mathbf{\Gamma}$ is assigned to each path, confounding the propagation effects. The polarimetric scattering matrix describes two effects. First, it describes the mapping of the transmitted polarisation vector to the received polarisation vector. These 4 mappings are defined in Tab. 3.2. Second, variations of the complex signal envelope and shift of the carrier frequency due to the propagation are described. Generally, the magnitude of the matrix entries correspondence to the transmission attenuation by scattering, diffraction or penetration [HTK06]. Assuming the propagation effects as linear, the scattering matrix entries can be modelled as LTI systems [Mor94, Tre01]. The polarimetric scattering matrix $\mathbf{\Gamma}(t)$ is given by

$$\mathbf{\Gamma}(t) = \begin{bmatrix} \gamma_{\varphi\varphi}(t) & \gamma_{\varphi\vartheta}(t) \\ \gamma_{\vartheta\varphi}(t) & \gamma_{\vartheta\vartheta}(t) \end{bmatrix} \quad . \tag{3.19}$$

Model for the Spatial Domain

According to the ray-based modelling, propagation paths can be modelled as infinite small and discrete events in the spatial/angular domain. Hence, the DoA $\mathbf{\kappa}_{\mathrm{Rx}_p}$ and the DoD $\mathbf{\kappa}_{\mathrm{Tx}_p}$ of the p-th propagation path can be modelled in terms of shifted Dirac impulses δ. The joint model for the spatial propagation at Tx and Rx side is

$$h\left(\mathbf{\kappa}_{\mathrm{Tx}}, \mathbf{\kappa}_{\mathrm{Rx}}\right) = \delta\left(\mathbf{\kappa}_{\mathrm{Tx}} - \mathbf{\kappa}_{\mathrm{Tx}_p}\right) \cdot \delta\left(\mathbf{\kappa}_{\mathrm{Rx}} - \mathbf{\kappa}_{\mathrm{Rx}_p}\right) \quad . \tag{3.20}$$

The normalised direction vectors $\mathbf{\kappa}_{\mathrm{Tx}}\left(\varphi_{\mathrm{Tx}}, \vartheta_{\mathrm{Tx}}\right)$ and $\mathbf{\kappa}_{\mathrm{Rx}}\left(\varphi_{\mathrm{Rx}}, \vartheta_{\mathrm{Rx}}\right)$ are the wave vectors at Tx and Rx side, respectively.

Model for the Delay

The delay τ_p of the p-th specular propagation path is modelled by a shifted Dirac impulse in the time domain [Hay94, Mol12].

$$h(\tau) = \delta\left(\tau - \tau_p\right) \tag{3.21}$$

Model of the Polarimetric Scattering Matrix

Lets neglect polarisation for the moment and consider the passband signal $s(t)$ and baseband signal $\tilde{s}(t)$ [Hay94].

$$s(t) = \Re \left\{ \tilde{s}(t) \cdot \exp \left\{ J2\pi f_c t \right\} \right\} \tag{3.22}$$

The signal is transmitted over the propagation channel $h(t)$.

$$h(t) = a(t)*\delta(t - \tau) \tag{3.23}$$

Response $a(t)$ is an LTI system and represents the combined impulse response of multiple propagation effects. As the propagation channel is assumed as linear, the output signal at passband is [Hay94, p. 99]

$$
\begin{aligned}
s(t)*h(t) &= \Re \left\{ \tilde{a}(t)* [\tilde{s}(t)*\delta(t - \tau)] \cdot \exp \left\{ J2\pi f_c t \right\} \cdot \exp \left\{ -J2\pi f_c \tau \right\} \right\} \\
&= \Re \left\{ [\gamma(t)*\tilde{s}(t - \tau)] \cdot \exp \left\{ J2\pi f_c t \right\} \right\} \quad .
\end{aligned}
\tag{3.24}
$$

Accordingly, the delayed baseband signal $\tilde{s}(t - \tau)$ is convolved with the complex-valued *path weight function* $\gamma(t) = \tilde{a}(t) \cdot \exp \left\{ -J2\pi f_c \tau \right\}$. Each entry of the scattering matrix is made up by such a path weight function.

LINE-OF-SIGHT   If a propagation path emerges from line of sight (LOS) connection between Tx and Rx, the scattering matrix and its respective entries are given by the Friis formula [Kil15].

$$\mathbf{\Gamma}(f) = \sqrt{\frac{c_0^2}{(4\pi)^2 \cdot f^2 \cdot R^2}} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \exp \left\{ -J2\pi f_c \, R / c_0 \right\} \tag{3.25}$$

Because no scattering occurs, the polarisation vector is not varied and the scattering matrix becomes a diagonal matrix.

SINGLE-BOUNCE SCATTERING   Single-bounce scattering are the majorly considered propagation effect in the radar scattering channel. For a single bounce scattering as depicted in Fig. 3.8, the bistatic radar equation can be utilised to calculate $\tilde{a}(t)$ [Sko90, Wil05]

$$\tilde{a}(t) \quad \circ\!\!\!-\!\!\!\bullet \quad \tilde{A}(f) = \sqrt{\frac{c_0^2}{(4\pi)^3 \cdot f^2 \cdot R_{\text{Ix}}^2 \cdot R_{\text{Sx}}^2}} \cdot \xi(f) \quad , \tag{3.26}$$

with $\xi(f)$ the *reflectivity* of the scatterer. In general, the reflectivity depends on the frequency, polarisation, and angles of incident (Ix) and scattering (Sx) [39, 43]. Figure 3.9 shows the respective geometry of single bounce scattering. Hence, the reflectivity $\xi(f)$ of the scatterer is additional a function of $\varphi_{\text{Ix}}, \vartheta_{\text{Ix}}, \varphi_{\text{Sx}}, \vartheta_{\text{Sx}}, q_{\text{Ix}}, q_{\text{Sx}}$. For sake of convenience, the dependency on the angles of incidence $(\varphi_{\text{Ix}}, \vartheta_{\text{Ix}})$ and the angles of scattering $(\varphi_{\text{Sx}}, \vartheta_{\text{Sx}})$ will be dropped subsequently. Considering all polarisation combinations of q_{Ix} and q_{Sx}, and concatenating them
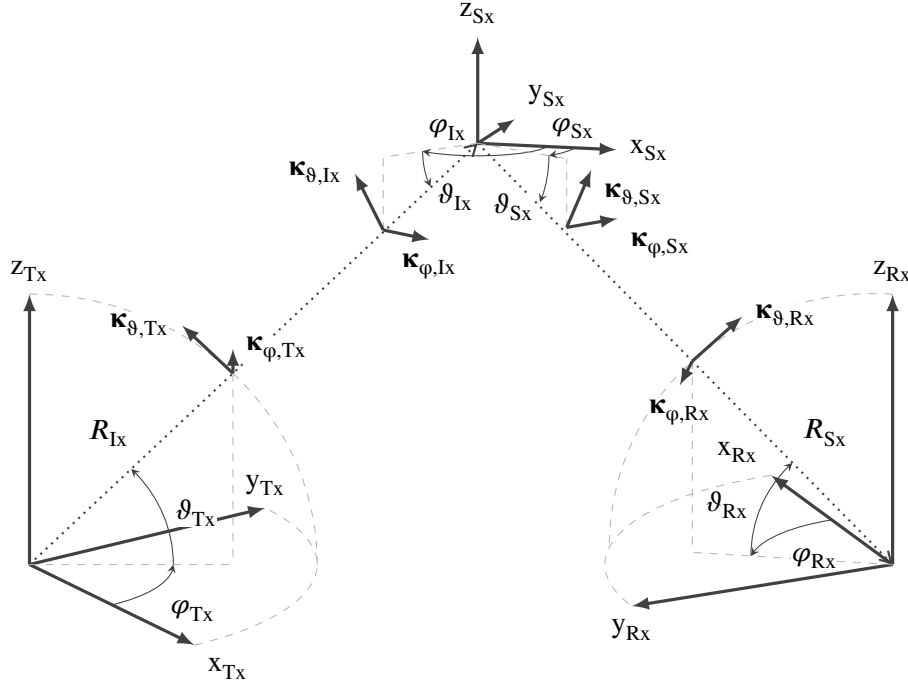
Figure 3.9: *Basic geometry of single-bounce scattering and the respective geometrical parameters. Picture taken from [43].*

in the matrix $\boldsymbol{\Xi}(f)$, the polarimetric scattering matrix becomes

$$\boldsymbol{\Gamma}(f) = \sqrt{\frac{c_0{}^2}{(4\pi)^3 \cdot f^2 \cdot R_{\text{Ix}}{}^2 \cdot R_{\text{Sx}}{}^2}} \cdot \boldsymbol{\Xi}(f) \cdot \exp\left\{-\jmath 2\pi f \left(R_{\text{Ix}} + R_{\text{Sx}}\right)\middle/ c_0\right\} \quad . \tag{3.27}$$

Frequency Independence So far, the matrix entries, i.e. the path weight functions $\gamma(t)$, are described as LTI systems. Assuming the path weight functions as a LTI system is slightly difficult, especially in the parameter estimation scheme. The path weight functions cannot be easily assumed as varying over the whole frequency band, because this may result in ambiguous or degraded estimation of the other parameters, or the parameter estimation problem will become under determined (more unknowns than measurement samples). Therefore, a model, which describes the path weight functions or their respective LTI system model in terms of a few parameters, is necessary. To the best of the authors knowledge, a general model for the path weight functions is not known from literature. Therefore, simplifications have to be stated.

Under narrowband assumption ($W \ll f_{\text{c}}$) the reflectivity and the free space attenuation can be assumed as frequency independent. The frequency independent reflectivity $\xi = \xi\left(f_{\text{c}}\right)$ will be denoted as *radar cross section (RCS)* [Sko90], which is a magnitude value here and not a power value. Accordingly, the scattering matrix can be approximated.

$$\boldsymbol{\Gamma}(f) \approx \boldsymbol{\Gamma}\left(f_{\text{c}}\right) = \boldsymbol{\Gamma} = \sqrt{\frac{c_0{}^2}{(4\pi)^3 \cdot f_{\text{c}}{}^2 \cdot R_{\text{Ix}}{}^2 \cdot R_{\text{Sx}}{}^2}} \cdot \boldsymbol{\Xi}\left(f_{\text{c}}\right) \cdot \exp\left\{-\jmath 2\pi f_{\text{c}} \left(R_{\text{Ix}} + R_{\text{Sx}}\right)\middle/ c_0\right\}$$
$$\tag{3.28}$$

Because the scattering matrix becomes a constant, the convolution in equation (3.24) simplifies to a product. Hence, the complex envelope of the transmit signal is not varied and undergoes amplitude and phase variations only [Tre01, p. 235]. The entries of the scattering matrix will be subsequently denoted as *path weight*.

Doppler Shift – Derivation and Modelling

So far, a time-invariant propagation channel has been assumed. However, if the observer or a scatterer is moving, the geometrical path parameters and the polarimetric scattering matrix may vary over time. In the simplest case, the scattering matrix entries and the angles stay fix and the delay (respectively the propagation path length) varies. This path length variation due to relative movements of the scatterer and the observer system can be approximately described as a frequency shift: the *Doppler shift*.

First, consider a static (non-moving) observation system, which transmits the signal $s(t)$. Furthermore, a static scatterer is considered and single bounce scattering at him is assumed. In case of a single interaction of the signal with the scatterer, a delayed version of the transmit signal is captured at the receiver: $s\left(t - \tau_0\right)$. The delay τ_0 is related to the length R_0 of the propagation path.

$$R_0 = c_0 \cdot \tau_0 \tag{3.29}$$

Next, consider a scatterer, which is relatively moving w.r.t. the Tx and/or Rx of the observer system. Furthermore, it is assumed, that the respective relative velocity is much smaller than the propagation velocity of the EMW, such that relativistic effects are negligible. Then, the path length becomes a function of time t.

$$r(t) = c_0 \cdot \tau(t) = R_{\mathrm{Ix}}(t) + R_{\mathrm{Sx}}(t) \tag{3.30}$$

The path length is determined by the position $\mathbf{d}_{\mathrm{Tx}}(t)$ and $\mathbf{d}_{\mathrm{Rx}}(t)$ of the scatterer w.r.t. the Tx and Rx, respectively: $R_{\mathrm{Ix}}(t) = \left\|\mathbf{d}_{\mathrm{Tx}}(t)\right\|$ and $R_{\mathrm{Sx}}(t) = \left\|\mathbf{d}_{\mathrm{Rx}}(t)\right\|$. Variations of the path length can be expressed by the time derivative [Sko90].

$$\frac{dr(t)}{dt} = c_0 \frac{d\tau(t)}{dt} = \frac{dR_{\mathrm{Ix}}(t)}{dt} + \frac{dR_{\mathrm{Sx}}(t)}{dt} \tag{3.31}$$

Now assume a monostatic observer system: $\mathbf{d}_{\mathrm{Tx}}(t) \equiv \mathbf{d}_{\mathrm{Rx}}(t)$ and $R_{\mathrm{Ix}} = R_{\mathrm{Sx}}$. Furthermore, a non-moving observer and a homogeneously moving scatterer will be assumed. Therefore, the path length variation is constant, i.e. no acceleration is present.

$$v_0 = \frac{dr(t)}{dt} \tag{3.32}$$

Variable v_0 denotes the rate of change of the path length over time, i.e. twice the relative velocity of the scatterer w.r.t. the monostatic observer. A negative v_0 accounts for an approaching scatterer and a positive v_0 accounts for an departing scatterer. The time-dependent path length
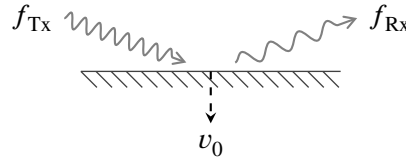
Figure 3.10: *Doppler effect for a monochromatic plane wave and relatively moving scatterer. If the scatterer approaches the observer ($v_0 < 0$), then $f_{Tx} < f_{Rx}$. If the scatterer departs from the observer ($v_0 > 0$), then $f_{Tx} > f_{Rx}$. (Picture adapted from [Fue09].)*

can be expressed as

$$r(t) = R_0 + v_0 \cdot t \quad . \tag{3.33}$$

Plugging into the transmit signal yields [CB93]

$$s\left(t - \tau(t)\right) = s\left(t - \frac{r(t)}{c_0}\right) = s\left(a \cdot t - \frac{R_0}{c_0}\right) = s\left(a \cdot t - \tau_0\right) \quad , \tag{3.34}$$

with $a = 1 + v_0/c_0$. Hence, the movement of the scatterer causes a temporal scaling (extension or compression) of the transmit signal, which is called *Doppler effect*. Because the signal energy is not allowed to change according to Parseval's Theorem, the scaling factor $\sqrt{a}$ is introduced [CB93]. Hence, the observed signal due to scatterer movement is $\sqrt{a} \cdot s(at)$, a time and amplitude scaled version of the transmit signal.

Figure 3.10 sketches the influence of the Doppler on a monochromatic signal. The Doppler causes a varying frequency of the received signal compared to the transmitted one.

Approximation of the Doppler Effect − The Doppler Shift Commonly, the Doppler effect is approximated, such that the temporal scaling becomes a constant frequency shift over the whole frequency band. This is only approximately true, because the Doppler effect results in a frequency dependent shift, and therefore is not constant over the whole frequency band. Consider the Fourier transformation of the observed signal, the frequency axis is scaled by the Doppler effect too.

$$\sqrt{a} \cdot s(at) \quad \circ\!\!-\!\!\bullet \quad \frac{1}{\sqrt{a}} \cdot S\left(\frac{f}{a}\right) \tag{3.35}$$

Take the first order Taylor series approximation of the frequency scaling for $|v_0| \ll c_0$.

$$\frac{f}{a} = \frac{f}{1 + v_0/c_0} \approx f \cdot \left(1 - \frac{v_0}{c_0}\right) = f - f\frac{v_0}{c_0} \tag{3.36}$$

Next, the narrowband conditions are assumed. The *frequency dependent* Doppler shift becomes a *bandwidth independent* Doppler shift v, which will be related to the carrier frequency f_c.

$$f\frac{v_0}{c_0} \approx f_c\frac{v_0}{c_0} \tag{3.37}$$

Summarised, the scaling of the frequency axis simplifies to a frequency shift [CB93].

$$\frac{f}{1 + v_0/c_0} \approx f - f_c \frac{v_0}{c_0} \qquad (3.38)$$

For sake of convenience, the Doppler shift at centre frequency is introduced: $v = f_c \, v_0/c_0$. It is worthwhile to note, that this approximation is the first order Taylor series expansion at the points $f = f_c$ and $v_0/c_0 = 0$. Plugging into the spectral representation of the transmit signal and taking $\sqrt{a} \approx 1$ into account yields

$$\frac{1}{\sqrt{a}} \cdot S\left(\frac{f}{a}\right) \approx S(f - v) = S(f) \cdot \delta(f - v) \qquad . \qquad (3.39)$$

The respective simplification in the time domain is

$$\sqrt{a} \cdot s\,(at) \approx s\,(t) \cdot \exp\{j2\pi vt\} \qquad . \qquad (3.40)$$

Summarised, the scaling of the transmit signal due to the Doppler effect can be approximated by the multiplication with a harmonic exponential in time domain. The respective propagation model for the p-th specular propagation path in time domain is

$$h\,(t) = \exp\{j2\pi v_p t\} \quad \circ\!\!-\!\!\bullet \quad H\,(v) = \delta\left(v - v_p\right) \qquad , \qquad (3.41)$$

with v_p the Doppler shift of the propagation path.

The narrowband approximation is valid, if the largest and smallest Doppler shift in the frequency band are approximately equal to the Doppler shift at carrier frequency.

$$\frac{v_0}{c_0} \cdot f_{\max} \approx \frac{v_0}{c_0} \cdot f_{\min} \approx \frac{v_0}{c_0} \cdot f_c \qquad (3.42)$$

Frequencies $f_{\max}$ and $f_{\min}$ are the upper and lower frequency of the transmit signal at passband, respectively. This approximation is valid, as long as the variation of the Doppler shift over the entire frequency band Δv is smaller than the frequency spacing $\Delta f = 1/T$. Variable T denotes the observation time.

$$\Delta v \ll \Delta f \qquad (3.43)$$

The variation of the Doppler shift over the entire frequency band is given by

$$\Delta v = \frac{v_0}{c_0} \cdot \left(f_{\max} - f_{\min}\right) = \frac{v_0}{c_0} \cdot W \qquad . \qquad (3.44)$$

Finally, the following criterion can be derived to validate the narrowband approximation [CB93].

$$\Delta v = \frac{v_0}{c_0} \cdot W \ll \Delta f = \frac{1}{T} \quad \rightarrow \quad T \cdot W \ll \frac{c_0}{v_0} \qquad (3.45)$$

Product $T \cdot W$ is the time-bandwidth product. Summarised, the narrowband approximation of

the Doppler shift is valid, if i) the time-bandwidth product is small $(T \cdot W \ll 1)$ and/or ii) the relative velocity is much smaller than the wave propagation velocity $\left(\left| v_0 \right| \ll c_0 \right)$.

The Multidimensional Path Function

Collecting the model for the delay domain, the spatial domain, the Doppler domain and the scattering matrix together yields the model of a specular propagation path: the *multidimensional path function* $\mathbf{H}_p$. The multidimensional path function in the complex baseband is given as the product of the Dirac functions in delay, spatial and Doppler domain, which is weighted by the scattering matrix [Ric05, TLS+04].

$$\mathbf{H}_p \left(\tau, \nu, \boldsymbol{\kappa}_{\mathrm{Tx}}, \boldsymbol{\kappa}_{\mathrm{Rx}} \right) = \boldsymbol{\Gamma}_p \cdot \delta \left(\tau - \tau_p \right) \cdot \delta \left(\boldsymbol{\kappa}_{\mathrm{Tx}} - \boldsymbol{\kappa}_{\mathrm{Tx}p} \right) \cdot \delta \left(\boldsymbol{\kappa}_{\mathrm{Rx}} - \boldsymbol{\kappa}_{\mathrm{Rx}p} \right) \cdot \delta \left(\nu - \nu_p \right) \quad (3.46)$$

The parameters of the propagation path p are collected in the vector $\boldsymbol{\varpi}_p$.

$$\boldsymbol{\varpi}_p = \begin{bmatrix} \varphi_{\mathrm{Tx}p} & \vartheta_{\mathrm{Tx}p} & \varphi_{\mathrm{Rx}p} & \vartheta_{\mathrm{Rx}p} & \tau_p & \nu_p & \Re\left\{ \mathrm{vec}\left\{ \boldsymbol{\Gamma}_p \right\} \right\}^{\mathrm{T}} & \Im\left\{ \mathrm{vec}\left\{ \boldsymbol{\Gamma}_p \right\} \right\}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \quad (3.47)$$

3.2.2 Multipath Propagation Channel Model

The multidimensional path function (3.46) describes the propagation of a single path. Commonly, the propagation channel constitutes of multiple propagation paths, arising due to e.g. multiple targets, multiple scatterings per target (extended target) or clutter. Because the propagation channel is assumed as linear, multipath propagation is modelled as the superposition of a finite number P of propagation paths [Czi07, Lan08, Ric05], resulting in the *multidimensional propagation channel model* [DEH19].

$$\mathbf{H} \left(\tau, \nu, \boldsymbol{\kappa}_{\mathrm{Tx}}, \boldsymbol{\kappa}_{\mathrm{Rx}} \right) = \sum_{p=1}^{P} \mathbf{H}_p \left(\tau, \nu, \boldsymbol{\kappa}_{\mathrm{Tx}}, \boldsymbol{\kappa}_{\mathrm{Rx}} \right) \quad (3.48)$$

For sake of convenience, the parameters of all propagation paths are collected into a single vector $\boldsymbol{\varpi}$.

$$\boldsymbol{\varpi} = \begin{bmatrix} \boldsymbol{\varpi}_1^{\mathrm{T}}, & \dots & , \boldsymbol{\varpi}_P^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \quad (3.49)$$

3.3 Summary

In this chapter, a model for the physical system under identification has been proposed. The considered physical system is the radar scattering channel, which is made up of propagating EMWs and scattering radar targets or clutter.

The propagation of EMWs has been described by ray optics or geometrical optics. The wave is treated as a ray, whose propagation direction coincides with the ray's pointing direction. Ray optic enables the assignment of specular propagation paths to scattering point targets at a radar target or clutter. Furthermore, geometrical parameters as e.g. directions in terms of angles or a

delay can be assigned to each specular propagation path, which are related to the radar (point-) targets and their spatial position. According to the geometrical optics a single propagation path is modelled by a multidimensional Dirac function. Because the radar scattering channel is made up of multiple point targets and clutter sources, multipath propagation is present. The respective model (3.48) represents multipath propagation as superimposed multidimensional Dirac functions. As pointed out in Section 1.1.1, a model of superimposed Dirac functions enables high-resolution of each contribution, here high-resolution estimation of propagation paths in the spatial, delay and Doppler domain.

CHAPTER 4

DEVICE MODEL – MODELLING OF THE RADAR SYSTEM

$\mathbf{S}$ OLVING the inverse problem necessitates a model $\mathcal{M}^D$ of the device, as discussed in Section 2.1. The device model describes i) the transmit to receive transfer characteristic of the observer, which will be denoted as transceiver response; and ii) models the received signal in terms of this transceiver response. Here, an observer equipped with antenna arrays at Tx and Rx side will be considered, in order to resolve directions of impinging or transmitted waves. Therefore, a parametric model of the antenna array in terms of the directions is required as well. Proper selection of the device model is highly important for the solution of the inverse problem and hence the quality of the parameter estimates, especially if high-resolution shall be achieved. In [LKT12] the impact of inaccurate device models on the parameter estimation quality, especially the correct antenna array model, has been discussed. As it has been shown, an inaccurate model will cause an increased estimation variance, bias or even false estimates. Hence, the device model has to be carefully assigned. Furthermore, the noise model $\mathcal{M}^N$ is required as well. The noise model describes the noise processes of the observer system and their influence on the observations. This results in a model on the distribution of the observations (the likelihood), which is required to solve the stochastic inverse problem by Bayes' theorem (see Section 2.3). In this chapter the device model for an exemplarily FMCW MIMO radar will be successively derived. A novel noise model for a co-located FMCW MIMO radar will be presented too.

Related Work and Own Contributions

There are many publications, which focus on the topic of modelling and calibrating sensor devices and multi-sensor devices, i.e. devices with arrays of sensors, see e.g. [LKT12, Suh+17, Vas+20, Whi+91]. However, each of them are somehow restricted to a certain device or device class, because the model and its parametrisation highly depends on the considered device itself. Especially the device bandwidth, properties of the hardware components or the size of the spatial aperture etc. determine the device model. Hence, literature-based device models cannot be applied out-of-the-box. As pointed out, a proper device model is required, which covers all aspects of the device sensing. The device model and its derivation will be divided into four parts:

- Model of the transceiver response

- Model of the antenna array

- Model of the sampling process

- Model of the noise processes.

Accordingly, modelling of a device spans a wide field of topics, because many different physical properties and aspects have to be accounted for. This fact in conjunction with the dependence of the device model on the device itself makes it hard to briefly describe the related work. Subsequently, the mentioned models will be derived for the FMCW radar presented in [2], which is exemplary considered here. A brief discussion on related work will be given in the respective section.

The basic principle and signal theory of a single input single output (SISO) FMCW radar will be reviewed and model for the baseband signal is derived in Section 4.1. Hardware impairments and noise will be neglected. Thereafter, novel transceiver response models for an FMCW radar are described in Section 4.2. A linear, dynamic radar device model will be developed first, which is later on refined to a non-linear, dynamic device model in order to account for non-linearities in the signal generation block of the hardware. Next, modelling of antennas and antenna arrays is discussed in Section 4.3. Two modelling approaches will be distinguished: *narrowband* and *wideband*. Necessary constraints in order to apply the presented modelling approaches are discussed. The model of the transceiver characteristic and the antenna arrays have to be parametrised from calibration measurements, which will be described in Section 4.4. If the observer is realised as an integrated on-chip antennas system, i.e. the array cannot be detached from the transceiver chain, a joint parametrisation of array and transceiver response model is necessary. A novel calibration method is proposed to accomplish this. As the inverse problem is solved using time sampled data, the sampling has to be modelled too. Especially the sampling in case of moving targets and the gathering of all MIMO channels by time-division multiplexing has to be accounted for. MIMO systems with TDM suffer from angle-Doppler coupling, which causes disturbances of the angle and Doppler estimation. A model to take into account angle-Doppler coupling in the description of the sampling will be presented. Hence, disturbances of angle and Doppler estimation can be circumvented. The sampling process will be described in Section 4.5. Furthermore, a novel method to cope with signal leakage in co-located FMCW radar systems will be presented. According to this method, a new noise source in FMCW radars has been established: leakage noise. Also, a novel model to describe the impact of phase noise of the radar hardware on the observations is derived. To the best of the authors knowledge, this is done for the first time in modelling of a FMCW MIMO radar. The modelling of the statistical characteristics of all considered noise sources will be discussed in Section 4.6. Also, the proposed noise model will be verified by measurements. Section 4.7 concludes the chapter.

The contributions in the field of modelling the response of FMCW radars have been published in [21, 5, 9]. The contributions in the area of modelling and calibration of antenna arrays for DoA estimation have been published in [7, 24, 26]. Publications in the are of array signal processing under broadband conditions are [23, 7, 25, 8]. The method to cope with transmit signal leakage in co-located FMCW radar systems has been published in [6]. The compensation of angle-Doppler coupling has been published in [10].

4.1 Signal Model of a Homodyning FMCW Radar

Understanding the functionality of a FMCW radar is fundamentally for the derivation of a device model. Therefore, a (theoretical) model of the baseband signal will be derived in this sections, which allows to understand the basics of FMCW radar, and their benefits and limits. Furthermore, the theoretical model allows to identify signal distortions, which occur in real FMCW radars.

4.1.1 Model of a FMCW Signal

A frequency modulated continuous wave (FMCW) signal is a linear frequency modulated (LFM) signal or chirp [CB93]. During the *modulation or transmit time* T_M the frequency is linearly increased (swept) from *start frequency* f_0 up to frequency $f_0 + W$, with W the *signal bandwidth*. Afterwards, the signal recovers from $f_0 + W$ back to starting frequency f_0, which takes the *recovery time* T_R. A new chirp is radiated after settling time T_W. The respective time-frequency diagram of an FMCW signal is depicted in Fig. 4.1. The signal is sampled by the analog-to-digital converter (ADC) only during the modulation time, such that T_R and T_W are idle times of the measurement capturing. Hence, the time $T_R + T_W$ is a guard interval to avoid inter symbol interference or disturbances due to antenna switching in e.g. TDM (see Section 4.5.3).

A FMCW signal during the transmit duration $0 \le t \le T_M$ can be represented in complex passband domain [CB93].

$$s(t) = \text{rect}\left(\frac{t - \frac{T_M}{2}}{T_M}\right) \cdot \exp\left\{j2\pi\left(f_0 \cdot t + \frac{W}{2T_M} \cdot t^2\right)\right\} \tag{4.1}$$

An *instantaneous frequency* $f(t)$ can be assigned to the FMCW signal [Pip93].

$$f(t) = \frac{W}{T_M} \cdot t + f_0 \tag{4.2}$$

The respective *instantaneous signal phase* $\phi(t)$ can be calculated from the instantaneous frequency $f(t)$ by integration.

$$\phi(t) = 2\pi \int_0^t f(t') \, dt' = 2\pi\left(f_0 t + \frac{W}{2T_M} t^2\right) - \phi_0 \tag{4.3}$$

Conversely, the instantaneous frequency can be calculated by differentiation of the instantaneous phase.

$$f(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} \tag{4.4}$$

For sake of convenience, the initial phase will be assumed as zero: $\phi_0 = 0$. In sensing applications the transmit signal is periodically radiated with *period time* T_P, required to capture
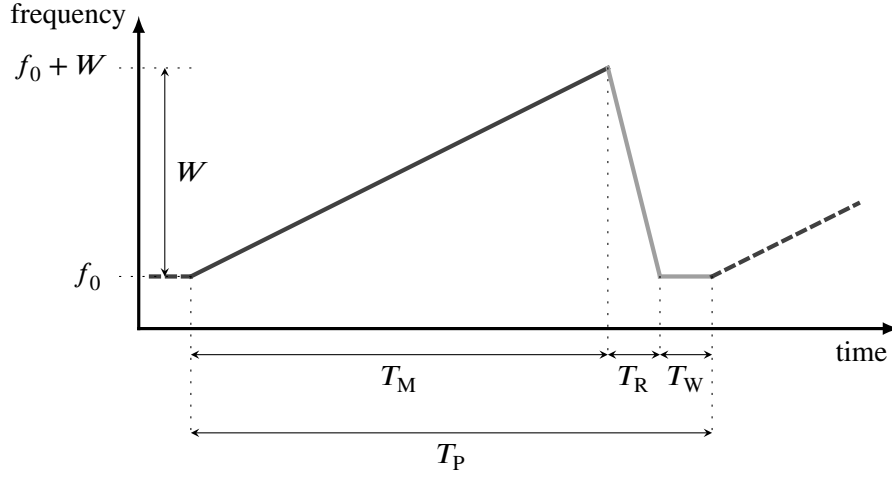
Figure 4.1: *Time-frequency diagram of a frequency modulated continuous wave (FMCW) signal, periodically radiated with $T_P = T_M + T_R + T_W$. Time $T_R + T_W$ is an idle time of the radar, while measurements are discarded or are not captured by the ADC.*

time-varying events (see Section 4.5).

$$s_P(t) = \sum_{n=-\infty}^{\infty} s\left(t - n \cdot T_P\right) \tag{4.5}$$

The signal period T_P and the modulation time T_M are related according to $T_P = T_M + T_R + T_W$.

4.1.2 Range Resolution

In radar applications the range resolution, i.e. the separability of two targets, is of significant importance. Considering a monostatic radar (see Section 4.3.4), the range resolution ΔR is

$$\Delta R = \frac{c_0}{2W} \quad . \tag{4.6}$$

In order to process the complete signal bandwidth, i.e. to have the full range resolution, the sampling frequency f_S of the ADC has to fulfil the Nyquist criterion.

$$f_S \geq W \tag{4.7}$$

If large bandwidths (e.g. several GHz) are used, appropriate ADCs are costly. Hence, *stretch processor or homodyning receivers* are commonly employed in FMCW radar systems [CB93, Ric13]. These receivers are a special realisation of matched filter receivers, suitable for processing of FMCW signals with a large time-bandwidth product [CB93].

4.1.3 Stretch Processing or Homodyning Receiver

A stretch processing [Ric13] or homodyning [CB93] receiver architecture (sometimes also denoted as de-ramping receiver) is sketched in Fig. 4.2. The echo signal, emerging from e.g.
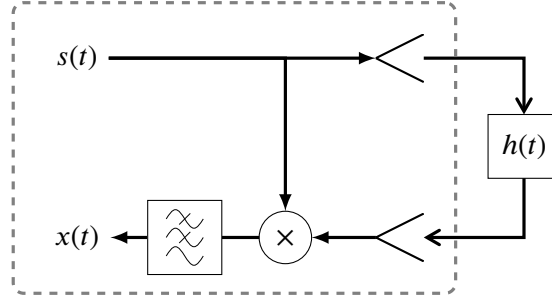
Figure 4.2: *Block diagram of a FMCW radar with stretch processor or homodyning receiver architecture. [21]*

target scattering, is down-converted to baseband by the transmit signal, using a mixer and a low-pass filter in the receiver stage. For the subsequent analysis, a single echo from a stationary target is assumed, which impinges with delay τ at the radar receiver. Neglecting attenuation due to propagation, the transfer function $h(t)$ of the propagation channel in Fig. 4.2 is

$$h(t) = \delta(t - \tau) \quad . \tag{4.8}$$

The delay causes the FMCW signal to be shifted along the time axis, see Fig. 4.3. Because the transmit signal is a LFM signal a one-to-one relationship of this delay to a frequency shift exists. Considering $s(t)$ in the complex passband (see equation (4.1)), the output signal $x(t)$ of the stretch processor is

$$x(t) = [s(t) * h(t)] \cdot s(t)^{\dagger} \quad . \tag{4.9}$$

Note, that, if real-valued signals are present, the integral due to the low-pass filter has to be taken into account. Plugging equation (4.1) into equation (4.9) yields the (ideal) baseband signal [FJ15, Ric13], while neglecting the rectangular function due to time limitation for a moment.

$$\begin{aligned} x(t) &= \exp\left\{ j2\pi \left(f_0 \cdot (t - \tau) + \frac{W}{2T_{\mathrm{M}}} \cdot (t - \tau)^2 \right) \right\} \cdot \exp\left\{ j2\pi \left(f_0 \cdot t + \frac{W}{2T_{\mathrm{M}}} \cdot t^2 \right) \right\}^{\dagger} \\ &= \exp\left\{ j2\pi \left(\frac{W}{2T_{\mathrm{M}}} \tau^2 - \frac{W}{T_{\mathrm{M}}} \tau t - f_0 \tau \right) \right\} \end{aligned} \tag{4.10}$$

Accordingly, the output $x(t)$ of the stretch processor is a harmonic [Ric13]. The frequency f_{b} of the harmonic, commonly denoted as *beat frequency*, equals the frequency shift between transmitted and received FMCW signal [Ric13].

$$f_{\mathrm{b}} = \frac{W}{T_{\mathrm{M}}} \tau \tag{4.11}$$

Typically, the *residual video phase* $W\tau^2/2T_{\mathrm{M}}$ can be neglected, if $W < f_{\mathrm{c}}$ holds [FJ15].

$$x(t) = \exp\left\{ j2\pi \left(\frac{W}{2T_{\mathrm{M}}} \tau^2 - f_{\mathrm{b}} t - f_0 \tau \right) \right\} \approx \exp\left\{ -j2\pi \left(f_{\mathrm{b}} t + f_0 \tau \right) \right\} \tag{4.12}$$
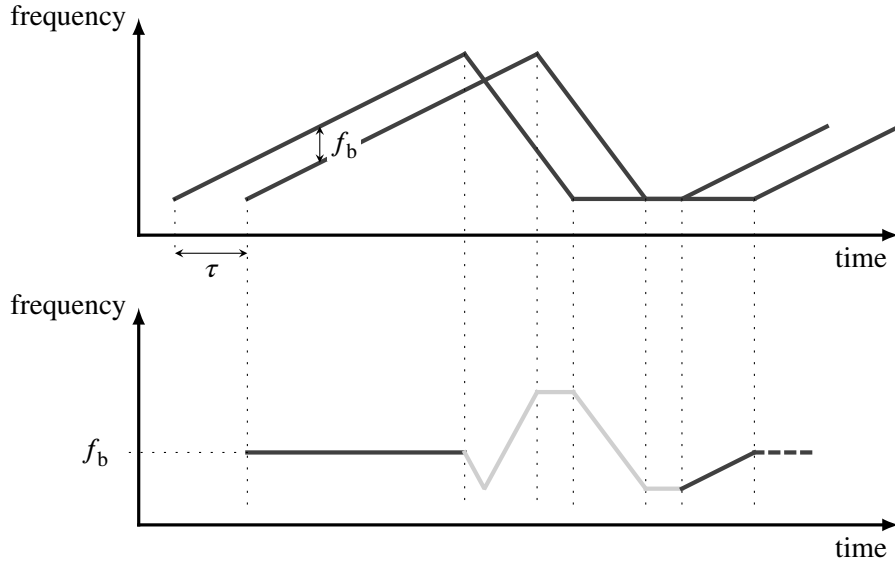
Figure 4.3: *Time-frequency diagram of a transmitted FMCW signal and the received, time shifted replica due the propagation delay τ. The baseband signal after homodyning is a harmonic signal with the beat frequency f_b. The fading curves are beat frequencies which are not sampled by the receiver.*

By spectral analysis of the baseband signal $x(t)$ the beat frequency and hence the delay can be estimated [CB93]. Summarised, the linear FMCW transmit signal in conjunction with the stretch processing receiver maps the propagation delay in a harmonic signal at baseband, whose frequency is proportional to the propagation delay.

Due to the stretch processing the received baseband signal is no longer a time shifted replica of the transmitted signal. Therefore, the stretch processor is not a classical matched filter, which is commonly used for pulse compression in radar [CB93]. If a matched filter would be employed, the baseband signal would be given by

$$x(t) = [s(t) * h(t)] * s(-t) \quad . \tag{4.13}$$

Therefore, the stretch processing does not accomplish a classical pulse compression, which would result in a pulse in time domain. It can be stated, that the pulse compression is accomplished by the Fourier transformation of the baseband signal after homodyning, which yields a peak in the spectral domain at the beat frequency location.

Advantages

The main benefit of the stretch processing is, that the ADC samples the beat frequency signal rather than the time-shifted transmit signal. A much lower ADC sampling frequency f_S becomes sufficient and therefore simpler and cheaper ADCs are usable. Because the beat frequency is directly related to the propagation delay, the largest ascertainable delay τ_{max} is

$$\tau_{\mathrm{max}} = \frac{T_{\mathrm{M}}}{W} \cdot \frac{f_{\mathrm{S}}}{2} \quad . \tag{4.14}$$

Disadvantages

Drawback of the stretch processing is the correlation without a delay line, resulting in a degradation of the range resolution [Pip95]. Reconsider the time limitation of the transmit signal, the baseband signal $x(t)$ is (see APPENDIX C.1)

$$x(t) \approx \mathrm{rect}\left(\frac{t - \frac{T_{\mathrm{M}}+\tau}{2}}{T_{\mathrm{M}} - \tau}\right) \cdot \exp\left\{j2\pi\left(\frac{W}{2T_{\mathrm{M}}}\tau^2 - \frac{W}{T_{\mathrm{M}}}\tau t - f_0\tau\right)\right\} \quad . \tag{4.15}$$

Hence, the beat frequency is observed over $T_{\mathrm{M}} - \tau$ only. The respective spectral density $|X(f)|^2$ of the baseband signal is

$$|X(f)|^2 \propto \left|\mathrm{sinc}\left(\left[T_{\mathrm{M}} - \tau\right] \cdot f\right) * \delta\left(f + f_{\mathrm{b}}\right)\right|^2 \quad . \tag{4.16}$$

The main lobe of the sinc-function is broadened compared to a full observation time of T_{M}, indicating a loss of resolution. Due to the resolution loss an effectively processed bandwidth can be stated [Pip95].

$$W' = \frac{W}{T_{\mathrm{M}}} \cdot (T_{\mathrm{M}} - \tau) = W - W \cdot \frac{\tau}{T_{\mathrm{M}}} = W \cdot \left(1 - \frac{\tau}{T_{\mathrm{M}}}\right) \tag{4.17}$$

Consequently, the spectral resolution and therefore the range resolution is degraded [Pip95].

$$\Delta R' = \frac{c_0}{2W'} = \frac{c_0}{2W \cdot \left(1 - \tau/T_{\mathrm{M}}\right)} = \frac{c_0}{2W} \cdot \frac{T_{\mathrm{M}}}{T_{\mathrm{M}} - \tau} \tag{4.18}$$

Considering the maximal observable beat frequency, which is given by $f_{\mathrm{S}}/2$, the max. range resolution loss can be given [FJ15].

$$\frac{T_{\mathrm{M}}}{T_{\mathrm{M}} - \tau_{\max}} - 1 = \frac{\tau_{\max}}{T_{\mathrm{M}} - \tau_{\max}} = \frac{1}{T_{\mathrm{M}}/\tau_{\max} - 1} = \frac{1}{2W/f_{\mathrm{S}} - 1} \tag{4.19}$$

The maximal range resolution loss is determined by the ratio of the sampling frequency f_{S} and the signal bandwidth W. Hence, by properly choosing the signal parameters and the sampling frequency, the resolution loss can be minimised.

Subsequently it will be assumed, that the modulation time is much larger than the delay: $T_{\mathrm{M}} \gg \tau$. Hence, the resolution loss is negligible and the range resolution is approximately given by the whole signal bandwidth.

4.1.4 Doppler Effect and Doppler Shift

Consider the Doppler effect formula (3.34) and the FMCW signal model (4.1), the signal at the receiver is

$$\sqrt{a} \cdot s\left(at - \tau_0\right) = \sqrt{a} \cdot \exp\left\{j2\pi\left[f_0 \cdot \left(at - \tau_0\right) + \frac{W}{2T_{\mathrm{M}}} \cdot \left(at - \tau_0\right)^2\right]\right\} \quad . \tag{4.20}$$

The respective baseband signal after homodyning is

$$
\begin{aligned}
x\left(t\right) &= \sqrt{a} \cdot s\left(t\right)^{\dagger} \cdot s\left(at - \tau_0\right) \\
&= \sqrt{a} \cdot \exp\left\{ j2\pi \left[\left(a - 1\right) f_0 t + \frac{W}{2T_{\mathrm{M}}} \left(a^2 - 1\right) t^2 + \frac{W}{2T_{\mathrm{M}}} \tau_0^2 - \frac{W}{T_{\mathrm{M}}} a\tau_0 t - f_0\tau_0 \right] \right\} \quad .
\end{aligned}
$$
(4.21)

Accordingly, the signal remains a chirp and is not a simple harmonic.

As pointed out in Section 3.2.1 the Doppler effect can be approximated, if $\left|v_0\right| \ll c_0$ holds and the narrowband assumption can be drawn. Hence, the Doppler induced scaling of the signal becomes negligible. The baseband signal for a single period is approximately (see APPENDIX C.3 for a detailed derivation)

$$
x\left(t\right) \approx \exp\left\{ -j2\pi \left(vt + \frac{W}{T_{\mathrm{M}}} \tau_0 t + f_0\tau_0 \right) \right\} \quad ,
$$
(4.22)

with $v = f_0\, v_0 \big/ c_0$ the Doppler shift.

4.2 Transceiver Model of a FMCW Radar

Real observers suffer from hardware impairments, which cause amplitude and phase distortions of the transmitted and received signal. Such hardware impairments and their influence have not been considered for the baseband signal derived in Section 4.1. They have to be accounted for by the device model $\mathcal{M}^{\mathrm{D}}$, which describes the transceiver (joint transmitter & receiver) characteristic of the device. In the following, the FMCW radar chip presented in [Hit+17] will be exemplary considered.

Modelling approaches for FMCW radars are mainly towards modulation errors of the frequency sweep as e.g. non-linear sweeps [Avi+12, Ayh+16, MHL07], variations in the sweep rate [Gro17] or sinusoidal distortions of the sweep [Pip95]. Such modulation errors cause degradation of the range resolution. However, modelling of the transceiver characteristic, which also degrades the range resolution, is less considered in literature. Description of the transfer characteristic of the radar can be model-based or not. In model-based description, which will be considered here, two principles can be distinguished: i) physical and ii) behavioural modelling [Gri14, 5, Sch09]. In physical modelling the device is modelled by the microscopic description of each hardware component and their interaction. Such modelling approaches are conceivable for individual hardware components or slightly complicated systems only, because the modelling effort scales with the device complexity. Behavioural modelling is a macroscopic description of the influence of the device characteristic on the transmit and/or receive signal. Hence, behavioural modelling is a black-box modelling useful for complicated devices, because hardware details are not accounted for or not of interest.

A separate measurement of all radar components is not possible due to the MMIC setup of the considered radar chip. Furthermore, the influence of the radar device on the received signal is of interest, rather than a detailed description of hardware related issues. Hence, a behavioural model for the transceiver characteristic of the radar system will be developed. The behavioural model will be derived by i) identifying the best suited model structure and ii) estimating the
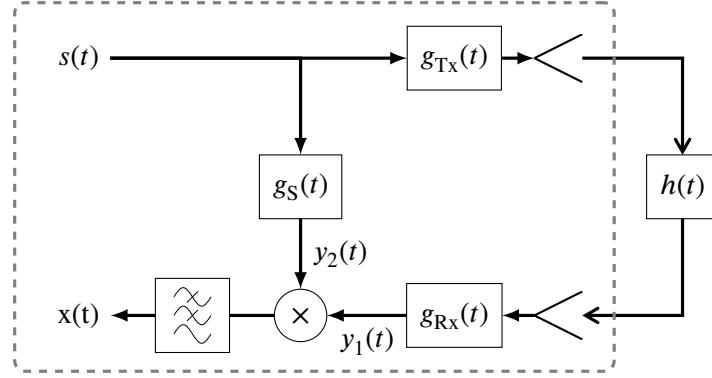
Figure 4.4: *Model of a real FMCW radar with LTI systems to account for respective hardware impairments. [21]*

corresponding model parameters from calibration measurements. Assuming the radar as time-invariant, the device model has to be structurally determined and parametrised from calibration measurements only once. First, a linear device model will be developed in Section 4.2.1. As this model is not sufficient to describe the response of the considered radar, a novel non-linear device model will be developed in Section 4.2.2. Both models attempt to describe the joint influence of the hardware impairments of the transmit and receive chain on the received signal. The developed device models have been published in [21] and [5], respectively.

4.2.1 Linear Device Model

May the hardware impairments of the radar be modelled as linear time-invariant systems [21]. Then, the radar model in Fig. 4.2 will be extended by incorporating LTI systems in the transmit and receive path to account for hardware impairments. The respectively advanced radar model is shown in Fig. 4.4. The following signals and system responses are introduced, which will be described by their respective passband representation.

- Transmit signal: $s(t) = U \exp\{j\phi(t)\}$ (amplitude $U \in \mathbb{R}$, phase $\phi(t)$ according to equation (4.3))

- LTI system in transmit path: $g_{\text{Tx}}(t) \circ\!\!-\!\!\bullet\ G_{\text{Tx}}(f)$

- LTI system in receive path: $g_{\text{Rx}}(t) \circ\!\!-\!\!\bullet\ G_{\text{Rx}}(f)$

- Propagation channel response: $h(t) \circ\!\!-\!\!\bullet\ H(f)$

- LTI system in the feedback path: $g_{\text{S}}(t) \circ\!\!-\!\!\bullet\ G_{\text{S}}(f)$

Lets introduce the respective input signals to the mixer of the stretch processing receiver.

$$y_1(t) = s(t) * g_{\text{Tx}}(t) * h(t) * g_{\text{Tx}}(t) \tag{4.23a}$$
$$y_2(t) = s(t) * g_{\text{S}}(t) \tag{4.23b}$$

Analytic signals are considered in the following, such that the low-pass filter in Fig. 4.4 can be neglected [Ric13]. The output signal of the stretch processor is

$$x(t) = y_1(t) \cdot y_2(t)^\dagger = \left[s(t) * g_{\mathrm{Tx}}(t) * h(t) * g_{\mathrm{Rx}}(t) \right] \cdot \left[s(t) * g_{\mathrm{S}}(t) \right]^\dagger \quad . \tag{4.24}$$

In order to proceed the following theorem is used [9].

Theorem 1. *Transmitting an FMCW signal over an LTI system and the FMCW signal fulfils* $T_M \cdot W \gg 1$, *the output signal is an FMCW signal whose amplitude and phase is varied according to the frequency response of the LTI system.*

Accordingly, the input signals to the mixer of the stretch processor are

$$y_1(t) \approx U \exp\{j\phi(t)\} \cdot G_{\mathrm{Tx}}(f(t)) \cdot H(f(t)) \cdot G_{\mathrm{Rx}}(f(t)) \tag{4.25a}$$

$$y_2(t) \approx U \exp\{j\phi(t)\} \cdot G_{\mathrm{S}}(f(t)) \quad , \tag{4.25b}$$

with $f(t)$ the instantaneous frequency (4.2). Plugged into the stretch processor model (4.24) the output signal of the stretch processor is

$$x(t) \approx U^2 \cdot H(f(t)) \cdot G_{\mathrm{Tx}}(f(t)) \cdot G_{\mathrm{Rx}}(f(t)) \cdot G_{\mathrm{S}}(f(t)) = U^2 \cdot H(f(t)) \cdot G(f(t)) \quad . \tag{4.26}$$

Summarised, the signal after the stretch processor is a multiplication of the system transceiver response $G(f(t))$ and the propagation channel response $H(f(t))$ [21]. Hence, the radar observations can be calibrated by deconvolution with the system transceiver response.

Model Calibration and Verification

The derived system model describes the transceiver response of the radar by the LTI system $G(f(t))$. This LTI system will be deduced from calibration measurements. The basic settings of the radar for the calibration measurements are summarised in Tab. 4.1. Because the radar is realised as a MMIC, back-to-back of connection the transmit and receive chain over a known reference standard cannot be accomplished. Instead, over-the-air calibration has to be applied, with the propagation channel as reference standard. Note, that in an over-the-air calibration setup, the response of the Tx and Rx antenna are included in the calibration measurement too. Because Tx and Rx are on the same MMIC (monostatic radar), transmitter and receiver cannot be placed in LOS connection. Therefore, a corner reflector (i.e. a radar target) is employed for the calibration measurements. The corner reflector is placed at a known distance R apart from the radar and has a known RCS. The respective propagation channel response $H(f)$, which is now the reference standard, is given by the radar equation [Sko90].

$$H(f) = \sqrt{\frac{c^2}{(4\pi)^3 f^2 R^4}} \cdot \xi \cdot \exp\left\{-j 2\pi f\, 2R / c_0\right\} \tag{4.27}$$

For the calibration and verification measurements, two corner reflectors with RCS of $\xi = 0\,\mathrm{dBsm}$ and $\xi = -10\,\mathrm{dBsm}$ were separately deployed at a distance of $R = 1.6\,\mathrm{m}$ and $R = 0.8\,\mathrm{m}$,

Table 4.1: *Basic settings of the radar device for the calibration measurements.*

Parameter	Setting	Parameter	Setting
f_0	147 GHz	W	20 GHz
T_{M}	1024 µs	f_{S}	4 MHz

respectively.

Because the radar is monostatic strong leakage between the Tx and Rx channel occurs, see Fig. 4.5(a). High-pass filtering is applied to suppress the corresponding beat frequencies. Afterwards, because the I-channel of the radar has been captured only and the baseband signal is therefore real-valued, the analytic signal is calculated. The analytic time domain signal is depicted in Fig. 4.5(b). Note, that the time scale was mapped to frequency by $f = W/T_{\mathrm{M}} \cdot t + f_0$. The range spectra for both deployed corner reflectors after pre-processing are shown in Fig. 4.6. The highest beat frequency peaks can be assigned to the ranges of the corner reflectors. Furthermore, additional beat frequency peaks are visible, indicating the presence of more scatterers. However, these beat frequencies cannot be assigned to a scatterer of the surrounding environment. Furthermore, their location varies, if the target distance is varied. This, however, is crucial, because the surrounding environment has not changed. It can be concluded, that these beat frequencies are not due to scattering. Instead, they are parasitic signal components due to the radar hardware and will be termed as *ghost targets*. Ghost targets will be defined as beat frequencies, which occur in conjunction with the beat frequency of a true target echo. It has been found, that the ghost targets are due to hardware impairments, which are not covered by the derived LTI device model [21]. In order to apply the derived system model, bandwidth limitation is conducted by reducing the frequency range to 155 GHz to 161 GHz. The range spectra after band limitation are shown in Fig. 4.6, indicating that the ghost targets vanished.

Summarised, the derived LTI device model does not properly describe the radar system. Ghost targets are present in the received signal, whose occurrence cannot be described by a LTI device model. The derived radar system model is applicable, if bandwidth limitation (reduced from 20 GHz to 6 GHz) is applied in a post-processing step. However, band limitation reduces the resolution, which is a major drawback.

4.2.2 Non-Linear Device Model

It has been pointed out in the previous section, that the considered radar suffers from parasitic signal components or ghost targets, which are not covered by the LTI system model. Because these signal components occur with large amplitudes, they cannot be distinguished from real targets in the range profile and hence the number of targets will be overestimated by a detector. The ghost targets can be mitigated by additional filtering of the transmit signal [5], but realisation of filters requires more place on the MMIC and hence increases the costs. Because the ghost targets are hardware caused and therefore predictable, they can be accounted for in the target detection step by e.g. excluding the respective peaks in the range profile [5]. However, if a true target peak is hidden or intersected by a ghost target peak, the true target cannot be detected. Hence, this approach may fail in case of multiple targets. In the previous section the
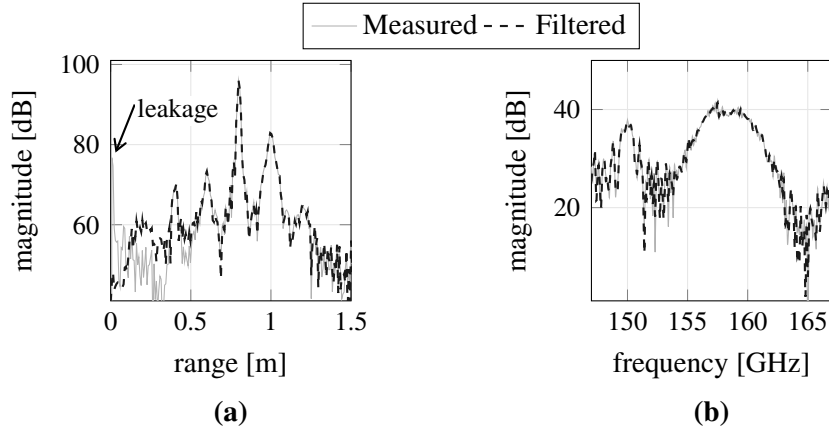
Figure 4.5: *Measurement with target at 0.8 m distance and radar cross section of −10 dBsm. The range spectrum before and after high-pass filtering to suppress the leakage is shown in (a). The respective time domain signals are shown in (b). The time axis has been mapped to frequency by $f = W/T_\mathrm{M} \cdot t + f_0$.*
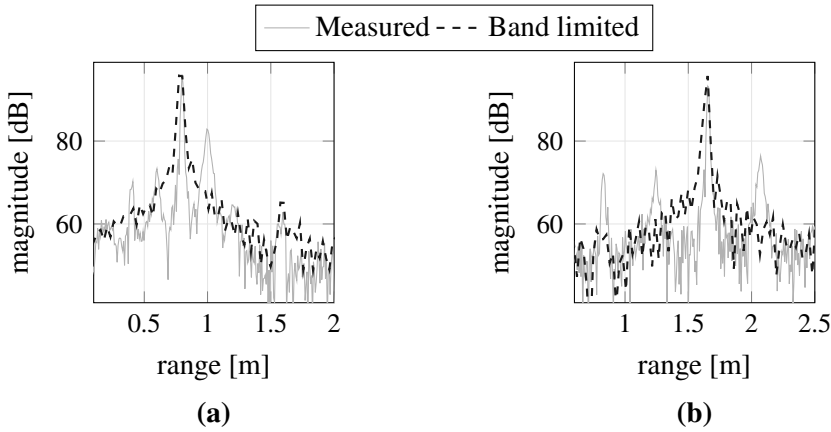
Figure 4.6: *Range spectra of the measured and band limited signal for two different measurements with a single target (a) at distance 0.8 m and RCS of −10 dBsm; (b) at distance 1.6 m and RCS of 0 dBsm. The bandwidth has been reduced from 20 GHz to 6 GHz. [21]*

occurrence of ghost targets has been tackled by limiting the bandwidth of the received signal, which decrease the resolution and dynamic range of the measurements.

A signal processing solution is to favour, which mitigates the ghost targets and preserves the resolution. Using signal processing to compensate or mitigate hardware impairments is conceptually known as *dirty RF* [Fet+05, Fet+07]. In [Gri+14] modelling of non-linear distortions in communication system receivers and their real-time mitigation has been discussed. However, dirty RF attempts to "clean" the baseband signal from hardware caused distortions. Here, estimation of model parameters from the baseband signal is focused and the hardware impairments are accounted for in the estimator, rather than attempting to mitigate them beforehand. Hence, both approaches are different in terms of how they tackle the hardware impairments.

Subsequently the transfer characteristic of the radar in conjunction with the transmit signal

generation will be considered, in order to derive a novel device model of an FMCW radar [5]. First, the source of ghost target causing hardware impairments has to be identified. Afterwards, their proper modelling and last their incorporation in an extended device model has to be clarified. Due to the new modelling approach, the device model $\mathcal{M}^{\mathrm{D}}$ and propagation (physical) model $\mathcal{M}^{\mathrm{P}}$ become coupled in the system model $\mathcal{M}^{\mathrm{S}}$, because the device model depends on the parameters of the propagation model. In classical system models for radar applications, as presented e.g. in [21], device and propagation model can be decoupled by deconvolution. This is no longer possible for the non-linear device model, which will be derived in the following. By estimating the propagation channel parameters, the corresponding device model is automatically selected and the ghost targets are suppressed, i.e. parameter estimation and mitigation are jointly done. Consequently, parameter estimation becomes a deconvolution aspect, as the effects of the hardware non-linearities are mitigated in the parameter estimation step.

Analysis of the Signal Synthesiser

A block diagram of the signal generator of the considered MMIC radar chip is shown in Fig. 4.7. The radar chip features a *frequency offset synthesiser* for signal generation, where a ramp oscillator (RO) signal, which is a FMCW signal at passband, is up-converted by mixing with a local oscillator (LO) signal [Hit+17]. The digitally generated RO signal is applied to a quadrupler, resulting in a signal at a higher passband range and four times the bandwidth. The low phase noise LO signal, generated by a push-push voltage controlled oscillator (VCO), is applied to a doubler to enhance its frequency. Both signals are subsequently mixed by a Gilbert cell mixer, which does not feature a sideband suppression [Hit+17]. Hence, the lower and the higher sideband are present at the output of the mixer. Due to the filter characteristics of the signal paths and the passband characteristic of the employed on-chip antennas, the lower sideband is attenuated [Hit+17]. In the following, the RO and LO signal are assumed as perfect, hence no distortions in the RO ramp are considered. The corresponding model for the RO and LO signal are

$$w_{\mathrm{RO}}(t) = U_{\mathrm{RO}} \cdot \cos\left(\pi\frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}}t^2 + 2\pi f_{\mathrm{RO}}t + \phi_{\mathrm{RO}}\right) \tag{4.28a}$$

$$w_{\mathrm{LO}}(t) = U_{\mathrm{LO}} \cdot \cos\left(2\pi f_{\mathrm{LO}}t + \phi_{\mathrm{LO}}\right) \quad . \tag{4.28b}$$

The parameters W_{RO} and T_{M} are bandwidth and modulation time of the ramp of the RO signal, respectively; and frequency f_{RO} is the centre frequency of the RO signal. Variables ϕ_{LO} and ϕ_{RO} are the initial phases of LO and RO signal, respectively. For sake of convenience, the phase terms are omitted in the following ($\phi_{\mathrm{LO}} = 0$ and $\phi_{\mathrm{RO}} = 0$).

Before the mixing, the LO and RO signal are up-converted by multipliers in order to enhance the bandwidth (RO signal) and increase the centre frequency (LO signal). Multipliers are non-linear devices, which input-output relation can be formulated in terms of a non-linear function $x_{\mathrm{out}} = f(x_{\mathrm{in}})$. Afterwards a mixer is used, which is a non-linear device too. A common model for the input-output relation of non-linear devices is the memory-less power series model [Sch09], which can be assumed as the Taylor series approximation of the non-
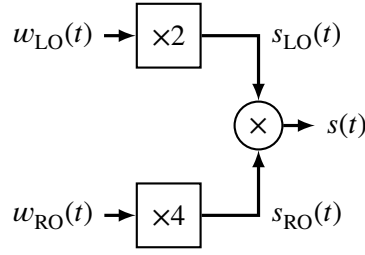
Figure 4.7: *Block diagram of the signal generator of the MMIC radar chip in [Hit+17]. [5]*

linear transfer function. The input-output relation is

$$f\left(x_{\text{in}}\left(t\right)\right) = x_{\text{out}}\left(t\right) = \sum_{k=0}^{K} c_k \cdot x_{\text{in}}^k\left(t\right) \quad, \tag{4.29}$$

with K the degree of the non-linearity and c_k the coefficients of the power series. Applying the power series model to the RO and LO signal yields

$$s_{\text{RO}}\left(t\right) = \sum_{m=1}^{M} a_m \cdot U_{\text{RO}}^m \cdot \cos\left(\pi m \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi m f_{\text{RO}} t\right) \tag{4.30a}$$

$$s_{\text{LO}}\left(t\right) = \sum_{n=1}^{N} b_n \cdot U_{\text{LO}}^n \cdot \cos\left(2\pi n f_{\text{LO}} t\right) \quad, \tag{4.30b}$$

with M and N denoting the degrees of non-linearity of the RO and LO multiplier, respectively. In the following, the summation limits will be omitted for notational convenience.

The resulting signals after the multipliers are harmonics of the input signals, rather than just the up-converted signals for $n = 4$ and $m = 2$. Assuming the mixer as perfect (i.e. performs an ideal multiplication), the transmit signal is given by

$$s\left(t\right) = \sum_n \sum_m a_m \cdot U_{\text{RO}}^m \cdot b_n \cdot U_{\text{LO}}^n \cdot \cos\left(2\pi t n f_{\text{LO}}\right) \cdot \cos\left(\pi m \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi m f_{\text{RO}} t\right) \quad. \tag{4.31}$$

The assumption of a perfect mixer can be relaxed for the subsequent investigations, because a non-perfect mixer just causes additional harmonics, which can be described by adjusting the degrees of non-linearity N and M of the power series models. Using trigonometric identity [1] yields

$$s\left(t\right) = \sum_n \sum_m c_{n,m} \cdot \left[\cos\left(\phi_{n,m}^+\left(t\right)\right) + \cos\left(\phi_{n,m}^-\left(t\right)\right)\right], \tag{4.32}$$

---

[1] $\cos x \cos y = 0.5 \cdot \left[\cos(x - y) + \cos(x + y)\right]$
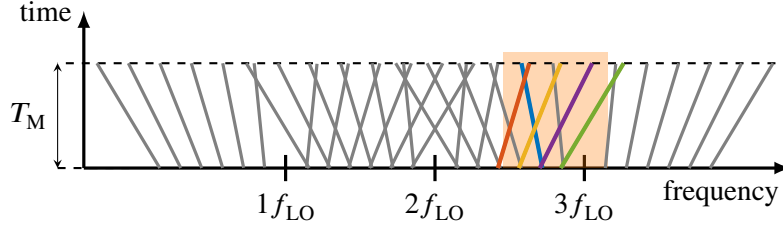
Figure 4.8: *Frequency-time diagram of the transmit signal for non-linearity degrees $M = 6$ and $N = 3$ for the RO and LO multiplier, respectively. The orange highlighted area denotes the band of operation of the radar system and the coloured ramps denote the actually transmitted ramps.* [5]

with

$$c_{n,m} = 0.5 \cdot a_m \cdot U_{\text{RO}}^m \cdot b_n \cdot U_{\text{LO}}^n \tag{4.33a}$$

$$\phi_{n,m}^+ (t) = 2\pi t n f_{\text{LO}} + \pi m \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t m f_{\text{RO}} \tag{4.33b}$$

$$\phi_{n,m}^- (t) = 2\pi t n f_{\text{LO}} - \pi m \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 - 2\pi t m f_{\text{RO}} \quad . \tag{4.33c}$$

Summarised, the transmit signal $s(t)$ constitutes of multiple FMCW signals, which operate at different frequency ranges and have different bandwidths. Figure 4.8 shows the frequency-time diagram of the transmit signal for the non-linearity degrees $M = 6$ and $N = 3$. The highlighted area and the coloured ramps indicate the operational bandwidth of the radar system and the actually transmitted ramps.

Non-linear Static Device Model

In order to derive a device model, the received baseband signal has to be analysed, while considering the transmit signal model and the propagation model in conjunction with the system architecture. As system architecture, the stretch processing receiver depicted in Fig. 4.2 is considered. Accordingly, the (real-valued) baseband signal is

$$x(t) = \int_t^{t+T} s(t') \cdot \left[s(t') * h(t') \right] \, dt' \quad , \tag{4.34}$$

with $h(t)$ denoting the transfer function of the propagation channel, and T the integration time of the low-pass filter. The propagation channel constitutes of the superposition of P propagation paths, which can be modelled in time domain as (see Section 3.2)

$$h(t) = \sum_{p=1}^{P} \gamma_p \cdot \delta(t - \tau_p) \quad . \tag{4.35}$$

The approximate baseband signal is given by (see APPENDIX C.2 for the derivation)

$$
x(t) \approx \sum_p \frac{\gamma_p}{2} \cdot \sum_{n_1} \sum_{m_1} \sum_{n_2} \sum_{m_2} c_{n_1,m_1} \cdot c_{n_2,m_2} \cdot \cos \left(\phi^+_{n_1,m_1}(t) - \phi^+_{n_2,m_2}(t - \tau_p) \right)
$$

$$
+ \sum_p \frac{\gamma_p}{2} \cdot \sum_{n_1} \sum_{m_1} \sum_{n_2} \sum_{m_2} c_{n_1,m_1} \cdot c_{n_2,m_2} \cdot \cos \left(\phi^-_{n_1,m_1}(t) - \phi^-_{n_2,m_2}(t - \tau_p) \right) \qquad (4.36)
$$

According to the summation indices, the baseband signal constitutes of the superposition of auto- and cross-terms.

AUTO-TERMS   For the auto-terms $\left(m = m_1 = m_2, n = n_1 = n_2 \right)$ the baseband signal reduces to

$$
x(t) \approx \sum_p \frac{\gamma_p}{2} \sum_n \sum_m c^2_{n,m} \cos \left(\phi^+_{n,m}(t) - \phi^+_{n,m}(t - \tau_p) \right)
$$

$$
+ \sum_p \frac{\gamma_p}{2} \sum_n \sum_m c^2_{n,m} \cos \left(\phi^-_{n,m}(t) - \phi^-_{n,m}(t - \tau_p) \right) \qquad (4.37)
$$

with the phase differences

$$
\phi^+_{n,m}(t) - \phi^+_{n,m}\left(t - \tau_p\right) = 2\pi \tau_p n f_{\mathrm{LO}} + \theta_{n,m}\left(t, \tau_p\right) \qquad (4.38a)
$$

$$
\phi^-_{n,m}(t) - \phi^-_{n,m}\left(t - \tau_p\right) = 2\pi \tau_p n f_{\mathrm{LO}} - \theta_{n,m}\left(t, \tau_p\right) \qquad . \qquad (4.38b)
$$

The resulting beat frequency at the stretch processor output is given by the derivative of the above phase terms.

$$
f_{\mathrm{b}} = \left| \pm \frac{1}{2\pi} \frac{\partial}{\partial t} \theta_{n,m}\left(t, \tau_p\right) \right| = m \frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}} \tau_p \qquad (4.39)
$$

Consequently, each delay τ_p causes multiple beat frequencies, which depend on the non-linearity degree m of the RO signal multiplier.

CROSS-TERMS   Focusing the cross-terms $\left(m_1 \neq m_2, n_1 \neq n_2 \right)$ the resulting beat frequency of the baseband signal is

$$
f_{\mathrm{b}}(t) = \pm \left(n_1 - n_2 \right) f_{\mathrm{LO}} \pm m_2 \frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}} \tau_p \pm \left(m_1 - m_2 \right) f_{\mathrm{RO}} \pm \left(m_1 - m_2 \right) \frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}} t \qquad . \qquad (4.40)
$$

As the resulting beat frequency depends on the time t it undergoes variations over time. Only a small portion will remain in the baseband, if the beat frequency exceeds the cut-off frequency of the low-pass filter. Hence, the cross-terms occur as short-time distortions, e.g. ramps with different and intersecting slopes. Because the cross-term beat frequency depends on the centre frequency of the RO and LO signal, which are much higher than the cut-off frequency of the low-pass filter, the cross-term have a minor influence on the baseband signal and therefore can be assumed as negligible.
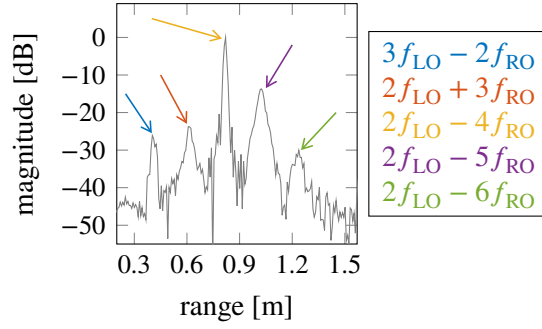
Figure 4.9: *Centre frequencies of the FMCW signals which cause the respective beat frequency peaks. The colours correspond to the colours of the FMCW ramps in Fig. 4.8.*

Summarised, the baseband signal constitutes of multiple beat frequencies for each target. This is caused by the multiple FMCW signals, which are due to the non-linear multipliers of the signal generator. Accordingly, a model for the baseband in case of P targets is

$$x(t) = \sum_{p=1}^{P} \gamma_p \cdot \sum_{m=1}^{M} c_m \cdot \cos\left(m \frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}} t\tau_p + \phi_m \right) \quad, \tag{4.41}$$

with c_m the power series coefficients, which are related to the non-linearity of the RO and LO multiplier. The phase term ϕ_m accounts for all yet unconsidered phase variations for each beat frequency.

Model Order and Degrees of Non-linearity   The degrees of non-linearity M and N (i.e. the order of the power series device models) have to be known, because they determine the number of observed ghost targets. Basically, they depend on the degree of the non-linearity of the multipliers, which are unknown in advance. According to the transmit signal model in equation (4.32), the generated FMCW signals feature different frequency ranges. Due to the passband characteristic of the radar system, caused by e.g. the antennas or amplifiers, these frequency ranges are differently attenuated. Hence, the radar system transmits only a subset of the generated FMCW signals. This subset of transmitted FMCW signals can be identified from calibration measurements. The transmitted FMCW signals and their corresponding peaks in the range profile are depicted in Fig. 4.9. Accordingly, the non-linearity degrees have been determined to be $N = 3$ and $M = 6$.

Estimation of Power Series Coefficients   Consider a single target. Then, the complex-valued baseband signal is

$$x(t) = \gamma \cdot \sum_{m=2}^{6} c_m \cdot \exp\left\{ j2\pi m \frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}} t\tau \right\} \quad. \tag{4.42}$$

Note, that c_m is complex-valued in order to account for the phase term ϕ_m: $c_m = |c_m| \exp\{j\phi_m\}$. The complex path weight γ accounts for the propagation attenuation given by the radar

equation [Sko90] and for remaining phase variations [Ric13].

$$\gamma\left(f_{\mathrm{c}}, W\right) = \sqrt{\frac{1}{(4\pi)^3 \, f_{\mathrm{c}}^2 c_0^2 \tau^4}} \cdot \xi \cdot \exp\left\{-\jmath 2\pi f_{\mathrm{c}}\tau\right\} \cdot \exp\left\{\jmath\pi\frac{W}{T_{\mathrm{M}}}\tau^2\right\} \qquad (4.43)$$

Because each generated FMCW signal operates at a different frequency range, attenuation and phase variation due to propagation differ. Hence, a separate path weight has to be considered for each beat frequency.

A model for the observation of K data samples, which are stacked in vector $\mathbf{y} \in \mathbb{C}^K$ is

$$\mathbf{y} = \mathbf{E}\left(\tau\right) \cdot \mathrm{diag}\left\{\boldsymbol{\gamma}\right\} \cdot \mathbf{c} + \mathbf{w} \quad . \qquad (4.44)$$

The respective matrix and vectors are given by

$$\mathbf{E}\left(\tau\right) = \exp\left\{\jmath 2\pi\frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}}\tau T_{\mathrm{S}} \cdot \mathbf{k} \cdot \mathbf{m}^{\mathrm{T}}\right\} \qquad (4.45\mathrm{a})$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma\left(3f_{\mathrm{LO}} - 2f_{\mathrm{RO}}, 2W_{\mathrm{RO}}\right) \\ \gamma\left(2f_{\mathrm{LO}} + 3f_{\mathrm{RO}}, 3W_{\mathrm{RO}}\right) \\ \gamma\left(2f_{\mathrm{LO}} + 4f_{\mathrm{RO}}, 4W_{\mathrm{RO}}\right) \\ \gamma\left(2f_{\mathrm{LO}} + 5f_{\mathrm{RO}}, 5W_{\mathrm{RO}}\right) \\ \gamma\left(2f_{\mathrm{LO}} + 6f_{\mathrm{RO}}, 6W_{\mathrm{RO}}\right) \end{bmatrix} \qquad (4.45\mathrm{b})$$

$$\mathbf{c} = \begin{bmatrix} c_2, c_3, c_4, c_5, c_6 \end{bmatrix}^{\mathrm{T}} \quad . \qquad (4.45\mathrm{c})$$

The operator $\mathrm{diag}\left\{\boldsymbol{\gamma}\right\}$ forms a diagonal matrix from the entries of vector $\boldsymbol{\gamma}$, and $\mathbf{m} = \begin{bmatrix} 2, \ldots, 6 \end{bmatrix}^{\mathrm{T}}$ and vector $\mathbf{k} = \begin{bmatrix} 0, \ldots, K-1 \end{bmatrix}^{\mathrm{T}}$. Vector $\mathbf{w}$ accounts for the additive measurement noise. Assuming proper complex, white and normally distributed noise, an estimate of the power series coefficients $\mathbf{c}$ is given by the best-linear-unbiased estimator (BLUE) [MS00].

$$\hat{\mathbf{c}} = \begin{bmatrix} \mathbf{E}\left(\tau\right) \cdot \mathrm{diag}\left\{\boldsymbol{\gamma}\right\} \end{bmatrix}^{+} \cdot \mathbf{y} \qquad (4.46)$$

Exemplary measurements with two corner reflectors as targets, which are placed at different distances apart from the radar and featuring different radar cross section (RCSs), are considered now. From the first measurement in Fig. 4.10(a), the power series coefficients are estimated. The estimated coefficients and the known measurement setup (target distance and RCS) are used to reconstruct the range spectrum of the second measurement, which is shown in Fig. 4.10(b). A good match of the peak height and peak position between measurement and model can be observed. However, the peak width does not match to the measurements. The peak width indicates the presence of dynamic effects in the system, which occur due to the frequency dependence of e.g. power amplifiers, transmission lines, splitters etc. [Sch09]. Frequency dependence or dynamic effects also indicate memory effects, hence the dependence of the output to the history of the input. Last, the peak width varies for each beat frequency, because the corresponding FMCW signals are at different frequency ranges and therefore undergo different variations of the radar system response.

Summarised, the power series model properly describes the occurrence of ghost targets, i.e.
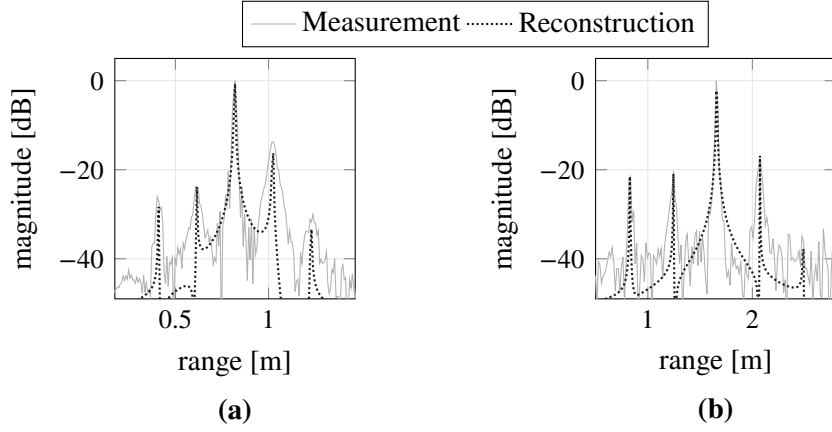
Figure 4.10: *Range profile of two example measurements (target parameters: (a) 0.8 m, 0.45 dBsm and (b) 1.6 m, 0.45 dBsm) compared to the reconstructed range profile according to the static power series model for (a) estimation and (b) reconstruction. [5]*

multiple beat frequencies for a single target. The previously introduced term "ghost targets" for the parasitically occurring beat frequencies is slightly misleading, because they are caused by the same target but are due to multiple transmit signals with different bandwidths. Hence, a more appropriate term would be beat frequency harmonics. The power series model is a *non-linear static* device model, such that system dynamics due to the frequency dependence of radar system components are not accounted for. Hence, a *non-linear dynamic* system model becomes necessary in order to properly incorporate the transfer characteristics of the radar system.

Advanced Radar Device Model

The joint consideration of dynamics and non-linearities in modelling of feedback-free systems can be accomplished by *Volterra or Wiener theory* [Nel01, Sch80]. Both theories describe the expansion of a non-linear functional (e.g. an input-output relation of a system) into a functional series, whereas memory or dynamic effects are accounted for. The truncated Volterra series for a non-linear, time-invariant, SISO system with linear memory is [Sch09]

$$x_{\text{out}}(t) = \sum_{n=0}^{N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_n(\tau_1, \ldots, \tau_n) \prod_{i=1}^{n} x_{\text{in}}(t - \tau_i) \, d\tau_i \quad , \tag{4.47}$$

with $g_n(\tau_1, \ldots, \tau_n)$ the n-th Volterra kernel and N the truncation order of the series.

Estimation of the Volterra kernels is complicated and requires multiple measurements [Sch80]. Therefore, simplifications with reduced effort are commonly employed [Nel01]. Assume, that the non-linearity and the dynamics are separable, the Volterra kernels are zero along their off-diagonals [Sch09].

$$g_n(\tau_1, \ldots, \tau_n) = \begin{cases} g_n(\tau) & \tau = \tau_1 = \ldots = \tau_n \\ 0 & \text{otherwise} \end{cases}$$
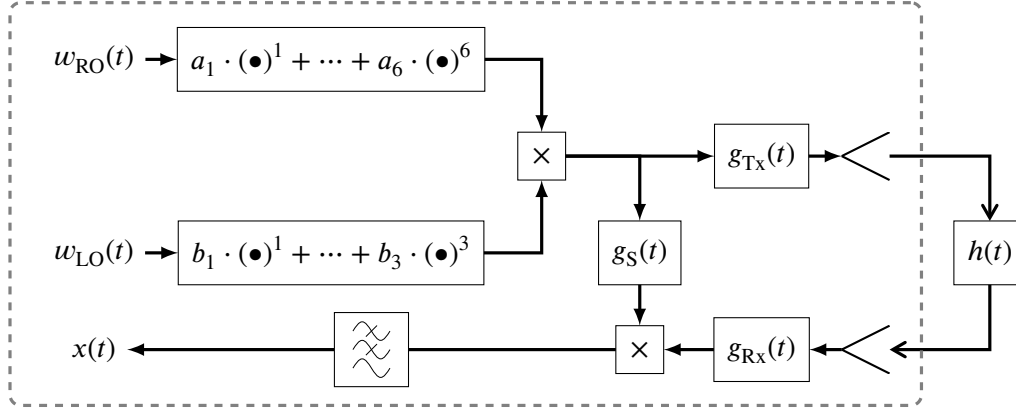
Figure 4.11: *Radar device model with Hammerstein model at passband domain. [5]*

Plugging into equation (4.47) yields the *generalised Hammerstein series*.

$$x_{\text{out}}(t) = \sum_{n=0}^{N} \int_{-\infty}^{\infty} g_n(\tau) \cdot x_{\text{in}}^n(t - \tau)\, d\tau = \sum_{n=0}^{N} g_n(t) * x_{\text{in}}^n(t) \tag{4.48}$$

System $g_n(t)$ denotes a LTI system. The Hammerstein series describes a *Hammerstein model*, which models the system as a serial cascade of two subsystems: a static, non-linear system followed by a dynamic, linear system [Hab99, Sch09]. The static, non-linear subsystem is described by a power series; whereas the dynamic, linear subsystem is a LTI system. If different LTI systems are considered for different powers as in equation (4.48), the model is called *generalised Hammerstein model* [Hab99]. The generalised Hammerstein modelling approach is closely related to the power series approach, where the series coefficients are time-variant. Separating the generalised Hammerstein model explicitly in the power series and LTI system model yields

$$x_{\text{out}}(t) = \sum_{n=0}^{N} g_n(t) * \left[c_n \cdot x_{\text{in}}^n(t) \right] \quad . \tag{4.49}$$

Subsequently, the Hammerstein model structure will be considered to describe the radar system at passband and baseband domain.

Hammerstein Model at Passband A Hammerstein model at passband domain is given by combining the power series model and the model in Section 4.2.1, which accounts for occurring system dynamics at passband level. The model structure is shown in Fig. 4.12. This model structure has a physical interpretation, because the model components can be related to hardware components and their behaviour in the real system. However, the measurements are at baseband domain such that the Hammerstein model in passband domain is hard to identify. Therefore, a baseband model, which equivalently describes the behaviour of the system, is necessary.

Generalised Hammerstein Model at Baseband An equivalent baseband model is defined as a model which results in the same baseband signal as the passband model would. In order to transform the passband model in Fig. 4.11 into an equivalent baseband representation,
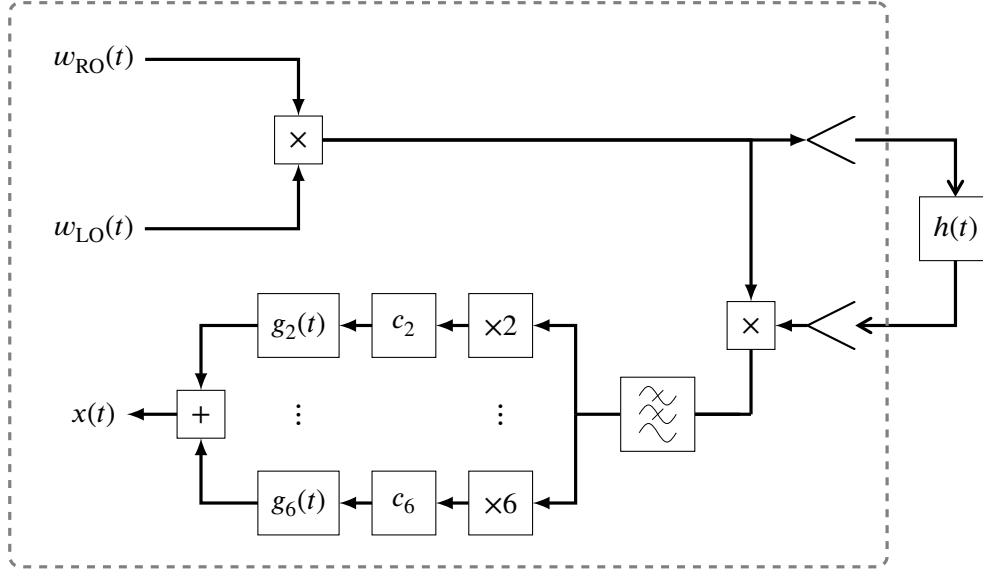
Figure 4.12: *Radar device model with generalised Hammerstein model at baseband. [5]*

several considerations have to be taken into account. First, the non-linearities at transmitter passband can be equivalently represented by non-linearities at receiver baseband, because only contributions of the auto terms of the multiple generated FMCW signals are considered at baseband, and cross terms are neglected. Second, the generated FMCW signals are at different frequency ranges (see Fig. 4.8) and therefore are influenced by different parts of the radar transfer function at passband. Hence, each beat frequency at baseband is affected by a different system dynamic. Last, the different system dynamics are assumed to be related to non-intersecting parts of the radar transfer function at passband. Hence, the system dynamics can be modelled independently. Summarised, an equivalent representation of the Hammerstein model at passband is given by a generalised Hammerstein model [Hab99] at baseband. The corresponding block chart is shown in Fig. 4.11. Disadvantageously, a generalised Hammerstein model features more unknown parameters than the Hammerstein model [Nel01].

According to [9], a convolution at passband simplifies to a multiplication in baseband for FMCW radars with stretch processor receiver architecture. The output of the generalised Hammerstein model in the complex baseband is

$$x\left(t\right) = \sum_{m=2}^{6} c_m \cdot \gamma_m \cdot g_m\left(t\right) \cdot \exp\left\{j2\pi m \frac{W_{RO}}{T_M} t\tau\right\} \quad . \tag{4.50}$$

Estimation of the Hammerstein model will be divided into the estimation of the power series coefficients c_m (i.e. identification of the static, non-linear subsystem) and determination of the LTI systems $g_m\left(t\right)$ (i.e. identification of the dynamic, linear subsystem). The power series coefficients are estimated by the BLUE approach from the previous section. Knowing the power series coefficients the input signal $z_m\left(t\right)$ to each LTI system $g_m\left(t\right)$ (i.e. the m-th dynamic, linear subsystem) is

$$z_m\left(t\right) = c_m \cdot \gamma_m \cdot \exp\left\{j2\pi m \frac{W_{RO}}{T_M} t\tau\right\} \quad . \tag{4.51}$$

The sum of the outputs $z_m(t)$ of all M LTI systems is the baseband signal $x(t)$, such that a multiple input single output (MISO) system is given [Nel01]. Hence, the data model for a noisy radar observation $y(t)$ is

$$y(t) = \mathbf{g}(t)^{\mathrm{T}} \cdot \mathbf{z}(t) + w(t) \quad , \tag{4.52}$$

with $\mathbf{g}(t) = \left[g_2(t), \dots, g_6(t)\right]^{\mathrm{T}}$ the vector of LTI system responses, $\mathbf{z}(t) = \left[z_2(t), \dots, z_6(t)\right]^{\mathrm{T}}$ the vector of input signals and $w(t)$ the proper complex, normally distributed measurement noisy. Estimation of the vector $\mathbf{g}(t)$ from the output $y(t)$ while knowing the input $\mathbf{z}(t)$ equals the identification of a MISO system with p-canonical model structure [Nel01]. The data model for K observed samples is

$$\mathbf{y} = \sum_{m=2}^{6} \mathbf{z}_m \odot \mathbf{g}_m + \mathbf{w} \quad , \tag{4.53}$$

with vectors $\mathbf{y}, \mathbf{z}_m, \mathbf{g}_m \in \mathbb{C}^K$. The vectors contain the samples of the m-th beat frequency signal $\mathbf{z}_m = \left[z_m(0), \dots, z_m\left((K-1)T_{\mathrm{S}}\right)\right]^{\mathrm{T}}$, and $\mathbf{g}_m = \left[g_m(0), \dots, g_m\left((K-1)T_{\mathrm{S}}\right)\right]^{\mathrm{T}}$ the samples of the corresponding system dynamic.

Summarised, the generalised Hammerstein model at baseband describes the response of the radar system as a MISO system of multiple LTI systems. Each LTI system gets a unique input signal, which is given by a harmonic of the respective beat frequency signal. Identification of the LTI systems $g_m(t)$ can be accomplished in frequency or time domain by separation technique or parametric modelling, respectively.

FREQUENCY DOMAIN SEPARATION   All beat frequencies and their respective system dynamic are superimposed in time domain. Taking the Fourier transform of equation (4.50) yields the spectrum $Y(f)$.

$$y(t) \quad \circ\!\!\!-\!\!\!\bullet \quad Y(f) = \sum_{m=2}^{6} c_m \cdot \gamma_m \cdot G_m\left(f - m\frac{W_{\mathrm{RO}}}{T_{\mathrm{M}}}\tau\right) + W(f) \tag{4.54}$$

The system dynamics are differently shifted in frequency domain according to their respective beat frequency. Assuming, that the dynamics do not superimpose in frequency domain (the peaks do not overlap), they can be separated by gating. A window function has to be determined for gating, i.e., the window type and the window width has to be selected. Also, the windows should be selected such that the complete spectrum part, which is occupied by the beat frequency components, is captured. A Hann window function with an overlapping of 50% percent between neighbour windows will be used to separate the system dynamics. The window functions are depicted in Fig. 4.13. Due to the window overlapping, the spectrum part with the beat frequency dynamics is completely captured. The estimated system dynamics are depicted in Fig. 4.14. The contribution of the dynamic of each beat frequency to the time domain measurement signal becomes obvious. In [21] the problem of multiple occurring beat frequencies has been tackled by limiting the time domain signal to the range of approx. 0.4 ms to 0.75 ms (correspondingly limiting the bandwidth). As a consequence, all parasitic beat frequencies were suppressed,
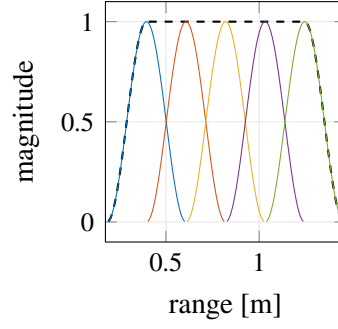
Figure 4.13: *Window functions to separate the system dynamics in frequency domain. The colours indicate the respective beat frequency (see Fig. 4.9) for which the window is applied. The black dashed line indicates the overall window function, resulting from the superposition of all window functions.*
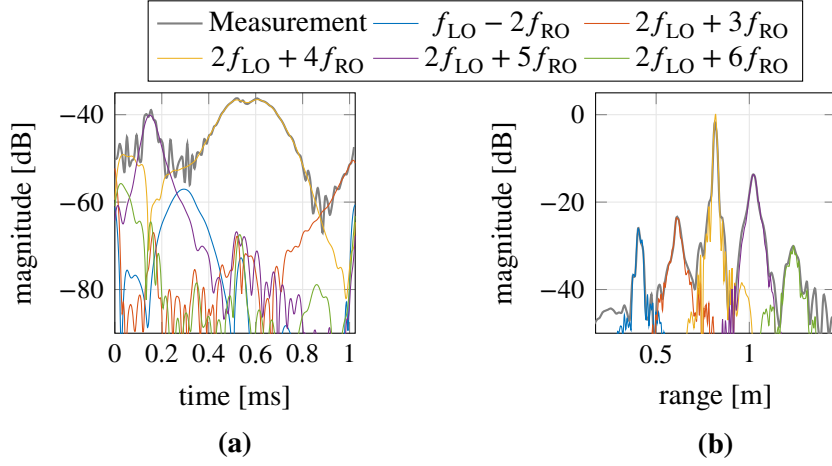
Figure 4.14: *Estimated system dynamics of each beat frequency in (a) time and (b) frequency domain. The dynamics were estimated by gating in frequency/range domain using the Hann window sequence depicted in Fig. 4.13. The coloured lines indicate the respective beat frequencies, influenced by their respective dynamics.*

because their contributions are low in this signal part as visible in Fig. 4.14(a).

Drawback of the gating approach is, that it is not a hands-on solution. For varying calibration scenarios the gating windows must be always adjusted accordingly. Furthermore, the estimated dynamics are confounded by the power series coefficients and the path weights, such that separation in post-processing is necessary. Therefore, an approach who directly solves the MISO identification problem is to favour.

TIME DOMAIN PARAMETRIC MODEL Because 5 beat frequencies were previously identified as significant, $5K$ samples would be necessary to estimate the system dynamic for all beat frequencies. However, if only one calibration measurement is taken into account, more unknowns than number of measurements would be present. Hence, a parametric model with a fewer number of parameters is required, in order to get a unique estimation result.

A parametric model for a system dynamic $g(t)$ can be given by the rational form [PS01]

$$g(t) = \frac{v_0 + \sum_{r=1}^{D} v_r \cdot \omega^r(t)}{1 + \sum_{r=1}^{F} f_r \cdot \omega^r(t)} \quad , \tag{4.55}$$

with model parameters (coefficients) $v_r, f_r \in \mathbb{C}$ and function $\omega(t) = j2\pi t$. Accounting for the superposition of multiple beat frequencies, each one described by a different LTI system, a MISO output-error model can be given [Nel01].

$$y(t) = \sum_{m=2}^{6} z_m(t) \cdot \frac{v_{0,m} + \sum_{r=1}^{V_m} v_{r,m} \cdot \omega^r(t)}{1 + \sum_{r=1}^{F_m} f_{r,m} \cdot \omega^r(t)} + w(t) = \sum_{m=2}^{6} z_m(t) \cdot \frac{\omega_{V_m}(t)^{\mathrm{T}} \cdot \mathbf{v}_m}{1 + \omega_{F_m}(t)^{\mathrm{T}} \cdot \mathbf{f}_m} + w(t) \tag{4.56}$$

The corresponding vectors of model parameters are

$$\mathbf{v}_m = \begin{bmatrix} v_{0,m}, \dots, v_{V_m,m} \end{bmatrix}^{\mathrm{T}} \qquad \in \mathbb{C}^{V_m+1} \tag{4.57a}$$

$$\boldsymbol{\omega}_{V_m}(t) = \begin{bmatrix} 1, \omega^1(t), \dots, \omega^{V_m}(t) \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{V_m+1} \tag{4.57b}$$

$$\mathbf{f}_m = \begin{bmatrix} f_{1,m}, \dots, f_{F_m,m} \end{bmatrix}^{\mathrm{T}} \qquad \in \mathbb{C}^{F_m} \tag{4.57c}$$

$$\boldsymbol{\omega}_{F_m}(t) = \begin{bmatrix} \omega^1(t), \dots, \omega^{F_m}(t) \end{bmatrix}^{\mathrm{T}} \quad \in \mathbb{C}^{F_m} \quad . \tag{4.57d}$$

The best suited numbers of numerator $V_1, \dots, V_M$ and denominator $F_1, \dots, F_M$ coefficients, and the corresponding model parameters $\mathbf{v}_1, \dots, \mathbf{v}_M, \mathbf{f}_1, \dots, \mathbf{f}_M$ have to be estimated from the calibration measurement. This, however, is a system identification problem. Estimation of the model parameters can be accomplished by maximum-likelihood (ML) estimation, where the square error between measurements and model is minimised [PS01]. Minimising the square error w.r.t. the numbers of coefficients is not practicable, because an increase of the model complexity always reduces the square error [PS01]. Hence, a more sophisticated criterion is necessary to validate the parameter estimates in conjunction with the numbers of coefficients.

The minimum description length (MDL) [PS01], which defines a trade-off between model accuracy (sum of squared errors) and model complexity (number of parameters), will be used as validation criterion. In order to jointly estimate the best suited numbers of coefficients and the corresponding model parameters, the MDL has to be minimised. Because minimisation of the MDL w.r.t. all unknowns is computationally cumbersome due to the high dimensionality of the problem, a sequential optimisation approach will be used. For given numbers of coefficients the model parameters are estimated and the MDL is calculated. Subsequently the MDL is minimised subject to the numbers of coefficients.

$$\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_M \atop \hat{\mathbf{f}}_1, \dots, \hat{\mathbf{f}}_M = \underset{\substack{V_1, \dots, V_M \\ F_1, \dots, F_M}}{\arg\min} \left\{ \ln(2K) \sum_{m=1}^{6} (F_m + V_m) + \underset{\substack{\mathbf{v}_1, \dots, \mathbf{v}_M \\ \mathbf{f}_1, \dots, \mathbf{f}_M}}{\arg\min} \sum_{k=0}^{K-1} \left| y(kT_{\mathrm{S}}) - \sum_{m=2}^{6} z_m(kT_{\mathrm{S}}) \frac{\omega_{F_m}(kT_{\mathrm{S}})^{\mathrm{T}} \mathbf{v}_m}{1 + \omega_{V_m}(kT_{\mathrm{S}})^{\mathrm{T}} \mathbf{f}_m} \right|^2 \right\} \tag{4.58}$$

The genetic algorithm from MatLab® global optimisation toolbox (Version 4.0) is used to minimise the MDL subject to the numbers of coefficients, whereas the output-error model function from MatLab® system identification toolbox (Version 9.9) is used to estimate the model
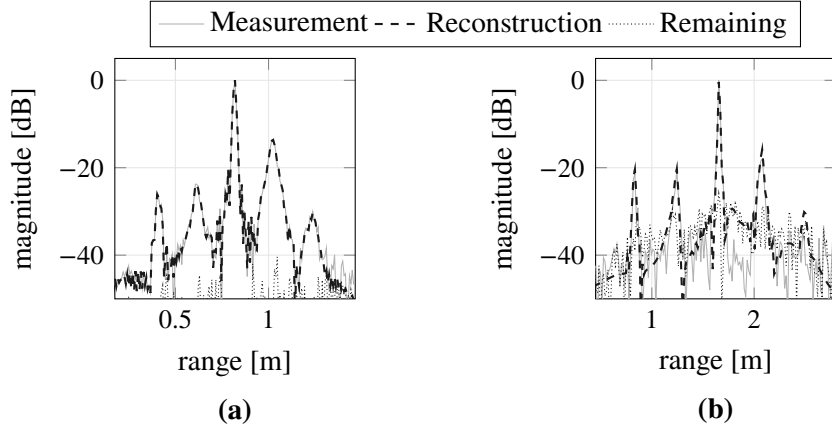
Figure 4.15: *Range profile of two example measurements (target parameters: (a) 0.8 m, 0.45 dBsm and (b) 1.6 m, 0.45 dBsm) compared to the reconstructed range profile according to the Hammerstein model at baseband. The remaining signal remains from the coherent subtraction of the reconstruction from the measurement.*

parameters for given numbers of coefficients. For unique parameter identification, the number of coefficients has to fulfil

$$\sum_{m=2}^{6} 5 F_m + V_m \leq K \quad , \tag{4.59}$$

which has to be considered as a constraint in the optimisation. Furthermore, integer optimisation can be applied, because the numbers of coefficients are $V_1, \ldots, V_M, F_1, \ldots, F_M \in \mathbb{N}$.

Model Calibration and Verification Figure 4.15(a) shows the range spectrum of the measurement, which is used to calibrate the proposed device model. Also, the range spectrum, calculated from the output of the calibrated device model, is shown too. A good match is visible which indicates, that the proposed device model properly describes the signal generation and transfer characteristics of the system.

For verification purpose, the calibrated Hammerstein model will be incorporated in an estimator. The ML estimator presented in [21] will be used, which estimates the complex weight γ and delay τ for each target. The device model of the radar system is incorporated in the estimator by extending the objective function accordingly.

$$\hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\tau}, \boldsymbol{\gamma}} \left\| \mathbf{y} - \sum_{m=2}^{6} c_m \cdot \mathbf{g}_m \odot \sum_{p=1}^{P} \gamma_{m,p} \cdot \mathbf{e}_m \left(\tau_p \right) \right\|^2$$

Vector $\mathbf{e}_m \left(\tau_p \right)$ represents the harmonic according to the m-th beat frequency for delay τ_p. The optimisation scheme presented in Section 5.3 is used to minimise the non-convex objective function subject to the parameters. The measured and reconstructed range profile, calculated from the estimation results and the data model, is shown Fig. 4.15(b). A good match of the reconstructed and measured range profile is obvious. Also, the remaining range profile is noise like, indicating a good match of measurement and reconstruction.
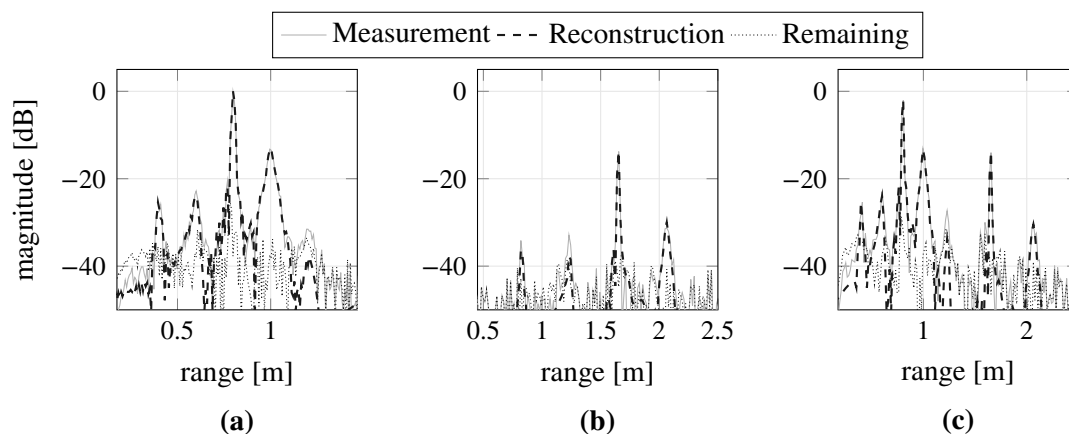
Figure 4.16: *Results for synthetic dual target data (target parameters: (a) 0.8 m, −10 dBsm and (b) 1.6 m, −10 dBsm). The plots show the measured, reconstructed and remaining range profiles for (a) the first target, (b) the second target and (c) the dual targets.*
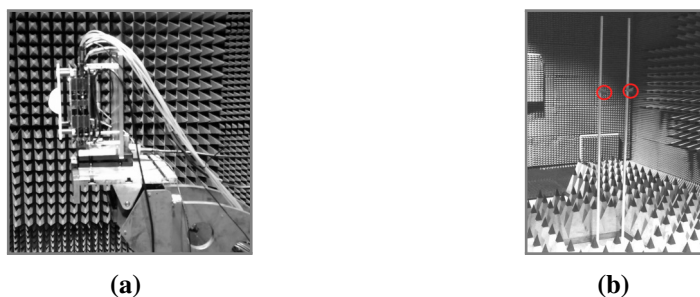
Figure 4.17: *Deployed radar system (a) and corner reflectors (b) in the anechoic chamber. [5]*

SYNTHETIC DATA   The proposed radar system model will be verified with measurements featuring a two target setup. The measurement is synthetically constructed by superimposing two single target measurements. The measurement was synthetically generated in order to check, how well the estimator in conjunction with the radar model can predict the two single target measurements. In both measurements, the target RCS was −10 dBsm, and the targets were placed at 0.8 m and 1.6 m, respectively. The measured range profile and the reconstructed range profile from the parameter estimates are depicted in Fig. 4.16. Obviously, the estimator is able to estimate the single target measurements. Furthermore, the remaining signal features no significant components and only noise remains. Hence, the ghost targets are successfully mitigated by the parameter estimator in conjunction with the device model.

MEASURED DATA   The proposed radar device model will be verified by measurements with multiple corner reflectors to set up multi-target scenarios. Pictures of the radar system and the deployed corner reflectors are shown in Fig. 4.17. Sequentially, three corner reflectors with varying RCS and distance are placed in front of the radar system, see Tab. 4.2 for a summary. Targets with a huge RCS are deployed, such that the ghost targets are properly present in order to verify the radar system model. As radar system the MIMO radar presented in [2] has been used, whereas only one channel has been captured and processed. The radar bandwidth and sweep time was set to 16 GHz and 1 ms, respectively. The scenario with a single target has been

Table 4.2: *Ranges, ghost ranges and of deployed targets; and estimated target ranges*

Deployed range [m]	Ghost ranges [m]	RCS [dBsm]	Estimated range [m]
3	1.52, 2.28 and 3.81	15.8	3.05
3	1.50, 2.25 and 3.75	15.8	3.05
3.81	1.91, 2.86 and 4.77	-10	3.82
2.87	1.44, 2.15 and 3.59	7	2.87
3.6	1.81, 2.73 and 4.52	-10	3.61
4.37	2.19, 3.28 and 5.46	7	4.37
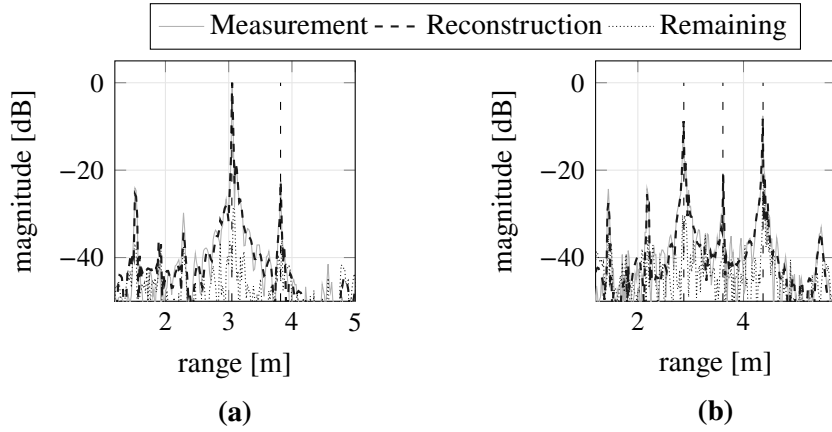
Figure 4.18: *Measured, reconstructed and remaining range profiles for the (a) two target and (b) three target scenario. The maximum-likelihood parameter estimator in conjunction with the device model is used for the reconstruction. The loosely dashed lines indicate the range of the deployed targets. [5]*

used to calibrate the radar device model, and the number of maximal occurring harmonics per target was found to be $M = 5$.

The measured, reconstructed and remaining range profile for the two and three target scenario are depicted in Fig. 4.18. First, the scenario with two targets will be considered. Obviously, the estimator is able to detect and separate the two targets properly and jointly mitigates the ghosts, such that the remaining range profile is noise like. Second, the three target scenario will be considered. Again, the parameter estimator in conjunction with the radar model properly detects all targets and mitigates the ghosts, such that the remaining range profile is noise like.

Last, the usability of the radar model for high-resolution will be verified. Therefore, a measurement with two targets of equal RCS and a separation of 12 mm is conducted. The measured range profile is shown in Fig. 4.19(a). According to the transmit signal bandwidth of 16 GHz the range resolution should be 9.4 mm. However, due to the transceiver characteristic of the radar, the bandwidth is not preserved and the range resolution is degenerated to approx. 15 mm. Hence, the target separation is below the Rayleigh resolution limit and the targets cannot be distinguished in the range profile, because only one peak is clearly visible. The parameter estimator in conjunction with the device model is able to resolve both targets, hence the radar model in conjunction with the ML estimator is suitable to achieve high-resolution. This is furthermore indicated by the proper fit of the measured and reconstructed range profile in Fig. 4.19(a). The

(a)

(b)
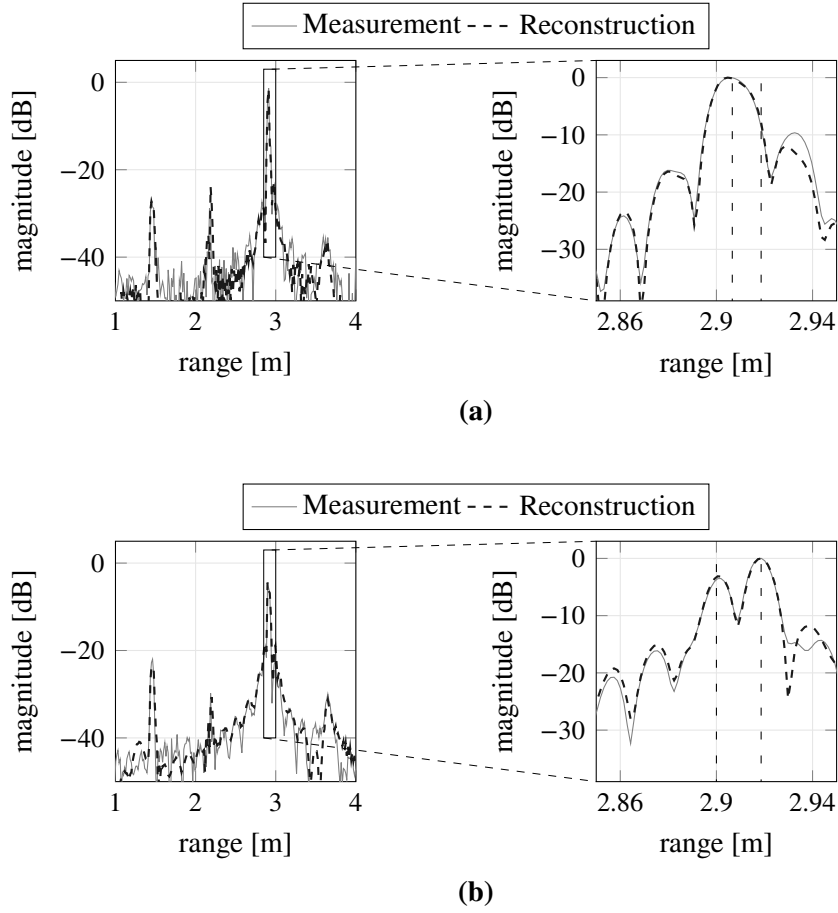
Figure 4.19: *Measured and model-based reconstructed range profile for measurements with two targets. The target ranges are indicated by loosely dotted lines. The targets are separated by approx. (a) 12 mm and (b) 18 mm. The range resolution of the radar is approx. 15 mm, such that the targets cannot be distinguished in the range profile in (a). The estimator properly detects and separates the targets in both cases, hence achieves a resolution better than the Rayleigh resolution,i.e., high-resolution.*

same measurement has been repeated with a target separation of 18 mm. The results are depicted in Fig. 4.19(b). Two peaks are clearly visible, such that the targets are separable by the range resolution of the radar. Also, the targets are detected and separated properly.

Summarised, a ML estimator in conjunction with the new radar device model properly detects the true targets and jointly mitigates the ghost targets. This has two advantages i) an overestimation of the number of targets is circumvented as the ghost targets are mitigated, and ii) targets hidden by ghost targets are detected. Hence, the estimator is robust in terms of target detection.

4.3 Antennas and Arrays of Antennas – Definition and Modelling

Estimation of direction, i.e. angles, of impingement and radiation will be accomplished by employing antenna arrays at Tx and Rx side. Hence, a MIMO system is employed. Exact knowledge of the array response is important for the direction estimation [7, Nai09, TF09]. As stated in literature, the antenna and array response depends on the direction, as well as on the polarisation of excitation [23, LKT12]. Furthermore, the array response is frequency dependent for wideband applications as e.g. for large signal bandwidths. Summarised, the array response has to be known in *angular*, *polarisation* and *frequency* domain. In this section wideband and narrowband modelling of the response of antennas and arrays of antennas are discussed for the purpose of direction estimation. Furthermore, the topology of a MIMO system is discussed too, whereas the case of *co-located* and *distributed* MIMO systems will be distinguished.

4.3.1 Modelling of the Antenna Output

Antennas will be described by means of a behavioural modelling approach. They will be considered as a linear time-invariant systems with a finite impulse response duration [WAS09]. The response of an antenna depends on the direction of impingement/radiation and polarisation of the impinging/radiated EMW. Direction and polarisation are vector quantities, whose definition necessitates a coordinate system. The Cartesian and spherical coordinate systems (see Fig. 3.1) can be considered therefore. It is convenient to describe the polarisation and direction of radiation of an EMW from an antenna in terms of spherical coordinates, because the number of required parameters is reduced compared to the number of parameters for a description in Cartesian coordinates. Hence, the characteristic of an antenna will be described in terms of the angles φ and ϑ, and polarisation vector $\mathbf{q} = \left[q_\varphi, q_\vartheta \right]^{\mathrm{T}}$.

Consider a transmit antenna. The electric field $\vec{\mathrm{E}}$ at distance R from the antenna, at radiation direction $\boldsymbol{\kappa}_{\mathrm{Tx}}$ and excitation signal $S(f)$ is given in [WAS09].

$$\vec{\mathrm{E}} \left(\boldsymbol{\kappa}_{\mathrm{Tx}}, f \right) = j 2 \pi f \cdot S(f) \cdot \sqrt{\frac{Z_0}{Z_{\mathrm{Tx}}(f)}} \cdot \frac{e^{j 2 \pi f \frac{R}{c_0}}}{2 \pi R c_0} \cdot \mathbf{h}_{\mathrm{Tx}} \left(\boldsymbol{\kappa}, f \right) \cdot \delta \left(\boldsymbol{\kappa} - \boldsymbol{\kappa}_{\mathrm{Tx}} \right) \qquad (4.60)$$

Vector $\mathbf{h}_{\mathrm{Tx}} \left(\boldsymbol{\kappa}, f \right) = \left[H_{\mathrm{Tx}} \left(\boldsymbol{\kappa}, f, q_\varphi \right), H_{\mathrm{Tx}} \left(\boldsymbol{\kappa}, f, q_\vartheta \right) \right]^{\mathrm{T}}$ is the *polarimetric radiation pattern* of the transmit antenna. It is the response of the antenna w.r.t. two EMWs of orthogonal polarisations q_ϑ and q_φ. The impedances Z_0 and $Z_{\mathrm{Tx}}(f)$ are the impedance of a plane wave in free space and the transmit antenna impedance, respectively. The phase is related to the phase centre of the antenna, from where the wave is assumed as radiated. Note, that $\mathbf{h}_{\mathrm{Tx}} \left(\boldsymbol{\kappa}, f \right)$ is a short hand notation for $\mathbf{h}_{\mathrm{Tx}} \left(\varphi, \vartheta, f \right)$.

Assume an EMW impinging under direction $\boldsymbol{\kappa}_{\mathrm{Rx}}$ at a receive antenna. The antenna output $X \left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f \right)$ is given by the polarimetric radiation pattern $\mathbf{h}_{\mathrm{Rx}} \left(\boldsymbol{\kappa}, f \right) = \left[H_{\mathrm{Rx}} \left(\boldsymbol{\kappa}, f, q_\varphi \right), H_{\mathrm{Rx}} \left(\boldsymbol{\kappa}, f, q_\vartheta \right) \right]^{\mathrm{T}}$

of the receive antenna, weighted with the electric field $\vec{E}(\boldsymbol{\kappa}, f)$ at the receive antenna [SW05].

$$X\left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f\right) = \sqrt{\frac{Z_{\mathrm{Rx}}(f)}{Z_0}} \cdot S(f) \cdot \left(\mathbf{h}_{\mathrm{Rx}}\left(\boldsymbol{\kappa}, f\right)^{\mathrm{T}} \cdot \vec{E}\left(\boldsymbol{\kappa}, f\right)\right) \cdot \delta\left(\boldsymbol{\kappa} - \boldsymbol{\kappa}_{\mathrm{Rx}}\right) \qquad (4.61)$$

Impedance $Z_{\mathrm{Rx}}(f)$ is the characteristic impedance of the receive antenna. Hence, the antenna output can be described as the convolution of the electric field vector with the polarimetric antenna response, sampled at the direction of impingement. For sake of convenience, the normalised polarimetric array pattern $\mathbf{a}_{\mathrm{Rx}}\left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f\right)$ and $\mathbf{a}_{\mathrm{Tx}}\left(\boldsymbol{\kappa}_{\mathrm{Tx}}, f\right)$ at Rx and Tx side are introduced, respectively.

$$\mathbf{a}_{\mathrm{Rx}}\left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f\right) = \sqrt{\frac{Z_{\mathrm{Rx}}(f)}{Z_0}} \cdot \mathbf{h}_{\mathrm{Rx}}\left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f\right) = \left[A_{\mathrm{Rx}}\left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f, q_\varphi\right), A_{\mathrm{Rx}}\left(\boldsymbol{\kappa}_{\mathrm{Rx}}, f, q_\vartheta\right)\right]^{\mathrm{T}} \quad (4.62\mathrm{a})$$

$$\mathbf{a}_{\mathrm{Tx}}\left(\boldsymbol{\kappa}_{\mathrm{Tx}}, f\right) = \frac{j2\pi f}{2\pi c_0} \cdot \sqrt{\frac{Z_0}{Z_{\mathrm{Tx}}(f)}} \cdot \mathbf{h}_{\mathrm{Tx}}\left(\boldsymbol{\kappa}_{\mathrm{Tx}}, f\right) = \left[A_{\mathrm{Tx}}\left(\boldsymbol{\kappa}_{\mathrm{Tx}}, f, q_\varphi\right), A_{\mathrm{Tx}}\left(\boldsymbol{\kappa}_{\mathrm{Tx}}, f, q_\vartheta\right)\right]^{\mathrm{T}}$$
$$(4.62\mathrm{b})$$

Summarised, the the output of an antenna and the radiated EMW is uniquely described by the polarimetric response of the antenna $\mathbf{a}(\boldsymbol{\kappa}, f)$. The polarimetric response is vectorial to account for two orthogonal polarisations, whereas each vector entry represents a LTI system.

4.3.2 Description of Antenna Arrays

An array of antennas is a composite of multiple antennas in a geometrical assembly [Bal97], which share a common reference point and whose feeding or output signals are coherently processed. The reference point of the array or the array centre is a point in space, typically chosen as the array's geometrical centre or centre of gravity. The array centre coincides with the origin of the spherical coordinate system and the array is described w.r.t. this point. Moreover, the propagation delay is defined to be the travelling time of the wave from the array geometrical centre at Tx to the array geometrical centre at Rx.

Array Geometry and Direction-Related Delay

The array geometry is the assembly of the antennas of the array. Several geometries are know, e.g. planar geometries like uniform linear array (ULA) or uniform rectangular array (URA), or conformal geometries as uniform circular array or stacked uniform circular array [Kil15]. The array geometry is realised by a mechanical construction or movement of a single antenna. The array geometry is described by means of a spatial vector $\mathbf{d}$ for each array antenna [Bal97, Kil15]. This vector describes the displacement of the phase centre of each antenna from the reference point of the array, see Fig. 4.20.

The spatial displacement $\mathbf{d}$ results in a time delay τ of an impinging/radiating wave w.r.t. the array reference point. Considering the Fraunhofer approximation, where far field conditions
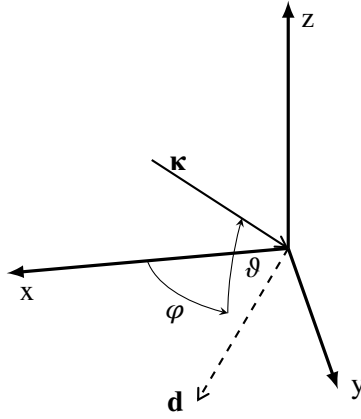
Figure 4.20: *Spatially displaced sampling at position* **d** *to transform the direction* **κ** *of impingement into a delay w.r.t. the coordinate system origin. (Picture adapted from [Hei09b].)*

and a plane wave (see Section 3.1.1) are assumed, this delay is given by [Bal97, Kil15]

$$\tau = \frac{1}{c_0} \mathbf{d}^{\mathrm{T}} \cdot \boldsymbol{\kappa}\left(\varphi, \vartheta\right) \quad , \tag{4.63}$$

with $\boldsymbol{\kappa}\left(\varphi, \vartheta\right)$ the normalised wave vector of the impinging/radiated wave. Accordingly, *the direction of impingement/radiation is mapped into a delay.* At receiver side this delay is

$$\tau_{\mathrm{Rx}} = \frac{1}{c_0} \mathbf{d}_{\mathrm{Rx}}^{\mathrm{T}} \cdot \boldsymbol{\kappa}_{\mathrm{Rx}}\left(\varphi_{\mathrm{Rx}}, \vartheta_{\mathrm{Rx}}\right) \quad , \tag{4.64}$$

and at transmitter side this delay is

$$\tau_{\mathrm{Tx}} = \frac{1}{c_0} \mathbf{d}_{\mathrm{Tx}}^{\mathrm{T}} \cdot \boldsymbol{\kappa}_{\mathrm{Tx}}\left(\varphi_{\mathrm{Tx}}, \vartheta_{\mathrm{Tx}}\right) \quad . \tag{4.65}$$

The overall delay due to the spatially displaced sampling at Tx and Rx side is

$$\tau = \frac{1}{c_0} \left[\mathbf{d}_{\mathrm{Rx}}^{\mathrm{T}} \cdot \boldsymbol{\kappa}_{\mathrm{Rx}}\left(\varphi_{\mathrm{Rx}}, \vartheta_{\mathrm{Rx}}\right) + \mathbf{d}_{\mathrm{Tx}}^{\mathrm{T}} \cdot \boldsymbol{\kappa}_{\mathrm{Tx}}\left(\varphi_{\mathrm{Tx}}, \vartheta_{\mathrm{Tx}}\right) \right] \quad . \tag{4.66}$$

Summarised, the array geometry relates the direction of radiation $\boldsymbol{\kappa}_{\mathrm{Tx}}$ or direction of impingement $\boldsymbol{\kappa}_{\mathrm{Rx}}$ into a delay w.r.t. the reference point of the array at transmitter or receiver side, respectively. This founds the basis for direction estimation with antenna arrays.

Antenna Response in the Array Composite

An antenna array consists of M antennas. Each antenna in the array compound is described by its response $A_m\left(\boldsymbol{\kappa}, f, q\right)$, which depends on the direction vector $\boldsymbol{\kappa}$, and polarisation q being q_φ or q_ϑ. Note, that $A_m\left(\boldsymbol{\kappa}, f, q\right)$ is now the response of the m-th antenna in the array composite, and not the characteristic of the antenna outside of the array! In the array composite, coupling and crosstalk between the array antennas occur, which change the characteristics of each antenna.

To proceed it will be assumed, that the Fraunhofer approximation holds for the whole array,

i.e. a plane wave is considered in the far field or to be propagating across the array. Hence, the delay of the observed wave depends on the displacement **d** of the antenna only and is not related to the distance R of the source or the far field reference point [Kil15]. A criterion to ensure, that the Fraunhofer approximation holds for the whole array, is [Kil15]

$$R \gg \frac{4\kappa_0}{\pi} D^2 = \frac{8D^2}{\lambda} \quad , \tag{4.67}$$

with D the radius of a sphere, which enclosures the array and whose centre coincides with the array reference point [Kil15]. Hence, in order to have the most compact representation of the array regarding the Fraunhofer approximation it is best practise to set the centre of gravity of the array as the array reference point, such that the sphere becomes smallest.

According to the Fraunhofer approximation the response of the m-th antenna to a plane wave from direction $\boldsymbol{\kappa}(\varphi, \vartheta)$ is

$$B_m(\boldsymbol{\kappa}, f, q) = A_m(\boldsymbol{\kappa}, f, q) \cdot \exp\left\{-j2\pi\frac{f}{c_0}\boldsymbol{\kappa}^{\mathrm{T}} \cdot \mathbf{d}_m\right\} \quad . \tag{4.68}$$

Because the antenna position $\mathbf{d}_m$ in the array is fixed and the wave vector $\boldsymbol{\kappa}$ depends on the direction only, the delay $\tau_m(\boldsymbol{\kappa}) = \tau_m(\varphi, \vartheta) = \frac{1}{c_0}\boldsymbol{\kappa}^{\mathrm{T}} \cdot \mathbf{d}_m$ is direction-dependent only. For sake of convenience the following abbreviation is introduced

$$B_m(\boldsymbol{\kappa}, f, q) = A_m(\boldsymbol{\kappa}, f, q) \cdot \exp\left\{-j2\pi f\tau_m(\boldsymbol{\kappa})\right\} \quad . \tag{4.69}$$

4.3.3 Modelling of an Antenna Array – Wideband and Narrowband Array Model

For practical application, realistic data of the antenna array response have to be considered for modelling, in order to incorporate antenna imperfections and mutual coupling. Typically, such data can be gained only on a finite measurement or simulation grid [24]. Sampled array responses in angular and frequency domain are not sufficient as model, as the array response has to be known at arbitrary sampling points [24, LD04]. Furthermore, storage saving compact representation of sampled array responses and the efficient reconstruction is quite important. Thus, an array model is necessary to continuously reconstruct or interpolate the discretised array responses. Such models can be derived from wave field modelling [DD94, 24]. There, the sampled responses of each array antenna are decomposed in a sampling matrix and a direction and frequency dependent basis vector, why this kind of array models are stated as *algebraic array models*.

Wideband Array Model

Equation (4.69) is the model of the (wideband) response of an antenna in the array compound. Stacking the responses of all M antennas into a vector yields the *array response vector*.

$$\mathbf{b}(\boldsymbol{\kappa}, f, q) = \left[B_1(\boldsymbol{\kappa}, f, q), \dots, B_M(\boldsymbol{\kappa}, f, q)\right]^{\mathrm{T}} \in \mathbb{C}^M \tag{4.70}$$

The array response vector $\mathbf{b}(\boldsymbol{\kappa}, f, q)$ is a LTI system of multiple input (receiver) or multiple output (transmitter) ports. Known algebraic models to continuously describe gridded wideband array data are the scalar spherical harmonics (SSH) with Padé approximation [AW03], vector spherical harmonics (VSH) with Slepian mode expansion (SME) [DR10], VSH with singularity expansion method (SEM) [Rob06] or the effective time-aperture distribution function (ETADF) [24]. Among the other approaches, the ETADF approach is the most promising one and will be used. The ETADF necessitates more coefficients and therefore does not provide the most compact description of the data. However, because simple Fourier basis functions are employed, the numerical complexity is lower compared to e.g. VSH or SSH basis functions. Furthermore, the first order derivative can be easily calculated.

Narrowband Array Model

Commonly, the narrowband approximation of the array response is used in array signal processing, see e.g. [Nai09, TF09, Vas+20]. Advantage is the reduced numerical complexity and the ability to apply numerical efficient algorithms for direction estimation as e.g. the multiple signal classification (MUSIC) or estimation of signal parameters via rotational invariance technique (ESPRIT).

Consider a point source in the far field, whose signal impinges as a plane wave at the array and having centre frequency f_c. Introduce the passband signals $\tilde{s}(t)$ and $\tilde{\mathbf{b}}(\boldsymbol{\kappa}, t, q)$ of the source signal and the array response, respectively.

$$\tilde{\mathbf{b}}(\boldsymbol{\kappa}, t, q) = \mathbf{b}(\boldsymbol{\kappa}, t, q) \cdot \exp\left\{ j2\pi f_c t \right\} \quad \circ\!\!-\!\!\bullet \quad \mathbf{b}(\boldsymbol{\kappa}, f - f_c, q) \tag{4.71a}$$

$$\tilde{s}(t) = s(t) \cdot \exp\left\{ j2\pi f_c t \right\} \quad \circ\!\!-\!\!\bullet \quad S(f - f_c) \tag{4.71b}$$

Signal $s(t)$ and $\mathbf{b}(\boldsymbol{\kappa}, q, t)$ are the respective complex envelops. The array output $\mathbf{x}(t)$ at baseband is given by the convolution of the array response and the source signal at passband, and subsequent down-conversion.

$$\mathbf{x}(t) = \exp\left\{ -j2\pi f_c t \right\} \cdot \left[\tilde{\mathbf{b}}(\boldsymbol{\kappa}, t, q) * \tilde{s}(t) \right] \tag{4.72}$$

Consider the frequency response $\tilde{B}_m(\boldsymbol{\kappa}, f, q)$ of the m-th antenna. Decompose the magnitude $K_m(\boldsymbol{\kappa}, f, q)$ into a constant part at centre frequency $K_m(\boldsymbol{\kappa}, f_c, q)$ and a remaining part $K'_m(\boldsymbol{\kappa}, f, q)$, and decompose the phase $\phi_m(\boldsymbol{\kappa}, f, q)$ into a linearly varying part $\phi'_m(\boldsymbol{\kappa}, f, q)$ and a remaining part $\phi''_m(\boldsymbol{\kappa}, f, q)$.

$$\begin{aligned} \tilde{B}_m(\boldsymbol{\kappa}, f, q) &= K_m(\boldsymbol{\kappa}, f, q) \cdot \exp\left\{ j\phi_m(\boldsymbol{\kappa}, f, q) \right\} \\ &= K'_m(\boldsymbol{\kappa}, f, q) \cdot \exp\left\{ j\phi''_m(\boldsymbol{\kappa}, f, q) \right\} \cdot K_m(\boldsymbol{\kappa}, f_c, q) \cdot \exp\left\{ j\phi'_m(\boldsymbol{\kappa}, f, q) \right\} \end{aligned} \tag{4.73}$$

Then, the output signal at passband in frequency domain, $\tilde{X}_m(f)$, is given by

$$\begin{aligned} \tilde{X}_m(f) &= S(f - f_c) \cdot K'_m(\boldsymbol{\kappa}, f, q) \cdot \exp\left\{ j\phi''_m(\boldsymbol{\kappa}, f, q) \right\} \cdot K_m(\boldsymbol{\kappa}, f_c, q) \cdot \exp\left\{ j\phi'_m(\boldsymbol{\kappa}, f, q) \right\} \\ &= S'_m(f - f_c) \cdot K_m(\boldsymbol{\kappa}, f_c, q) \cdot \exp\left\{ j\phi'_m(\boldsymbol{\kappa}, f, q) \right\} \quad . \end{aligned} \tag{4.74}$$

The linear part $\phi'_m(\boldsymbol{\kappa}, f, q)$ of the phase is mainly determined by the direction dependent delay $\tau_m(\boldsymbol{\kappa})$, but also by the delaying properties of the antenna. It can be determined by the Taylor

series expansion of the phase $\phi_m(\boldsymbol{\kappa}, f, q)$ at centre frequency.

$$\phi'_m(\boldsymbol{\kappa}, f, q) = \phi_m(\boldsymbol{\kappa}, f_c, q) + \left.\frac{\partial \phi_m(\boldsymbol{\kappa}, f, q)}{\partial f}\right|_{f=f_c} \cdot (f - f_c) \tag{4.75}$$

Introduce the *phase delay* $\tau_{p_m}(\boldsymbol{\kappa}, q)$ and the *group delay* $\tau_{g_m}(\boldsymbol{\kappa}, q)$.

$$\tau_{p_m}(\boldsymbol{\kappa}, q) = -\frac{\phi(\boldsymbol{\kappa}, f_c, q)}{2\pi f_c} \tag{4.76a}$$

$$\tau_{g_m}(\boldsymbol{\kappa}, q) = -\frac{1}{2\pi} \left.\frac{\partial \phi(\boldsymbol{\kappa}, f, q)}{\partial f}\right|_{f=f_c} \tag{4.76b}$$

Thus, the output at baseband in time domain can be given by [Hay94]

$$x_m(t) = s'_m\left(t - \tau_{g_m}(\boldsymbol{\kappa}, q)\right) \cdot K_m(\boldsymbol{\kappa}, f_c, q) \cdot \exp\left\{-\jmath 2\pi f \tau_{p_m}(\boldsymbol{\kappa}, q)\right\} \quad . \tag{4.77}$$

Signal $s'_m\left(t - \tau_{g_m}(\boldsymbol{\kappa}, q)\right)$ is the complex envelope of the transmit signal after transmission over the antenna. Assume, that the product of signal bandwidth W and the largest occurring group delay fulfils

$$W \cdot \max\left\{\left|\tau_{g_m}(\boldsymbol{\kappa}, q)\right|\right\} \ll 1 \quad . \tag{4.78}$$

As a result, the group delay of the antenna response becomes negligible. Furthermore, the complex envelope of the source signal shall not be significantly changed by the antenna response. Summarised, approximation $s'_m\left(t - \tau_{g_m}(\boldsymbol{\kappa}, q)\right) \approx s(t)$ holds [TF09], and the source signal is said to be *narrowband* w.r.t. the array. Thus, the baseband signal can be simplified.

$$x_m(t) \approx s(t) \cdot K_m(\boldsymbol{\kappa}, f_c, q) \cdot \exp\left\{-\jmath 2\pi f_c \tau_{p_m}(\boldsymbol{\kappa}, q)\right\} \tag{4.79}$$

Summarised, narrowband approximation assumes the array response as *bandwidth independent*, such that the direction dependent delay is approximated as a phase change at centre frequency. Hence, the response of each antenna in the array composite can be approximated by a constant amplitude and phase at centre frequency.

Consider the Taylor series expansion [AS72] of the passband response of the m-th antenna at centre frequency.

$$\tilde{B}_m(\boldsymbol{\kappa}, f, q) = \sum_{k=0}^{\infty} \frac{1}{k!} \cdot \left.\frac{\partial^k \tilde{B}_m(\boldsymbol{\kappa}, f, q)}{\partial f^k}\right|_{f=f_c} \cdot (f - f_c)^k \tag{4.80}$$

Then, the narrowband approximation of $\tilde{B}_m(\boldsymbol{\kappa}, f, q)$ is given by the zeroth-order Taylor series coefficient $\tilde{B}_m(\boldsymbol{\kappa}, f_c, q)$, which can be expressed as

$$\tilde{B}_m(\boldsymbol{\kappa}, f_c, q) = \tilde{A}_m(\boldsymbol{\kappa}, f_c, q) \cdot \exp\left\{-\jmath 2\pi f_c \tau_m(\boldsymbol{\kappa})\right\} \quad . \tag{4.81}$$

Stacking the narrowband response approximations of all M array antennas into a vector yields the narrowband array response at centre frequency, commonly stated as *array steering vector*.

$$\mathbf{b}\left(\boldsymbol{\kappa}, q\right) = \left[\tilde{B}_1\left(\boldsymbol{\kappa}, f_{\mathrm{c}}, q\right), \ldots, \tilde{B}_M\left(\boldsymbol{\kappa}, f_{\mathrm{c}}, q\right)\right]^{\mathrm{T}} \tag{4.82}$$

Known algebraic models to continuously describe the steering vector from gridded array data are e.g. the effective aperture distribution function (EADF) [LD04], the SSH [Del+06, SKK12] or the VSH [Koi04]. Here, the EADF will be used, because of its numerical simplicity in terms of the used basis functions, and because the first order derivative w.r.t. the angles can be easily calculated.

Narrowband Validation Criterion   A common criterion for narrowband is a small relative bandwidth W/f_{c} [Ric05, Zat97]. However, this definition is crucial, because only the signal bandwidth is taken into account. Frequency dependence of the array antennas or the influence of the array geometry is neglected. Hence, it cannot be ensured that the narrowband assumption is fulfilled by a small relative signal bandwidth.

A criterion which includes antenna and signal properties is stated in [Lan08]: The narrowband assumption is fulfilled, if the frequency response of each antenna is constant in magnitude and phase over the signal bandwidth. A method to verify this criterion is missed.

Another narrowband criterion can be derived from communication system theory, where a communication signal is assigned to undergo a frequency selective or frequency flat radio channel. A communication signal with a bandwidth smaller than the coherence bandwidth of the radio channel is designated as narrowband, and the radio channel can be assumed as frequency flat (frequency independent) [Mol12]. Such a coherence bandwidth can be similarly assigned to an antenna array. Depending on the direction, each array antenna introduces a delay on the impinging wave w.r.t. the array phase centre. Summing up all delayed waves leads to fading in the frequency domain (frequency selectivity). This frequency selective behaviour is described by the array transfer function $H(\boldsymbol{\kappa}, f, q)$, which is the average frequency response over all array antennas [Nai09].

$$H\left(\boldsymbol{\kappa}, f, q\right) = \frac{1}{M} \cdot \sum_{m=1}^{M} B_m\left(\boldsymbol{\kappa}, f, q\right) \tag{4.83}$$

Knowing the array transfer function, the direction dependent coherence bandwidth $W_{\mathrm{coh}}(\varphi, \vartheta)$ can be estimated. In general, the smallest coherence bandwidth over the unit sphere should be used to address the coherence bandwidth of the array

$$W_{\mathrm{coh}} = \min\left\{W_{\mathrm{coh}}\left(\varphi, \vartheta\right)\right\} \quad . \tag{4.84}$$

Accordingly, a new narrowband condition can be stated [26].

Definition 1. An antenna array can be sufficiently described by a narrowband model, if the signal bandwidth W is much smaller than the array coherence bandwidth W_{coh} in the frequency band under consideration.

For directive and non-isotropic antenna characteristics the coherence bandwidth increases, because only neighbourhood antennas receive the impinging wave [Jac+15]. In Fig. 4.21 the
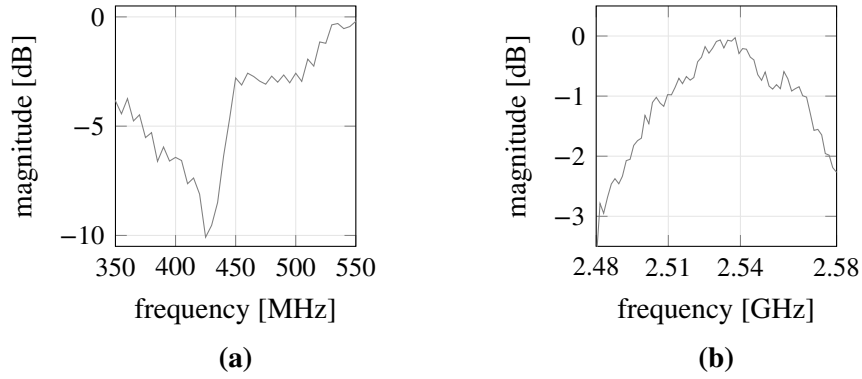
Figure 4.21: *Normalised magnitude of transfer function of (a) circular dipole array [23] and (b) circular patch array [37].*

transfer function of two (UCAs), consisting of dipole or patch antennas, are depicted. Note, that both arrays are calibrated at different centre frequencies. The UCA with patch antennas shows less variations of the magnitude compared to the dipole array. Therefore, the frequency selectivity is less severe and the coherence bandwidth is larger.

Constraints for Antenna Spacing Under Narrowband and Wideband Modelling

According to the modelling of the array by a narrowband or wideband approach and consideration of these models for direction estimation, the spatial displacement of the array antennas becomes constrained. As stated previously, estimation of the angles coincides with a time difference estimation, whereas in the narrowband case the time difference transforms into a phase difference only. The delay difference $\Delta\tau$ between array antennas 1 and 2 can be expressed in terms of their baseband signals as

$$x_1(t) = x_2(t - \Delta\tau) \quad \circ\!\!-\!\!\bullet \quad X_1(f) = X_2(f) \cdot \exp\{-j2\pi f \Delta\tau\} \quad . \tag{4.85}$$

The delay difference is given by

$$\Delta\tau = \frac{1}{c_0}\boldsymbol{\kappa}^{\mathrm{T}} \cdot \left(\mathbf{d}_1 - \mathbf{d}_2\right) = \frac{1}{c_0}\boldsymbol{\kappa}^{\mathrm{T}} \cdot \Delta\mathbf{d} \quad . \tag{4.86}$$

In the following, constraints for the antenna displacement $\Delta\mathbf{d}$ in order to ambiguous free estimate the delay difference and hence the angles φ and ϑ will be discussed.

Narrowband Model In the narrowband case, the phase term $\exp\{-j2\pi f \Delta\tau\}$ is examined at the centre frequency only. For an unambiguous phase term the argument must be in the range $]-\pi, \pi]$, such that

$$\left|2\pi f_{\mathrm{c}}\Delta\tau\right| = \left|\frac{2\pi f_{\mathrm{c}}}{c_0}\boldsymbol{\kappa}^{\mathrm{T}} \cdot \Delta\mathbf{d}\right| < \pi \tag{4.87}$$

must hold. Considering the Cauchy–Schwartz inequality[2] the magnitude becomes

$$\left| \frac{2\pi f_{\mathrm{c}}}{c_0} \boldsymbol{\kappa}^{\mathrm{T}} \cdot \Delta \mathbf{d} \right| = \frac{2\pi f_{\mathrm{c}}}{c_0} \left| \boldsymbol{\kappa}^{\mathrm{T}} \cdot \Delta \mathbf{d} \right| \leq \frac{2\pi f_{\mathrm{c}}}{c_0} \| \boldsymbol{\kappa} \| \cdot \| \Delta \mathbf{d} \| = \frac{2\pi f_{\mathrm{c}}}{c_0} \cdot \| \Delta \mathbf{d} \| \quad . \tag{4.88}$$

Hence, the following well known constraint for the antenna displacement is given [Nai09, TF09]

$$\| \Delta \mathbf{d} \| \leq \frac{\lambda_c}{2} \quad . \tag{4.89}$$

Wideband Model   In the wideband case multiple frequencies are considered instead of the centre frequency only. Assuming a discretised spectrum with spacing $\Delta f = 1/T$ and the observation duration T. The consecutive frequency bins f_1 and $f_2 = f_1 + \Delta f$ feature a directional dependent phase difference, which is

$$\exp\left\{ -j2\pi f_2 \Delta \tau \right\} = \exp\left\{ -j2\pi \left(f_1 + \Delta f \right) \Delta \tau \right\} = \exp\left\{ -j2\pi f_1 \Delta \tau \right\} \cdot \exp\left\{ -j2\pi \Delta f \Delta \tau \right\} \quad . \tag{4.90}$$

For an unambiguous phase difference the following criterion must be fulfilled

$$\left| \frac{2\pi \Delta f}{c_0} \boldsymbol{\kappa}^{\mathrm{T}} \cdot \Delta \mathbf{d} \right| = \frac{2\pi \Delta f}{c_0} \left| \boldsymbol{\kappa}^{\mathrm{T}} \cdot \Delta \mathbf{d} \right| \leq \pi \quad . \tag{4.91}$$

Considering again the Cauchy–Schwartz inequality the antenna displacement is constrained by [CG92]

$$\| \Delta \mathbf{d} \| \leq c_0 \cdot \frac{T}{2} \quad . \tag{4.92}$$

Summarised, the antenna displacement is differently constrained if a narrowband or wideband modelling approach is considered for direction estimation. This must be taken into account, if one has to chose between both modelling approaches. Note, that the displacement constraints can be relaxed, if directive array antennas are employed [Jac+15].

4.3.4 MIMO Topology – Displacement of Tx and Rx Array

For MIMO systems as depicted in Fig. 4.22 not only the array geometry is of importance. Also the arrangement of the transmit and receive array w.r.t. to each other, which will be denoted as MIMO topology, is important. Especially the spatial separation of transmit and receive array is of importance to describe the spatial observation of a target by the MIMO radar system. Because an array consists of spatially distributed antennas, it can be stated that each antenna focuses the target under *different angles*. Under far field conditions it is assumed, that all array antennas focus the target under the *same angles* as a plane wave travels across the array. Subsequently,

---

[2]$\left| \mathbf{x}^{\mathrm{T}} \cdot \mathbf{y} \right| \leq \| \mathbf{x} \| \cdot \| \mathbf{y} \|$
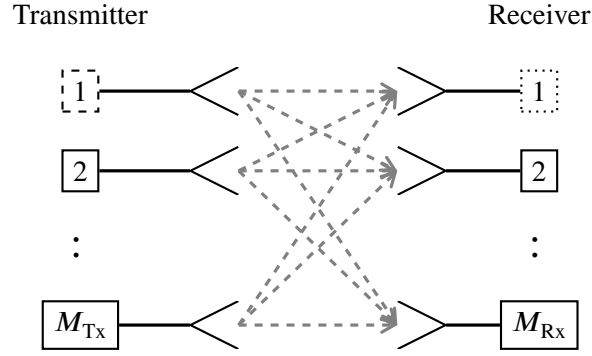
Figure 4.22: *Sketch of MIMO system, with a Tx channel (dashed box), a Rx channel (dotted box) and a MIMO channel (solid box).*

far field condition (i.e. plane waves) for the array at Tx and Rx side will be assumed, such that all antennas at Tx focus the target under the same angle and all antennas at Rx focus the target under the same angles. However, the angles at Tx and Rx side are not necessarily equal and remain subject to clarification, which depends on the array topology. Two topologies will be distinguished according to the spacing of transmit and receive array: *monostatic* and *bistatic*, see Fig. 4.23. Both topologies are different in terms of the separation of Tx and Rx array and the resulting focusing angles w.r.t. the target. Subsequently, both topologies will be briefly described and their influence on the spatial delay (4.66) will be discussed.

Bistatic or Distributed Array Topology

For the bistatic or distributed topology sketched in Fig. 4.23(b), transmit and receive array are widely separated [HBC08], such that the focusing angles and ranges to the target are "significantly different", as defined in the "IEEE Standard Radar Definitions" [Pan17] or see [Wil05]. Therefore, the far field condition for each array holds, but not for the complete MIMO system. Hence, Tx and Rx have different focusing angles to the target. Figure 4.20 shows a bistatic arrangement. Advantage of widely separated arrays is, that the target is focused under different angles and hence different parts of the RCS of the target are viewed, see Fig. 3.9 for a respective sketch of this aspect. If the RCS of the target is spatially diverse, hence varies w.r.t. the focusing angles, the possibility of target detection might be increased [HBC08]. The delay by joint processing of the Tx and Rx channel is

$$\tau_{\text{bi}} = \frac{1}{c_0} \left[\mathbf{d}_{\text{Rx}}^{\text{T}} \cdot \boldsymbol{\kappa}_{\text{Rx}} \left(\varphi_{\text{Rx}}, \vartheta_{\text{Rx}} \right) + \mathbf{d}_{\text{Tx}}^{\text{T}} \cdot \boldsymbol{\kappa}_{\text{Tx}} \left(\varphi_{\text{Tx}}, \vartheta_{\text{Tx}} \right) \right] \quad . \tag{4.93}$$

Hence, a virtual array is also present, which is not as easily understandable as in the monostatic case, see [Yan+13] for more details. From a modelling point of view, a bistatic topology requires a separate modelling of Tx and Rx array, because of the different focusing angles.

Monostatic or co-located Array Topology

For the monostatic or co-located topology sketched in Fig. 4.23(a), Tx and Rx array are closely located and their reference points are assumed to coincide. Hence, the term co-located

Figure 4.23: *MIMO topology (a) monostatic (co-located Tx and Rx) and (b) bistatic (distributed Tx and Rx) [Sko61]*

arrays [HV10, LS07, LS08] is often stated. co-located arrays focus the target under the same angles, such that $\varphi = \varphi_{\mathrm{Tx}} = \varphi_{\mathrm{Rx}}$ and $\vartheta = \vartheta_{\mathrm{Tx}} = \vartheta_{\mathrm{Rx}}$ holds. By joint processing of the Tx and Rx channels the delay in equation (4.66) becomes

$$\tau_{\mathrm{mono}} = \frac{1}{c_0} \left[\mathbf{d}_{\mathrm{Rx}} + \mathbf{d}_{\mathrm{Tx}}\right]^{\mathrm{T}} \cdot \boldsymbol{\kappa}\left(\varphi, \vartheta\right) \quad . \tag{4.94}$$

Hence, because Tx and Rx array are co-located, the spatial sampling by the MIMO array is conducted at the virtual point $\mathbf{d}_{\mathrm{Rx}} + \mathbf{d}_{\mathrm{Tx}}$. A virtual sampling point exists for each MIMO channel, such that a virtual array of increased size [LS07] is present. The virtual array aperture offers higher angular resolution, the ability to resolve more targets and better parameter identifiability [LS07, LS08]. From a modelling point of view, a monostatic topology enables a joint modelling of Tx and Rx array, because of the unique focusing angles. Hence, the MIMO channels can be described as the channels of a single, virtual array.

A note w.r.t. the joint processing of Tx and Rx array is necessary. If both arrays shall be jointly processed to get the virtual array (co-located case), the Tx and Rx have to be *coherent* to each other. Coherence can be ensured by the radar hardware or by calibration [Dür+20]. Here, a co-located MIMO radar is considered and a fully coherent system will be assumed. If both arrays are separately processed as e.g. separate angles shall be estimated at both sides (distributed case), the measurements over the Rx and the measurements over the Tx have to be coherent each. However, coherence of Tx and Rx w.r.t. each other is not demanded.

4.4 Joint Parametrisation of Antenna Array and Transceiver Model

The model of the antenna array (see Section 4.3) as well as the response model of the radar (see Section 4.2) has to be calibrated from measurements. Theoretical or simulation-based determination of the array or device response is not suitable under practical considerations, because of mechanical and electrical imperfections, which are not account for in the simulations

(a)　　　　　　　　　　　　　　　　　**(b)**

Figure 4.24: *Mounting of (a) the array in [25] and (b) the reference antenna in an anechoic chamber for array calibration measurements. (Pictures by Michael Huhn, RF and Microwave Research Group, TU Ilmenau)*

or the theoretical models [7]. Therefore, calibration measurements will be employed [Dür+19, Vas+20].

Commonly, the array response and the device response are measured separately by detaching the array from the device [25]. Array calibration can be conducted in an anechoic chamber or a test range using a well known calibration system [7]. The array is attached to a 3-D positioner system and a fix reference antenna is set up. Figure 4.24 shows the setup for the calibration of an 30 GHz antenna array [25]. The spatial characteristic is captured according to the azimuth and elevation angle by moving the array w.r.t. the origin of the coordinate system of the 3-D positioner. Polarimetric properties are captured by employing two orthogonal polarisations at the reference antenna. The frequency characteristic of the array is captured by wideband measurements. The response of the device can be captured by back-to-back measurements using a well known reference or by over-the-air (OTA) calibration with the LOS propagation channel as reference.

However, if the array cannot be detached from the device as e.g. in case of an integrated MIMO radar [2], array and device response are jointly captured by the calibration measurements and other processing techniques are required to separate them. In Section 4.4.1 the separation of array and device response under narrowband approximation is discussed. Because the narrowband approximation is crucial in case of a high signal bandwidth and/or large arrays [26], a wideband array model and its parametrisation from calibration measurements is discussed in Section 4.4.2. The wideband and narrowband array model will be compared in Section 4.4.3.

The monostatic MIMO radar presented in [2] is exemplary considered in the following. The calibration measurements have been conducted by university Ulm. Due to the monostatic setup, Tx and Rx array are co-located and a virtual array is present, see Section 4.3.4. The response of this virtual array depends only on two angles: azimuth and elevation, which are equal for Tx and Rx. Therefore, array calibration has to be conducted according to these two angles only. Contrarily, the joint array response of a bistatic MIMO radar depends on four angles: azimuth and elevation at Tx and at Rx. Therefore, calibration measurements have to be conducted according to these four angles. The radar has been mounted on a positioner, and a corner reflector has been set as target. Because Tx and Rx array are linear arrays and are arranged in the same plane, the virtual array is a linear array too. Hence, calibration measurements in the azimuth plane only have been conducted, by rotating the radar using the positioner device.

4.4.1 Joint Estimation of Narrowband Array Response and Transceiver Response

The separation of the system and array response from one set of calibration measurements necessitates a model describing the embedding of the antenna and observer response in the measurements. The joint response of array and device is measured over multiple angles $\kappa(\varphi, \vartheta)$ and frequencies f, in order to properly characterise the array. Note, that only a single polarisation will be considered in the following. In total, the calibration measurements are conducted at N different angles, K samples are captured per channel and the array constitutes of M antennas.

Subsequently, narrowband approximation of the antenna array response will be considered, see Section 4.3. This approximation is only meaningful as long as the signal bandwidth is smaller than the coherence bandwidth of the array [26], such that (4.78) is fulfilled. This will be assumed in the following. Furthermore, it is assumed, that each array antenna is driven by an individual transmitting/receiving chain of the observer system. Hence, the calibration measurement of each antenna m is confounded by a different frequency response $g_m(f)$, on the contrary to switched arrays connected to a SISO system. The responses are assumed to undergo variations over the considered frequency band. For sake of convenience, the frequency response of all transmitting/receiving channels is concatenated in vector $\mathbf{g}(f) \in \mathbb{C}^M$. Furthermore, coupling between the array antennas is assumed, which will be described by the matrix $\mathbf{C} \in \mathbb{C}^{M \times M}$, which is independent on the frequency and the angles of incidence/radiation. Therefore, the coupling matrix describes the average, global coupling between the antennas and the transmitting/receiving channels of the observer. Measurement errors and noise are accounted for by an additive term $\mathbf{w}(f) \in \mathbb{C}^M$, which describes a normally distributed and proper complex random process with zero mean. Summarised, the data model for a calibration measurement $\mathbf{y}(\kappa, f)$ at certain frequency and angles is [26]

$$\mathbf{y}(\kappa, f) = \mathbf{C} \cdot \mathrm{diag}\{\mathbf{g}(f)\} \cdot \mathbf{b}(\kappa) + \mathbf{w}(f) \quad . \tag{4.95}$$

The steering vector $\mathbf{b}(\kappa)$ and the observer response $\mathbf{g}(f)$ has to be estimated from the calibration measurements $\mathbf{y}(\kappa, f)$, while accounting for the coupling $\mathbf{C}$. The estimation will be accomplished by tensor decomposition of the calibration data. For sake of convenience the calibration measurements are concatenated in the tensor $\mathcal{Y} \in \mathbb{C}^{M \times K \times N}$. As published in [26] tensor decomposition methods can be used to identify $\mathbf{g}(f)$, $\mathbf{b}(\kappa)$ and $\mathbf{C}$ from the measurements $\mathcal{Y}$.

Joint Estimation by Tensor Decomposition

Tensor decomposition is a processing method to analyse multidimensional data sets [Bro+09, DDV00, KB09]. According to a model of the multidimensional data set, the data set features an underlying, hidden structure. Tensor decomposition methods take advantage of the data model in order to identify this structure. For example, the data model in equation (4.95) describes a multidimensional outer product, which can be identified by rank decomposition methods.

The parallel factors (PARAFAC) decomposition [KB09] is a widely used tensor rank decomposition, which describes the tensor as a sum of L rank-one tensors. A rank-one tensor is given by the multidimensional outer product of several vectors. These vectors are stacked into
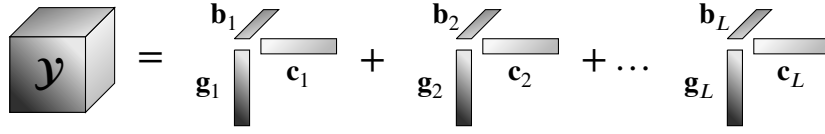
Figure 4.25: *PARAFAC decomposition of a three-way data tensor $\mathcal{Y}$ into a sum of L rank-one tensors.*

matrices yielding the set of decomposition or loading matrices $\mathbf{D}_i$ according to the number of tensor dimensions. The PARAFAC decomposition is unique, i.e., the resulting decomposition matrices are unique up to a scaling and permutation of their columns, if $\sum_i \text{k-rank} \{\mathbf{D}_i\} \geq 2L + 2$ holds [Kru77, Ste09]. L is the considered rank of the tensor for the decomposition and k-rank denotes Kruskal's rank of a matrix. The k-rank of a matrix is defined as the maximum number of columns of the matrix in a subset of them, such that every subset is linearly independent [Ste09].

Tensor decomposition of Tucker type are conceivable, e.g. the higher-order singular value decomposition [DDV00]. On the contrary to the PARAFAC decompositions, Tucker decompositions are more flexible w.r.t. the data, because the core tensor is not restricted to be diagonal and unity. Drawback of a Tucker decomposition for the problem under consideration is, that the decomposition matrices cannot be directly related to the array or observer response due to the arbitrary core tensor. Consequently, the array and observer response cannot be gained from a Tucker decomposition.

The PARAFAC decomposition has been proposed in [26] to jointly estimate the steering vector and the device response from the three-way calibration data tensor $\mathcal{Y}$. The respective PARAFAC decomposition is graphically sketched in Fig. 4.25, whereas the loading matrices are

- $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_L] \in \mathbb{C}^{K \times L}$: matrix of observer responses

- $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_L] \in \mathbb{C}^{N \times L}$: matrix of steering vectors

- $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_L] \in \mathbb{C}^{M \times L}$: coupling matrix   .

Hence, by applying PARAFAC decomposition to the tensor of calibration measurements, the array and device response can be gained from the loading matrices. In order to estimate the response of all device transceiver channels and their attached antennas, the tensor rank for the decomposition is set to $L = M$. According to the criterion for a unique PARAFAC decomposition, it must hold

$$\text{k-rank} \{\mathbf{G}\} + \text{k-rank} \{\mathbf{B}\} + \text{k-rank} \{\mathbf{C}\} \geq 2M + 2 \quad . \tag{4.96}$$

This criterion is fulfilled, if the matrices are of full column rank. The matrices are of full column rank, if the number of effective antennas, hence the number of antennas which contribute information from the observed scenario, is equal to the number of physical antennas. The number of effective antennas can be deteriorated, if the coupling between the antennas is large, such that one or multiple antennas do not contribute any additional information.

Example 4.1. Consider a theoretical array of M antennas, whereas all antennas are located at the same spatial position and have the same response. Hence, the number of effective antennas
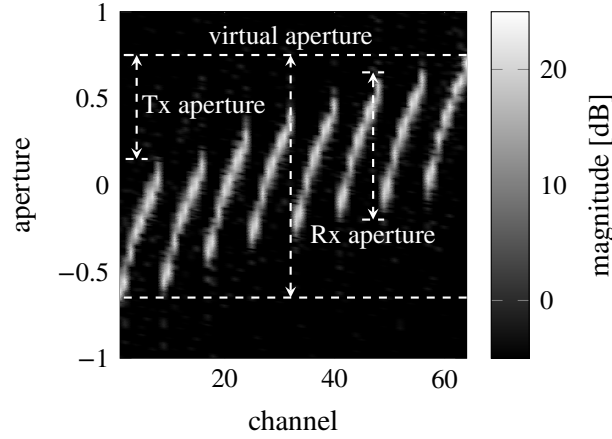
Figure 4.26: *Aperture of the* 8×8 *MIMO radar in [2]. Because Tx and Rx array are co-located, their respective apertures combine to a virtual aperture of larger size.*

should be 1, because all other $M - 1$ antennas do not contribute any additional information of the observed scenario.

The PARAFAC decomposition is applied to calibration measurements of two MIMO radars. For the decomposition it is assumed that the number of effective antennas is equal to the number of physical antennas. First, calibration measurements of a monostatic MIMO radar in [2] are considered. The estimated aperture (Fourier transform of the steering vectors over azimuth domain) for all MIMO channels is depicted in Fig. 4.26. The aperture of the Rx array only is represented by eight single maxima. Because Tx and Rx array are co-located and the respective channels are coherent, the apertures of both arrays are "convolved" with each other. This results in eight shifts of the Rx aperture in the aperture domain. The shifted Rx aperture gives the resulting virtual aperture, whose size (equal to a bandwidth) is larger than the size of the Rx only aperture. Hence, the angular resolution is increased. Second, calibration measurements of a 16x16 binocular MIMO radar are considered. The the narrowband response is again estimated by the PARAFAC decomposition. The aperture domain is shown in Fig. 4.27. The binocular radar consists of two monostatic and two bistatic MIMO radars, which all operate coherently. The aperture domain shows 4 blocks, whereas each block corresponds to the virtual aperture of a monostatic or bistatic MIMO radar. For a coherent processing of the observations of all 4 MIMO radars the resulting virtual aperture is sketched in Fig. 4.27. The derived narrowband array response is provided to the EADF approach [LD04], in order to get a continuous description of the array response in the angular domain.

4.4.2 Joint Estimation of Wideband Array Response and Transceiver Response

If a wideband array model is considered, hence the array response is assumed as *bandwidth dependent*, separation of the overall response into the array and observer response is not necessary. The observer response can be swapped to the wideband array model [25], if the device response is describable as an LTI system. Consequently, the calibration measurements itself are the wideband array description. The calibration measurements are provided to the
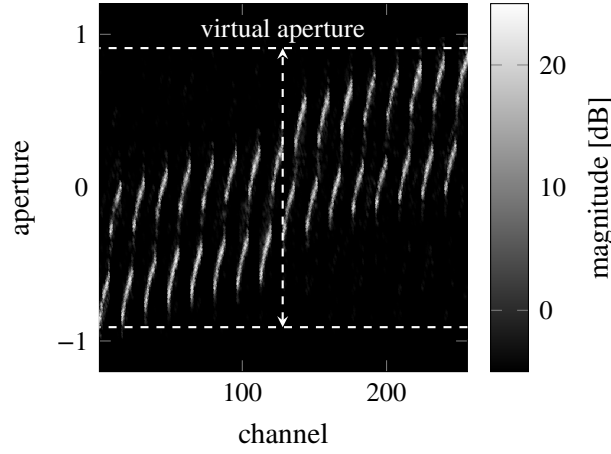
Figure 4.27: *Aperture of a* 16×16 *binocular MIMO radar (see Section 7.4.2). The apertures of the two monostatic (lower left and upper right) and the two bistatic (upper left and lower right) MIMO arrays are visible.*

ETADF approach [24] to get a grid free description of the response in angular and frequency domain.

However, if an estimate of the response of the observer is of necessity due to e.g. modelling purposes, a reference response has to be defined. Here, the response of the observer for 0° impingement/radiation direction will be used as reference response. Note, that this response is not the response of the transmitting/receiving chains of the observer only, as the antennas are still included. The respective measurement is taken from the calibration data, and the calibration data w.r.t. all other directions are deconvolved by this measurement. As a result the calibration data are now related to the response of the observer to an impinging/radiated wave at 0°.

Denoising by Singular Value Decomposition

Calibration measurements are confounded by additive errors as e.g. thermal noise or parasitic reflections from the mechanical measurement setup. In order to reduce such errors, decomposition of the calibration data as described in [26] can be applied. In the following, the MIMO radar in [2] will be considered, which has a linear array at Tx and Rx side. The angle dependent responses of all MIMO channels are measured in an anechoic chamber by placing a corner reflector w.r.t. the azimuth plane in front of the radar. Because the radar features linear arrays, only azimuth angles are varied by rotating the radar on a turntable. Note, that the Tx and Rx channels are jointly measured and not separately. Because a wideband array model, i.e. a frequency dependent array response, is considered, joint processing of all calibration data as described in Section 4.4.1 is not applicable. The processing has to be carried out for each frequency and angle separately. A model of the frequency and angle dependent measurements is given by

$$\mathbf{y}\left(\varphi_{\mathrm{Tx}}, \varphi_{\mathrm{Rx}}, f\right) = \mathbf{b}_{\mathrm{Tx}}\left(\varphi_{\mathrm{Tx}}, f\right) \otimes \mathbf{b}_{\mathrm{Rx}}\left(\varphi_{\mathrm{Rx}}, f\right) + \mathbf{w}\left(f\right) \quad . \tag{4.97}$$

The considered radar has co-located Tx and Rx array. Hence, the angles are equal $\varphi_{\mathrm{Tx}} \equiv \varphi_{\mathrm{Rx}} \equiv \varphi$, and the measurements become $\mathbf{y}\left(\varphi_{\mathrm{Tx}}, \varphi_{\mathrm{Rx}}, f\right) \Rightarrow \mathbf{y}\left(\varphi, f\right)$. The measurement vector

$\mathbf{y}(\varphi, f) \in \mathbb{C}^{M_{\text{Tx}} \cdot M_{\text{Rx}}}$ will be reshaped in a matrix $\mathbf{Y}(\varphi, f) \in \mathbb{C}^{M_{\text{Tx}} \times M_{\text{Rx}}}$.

$$\mathbf{y}(\varphi, f) = \text{vec}\{\mathbf{Y}(\varphi, f)\} \quad \Longrightarrow \quad \mathbf{Y}(\varphi, f) = \mathbf{b}_{\text{Rx}}(\varphi, f) \cdot \mathbf{b}_{\text{Tx}}^{\text{T}}(\varphi, f) + \mathbf{W}(f) \tag{4.98}$$

Accordingly, the measurement matrix $\mathbf{Y}$ is of rank-one, which can be utilised to reduce the errors $\mathbf{W}$. The singular value decomposition (SVD) will be utilised therefore.

Definition 2. The singular value decomposition of a $M \times N$ complex-valued matrix $\mathbf{Z}$ of rank L is given by [MS00]

$$\mathbf{Z} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^{\text{H}} \quad . \tag{4.99}$$

The matrices $\mathbf{U} \in \mathbb{C}^{M \times M}$ and $\mathbf{V} \in \mathbb{C}^{N \times N}$ are unitary matrices, whose columns are orthonormal. Matrix $\mathbf{S} \in \mathbb{R}_+^{M \times N}$ is the matrix of the singular values $s_1 \geq \ldots \geq s_L > 0$.

$$\mathbf{S} = \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & \mathbf{0} \\ & & s_L & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array} \right] \tag{4.100}$$

If the matrix is of rank $L = 1$ and confounded by errors, the singular values are given by $s_1 + w_1 \geq w_2 \geq \ldots$, whereas w are singular values of the errors.

Consider the SVD of matrix $\mathbf{Y}(\varphi, f)$ and select the columns of $\mathbf{U}$ and $\mathbf{V}$ which correspond to the largest singular value s_1: $\mathbf{u}_1$ and $\mathbf{v}_1$. Then, the calibration matrix can be approximated.

$$\mathbf{Y}(\varphi, f) \approx \mathbf{u}_1 \cdot s_1 \cdot \mathbf{v}_1^{\text{H}} \tag{4.101}$$

Because the contributions of all other singular values, which are assigned to the errors only, are discarded, the measurements are denoised. Note, that by considering the largest singular value only it is inherently assumed, that the reflection from the calibration target (the corner reflector) always supersedes the errors in the calibration data, i.e. are larger in amplitude. This is ensured as long as the target is in the vicinity of the broad fire direction, because the antennas of the radar have their highest gain in that direction [Hit+18, Hit+17]. However, if the target direction becomes closer to end fire, the amplitude of the target reflection in the data decreases. As a consequence the largest singular value may now correspond to a parasitic reflection or leakage and no longer to the target reflection. Therefore, gating of the peak corresponding to the target reflection is carried out prior to the denoising.

The array response vectors of the Tx and Rx array $\mathbf{b}_{\text{Tx}}(\varphi, f)$ and $\mathbf{b}_{\text{Rx}}(\varphi, f)$ can be identified from the decomposition results as well, and $\mathbf{Y}(\varphi, f)$ is the matrix of calibration data.

$$\mathbf{b}_{\text{Rx}}(\varphi, f) = \sqrt{s_1} \cdot \mathbf{u}_1 \tag{4.102a}$$

$$\mathbf{b}_{\text{Tx}}(\varphi, f) = \sqrt{s_1} \cdot \mathbf{v}_1^{\dagger} \tag{4.102b}$$

This is especially useful for bistatic (distributed) MIMO arrays, where the Tx and Rx array cannot be separately calibrated and φ_{Tx} and φ_{Rx} differ. From the decomposition the respective
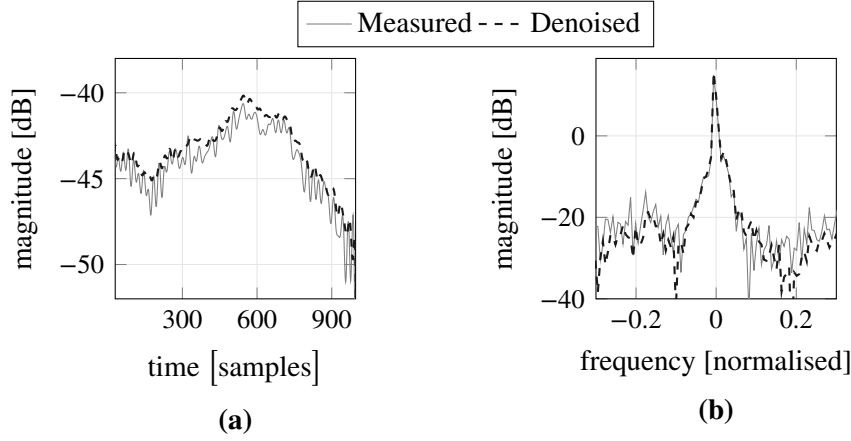
Figure 4.28: *Measured and resulting signal after the denoising. The signals are shown for (a) time and (b) frequency domain. Calibration measurements of the 10th channel of the FMCW MIMO radar in [2] are used. The direction is 0°.*

array response vectors can be assigned.

$$\mathbf{b}_{\text{Rx}}\left(\varphi_{\text{Rx}}, f\right) = \sqrt{s_1} \cdot \mathbf{u}_1 \tag{4.103a}$$

$$\mathbf{b}_{\text{Tx}}\left(\varphi_{\text{Tx}}, f\right) = \sqrt{s_1} \cdot \mathbf{v}_1^{\dagger} \tag{4.103b}$$

Figure 4.28 shows the time and frequency domain signal of a calibration measurement before and after application of the SVD-based denoising. The direction of radiation/impingement is chosen to be 0°. From the frequency domain signal a reduction of the noise floor becomes obvious, and the time domain signal becomes less noisy.

4.4.3 Comparison of Narrowband and Wideband Array Model

The calibrated narrowband and wideband array model for the MIMO radar presented in [2] will be compared in the following. The array ambiguity function will be utilised as figure of merit for comparison. The array ambiguity function is commonly used to investigate array ambiguities in direction estimation and to investigate the resolution capabilities of an array. First, a definition of an ambiguity in direction estimation with antenna arrays will be given.

Definition 3. An ambiguity is defined as (at least) two different wave vectors $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$, which cause the same array response.

$$\mathbf{b}\left(\boldsymbol{\kappa}_1, t, q\right) \equiv \mathbf{b}\left(\boldsymbol{\kappa}_2, t, q\right) \quad ; \quad \boldsymbol{\kappa}_1 \neq \boldsymbol{\kappa}_2 \tag{4.104}$$

The ambiguity function is the correlation of the array response for two different directions. The ambiguity function features a maximum for equal directions and, in case of ambiguities, multiple maxima for different directions. The normalised wideband array ambiguity function

$\boldsymbol{\Psi}_{\text{wideband}}$ for reference direction $\boldsymbol{\kappa}$ and testing direction $\hat{\boldsymbol{\kappa}}$ is [SM11]

$$\Psi_{\text{wideband}} (\boldsymbol{\kappa}, \hat{\boldsymbol{\kappa}}) = \frac{\left| \sum_{k=0}^{K-1} \mathbf{b} \left(\boldsymbol{\kappa}, f_0 + k\Delta f \right)^{\text{H}} \cdot \mathbf{b} \left(\hat{\boldsymbol{\kappa}}, f_0 + k\Delta f \right) \right|}{\sum_{k=0}^{K-1} \left\| \mathbf{b} \left(\hat{\boldsymbol{\kappa}}, f_0 + k\Delta f \right) \right\|^2} \quad . \tag{4.105}$$

The exemplary considered MIMO radar can resolve azimuth angles only. Hence, investigations of array ambiguities are limited to the azimuth plane. A measured array response $\mathbf{y} \left(\varphi, f_0 + k\Delta f \right)$ for a certain azimuth angle φ is taken from the calibration measurements. Afterwards, the ambiguity function is determined using the calculated array response $\mathbf{b} \left(\hat{\varphi} \right)$ or $\mathbf{b} \left(\hat{\varphi}, f_0 + k\Delta f \right)$ from the narrowband or wideband array model, respectively. The ambiguity function for the wideband array model is given by

$$\Psi_{\text{wideband}} (\varphi, \hat{\varphi}) = \frac{\left| \sum_{k=0}^{K-1} \mathbf{y} \left(\varphi, f_0 + k\Delta f \right)^{\text{H}} \cdot \mathbf{b} \left(\hat{\varphi}, f_0 + k\Delta f \right) \right|}{\sum_{k=0}^{K-1} \left\| \mathbf{b} \left(\hat{\varphi}, f_0 + k\Delta f \right) \right\|^2} \quad . \tag{4.106}$$

To get the ambiguity function for the narrowband model, the selected response $\mathbf{y} \left(\varphi, f_0 + k\Delta f \right)$ has to be calibrated using the decomposition results from Section 4.4.1. Accordingly, the ambiguity function for the narrowband array model is

$$\Psi_{\text{narrowband}} (\varphi, \hat{\varphi}) = \sum_{k=0}^{K-1} \frac{\left| \left[\left(\mathbf{C}^{\text{-1}} \cdot \mathbf{y} \left(\varphi, f_0 + k\Delta f \right) \right) \oslash \mathbf{g} \left(f_0 + k\Delta f \right) \right]^{\text{H}} \cdot \mathbf{b} \left(\hat{\varphi} \right) \right|}{K \cdot \left\| \mathbf{b} \left(\hat{\varphi} \right) \right\|^2} \quad . \tag{4.107}$$

Operator $\oslash$ denotes the element-wise division. Note, that the frequency dependence of the array response is dropped due to the narrowband modelling.

The ambiguity functions for the narrowband and wideband array, and for different azimuth angles are shown in Fig. 4.29. Note, the virtual array is a linear array and $0°$ correspondence to the broad fire direction. Obviously, the narrowband array model starts to fail towards the end fire direction, because the maximum of the ambiguity function is shifted apart from the true angle. Hence, azimuth estimates tend to become biased. However, the maximum of the ambiguity function for the wideband array model stays at the true azimuth direction.

An explanation for the failure of the narrowband model is the size of the virtual array, which violates the narrowband criterion in (4.78). Furthermore, the antenna spacing in the virtual array is sometimes larger than $\lambda_{\text{c}}/2$ [2], which violates the narrowband criterion of antenna spacing, see Section 4.3.3. Figure 4.30 shows the spectrum of 8 channels taken from calibration measurements and a target angle of $-30°$. Obviously, the peaks of the beat frequencies are at different frequencies. Hence, the target echo returns arrive at different delays, resulting in a delay spread across the array. The virtual array is a linear array, and the delay τ_m observed at the virtual array position r_m and for azimuth angle φ is

$$\tau_m (\varphi) = \frac{r_m}{c_0} \sin (\varphi) \quad . \tag{4.108}$$

The size of the virtual array is approximately $70\lambda_{\text{c}}$ [2], such that the largest observed delay for
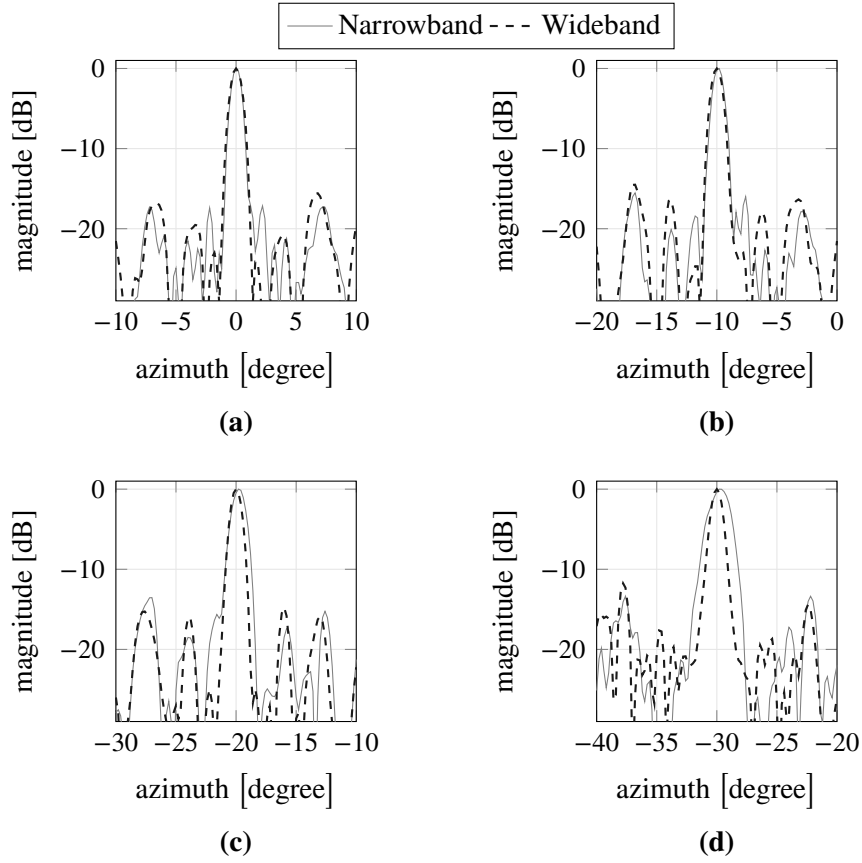
Figure 4.29: *Array ambiguity function of the MIMO radar in [2] for the wideband and narrowband array model and various directions $\hat{\varphi}$: (a) $\hat{\varphi} = 0°$, (b) $\hat{\varphi} = -10°$, (c) $\hat{\varphi} = -20°$ and (d) $\hat{\varphi} = -30°$.*

direction φ is

$$\tau_{\max}(\varphi) = \frac{70\lambda_c}{c_0} \sin(\varphi) = \frac{70}{f_c} \sin(\varphi) \quad . \tag{4.109}$$

Plugging into equation (4.78) yields the following narrowband criterion for the virtual array

$$\frac{W}{f_c} \ll \frac{1}{70 \cdot \sin(\varphi)} \quad . \tag{4.110}$$

Accordingly, the narrowband criterion depends on the aperture of the array, which is faced by an impinging wave. For the considered case, the narrowband criterion is always fulfilled for broad fire direction ($\varphi = 0°$), and becomes more severe towards end fire direction ($\varphi = \pm 90°$). Because the radar operates with a bandwidth of 10 GHz at a centre frequency of around 155 GHz, the narrowband assumption is severely violated towards end fire. This finding is proven in Fig. 4.30, showing the measured frequency spectrum of 8 MIMO channels for $\varphi = -30°$. The peaks in the frequency spectrum are not at the same frequency bin. Hence they do not have the same beat frequency, indicating the occurrence of a delay spread across the array. Accordingly, the narrowband assumption is violated [26] and narrowband modelling
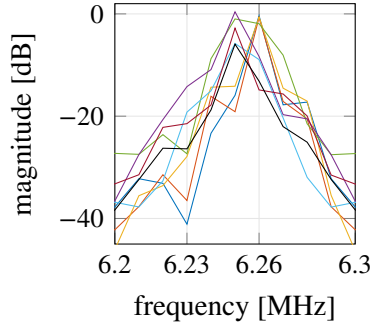
Figure 4.30: *Signal spectrum of calibration measurements of 8 channels of the FMCW MIMO radar presented in [2]. The azimuth angle is chosen to be −30°.*

of the array response is not applicable.

4.5 Model of the Sampling Process

So far, gathering of the observations by ADC sampling of the baseband signal has not been considered in the previous device modelling steps. Also, the process of gathering all MIMO channels has not been discussed yet. However, as the parametric system identification is done from sampled data the sampling process, i.e. the process of gathering the observations, has to be taken into consideration. A time-invariant channel has been considered for the device model derivations. However, if the channel becomes time-variant due to e.g. moving scatterer, or time-division multiplexing (TDM) is employed to capture all MIMO channels, a model for the observation and the respective sampling process becomes necessary. Basically, the propagation path parameters direction and delay are determined from time delays in a respective domain. Therefore, baseband signals are time sampled. In order to relate the propagation parameters to a time delay measurement, while jointly considering time-variance and the TDM accessing scheme, a model of the sampling process over the time axis is required.

4.5.1 Observation of Time-Invariant Systems

If the propagation channel is time-invariant, hence the scatterer nor the observer is moving, the delay τ of the echo signal is a constant.

$$\tau = \tau_0 \tag{4.111}$$

An observation of a single ramp of the FMCW is sufficient to determine this delay. The time axis t is discretised by sampling of the observations by the ADC.

$$t = k \cdot T_\mathrm{S} \tag{4.112}$$

Variable $k \in \mathbb{Z}, k = 0 \dots K - 1$ indicates the captured samples. This sampling can be schematically sketched as a 1-D sampling of a signal (see Fig. 4.31).
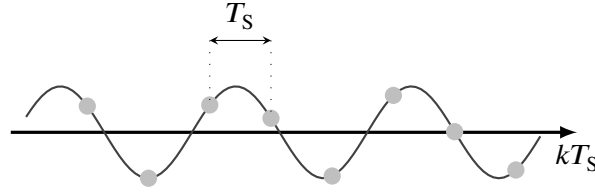
Figure 4.31: *Discretisation of the time axis by ADC sampling time T_S*
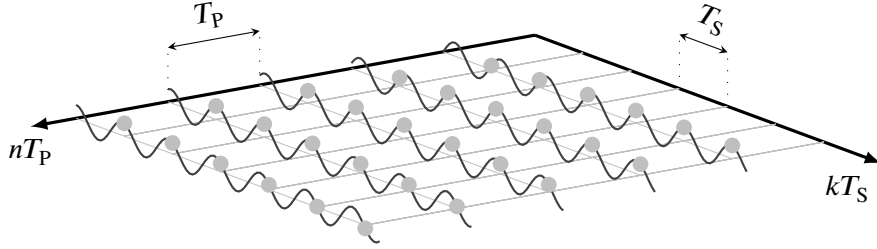
Figure 4.32: *Discretisation of the time axis by ADC sampling time T_S and signal period T_P*

FMCW Radar and Delay

Considering the model (4.12) of the baseband signal and plugging in the discretised time axis, the sampled phase of the baseband signal becomes

$$-\frac{1}{2\pi}\phi\left(kT_S\right) \approx \frac{W}{T_M}\tau_0 kT_S + f_0\tau_0 \quad . \tag{4.113}$$

4.5.2 Observation of Time-Variant Systems

Considering a time-variant channel due to e.g. a moving scatterer, the delay τ of the echo signal becomes a time dependent quantity (see Section 3.2.1). Assume a homogeneously moving scatterer, which moves with relative velocity $v_0 = dR(t)/dt$ w.r.t. the observer. The time-variant delay is

$$\tau(t) = \tau_0 + \frac{v_0}{c_0}t \quad . \tag{4.114}$$

Repeating observations are necessary in order to capture sufficient information on this time-varying behaviour. Consider a periodically radiated transmit signal with repetition/period time T_P. Then, the sampling of the time axis can be split into the sampling by the ADC and the sampling by the signal period. Accordingly, the time axis t is discretised as

$$t = k \cdot T_S + n \cdot T_P \quad , \tag{4.115}$$

with $n \in \mathbb{Z}, n = 0 \ldots N-1$ the number of captured/radiated signal periods, and $(K-1)\cdot T_S < T_P$ should hold. Accordingly, a 2-D sampling scheme of the time axis can be established, see Fig. 4.32. Note, that this 2-D sampling scheme is only virtually present, because there is still *one unique time axis*.
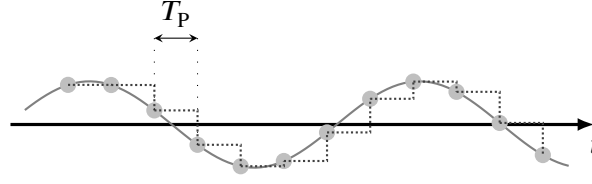
Figure 4.33: *Motion induced Doppler harmonic (solid line) and their approximation (dotted line) by neglecting the range-Doppler coupling.*

FMCW Radar and Doppler Shift

Consider model (4.22) of the baseband signal and plug in the discretised time axis (4.115). The sampled phase of the baseband signal becomes

$$-\frac{1}{2\pi}\phi\left(kT_{\mathrm{S}}+nT_{\mathrm{P}}\right) \approx f_0\frac{v_0}{c_0}\left(kT_{\mathrm{S}}+nT_{\mathrm{P}}\right) + \frac{W}{T_{\mathrm{M}}}\tau_0\left(kT_{\mathrm{S}}+nT_{\mathrm{P}}\right) + f_0\tau_0 \quad . \tag{4.116}$$

Term $f_0 kT_{\mathrm{S}}\, v_0/c_0$ accounts for *range-Doppler coupling*, i.e. the Doppler shift during the observation of the FMCW ramp. Hence, the Doppler shift causes a variation of the baseband signal phase during the observation of the FMCW ramp. This phase variation cannot be easily distinguished from a delay in FMCW radar with stretch processing receiver. A delay causes a shift of the FMCW signal along the time axis, and a Doppler causes a shift of the FMCW signal along the frequency axis, see Fig. 4.1. In both cases a beat frequency occurs at baseband, which cannot be unambiguously assigned to the Doppler or the delay or a mix of both. Consider a steep slope W/T_{M} of the FMCW transmit signal due to e.g. a short modulation time. Then, $\tau_0 \cdot W/T_{\mathrm{M}} \gg f_0 \cdot v_0/c_0$ holds and the range-Doppler coupling can be neglected, because only a minor influence of the Doppler shift on the phase variation during a single FMCW ramp is observed. A periodic FMCW signal with a steep slope is denoted as *chirp sequence signal* [FJ15]. Accordingly, the phase simplifies [FJ15].

$$-\frac{1}{2\pi}\phi\left(kT_{\mathrm{S}}+nT_{\mathrm{P}}\right) \approx f_0\frac{v_0}{c_0}nT_{\mathrm{P}} + \frac{W}{T_{\mathrm{M}}}\tau_0 kT_{\mathrm{S}} + f_0\tau_0 \tag{4.117}$$

Figure 4.33 shows a Doppler caused harmonic, i.e. the differentiation of the phase of the baseband signal, and their corresponding approximation by the above model. According to equation (4.117) the delay related phase variation is observed over one signal period, whereas the phase variation due to the Doppler shift is observed over multiple periods. As the Doppler related phase variation is sampled over the period time T_{P}, the period time has to fulfil the Nyquist criterion w.r.t. the Doppler harmonic.

4.5.3 Observation by MIMO Systems

So far, a SISO system has been considered, which is capable to resolve delay and Doppler. If directional information are of interest too, antenna arrays (see Section 4.3) at Tx and/or Rx have to be employed. On the contrary to phased-array radar, where beamforming is applied [Bro03, Fri09], MIMO radars exploit the transmission of orthogonal waveforms and the additional degrees of freedom are exploited [BF03]. Such radar system are termed as
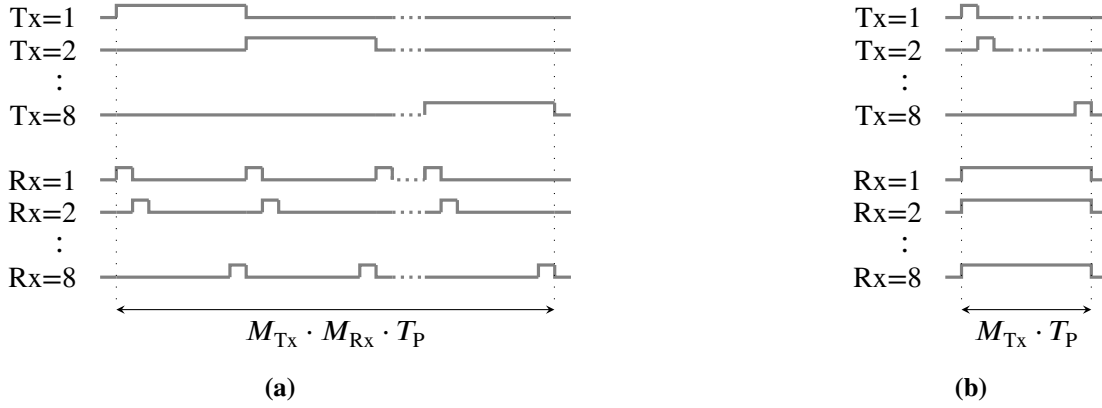
Figure 4.34: *Timing diagrams for a single MIMO snapshot in a TDM system. Eight sequentially radiating transmitters and eight (a) sequentially captured or (b) parallel captured receivers are sketched. The high level indicates the activity of the respective Tx or Rx channel.*
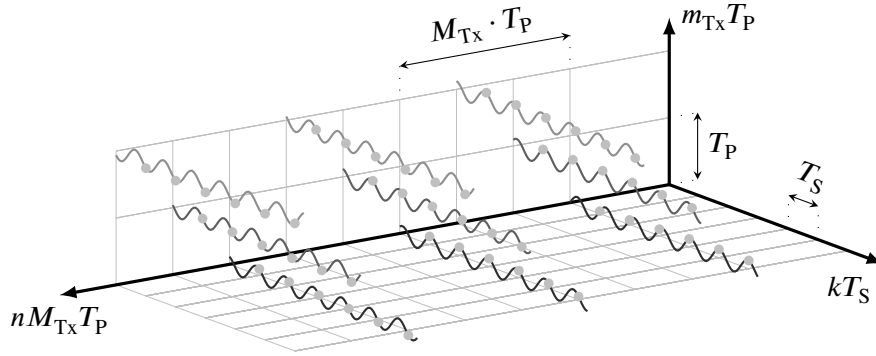
Figure 4.35: *Discretisation of the time axis by ADC sampling time T_S, signal period T_P and MIMO snapshot duration $M_{Tx}T_P$ for TDM at transmitter.*

MIMO radar [Fis+06]. Considering MIMO measurements, the delay of the received signal for a single MIMO channel observation is given by the summed delays in equation (4.114) and equation (4.93).

$$\tau(t) = \tau_0 - \frac{1}{c_0}\left[\mathbf{d}_{Rx}^T \cdot \boldsymbol{\kappa}_{Rx}\left(\varphi_{Rx}, \vartheta_{Rx}\right) + \mathbf{d}_{Tx}^T \cdot \boldsymbol{\kappa}_{Tx}\left(\varphi_{Tx}, \vartheta_{Tx}\right)\right] + \frac{v_0}{c_0}t \qquad (4.118)$$

In order to gather all MIMO channels free of self interference, the transmitted waveforms have to be *orthogonal*. Hence, multiplexing is required. Known multiplexing techniques for MIMO measurements are time-division multiplexing (TDM), frequency-division multiplexing or code-division multiplexing [SBL14, UP09]. In the following a TDM scheme will be considered, because it provides the best trade-off in terms of hardware effort, SNR and measurement duration. In TDM the MIMO channels are captured by successively switching i) the transmit and receive channels [Tho+01] or ii) the transmit channels [2], if parallel receiver channels are employed. Hence, a MIMO snapshot is realised as multiple SISO or single input multiple output (SIMO) measurements. The timing diagrams for both TDM schemes are shown in Fig. 4.34. Consider M_{Tx} antennas at the Tx side and M_{Rx} antennas at Rx side. Then, the sampling points for the $m_{Tx} = 0 \ldots M_{Tx} - 1$ transmit channel and the $m_{Rx} = 0 \ldots M_{Rx} - 1$

receive channel are

$$t = k \cdot T_S + \left(m_{Rx} + m_{Tx} M_{Rx} \right) \cdot T_P + n \cdot M_{Tx} M_{Rx} T_P \tag{4.119}$$

for sequentially captured MIMO channels, and

$$t = k \cdot T_S + m_{Tx} \cdot T_P + n \cdot M_{Tx} T_P \tag{4.120}$$

for parallelly captured Rx channels. Figure 4.35 shows the respective 3-dimensional sampling of the time axis according to the ADC sampling time, the period time and the MIMO snapshot time.

FMCW TDM-MIMO Radar

Lets consider an FMCW MIMO radar with parallel receivers and TDM scheme for the transmit channels, e.g. [2]. Then, the phase of the baseband signal of the m_{Tx}-m_{Rx} MIMO channel, considering the approximation in equation (4.117) and the delay given by equation (4.118), can be stated.

$$-\frac{1}{2\pi}\phi \left(kT_S + m_{Tx}T_P + nM_{Tx}T_P \right) \approx f_0\tau_0 + \frac{W}{T_M}\tau_0 \cdot kT_S + v \cdot m_{Tx}T_P + v \cdot nM_{Tx}T_P$$

$$-\frac{W}{T_M} \left(\frac{1}{c_0}\mathbf{d}_{Rx}{}^T\boldsymbol{\kappa}_{Rx} + \frac{1}{c_0}\mathbf{d}_{Tx}{}^T\boldsymbol{\kappa}_{Tx} \right) kT_S \tag{4.121}$$

Variable $v = f_0\, v_0/c_0$ is the Doppler shift.

Angle-Doppler Coupling − Compensation and Exploitation

The Tx channels are accessed at different time instances due to the TDM. Consequently, the Doppler induced harmonic in the baseband signal is observed at different time instances for each active Tx channel, resulting in different phase constellations due to a phase shift. Figure 4.36 shows the time shifted sampling of the Doppler induced harmonic for a MIMO radar system with parallel receivers and TDM of the transmitters. The time shifted observation of the Doppler causes *angle-Doppler coupling*, which deteriorates the Doppler and angle estimation [Som+02]. The phase shift $2\pi v m_{Tx}T_P$ represents the angle-Doppler coupling for this radar system.

Typically, the Doppler shift v is estimated over the N gathered MIMO snapshots, which requires each MIMO snapshot to be time aligned [Som+02]. Due to the angle-Doppler coupling, the MIMO measurements are not time aligned. A time alignment and hence compensation of the angle-Doppler coupling can be achieved by pre-processing of the measurements using Fourier interpolation [BRW17, Som+02]. This interpolation requires the fulfilment of the Nyquist-criterion w.r.t. the Doppler shift, such that the maximal allowed Doppler shift is: $\max\{|v|\} \leq 1/2T_P$. Afterwards, the Doppler shift is estimated over the gathered MIMO snapshots, such that the maximal resolvable Doppler shift is: $\max\{|v|\} \leq 1/2M_{Tx}T_P$. However, the angle-Doppler coupling can also be exploited by proper consideration in the model of the parameter estimator. Consequently, a maximal Doppler shift
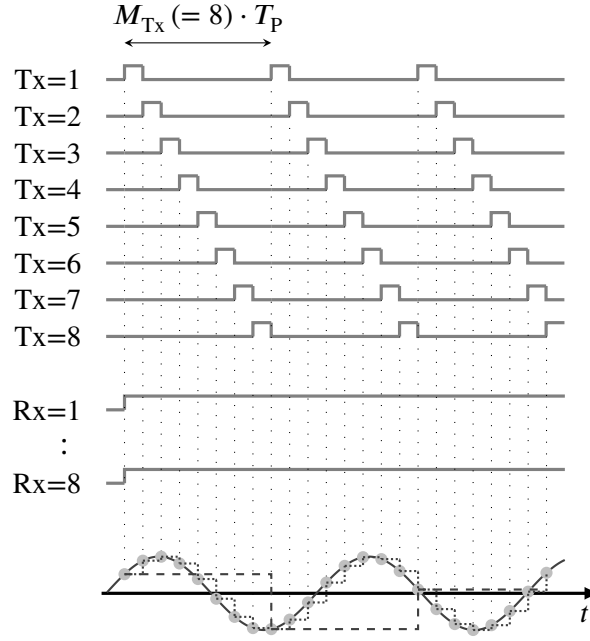
Figure 4.36: *Sampling of the motion induced Doppler harmonic (solid line) by a 8×8 MIMO radar with parallelly captured receive channels and TDM of the transmit channels. For each active Tx channel the Doppler induced harmonic is sampled at a different time instance. The densely dotted and dashed line indicates the model-based approximation of the Doppler harmonic, if their temporal evolution over the duration of the MIMO snapshot is considered or ignored in the model, respectively. [10]*

of max $\{|v|\} \leq 1/2T_{\mathrm{P}}$ can be estimated and a single MIMO snapshot becomes sufficient for Doppler estimation.

Ambiguity Function The ambiguity function, which is actually the correlation of the simulated data and the considered data model in the noise-free case, for the velocity dimension will be investigated. As data model equation (4.121) is considered. The ambiguity function is calculated for the case of compensation and exploration of the angle-Doppler coupling. For the case of compensation, the simulated data are time aligned [Som+02] and the angle-Doppler coupling term $2\pi v m_{\mathrm{Tx}} T_{\mathrm{P}}$ is neglected in the data model. The ambiguity function is calculated for the case of 1 and 8 MIMO snapshots. A 8×8 monostatic FMCW MIMO radar with parallel receivers and TDM accessing for the transmitters will be considered. The receive and transmit array are ULAs with isotropic antennas, which are spaced by $\lambda_{\mathrm{c}}/2$ and λ_{c}, respectively. The parameters of the radar are: $W = 20\,\mathrm{GHz}$, $f_0 = 150\,\mathrm{GHz}$, $T_{\mathrm{M}} = 1024\,\mu\mathrm{s}$, $T_{\mathrm{P}} = 1124\,\mu\mathrm{s}$ and $T_{\mathrm{S}} = 500\,\mathrm{ns}$. A single target at a distance of 2 m and azimuth angle of $0°$ is considered, which moves with a velocity of 0.025 m/s apart from the radar. The model (4.121) is used to calculate the simulated data for each MIMO channel.

The calculated ambiguity functions are shown in Fig. 4.37. From Fig. 4.37(a) it becomes obvious, that the velocity (respectively the Doppler shift) can be estimated over a single MIMO snapshot, if the angle-Doppler coupling is considered in the data model. Furthermore, Fig. 4.37(b) indicates that exploitation of the angle-Doppler coupling increases the Doppler bandwidth compared to the compensation approach.
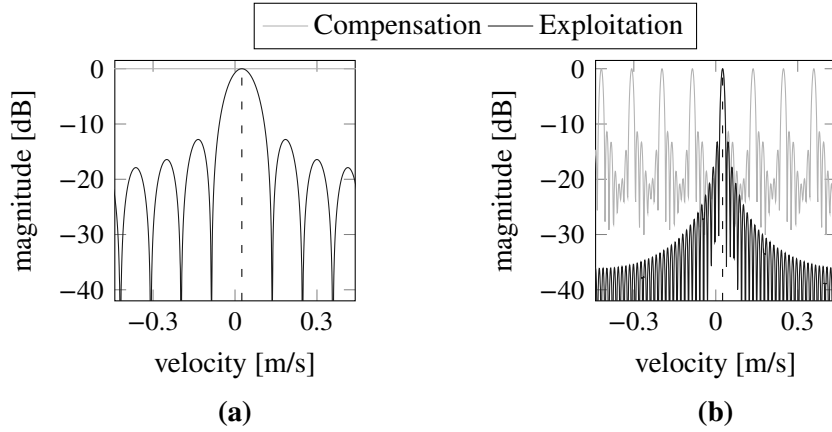
Figure 4.37: *Ambiguity function in the velocity dimension for (a) 1 MIMO snapshot and (b) 8 MIMO snapshots. The ambiguity function is calculated for compensation and exploitation of the angle-Doppler coupling. The black dashed line indicates the velocity of the target. [10]*

4.6 Model of Noise Sources in FMCW MIMO Radar

The models so far consider perfect, i.e. noise-less measurements. Apart from influences due to hardware impairments, measurements are always confounded by unwanted disturbance signals. Such disturbances can be interference from e.g. other observers, which will not be considered here, or disturbance signals from the observer system itself. Such disturbances will be considered in the following and are termed as *noise*. Noise origins from various sources such as thermal caused motion of electrons in electrical devices or vibrations in mechanical setups. Various noise processes in electrical devices are known from literature, e.g. *shot noise*, *thermal noise*, *phase noise* or *flicker noise* [SSR06].

Noise causes uncertainties on the measurements and limits the measurement accuracy. In order to account for these uncertainties in the interpretation of the measurements or the parameter estimation, a model $\mathcal{M}^{\mathrm{N}}$ of the noise process is necessary. However, description of the time-dependent behaviour of the noise by a deterministic mathematical model is generally hard up to impossible, because limited knowledge on the noise source is available. Hence, modelling on a microscopic level is not taken into account, such that modelling on a macroscopic level is employed. There, the noise is modelled as a *random process* according to the law of large numbers [SSR06].

Three noise sources in FMCW radars will be considered in the following: *leakage noise*, *thermal noise* and *phase noise*. Statistical models by means of a probability density function (PDF) will be derived for these noise sources, in order to account for the uncertainties which are introduced in the observations by these noise signals.

4.6.1 Thermal Noise Model

Thermal noise or sometimes denoted as Johnson–Nyquist noise is due to thermal agitation of the electrons within an electrical conductor of finite conductivity [SSR06]. Thermal noise $w(\tau)$ is

commonly modelled as a circularly normal distributed random process in the complex baseband. Circularity of a normally distributed process implies propriety and also zero-mean [SS10].

$$\mathrm{E}\left\{w\left(\tau\right)\right\} = 0 \tag{4.122}$$

Thermal noise is assumed as wide-sense stationary (WSS) (see APPENDIX B.2.1 for a definition of wide-sense stationarity) and white, i.e., uncorrelated process. The respective second-order statistic is given by the auto-correlation function (ACF) $\psi_{\mathrm{w}}\left(\tau\right)$ of the process (the expectation is taken w.r.t. the delay τ').

$$\psi_{\mathrm{w}}\left(\tau\right) = \mathrm{E}\left\{w\left(\tau' + \tau\right)^{\dagger} \cdot w\left(\tau'\right)\right\} = \eta \cdot \delta\left(\tau\right) \tag{4.123}$$

η is the noise power. Due to the assumption of a white process, the power spectral density of the noise process is uniform. The respective PDF of the noise is

$$w\left(\tau\right) \sim \mathcal{CN}\left(0, \eta \cdot \delta\left(\tau\right)\right) \quad . \tag{4.124}$$

Thermal noise will be assumed as independent and identical distributed (i.i.d.) process, such that individual realisations are mutually independent but have the same distribution. Accordingly, the ACF in time domain $\psi_{\mathrm{w}}\left(t\right)$ is

$$\psi_{\mathrm{w}}\left(t\right) = \mathrm{E}\left\{w\left(t' + t\right)^{\dagger} \cdot w\left(t'\right)\right\} = \delta\left(t\right) \quad . \tag{4.125}$$

As the thermal noise is assumed as WSS and i.i.d., the second-order statistic of the thermal noise process in the delay-time domain $\psi_{\mathrm{w}}\left(\tau, t\right)$ is

$$\psi_{\mathrm{w}}\left(\tau, t\right) = \mathrm{E}\left\{w\left(\tau' + \tau, t' + t\right)^{\dagger} \cdot w\left(\tau', t'\right)\right\} = \eta \cdot \delta\left(\tau\right) \cdot \delta\left(t\right) \quad . \tag{4.126}$$

Summarised, the thermal noise in delay-time domain is modelled as an i.i.d. random process with PDF

$$w\left(\tau, t\right) \sim \mathcal{CN}\left(0, \psi_{\mathrm{w}}\left(\tau, t\right)\right) \quad . \tag{4.127}$$

Figure 4.38 shows the average spectrum of two measurements with a single target, and the estimated power spectral density (PSD) of the thermal noise process.

Last, the statistic of the thermal noise in MIMO radar has to be clarified. The thermal noise among the parallel Rx channels is uncorrelated and heteroscedastic (has different power), because the Rx channels have different receiver chains. As thermal noise is assumed as a white process over the time domain, thermal noise among the dimension of the Tx channels is also uncorrelated, if a TDM accessing scheme is employed to access the different Tx channels. Because thermal noise occurs only in the receiver channels it can be assumed, that the noise power is equal for MIMO channels sharing the same receiver channel. However, due to e.g. different phase noise of the transmit and homodyning signal, the thermal noise power slightly varies among the different Tx channels. Hence, the MIMO channels are assumed to be subject to different thermal noise powers. Summarised, thermal noise is assumed as *uncorrelated and heteroscedastic among the MIMO channels*. Hence, the MIMO covariance matrix of the

Figure 4.38: *Average power spectral density (averaged over 64 measurements) of two measurements with a FMCW radar, and the fitted power spectral density of the model of the thermal noise.*

thermal noise is modelled as a matrix with diagonal structure.

4.6.2 Mitigation of Leakage and Leakage Noise Model

In a co-located radar setup as e.g. a monostatic radar, transmitter and receiver are close to each other. Hence, leakage or cross-talk occurs [DSW19], which disturbs the received signal. Besides from leakage due to the limited isolation between the transmit and receive chain [Sto92], short-range leakage may by present [DSW19, Mel+15, Mel+17]. Short-range leakage can be due to undesired targets in the radars vicinity (e.g. the bumper of a car) or the radar setup (e.g. a radome or lens). Hence, short-range leakage is composed of many individual propagation paths. Leakage and short-range leakage cause the signal spectrum in a FMCW radar system to be disturbed and the sensitivity of the radar is lowered. Figure 4.39 shows the spectrum of a measurement with the FMCW radar in [2]. The radar is equipped with a focusing lens, which is approx. 2 cm apart from the antennas. Short-range leakage and leakage occur due to the lens and the close location of Tx and Rx, respectively. As a result, the lower parts of the spectrum of the receive signal are disturbed, such that beat frequencies of close target reflections are hard to identify in the spectrum. Hence, the radar is possibly "blind" w.r.t. close targets. Solutions in hardware [Bea+90, Dür+18, DSW19, Mel+15] or digital pre-processing [Kia+18, LMY04, Mel+17, Par+19] are proposed in literature to reduce the influence of the leakage. In the following, a novel signal processing solution will be presented, which has been published in [6]. It can be easily incorporated in a ML parameter estimation framework, such that hardware changes are not necessary.

Furthermore, leakage affects the development of a parameter estimator too. An estimator would attempt to resolve the individual path contributions of the short-range leakage as they occur with a large power. Thus an estimator can be employed to suppress the leakage by resolving its individual contributions. As the leakage is a distinct process for each MIMO channel, a joint processing over all MIMO channels is prohibited. Hence, the estimator has to account for distinct models for each MIMO channel, which severely increases the computational complexity. Therefore, resolution of the leakage by an estimator is not adorable.
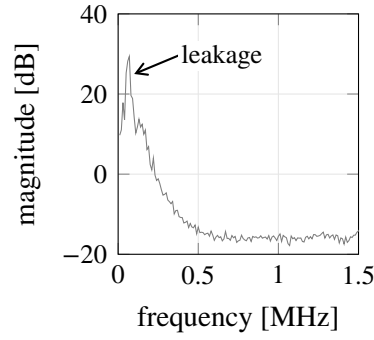
Figure 4.39: *Average power spectral density (averaged over 64 measurements) of a measurement with a FMCW radar featuring co-located transmit and receive channel. The radar is equipped with a lens (2 cm apart from the antennas), such that short-range leakage occurs.*

Background Subtraction

Leakage will be assumed as a disturbance of the received signal due to the radar setup. As long as the system hardware remains unchanged, the leakage can be assumed as static. Hence, *background subtraction*, i.e. the coherent subtraction of an estimate of the leakage from the measurement, can be used. An estimate of the static leakage can be gained from a calibration measurement with no target, or a measurement with an apart target, which can be excluded by gating. In Fig. 4.40 the spectrum of a time-limited measurement before and after background subtraction is shown. Obviously, the leakage is reduced and the peak of the target reflection becomes more present. A small signal portion remains after the background subtraction, which will be termed as *leakage noise*. The leakage noise can be caused by phase variations in the radar hardware. Due to phase variations the estimated leakage and the leakage in the measurement slightly differ, such that the leakage is not totally mitigated by background subtraction. A solution would be an online estimation of the leakage, such that phase variations of the system are accounted for. Another solution would be a probabilistic modelling approach to suppress the leakage noise, which will be considered subsequently.

Leakage Noise Model

First, it will be assumed that each MIMO channel is subject to different leakage noise. Hence, the leakage noise is assumed as *uncorrelated among the MIMO channels*. Accordingly, derivation of a model for leakage noise in a SISO radar will be sufficient. An extension to the MIMO case is given by considering different model parameters for each MIMO channel and a diagonal structure for the respective MIMO covariance matrix.

Leakage noise will be assumed as a coloured random process, as it is caused by random variations in the radar hardware like phase noise or flicker noise. Leakage noise in delay domain $l(\tau)$ will be modelled as a zero mean, proper complex normally distributed random process.

$$l(\tau) \sim \mathcal{CN}\left(0, \psi_1\left(\tau_1, \tau_2\right)\right) \tag{4.128}$$
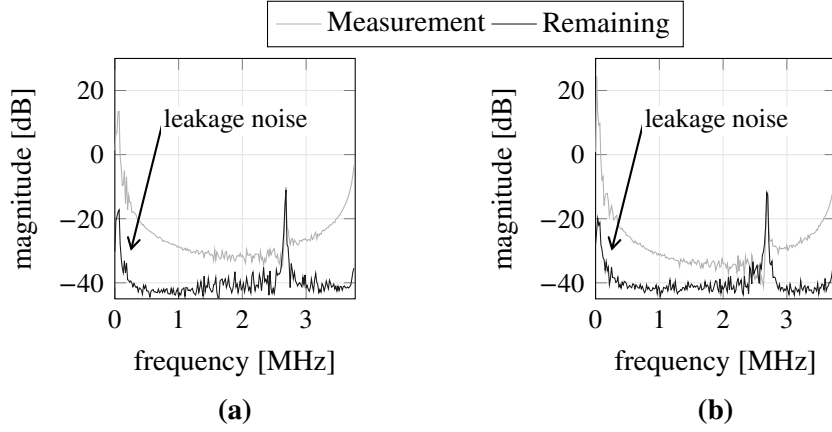
Figure 4.40: *Average power spectral density (averaged over 64 measurements) of a measurement before and after background subtraction. The short range leakage is reduced and the peak of the target reflection becomes more obvious. A signal portion remains after background subtraction, which is termed as leakage noise.*

Function $\psi_1\left(\tau_1, \tau_2\right)$ denotes the correlation of the leakage noise at the lags τ_1 and τ_2.

$$\psi_1\left(\tau_1, \tau_2\right) = \mathrm{E}\left\{l\left(\tau_1\right) \cdot l\left(\tau_2\right)^\dagger\right\}$$

The leakage noise will be assumed as a WSS process in delay domain, such that observations of the process at different frequency bins f_1 and f_2 are uncorrelated.

$$\mathrm{E}\left\{L\left(f_1\right) \cdot L\left(f_2\right)^\dagger\right\} = 0, \quad \forall f_1 \neq f_2$$

The PSD of the leakage noise is assumed as exponentially decaying, see Fig. 4.40. Accounting for the WSS assumption, a model for the correlation of the leakage noise at frequency bins f_1 and f_2 is

$$\Psi_1\left(f_1, f_2\right) = \varsigma \cdot \exp\left\{-\varepsilon\left(f_1 - \varrho\right)\right\} \cdot \mathrm{H}\left(f_1 - \varrho\right) \cdot \delta\left(f_1 - f_2\right) \tag{4.129}$$

Function $\mathrm{H}\left(\bullet\right)$ is the Heaviside step function, which is used to model the limitation of the leakage noise process in frequency domain. Offset frequency ϱ accounts for the offset of the slope of the leakage noise from the $0\,\mathrm{Hz}$ frequency. Variable ε defines the slope of the noise, which actually describes the coherence time of the leakage noise; and ς is the power. The correlation function in delay domain is given by the inverse Fourier transform of equation (4.129).

$$\Psi_1\left(f_1, f_2\right) \quad \overset{f_1 \to \tau_1}{\underset{f_2 \to \tau_2}{\bullet\!\!-\!\!\circ}} \quad \psi_1\left(\tau_1, \tau_2\right)$$

Because the leakage noise is assumed as WSS, only the lag difference $\tau = \tau_1 - \tau_2$ is relevant,

which yields the ACF: $\psi_1\left(\tau_1, \tau_2\right) \Rightarrow \psi_1\left(\tau\right)$.

$$\psi_1\left(\tau\right) = \frac{\varsigma}{\varepsilon - j2\pi\tau} \cdot \exp\left\{j2\pi\varrho\tau\right\} \tag{4.130}$$

The observation time of the measurements is limited. Hence, model (4.129) is not the correlation of the observed leakage noise in frequency domain, because this model assumes a unlimited observation time. The time limitation causes a convolution with a sinc-function in frequency domain.

$$\Psi_1\left(f_1, f_2\right) = \left[\varsigma\exp\left\{-\varepsilon\left(f_1 - \varrho\right)\right\} H\left(f_1 - \varrho\right)\delta\left(f_1 - f_2\right)\right] * \operatorname{sinc}\left(T_{\mathrm{M}}f_1\right) * \operatorname{sinc}\left(T_{\mathrm{M}}f_2\right) \tag{4.131}$$

The leakage noise will be assumed as an i.i.d. process over time t. Accordingly, the leakage noise is uncorrelated at different time lags t_1 and t_2. Furthermore, the process is assumed as WSS, such that the lag difference $t = t_1 - t_2$ is relevant.

$$\psi_1\left(t_1, t_2\right) = \mathrm{E}\left\{l\left(t_1\right)^{\dagger} \cdot l\left(t_2\right)\right\} = \delta\left(t_1 - t_2\right) = \delta\left(t\right) \tag{4.132}$$

Knowing the second-order statistic of the leakage noise in delay τ and time t, the second-order statistic $\psi_1\left(\tau, t\right)$ in the delay-time domain can be derived.

$$\psi_1\left(\tau_1, \tau_2; t_1, t_2\right) = \mathrm{E}\left\{l\left(\tau_1, t_1\right)^{\dagger} \cdot l\left(\tau_2, t_2\right)\right\} = \psi_1\left(\tau\right) \cdot \delta\left(t\right) \tag{4.133}$$

Summarised, leakage noise in delay-time domain is an i.i.d. random process with the probability density function

$$l\left(\tau, t\right) \sim \mathcal{CN}\left(0, \psi_1\left(\tau, t\right)\right) \quad . \tag{4.134}$$

Figure 4.41 shows the average PSD of two exemplary measurements and the estimated PSD of the leakage noise process, see Section 5.2 for details of the estimation procedure.

4.6.3 Phase Noise Model

Phase noise is a random variation of the signal phase, which can be caused by noise occurring in the phase locked loops (PLLs), oscillators, amplifiers etc. In FMCW radar phase noise is of major concern [BB93, 2]. Due to the presence of phase noise in the transmit and homodyning signal range-correlation effects occur [BB93], such that the noise floor in the received signal is not a temporarily white process. Hence, assuming only white noise in the observation will cause a model error, which deteriorates the parameter estimation results. Therefore, a model to account for the phase noise in the parameter estimator is necessary.

Phase Noise in FMCW SISO Radar

A model of phase noise occurrence in a FMCW SISO radar is sketched in Fig. 4.42. In general, the transmit $s_{\mathrm{Tx}}\left(\tau\right)$ and homodyning $s_{\mathrm{Rx}}\left(\tau\right)$ signal are affected by separate phase noise sources
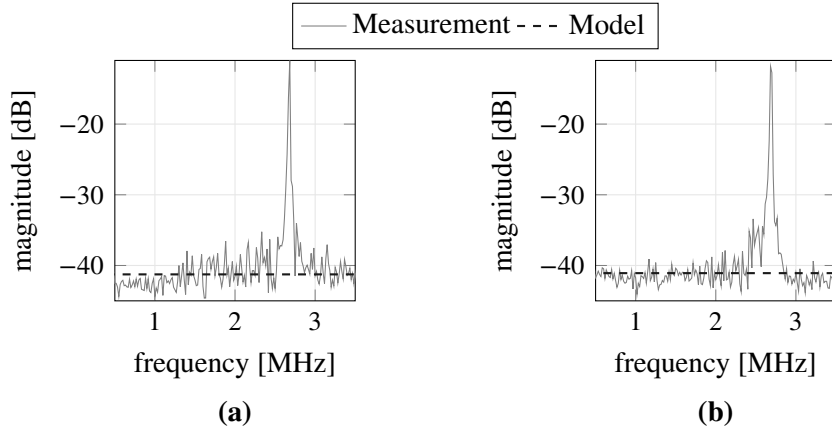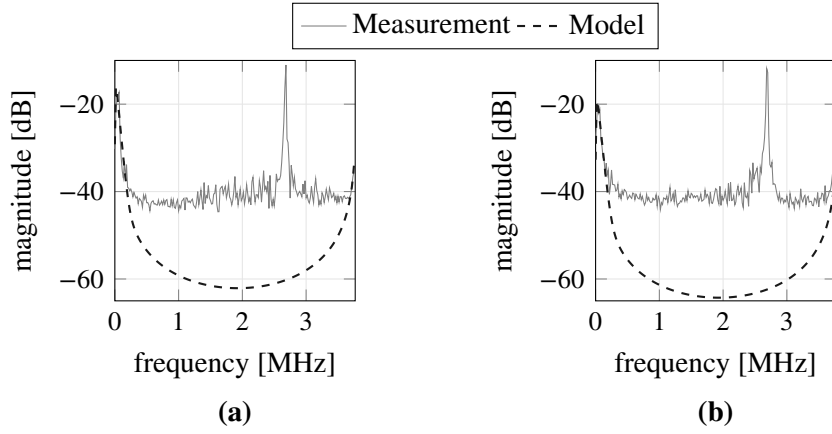
Figure 4.41: *Average power spectral density (averaged over 64 measurements) of two measurements with a FMCW radar, and the fitted model of the power spectral density of the leakage noise.*
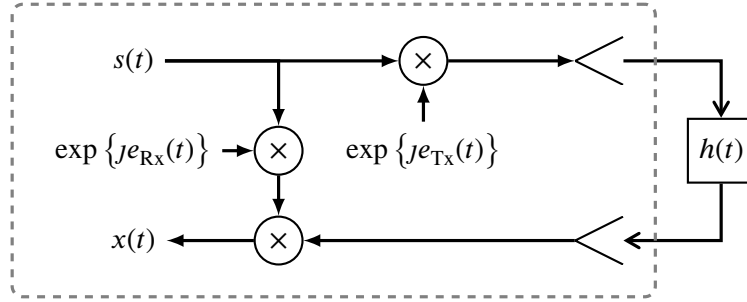
Figure 4.42: *Sketch of an FMCW radar with stretch processor receiver architecture. The phase noise processes $e_{Tx}(t)$ and $e_{Rx}(t)$ affect the transmit and homodyning signal, respectively.*

as e.g. both signals are driven by different PLLs, see [2]. The respective signal models are

$$s_{\mathrm{Tx}}(\tau) = \exp\{J\varphi(\tau)\} \cdot \exp\{Je_{\mathrm{Tx}}(\tau)\} \tag{4.135a}$$

$$s_{\mathrm{Rx}}(\tau) = \exp\{J\varphi(\tau)\} \cdot \exp\{Je_{\mathrm{Rx}}(\tau)\} \quad, \tag{4.135b}$$

with $e_{\mathrm{Tx}}(\tau)$ and $e_{\mathrm{Rx}}(\tau)$ the *phase noise process* at Tx and Rx, respectively. Phase function $\varphi(\tau)$ is the phase of the FMCW signal, see Section 4.1. The transmit signal is passed over a transmission channel, which constitutes of the propagation channel featuring a single propagation path with delay τ_p and path weight γ_p, and the response of the radar system $G(f(\tau))$. The received signal after stretch processing (homodyning,de-ramping) is

$$x(\tau) = \gamma_p \cdot G(f(\tau)) \cdot \exp\{J\varphi(\tau_p)\} \cdot \exp\{J\Delta_p(\tau)\} \quad. \tag{4.136}$$

Variable $\Delta_p(\tau)$ denotes the *phase noise difference*.

$$\Delta_p(\tau) = e_{\mathrm{Tx}}(\tau - \tau_p) - e_{\mathrm{Rx}}(\tau) \tag{4.137}$$

In case of multiple propagation paths, the received signals becomes

$$x(\tau) = \sum_p h_p(\tau) \cdot \exp\{J\Delta_p(\tau)\} \quad , \tag{4.138}$$

with $h_p(\tau) = \gamma_p \cdot G(f(\tau)) \cdot \exp\{J\varphi(\tau_p)\}$ the channel transfer function for the p-th propagation path and $\Delta_p(\tau)$ the corresponding phase noise difference.

The phase noise $e(\tau)$ at Tx and Rx will be modelled as an *Ornstein-Uhlenbeck (OU)* process [BTB19, HK17]. An OU process is described by means of a stochastic differential equation.

$$\frac{d}{d\tau}e(\tau) = \frac{1}{\beta} \cdot [\mu - e(\tau)] + \sqrt{\frac{2\alpha}{\beta}} \cdot \frac{d}{d\tau}w(\tau) \quad , \quad \beta > 0 \, , \; \alpha \geq 0 \tag{4.139}$$

Process $w(\tau)$ is a Wiener process or standard Brownian motion, β is the correlation time, α the diffusion constant and μ the drift term. An OU is a normal distributed random process, whose first-and second-order statistic is given by solving the respective stochastic differential equation using the Itô calculus.

$$e(\tau) \sim \mathcal{N}\left([e(0) - \mu] \cdot \exp\left\{-\frac{t}{\beta}\right\} + \mu, \alpha \cdot \left[1 - \exp\left\{-2\frac{t}{\beta}\right\}\right]\right) \tag{4.140}$$

The phase noise difference $\Delta_p(\tau)$ will be assumed as zero mean. This assumption is justified by the fact, that the considered radar (see [2]) operates in a coherent fashion, because a common reference signal is distributed to all transmit and receive channels. Hence, Tx and Rx channel are facing the same drift, such that the phase noise difference is drift free. Therefore, $\mu = 0$ can be assumed. Furthermore, the initial phase noise difference will be assumed as $e(0) = 0$, such that the OU process becomes a stationary process [HK17]. Figure 4.43 shows four realisations of such a OU process for various diffusion constants. The PDF of $e(\tau)$ becomes [BTB19]

$$e(\tau) \sim \mathcal{N}\left(0, \psi_e(\tau)\right) \quad , \tag{4.141}$$

with $\psi_e(\tau)$ the ACF of the process.

$$\psi_e(\tau) = \alpha \cdot \exp\left\{-\frac{|\tau|}{\beta}\right\} \tag{4.142}$$

As the phase noise will be assumed as a WSS process, the respective PSD $\Psi_e(f)$ is given by the Fourier transform of the ACF according to the Wiener–Khinchin theorem.

$$\Psi_e(f) = \sqrt{\frac{2}{\pi}} \cdot \frac{\alpha \cdot \beta}{1 + (2\pi f\beta)^2} \tag{4.143}$$

Because the phase noise processes $e_{Tx}(\tau)$ and $e_{Rx}(\tau)$ at Tx and Rx are modelled as normally distributed, the distribution of the phase noise difference $\Delta(\tau)$ is normally distributed too.

For parameter estimation the distribution of observation $x(\tau)$ in (4.138) has to be known. Because $h_p(\tau)$ is deterministic and the summation is a linear operation, the distribution of $\exp\{J\Delta_p(\tau)\}$ has to be determined. However, the distribution of $\exp\{J\Delta_p(\tau)\}$ is not easily
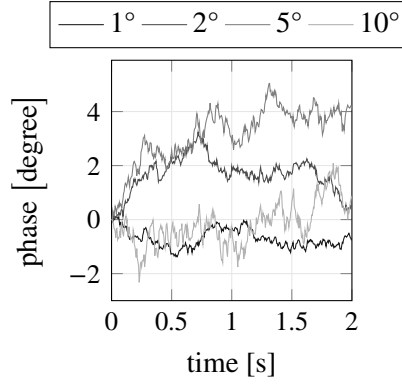
Figure 4.43: *Exemplary realisations of an Ornstein-Uhlenbeck process with different diffusion constants α; and parameters $\mu = 0$, $e(0) = 0$ and $\beta = 2$.*

given, as it includes the non-linear transformation of the Gaussian random variable $\Delta_p(\tau)$. Here, a Gaussian approximation of the non-linear transformation of a Gaussian random variable will be considered [GH08]. Assuming the phase noise difference as a process with minor variations, the first-order approximation can be used [BB93].

$$\exp\{J\Delta_p(\tau)\} \approx 1 + J\Delta_p(\tau) \tag{4.144}$$

Plugging into equation (4.138) the phase noise becomes *additive amplitude noise $p(\tau)$.*

$$x(\tau) \approx \sum_p h_p(\tau) \cdot \left[1 + J\Delta_p(\tau)\right] = \sum_p h_p(\tau) + J \sum_p h_p(\tau) \cdot \Delta_p(\tau) = \sum_p h_p(\tau) + p(\tau) \tag{4.145}$$

The received signal $x(\tau)$ can be separated into the deterministic part $\sum_p h_p(\tau)$ and the stochastic part $p(\tau) = J \sum_p h_p(\tau) \cdot \Delta_p(\tau)$.

According to the Gaussian approximation $x(\tau)$ is normal distributed, and the first- and second-order statistics have to be determined. For the first-order statistic it has to be noted, that the channel $h_p(\tau)$ is deterministic and therefore a constant w.r.t. the expectation operator.

$$\mathrm{E}\{x(\tau)\} = \mathrm{E}\left\{\sum_p h_p(\tau) + p(\tau)\right\} = \sum_p h_p(\tau) + \mathrm{E}\{p(\tau)\} \quad . \tag{4.146}$$

The phase noise is assumed as zero mean, such that the first-order statistic of $p(\tau)$ vanishes.

$$\mathrm{E}\{p(\tau)\} = \mathrm{E}\left\{J \sum_p h_p(\tau) \cdot \Delta_p(\tau)\right\} = J \sum_p h_p(\tau) \cdot \mathrm{E}\{\Delta_p(\tau)\} = 0 \tag{4.147}$$

Accordingly, the second-order statistic $\psi_{\mathrm{x}}(\tau_1, \tau_2)$ can be given.

$$\psi_{\mathrm{x}}(\tau_1, \tau_2) = \mathrm{E}\left\{\left[x(\tau_1) - \mathrm{E}\{x(\tau_1)\}\right] \cdot \left[x(\tau_2) - \mathrm{E}\{x(\tau_2)\}\right]^{\dagger}\right\}$$

$$= \mathrm{E}\left\{p(\tau_1) \cdot p(\tau_2)^{\dagger}\right\} = \psi_{\mathrm{p}}(\tau_1, \tau_2) \tag{4.148}$$

The second-order statistic $\psi_{\mathrm{p}}\left(\tau_1, \tau_2\right)$ of the amplitude phase noise $p\left(\tau\right)$ is given by

$$
\begin{aligned}
\psi_{\mathrm{p}}\left(\tau_1, \tau_2\right) &= \mathrm{E}\left\{\left[J \sum_p h_p\left(\tau_1\right) \cdot \varDelta_p\left(\tau_1\right)\right]^{\dagger} \cdot \left[J \sum_p h_p\left(\tau_2\right) \cdot \varDelta_p\left(\tau_2\right)\right]\right\} \\
&= \mathrm{E}\left\{\sum_{p_1} \sum_{p_2} h_{p_1}\left(\tau_1\right)^{\dagger} \cdot \varDelta_{p_1}\left(\tau_1\right) \cdot h_{p_2}\left(\tau_2\right) \cdot \varDelta_{p_2}\left(\tau_2\right)\right\} \\
&= \sum_{p_1} \sum_{p_2} \mathrm{E}\left\{h_{p_1}\left(\tau_1\right)^{\dagger} h_{p_2}\left(\tau_2\right) \cdot \varDelta_{p_1}\left(\tau_1\right) \varDelta_{p_2}\left(\tau_2\right)\right\} \quad .
\end{aligned}
\tag{4.149}
$$

Again, the channel $h_p\left(\tau\right)$ is deterministic and hence a constant w.r.t. the expectation operator, such that the second-order statistic can be simplified.

$$
\begin{aligned}
\psi_{\mathrm{p}}\left(\tau_1, \tau_2\right) &= \sum_{p_1} \sum_{p_2} \mathrm{E}\left\{h_{p_1}\left(\tau_1\right)^{\dagger} h_{p_2}\left(\tau_2\right)\right\} \cdot \mathrm{E}\left\{\varDelta_{p_1}\left(\tau_1\right) \varDelta_{p_2}\left(\tau_2\right)\right\} \\
&= \sum_{p_1} \sum_{p_2} \psi_{h_{p_1,p_2}}\left(\tau_1, \tau_2\right) \cdot \psi_{\varDelta_{p_1,p_2}}\left(\tau_1, \tau_2\right)
\end{aligned}
\tag{4.150}
$$

The second-order statistic $\psi_{\varDelta_{p_1,p_2}}\left(\tau_1, \tau_2\right)$ is given by

$$
\psi_{\varDelta_{p_1,p_2}}\left(\tau_1, \tau_2\right) = \mathrm{E}\left\{\left[e_{\mathrm{Tx}}\left(\tau_1 - \tau_{p_1}\right) - e_{\mathrm{Rx}}\left(\tau_1\right)\right] \cdot \left[e_{\mathrm{Tx}}\left(\tau_2 - \tau_{p_2}\right) - e_{\mathrm{Rx}}\left(\tau_2\right)\right]\right\} \quad .
\tag{4.151}
$$

The phase processes $e_{\mathrm{Rx}}\left(\tau\right)$ and $e_{\mathrm{Tx}}\left(\tau\right)$ are assumed as WSS, such the second-order statistic of the phase noise difference depends on the time difference $\tau = \tau_1 - \tau_2$. Hence, the second-order statistic can be simplified: $\psi_{\varDelta_{p_1,p_2}}\left(\tau_1, \tau_2\right) \Rightarrow \psi_{\varDelta_{p_1,p_2}}\left(\tau\right)$.

$$
\psi_{\varDelta_{p_1,p_2}}\left(\tau\right) = \psi_{e_{\mathrm{Tx,Tx}}}\left(\tau + \tau_{p_2} - \tau_{p_1}\right) - \psi_{e_{\mathrm{Tx,Rx}}}\left(\tau - \tau_{p_1}\right) - \psi_{e_{\mathrm{Tx,Rx}}}\left(\tau + \tau_{p_2}\right) + \psi_{e_{\mathrm{Rx,Rx}}}\left(\tau\right)
\tag{4.152}
$$

$\psi_{e_{\mathrm{Tx,Tx}}}\left(\tau\right)$ is the ACF of the phase noise process at Tx, $\psi_{e_{\mathrm{Rx,Rx}}}\left(\tau\right)$ is the ACF of the phase noise process at Rx, and $\psi_{e_{\mathrm{Tx,Rx}}}\left(\tau\right)$ is the cross-correlation function (CCF) between these noise processes. The ACF and the CCF of the phase noise processes at Tx and Rx depend on the signalling of the Tx and the Rx. Two signalling variants are distinguished: same signal at Tx and Rx (*common signalling*) or different signals at Tx and Rx (*disparate signalling*).

Common Signalling   If the same signal is used for transmission at the transmitter and homodyning at the receiver (e.g. in monostatic radar applications), the phase noise process at Tx and Rx can be assumed as equal: $e_{\mathrm{Tx}}\left(\tau\right) \equiv e_{\mathrm{Rx}}\left(\tau\right)$. Accordingly, the following relation for the ACF and the CCF holds.

$$
\psi_{e_{\mathrm{Tx,Tx}}}\left(\tau\right) = \psi_{e_{\mathrm{Rx,Rx}}}\left(\tau\right) = \psi_{e_{\mathrm{Tx,Rx}}}\left(\tau\right)
\tag{4.153}
$$

Disparate Signalling If the transmission and homodyning signal differ (e.g. in bistatic radar applications), then $e_{\text{Tx}}(\tau) \neq e_{\text{Rx}}(\tau)$ and the ACFs are $\psi_{e_{\text{Tx,Tx}}}(\tau) \neq \psi_{e_{\text{Rx,Rx}}}(\tau)$. Assuming the phase noise processes at Tx and Rx as uncorrelated, the CCF becomes $\psi_{\text{Rx,Tx}}(\tau) = 0$. Hence, the correlation of the phase noise difference simplifies.

$$\mathrm{E}\left\{\Delta_{p_1}\left(\tau' + \tau\right)\Delta_{p_2}\left(\tau'\right)\right\} = \psi_{e_{\text{Tx,Tx}}}\left(\tau + \tau_{p_2} - \tau_{p_1}\right) + \psi_{e_{\text{Rx,Rx}}}(\tau) \tag{4.154}$$

Phase Noise Model in FMCW MIMO Radar

If a FMCW MIMO radar is considered, phase noise correlation among the MIMO channels has to be taken into account. Consider the p-th propagation path with propagation delay τ_p. Due to the direction of transmission and impingement the delays $\boldsymbol{\tau}_{\text{Tx}_p}$ and $\boldsymbol{\tau}_{\text{Rx}_p}$ across the Tx and Rx array occur (see Section 4.3).

$$\boldsymbol{\tau}_{\text{Tx}_p} = \left[\tau_{1_{\text{Tx}_p}}, \ldots, \tau_{M_{\text{Tx}_p}}\right]^{\text{T}}$$

$$\boldsymbol{\tau}_{\text{Rx}_p} = \left[\tau_{1_{\text{Rx}_p}}, \ldots, \tau_{M_{\text{Rx}_p}}\right]^{\text{T}}$$

Also, lets introduce the vectors $\mathbf{e}_{\text{Tx}}(\tau)$ and $\mathbf{e}_{\text{Rx}}(\tau)$, which encompass the phase noise processes for each channel at Tx and Rx side, respectively.

$$\mathbf{e}_{\text{Tx}}(\tau) = \left[e_{\text{Tx}_1}, \ldots, e_{\text{Tx}_{M_{\text{Tx}}}}\right]^{\text{T}}$$

$$\mathbf{e}_{\text{Rx}}(\tau) = \left[e_{\text{Rx}_1}, \ldots, e_{\text{Rx}_{M_{\text{Rx}}}}\right]^{\text{T}}$$

Then, the phase noise difference for the p-th path and the m-th MIMO channel (formed by the m_{Tx}-Tx channel and the m_{Rx}-Rx channel) is given by

$$\Delta_{m_p}(\tau) = e_{\text{Tx}_{m_{\text{Tx}}}}\left(\tau - \tau_{m_{\text{Tx}_p}} - \tau_p + \tau_{m_{\text{Rx}_p}}\right) - e_{\text{Rx}_{m_{\text{Rx}}}}(\tau) \quad . \tag{4.157}$$

The vector of the MIMO transfer function $\mathbf{h}_p(\tau) \in \mathbb{C}^M$ is introduced too, which confounds the transfer function for each MIMO channel w.r.t. the p-th path. According to model (4.145) for a SISO observation, the model for a MIMO observation $\mathbf{x}(\tau)$ under the first-order approximation can be stated.

$$\mathbf{x}(\tau) \approx \sum_p \mathbf{h}_p(\tau) + \jmath \sum_p \mathbf{h}_p(\tau) \odot \boldsymbol{\Delta}_p(\tau) = \sum_p \mathbf{h}_p(\tau) + \mathbf{p}(\tau) \tag{4.158}$$

The PDF of $\mathbf{x}(\tau)$ is given by a multidimensional normal distribution, which mean vector and covariance matrix are subject to clarification.

The multidimensional phase noise process at Tx and Rx side will be assumed as zero mean, because a common signal is distributed to all Tx and Rx channels. Hence, the first-order statistic

of the MIMO observation $\mathbf{x}(\tau)$ is

$$E\{\mathbf{x}(\tau)\} = E\left\{\sum_p \mathbf{h}_p(\tau) + \mathbf{p}(\tau)\right\} = \sum_p \mathbf{h}_p(\tau) + \mathbf{0}_{M_{\text{Tx}} M_{\text{Rx}}} \quad . \tag{4.159}$$

Accordingly, the second-order statistic $\boldsymbol{\Psi}_{\mathbf{x}}(\tau_1, \tau_2)$ of the MIMO observation simplifies.

$$\boldsymbol{\Psi}_{\mathbf{x}}(\tau_1, \tau_2) = E\left\{[\mathbf{x}(\tau_1) - E\{\mathbf{x}(\tau_1)\}] \cdot [\mathbf{x}(\tau_2) - E\{\mathbf{x}(\tau_2)\}]^{\text{H}}\right\} = E\left\{\mathbf{p}(\tau_1) \cdot \mathbf{p}(\tau_2)^{\text{H}}\right\} = \boldsymbol{\Psi}_{\mathbf{p}}(\tau_1, \tau_2) \tag{4.160}$$

Hence, the second-order statistic depends on the phase noise only, as in the SISO case. Straight forward extension of the covariance model (4.150) for a SISO system gives the covariance of a MIMO system.

$$\begin{aligned}
\boldsymbol{\Psi}_{\mathbf{p}}(\tau_1, \tau_2) &= \sum_{p_1}\sum_{p_2} E\left\{\left[\mathbf{h}_{p_1}(\tau_1) \cdot \mathbf{h}_{p_2}(\tau_2)^{\text{H}}\right] \odot \left[\boldsymbol{\Delta}_{p_1}(\tau_1) \cdot \boldsymbol{\Delta}_{p_2}(\tau_2)^{\text{T}}\right]\right\} \\
&= \sum_{p_1}\sum_{p_2} E\left\{\mathbf{h}_{p_1}(\tau_1) \cdot \mathbf{h}_{p_2}(\tau_2)^{\text{T}}\right\} \odot E\left\{\boldsymbol{\Delta}_{p_1}(\tau_1) \cdot \boldsymbol{\Delta}_{p_2}(\tau_2)^{\text{T}}\right\} \\
&= \sum_{p_1}\sum_{p_2} \boldsymbol{\Psi}_{\mathbf{h}_{p_1, p_2}}(\tau_1, \tau_2) \odot \boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau_1, \tau_2)
\end{aligned} \tag{4.161}$$

The second-order statistic $\boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau_1, \tau_2)$ of the phase noise difference $\Delta(\tau)$ is given by

$$\boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau) = E\left\{\left[\mathbf{e}_{\text{Tx}}\left(\tau_1 - \tau_{p_1}\right) - \mathbf{e}_{\text{Rx}}(\tau_1)\right] \cdot \left[\mathbf{e}_{\text{Tx}}\left(\tau_2 - \tau_{p_2}\right) - \mathbf{e}_{\text{Rx}}(\tau_2)\right]\right\} \quad . \tag{4.162}$$

The MIMO phase noises $\mathbf{e}_{\text{Tx}}(\tau)$ and $\mathbf{e}_{\text{Rx}}(\tau)$ are assumed as a WSS process, such that the covariance matrix $\boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau_1, \tau_2)$ depends on the time difference $\tau = \tau_1 - \tau_2$ only: $\boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau_1, \tau_2) \Rightarrow \boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau)$.

$$\boldsymbol{\Psi}_{\boldsymbol{\Delta}_{p_1, p_2}}(\tau) = \boldsymbol{\Psi}_{\mathbf{e}_{\text{Tx,Tx}}}\left(\tau + \tau_{p_2} - \tau_{p_1}\right) - \boldsymbol{\Psi}_{\mathbf{e}_{\text{Tx,Rx}}}\left(\tau - \tau_{p_1}\right) - \boldsymbol{\Psi}_{\mathbf{e}_{\text{Tx,Rx}}}\left(\tau + \tau_{p_2}\right) + \boldsymbol{\Psi}_{\mathbf{e}_{\text{Rx,Rx}}}(\tau) \tag{4.163}$$

Simplification of the MIMO Phase Noise Model Consideration of model (4.161) for the covariance matrix of the measurement data is difficult and computational effortable in a parameter estimator, because the dimension of the covariance matrix (which is $MK \times MK$ in a practical implementation, see Section 5.1) can become quite large. Hence, simplifications, which are mathematically instead of physically driven, will be stated in the following.

In order to simplify the phase noise model it will be assumed, that a common signal is distributed. Furthermore, the phase noise is assumed as an approximately white process over time. Assuming all propagation paths to be differently delayed, the covariance matrix for different paths vanishes. Furthermore, correlation of the phase noise difference among the MIMO channels will be neglected. This assumption is not valid from a physical viewpoint, as some MIMO channels share the same Tx or Rx channel, such that their phase noise processes are inherently correlated. Lets consider time-division multiplexing at Tx side and parallel

receiver channels at Rx side (see Section 4.5.3), and assume the phase noise processes as i.i.d. process in time domain. Thus it can be assumed, that the phase noise is uncorrelated among the Tx channels, as the respective phase noise processes are observed at different times. However, correlations among the Rx channels are still present, such that M_{Tx} covariance matrices of size $M_{\mathrm{Rx}}K \times M_{\mathrm{Rx}}K$ remain. The Cholesky decompositions [PP12] of these covariance matrices have to be calculated by the parameter estimator (see Section 5.3), which has an overall computationally complexity of $\mathcal{O}\left(M_{\mathrm{Tx}}M_{\mathrm{Rx}}{}^3K^3\right)$. If the phase noise among the MIMO channels (combination of Tx and Rx channels) are assumed as uncorrelated processes, $M_{\mathrm{Tx}}M_{\mathrm{Rx}}$ Cholesky decompositions of matrices of size $K \times K$ have to be calculated by the parameter estimator. The overall computational complexity in that case is $\mathcal{O}\left(M_{\mathrm{Tx}}M_{\mathrm{Rx}}K^3\right)$. Hence, assumption of uncorrelated phase noise difference among the MIMO channels severely reduces the computational complexity of the parameter estimator.

Taking into account the above simplifications the covariance matrix of the phase noise difference becomes

$$\mathbf{\Psi}_{\mathbf{\Delta}_{p_1,p_2}}(\tau) \approx \begin{cases} \mathbf{0} & , \ p_1 \neq p_2 \\ \mathrm{diag}\left\{\mathbf{\psi_e}(\tau)\right\} & , \ p_1 = p_2 = p \end{cases} . \tag{4.164}$$

The entries of vector $\mathbf{\psi_e}(\tau)$ are given by the ACF (4.142), whereas each entry is described in terms of another set of parameters a and β. Plugging the simplified covariance (4.164) into (4.161) yields the second-order statistic of the MIMO phase noise difference process $\mathbf{p}(\tau)$.

$$\mathbf{\Psi_p}(\tau) \approx \mathrm{diag}\left\{\mathbf{\psi_e}(\tau)\right\} \odot \sum_p \mathbf{\Psi}_{\mathbf{h}_p}(\tau) \tag{4.165}$$

According to this model the phase noise statistic is independent on the delay of a received path. This is a tough assumption, as range correlation effects are therefore neglected. However, the model complexity is severely reduced by this assumption. Summarised, the MIMO phase noise $\mathbf{p}(\tau)$ in delay domain is modelled as an additive, proper complex, normally distributed process.

$$\mathbf{p}(\tau) \sim \mathcal{CN}\left(\mathbf{0}, \mathrm{diag}\left\{\mathbf{\psi_e}(\tau)\right\} \odot \sum_p \mathbf{\Psi}_{\mathbf{h}_p}(\tau)\right) \tag{4.166}$$

Note, that also the modelling as proper complex process is an approximation due to numerical complexity reasons.

In order to derive the second-order statistic in time domain $\mathbf{\Psi_p}(t)$, a statement regarding the propagation channel is necessary. The propagation channel is time-variant, such that propagation paths can appear or vanish over time t. Therefore, the number of propagation paths becomes a time-dependent quantity $P(t)$. However, it will be assumed, that the number of propagation paths stays fix during the observation time, hence $P(t) = P$. Assuming the phase noise process $\mathbf{e}(t)$ as i.i.d. in the time domain, the second-order statistics in time domain is

$$\mathbf{\Psi_p}(t) = \sum_p \mathbf{\Psi}_{\mathbf{h}_p}(t) \odot \mathrm{E}\left\{\mathbf{\Delta}_p\left(t'+t\right) \cdot \mathbf{\Delta}_p\left(t'\right)^{\mathrm{T}}\right\} = \delta(t) \cdot \mathbf{I} \odot \sum_p \mathbf{\Psi}_{\mathbf{h}_p}(t) \quad . \tag{4.167}$$

Accordingly, the second-order statistic $\mathbf{\Psi_p}(\tau, t)$ in the delay-time domain is

$$\mathbf{\Psi_p}(\tau, t) = \mathrm{E}\left\{\mathbf{p}\left(\tau' + \tau, t' + t\right) \cdot \mathbf{p}\left(\tau', t'\right)^{\mathrm{H}}\right\} = \delta(t) \cdot \mathrm{diag}\left\{\boldsymbol{\psi_e}(\tau)\right\} \odot \sum_p \mathbf{\Psi_{h_p}}(\tau, t) \quad (4.168)$$

and the delay-time correlation function $\mathbf{\Psi_{h_p}}(\tau, t)$ of the p-th path is given by

$$\mathbf{\Psi_{h_p}}(\tau, t) = \mathrm{E}\left\{\mathbf{h}_p\left(\tau' + \tau, t' + t\right) \cdot \mathbf{h}_p\left(\tau', t'\right)^{\mathrm{H}}\right\} \quad , \quad (4.169)$$

with $\mathbf{h}_p(\tau, t)$ the delay-time model of the p-th propagation path.

Summarised, the MIMO phase noise in delay-time domain is modelled as an additive, normally distributed random process with probability density function

$$\mathbf{p}(\tau, t) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Psi_p}(\tau, t)\right) \quad . \quad (4.170)$$

4.6.4 Noise Model and Verification

Complete Noise Model

The presented noise models for a FMCW MIMO radar account for three noise sources: phase noise, leakage noise and thermal noise. Because all noise sources are additive disturbances of the receive signal an overall noise process $n(\tau, t)$ in the delay-time domain can be defined.

$$n(\tau, t) = w(\tau, t) + l(\tau, t) + p(\tau, t) \approx w(\tau, t) + l(\tau, t) + je(\tau, t) \cdot h(\tau, t) \quad (4.171)$$

As each of the noise processes is modelled as normally distributed the overall noise process is normally distributed too. Hence, the first-order and central second-order moment of $n(\tau, t)$ have to be determined. Because the overall noise process is given by the summation of multiple noise processes and these noise processes are zero-mean, the first-order moment is

$$\mathrm{E}\left\{n(\tau, t)\right\} = 0 \quad . \quad (4.172)$$

The phase noise, thermal noise and leakage noise are WSS processes and uncorrelated w.r.t. each other. Accordingly, the central second-order moment $\psi_{\mathrm{n}}(\tau, t)$ of the overall noise process can be stated.

$$\psi_{\mathrm{n}}(\tau, t) = \mathrm{E}\left\{n\left(\tau' + \tau, t' + t\right)^{\dagger} n\left(\tau', t'\right)\right\}$$

$$= \delta(t) \cdot \left[\eta \cdot \delta(\tau) + \psi_l(\tau) + \psi_e(\tau) \cdot \sum_p^P \psi_{\mathrm{h}_p}(\tau, t)\right] \quad (4.173)$$

The respective PDF of the overall noise process in the delay-time domain is

$$n(\tau, t) \sim \mathcal{CN}\left(0, \delta(t) \cdot \left[\eta \cdot \delta(\tau) + \psi_l(\tau) + \psi_e(\tau) \cdot \sum_p \psi_{\mathrm{h}_p}(\tau, t)\right]\right) \quad . \quad (4.174)$$
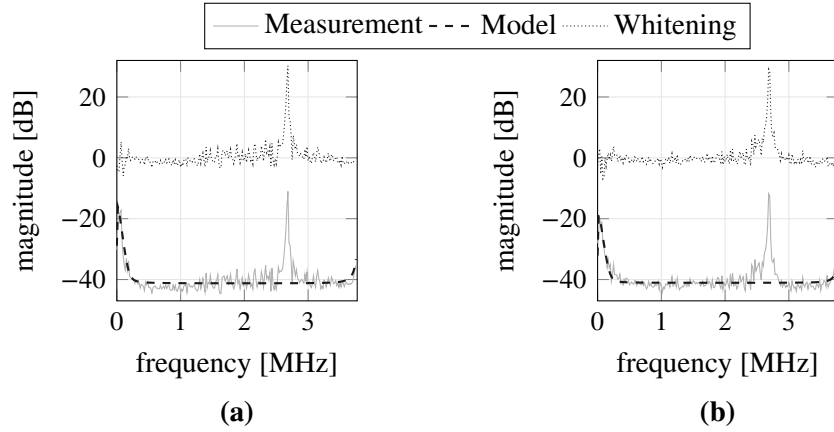
Figure 4.44: *Average power spectral density (average over 64 periods) of measurements with a single target, the fitted thermal and leakage noise model, and the resulting signal after whitening.*

The parameters of the covariance model are unknown in advance and hence must be estimated from measurements too, see Chapter 5 for details on the estimator. As the covariance model parameters are not of immediate interest, rather than the parameters of a propagation path, they enter the estimator as so called *nuisance parameters*.

Experimental Verification of the Noise Model

Figure 4.44 shows the range spectra of two measurements with an FMCW radar. For both measurements the parameters of the leakage and thermal noise model have been estimated and the respective PSDs are shown in the figures. The phase noise model will be neglected for the moment. The covariance matrix is constructed from the estimated leakage and thermal noise model, and this covariance matrix has been used to whiten (multiplication of the measurements with the inverse Cholesky factor of the covariance matrix) the measurements. After whitening the contribution from the leakage noise is suppressed, such that the overall noise floor becomes more uniform.

Last, the influence of the phase noise model will investigated. A measurement with a single corner reflector as target will be considered therefore, see Fig. 4.45. The physical path emerging from the reflection at the corner reflector is estimated by the approach presented in Chapter 5. Afterwards, the residual signal is calculated by coherently subtracting the path model (i.e. the channel model corresponding to the estimated path) from the measurements. A small signal portion remains in the residual signal. This signal portion can be due to model errors or phase variations in the measurements. A parameter estimator may attempt to cover them by additional propagation paths in the model. Hence, an overestimation of the number of propagation paths can occur. The residuals are used to estimate the parameters of the covariance model (4.173). For comparison, the parameters of the covariance model have been estimated for the case of considering phase noise (see Fig. 4.45(b)) and neglecting phase noise (see Fig. 4.45(a)) in the model. If phase noise is neglected the estimated PSD of the noise process does not fit well with the noise profile of the residuals, such that the whitening does not suppress the leakage noise. An explanation for the bad fit is the residual signal portion from the path model subtraction, which is not covered by the noise model. However, if phase noise is taken into account in
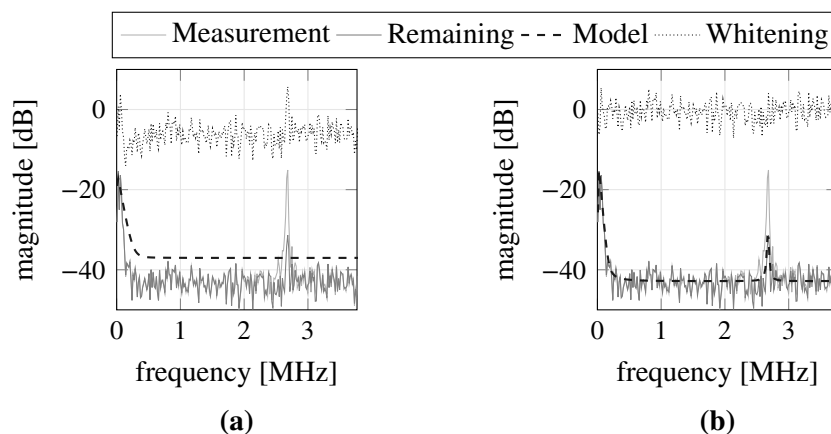
Figure 4.45: *Average power spectral density (averaged over 64 periods) of the measurements, the residuals after subtraction of the estimated path, the fitted model of the power spectral density of the noise, and the whitening result of the residuals. Phase noise has been (a) neglected or (b) considered in the noise model.*

the covariance model, the fit of the PSD with the residuals is much better, as the covariance model covers the residual signal portion. Consequently, the whitened residual is white noise like, which indicates a proper choice of the noise model in that case. Hence, the consideration of phase noise in the covariance model is also advantageously to cover model errors.

4.7 Summary

In this chapter, a detailed model of the response and noise sources of the considered FMCW MIMO radar has been developed.

The basic signal theory and hardware setup of an FMCW with homodyning receiver has been reviewed. A model for the received signal with neglected hardware effects has been presented. In order to take hardware effects into account, a model of the response of the radar hardware has been developed. The response of the MIMO radar system has been divided into the response of the transceiver chains and the response of the MIMO antenna array. The joint calibration of the model of the transceiver chain and the model of the antenna array using calibration data has been presented.

Hardware impairments are described by a behavioural transceiver model. It has been shown that, due to non-linearities in the signal generator, the transmit signal of the considered FMCW radar constitutes of multiple signals. As a result, the received baseband signal is distorted, because multiple beat frequencies for a single target occur. This may cause false detections. In order to account for these signal disturbances, a non-linear, dynamic response model of the transceiver chain has been developed using a generalised Hammerstein model. The developed response model allows the de-embedding (see Section 1.1.1) and mitigation of the signal disturbances, which improves the target detection performance.

In order to model the array response narrowband as well as wideband modelling has been discussed. Narrowband models are the de facto state-of-the-art. It turned out, that the narrowband array model is not suitable to describe the array response, because narrowband

modelling assumptions are violated. The radar is of broadband type (huge signal bandwidth and large array aperture), such that angle-delay coupling occurs, which is not coped by a narrowband array model. The wideband model remains useful, as it copes with angle-delay coupling. On the downside, the wideband model is much more complicated and hence causes a much higher numerical effort of a parameter estimator.

The sampling of the receive signals in conjunction with the interference free accessing of all MIMO channels by TDM has been discussed as well. It has been pointed out, that TDM accessing of a time-variant channel causes angle-Doppler coupling. A model, which accounts for angle-Doppler coupling has been developed. By exploiting this model in a parameter estimator, the angle-Doppler coupling is compensated and the Doppler bandwidth of a traditional TDM system is enhanced.

Three occurring noise sources have been considered: thermal noise, leakage noise and phase noise. For each noise sources a parametric model in the delay and time domain, and dimension of MIMO channels has been derived. The leakage and phase noise model are novelties for FMCW MIMO radars. The consideration of leakage noise enables the detection of close and weak targets. The consideration of phase noise is advantageously to cope with model errors. Model errors cause residual signals, which deteriorate the estimation quality and result in an overestimation of the number of propagation paths. The phase noise model accounts for such residual signals and suppresses them. Hence, an overestimation of the number of propagation paths is circumvented. The parameters of the noise models are unknown, and hence are subject to the parameter estimation problem as well.

PARAMETER ESTIMATION FROM OB-SERVATIONS BY AN FMCW MIMO RADAR

THE physical model and the radar device model from the previous two chapters, and the MAP estimator will be considered for the development of a model-based signal processing scheme. In this chapter, the implementation of a parameter estimator for a broadband, monostatic FMCW MIMO radar will be outlined. The resulting estimator can be applied to the radar data cube in Fig. 5.1. Target parameters as delay, angle and Doppler will be estimated, which can be used for e.g. target localisation.

Related Work and Own Contributions

Model-based signal processing is not a new application to radar signal processing, and there is much work done already [DEH19, LS08, Ric13]. Also, the idea of compressive sensing [CW08, ME11] has been transferred to radar applications [Coh+18, DEH19, YPP10]. Furthermore, space-time adaptive processing (STAP) methods have been introduced to MIMO radar [LS08, Ric13], in order to improve the robustness of target detection against interference as clutter, jamming or multipath by maximising the signal-to-interference-plus-noise ratio (SINR) at the processor output [Mel04]. Standard STAP processing is conducted for each acquired delay bin and each target separately. Here, joint processing of the space- (spatial-), delay and time-domain of MIMO radar observations is considered. Also, it will not be distinguished between target scattering and clutter in the estimator. Instead, every received propagation contribution will be resolved by the estimator. Separation of the resolved contributions into scattering from a true target and clutter will not be accomplished here. Hence, target detection and clutter separation is not considered.

A variety of model-based parameter estimators have been proposed for radars, ranging from fast Fourier transform (FFT)-based processing schemes [JAA17, Pay07, Sam+12], to subspace-based processing schemes as MUSIC [BRH14, Zha+10] and ESPRIT [DBG08, LZG19, NS10], to maximum-likelihood (ML) type estimators [LS08, SM11, SS98, XLS08]. ML methods are mainly criticised due to their high computational complexity, such that a real-time application of the estimator is hardly feasible. However, a ML estimator is very flexible regarding the data model. On the other hand, ESPRIT and FFT-based algorithms are computationally efficient
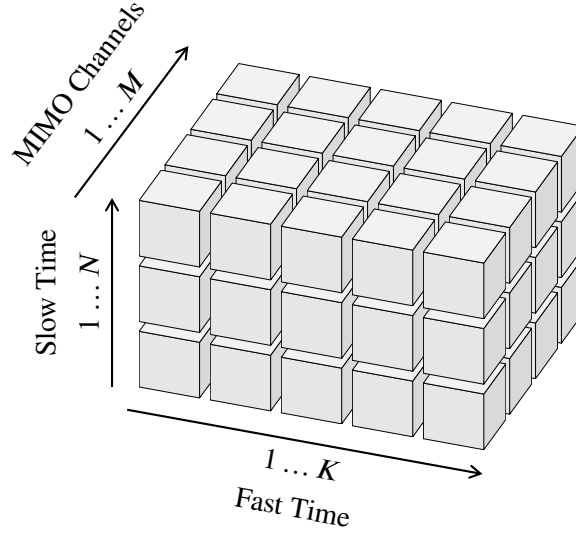
Figure 5.1: *Radar data cube $\mathcal{Y}$, which concatenates all observations in spatial-, time- and delay domain into a 3-dimensional tensor. The spatial domain is sampled by the MIMO channels, which are spanned by the Tx and Rx array. (Picture adapted from [Sta19].)*

methods, but demand a certain structure of the data model (Kronecker structure and shift-invariant observations in each dimension). Apart from the radar community, model-based or parametric signal processing is also well known in channel sounding, in order to derive the parameters of propagation paths from measurement data: SAGE [Fle+99], RARE [PMB04], ESPRIT [HTR03, 8] or RIMAX [LKT12, Ric05]. The close link between radar and channel sounding has been pointed out in Section 1.1, such that these algorithms are also conceivable for estimating target parameters.

Basically, each one of the target parameters τ, φ and v can be assigned to one dimension of the radar data cube in Fig. 5.1. For example, the delay relates to the fast time, the Doppler to the slow time, and the azimuth relates to the MIMO channels. Common signal processing algorithms for the radar data cube exploit this fact by processing each dimension separately, which is computationally very attractive. However, the separate processing lowers the resolution capability. Furthermore, the separate processing necessitates a certain structure of the system model, the so called *Kronecker structure* [Ric05]. Under this model structure, each parameter causes only variations of the measurements in his respective observation dimension. Hence, decoupled observation dimensions are assumed in the system model. For example, variations of the azimuth angle causes variations over the dimension of the MIMO channels only. This is only true, if a narrowband model of the array response is assumed. The Kronecker model structure is not be given by the system model, which will be considered in the following. Because angle-delay coupling is considered due to the wideband array model, and angle-Doppler coupling due to the TDM of the Tx channels is considered, the system model has no Kronecker structure. Furthermore, joint processing of all dimensions shall be applied to achieve high-resolution. Summarised, the common and computational attractive estimators in radar signal processing cannot be applied, because of the considered system model. Instead, a MAP estimator has to be applied, because this kind of estimators are not restricted by a certain model structure.

First, the statistical model of the noisy radar observations, stacked into the radar data cube

Fig. 5.1, will be developed in Section 5.1. There, the physical model $\mathcal{M}^{\mathrm{P}}$ of the propagation (see Chapter 3) and the device model $\mathcal{M}^{\mathrm{D}}$ of the MIMO radar from (see Chapter 4) are considered, which gives a novel system model $\mathcal{M}^{\mathrm{S}}$. Next, the statistical model, which considers the system model and the noise sources, will be derived. The noise sources presented in Section 4.6 including leakage noise and phase noise will be taken into account as noise disturbances. Based on the statistical model the likelihood of the observations in terms of the parameters of interest (the parameters of the physical model) and the nuisance parameters (parameters of the noise model) will be given. Afterwards, the statistical inversion using Bayes' rule is described and the required prior distribution of the parameters is stated using the principle of indifference. The MAP estimator will be used as point estimator to derive parameter values from the posterior distribution [28]. As the MAP is considered and the prior is derived from the principle of indifference, the parameter estimator becomes a maximum-likelihood estimator with box constraints on the parameters. Theoretic properties of ML estimators will be given, and how to calculate asymptotic values of the variance of the parameter values. The ML estimator results in an objective, such that an optimisation scheme is required, which takes box constraints into account. As the objective is of non-convex nature (see APPENDIX B) an iterative optimisation scheme is proposed, which includes a coarse global search and a subsequent gradient-based optimisation algorithm for refinement of the parameter estimates. The optimisation scheme will be described in Section 5.3. The complete parameter estimation framework will be outlined in Section 5.4. Section 5.5 concludes the chapter.

Parts of this chapter have been published in [6, 25, 10, 28].

5.1 Statistical Model – Model of the Noisy Observations

The statistical model describes the PDF of the observations in terms of the parameters of interest. This PDF is the likelihood. Subsequently, a statistical model for the radar observations, i.e. the radar data cube, will be derived. This model will be later on used to infer the parameters of interest from the measurements. The statistical model will be derived for observations with an FMCW MIMO radar with parallel receiver channels and TDM accessing of the transmit channels, as presented in [2]. First, some termini have to be clarified: a MIMO observation denotes the measurement of multiple MIMO snapshots, whereas one MIMO snapshot is a single measure of all MIMO channels. The radar data cube can be represented by a 3-way tensor $\mathcal{Y} \in \mathbb{C}^{K \times N \times M}$, with $M = M_{\mathrm{Tx}} \cdot M_{\mathrm{Rx}}$ denoting the number of MIMO channels, K the number of samples per channel, and N the number of MIMO snapshots. The considered MIMO radar requires the time $T = N \cdot M_{\mathrm{Tx}} \cdot T_{\mathrm{P}}$ to capture the radar data cube.

In case of a moving target the observation angle and the target range varies during the capturing time of $\mathcal{Y}$. Hence, these parameters are basically time-variant [27]. Commonly, time-variations of the delay are considered only, which results in Doppler (see Section 3.2.1); whereas the angle or path weight are assumed as stationary. However, the Doppler model in Section 3.2.1 assumes a homogeneously and linearly moving target (neither the moving direction nor the velocity changes). Hence, the Doppler shift is assumed as stationary too. It will be assumed, that the target is slowly moving compared to the capturing time of $\mathcal{Y}$, such that the path parameters

(angle, path weight and the Doppler shift) are fix during the observation time *T*. In order to ensure this assumption, the observation time must be chosen sufficiently small [27], e.g. *N* must be small. If the target is rapidly moving, such that static parameters are not to be considered, a dynamic system model to describe the time-variant behaviour of the parameters have to be employed [27].

5.1.1 The System Model – Model of Observations in the Noiseless Case

In the noiseless case, only the model of the physical system under identification $\mathcal{M}^{\mathrm{P}}$ (see Section 3.2), and the model of the response of the observer $\mathcal{M}^{\mathrm{D}}$ (see Section 4.2 and Section 4.3) are employed to describe the observations. The resulting model is the system model $\mathcal{M}^{\mathrm{S}}$, i.e., the system model describes the noiseless observations.

Model of a SISO Observation

First, a SISO observation will be considered. The presence of an observation device, i.e. the influence of the device response as well as the response of the Tx and Rx antenna, will be neglected for the moment. An observation of a single specular propagation path *p* in the noiseless case is given by (4.117), taking into account the narrowband, polarimetric scattering matrix $\mathbf{\Gamma}_p\left(f_{\mathrm{c}}\right)$ from Section 3.2.1. Resolution of the polarimetric scattering matrix $\mathbf{\Gamma}_p\left(f_{\mathrm{c}}\right)$ requires polarimetric measurements or at least polarimetric calibration data of the employed antennas [23, Ric05]. Because only single polarised antennas are employed at the considered radar and polarimetric calibration data of the antenna array are not available, polarimetric properties of the antennas and the propagation channel must be neglected.

The considered FMCW radar gathers real-valued baseband signals only, because no IQ stretch processor is employed. For the data model derivation the analytic signal (right side part of the signal spectrum) will be considered. Signal model (4.117) describes the mirrored version of the analytic signal, such that a time reversion $-t$ has to be considered as $u(-t) \circ\!\!-\!\!\bullet U(-f)$ holds for complex signals. According to equation (4.117), the *k*-th sample of the complex baseband signal *x* for the *n*-th FMCW ramp (signal period) is modelled as

$$x\left(k,n\right) \approx \gamma'_p \cdot \exp\left\{\jmath 2\pi \frac{W}{T_{\mathrm{M}}}\tau'_p k T_{\mathrm{S}}\right\} \cdot \exp\left\{\jmath 2\pi v'_p n T_{\mathrm{P}}\right\} \cdot \exp\left\{-\jmath 2\pi f_0 \tau'_p\right\} \quad , \qquad (5.1)$$

with $k = 0 \ldots K - 1$ and $n = 0 \ldots N - 1$. The delay and the Doppler shift of the *p*-th path are denoted by τ'_p and v'_p. The phase term $2\pi f_0 \tau_p$ is a constant. Therefore, the term can be swapped to the path weight and the model becomes

$$x\left(k,n\right) \approx \gamma_p \cdot \exp\left\{\jmath 2\pi \frac{W}{T_{\mathrm{M}}}\tau'_p k T_{\mathrm{S}}\right\} \cdot \exp\left\{\jmath 2\pi v'_p n T_{\mathrm{P}}\right\} \quad . \qquad (5.2)$$

Delay and Doppler shift have a difference in their order of magnitudes of up to 10^9, e.g. a delay in *ns* and a Doppler shift in *Hz* is commonly given for the considered scenarios (slowly moving targets). Thus, numerical instabilities in the parameter estimation framework

can occur [Ric05]. In order to avoid numerical instability the normalised delay τ_p and Doppler shift v_p are introduced.

$$\tau_p = 2\pi \frac{W}{T_{\mathrm{M}}} \tau_p' T_{\mathrm{S}} \tag{5.3a}$$

$$v_p = 2\pi v_p' T_{\mathrm{P}} \tag{5.3b}$$

Due to this normalisation the parameter range is now: $\tau_p = [0, 2\pi[$ and $v_p = [-\pi, \pi[$. Hence, numerical instabilities are circumvented as the parameters have now a similar range. After the normalised delay and Doppler shift have been estimated normalisation (5.3) has to be inverted in order to get the true delay and Doppler values. Plugging the normalisations into the model yields

$$x(k, n) \approx \gamma_p \cdot \exp\left\{J\tau_p k\right\} \cdot \exp\left\{Jv_p n\right\} \quad . \tag{5.4}$$

Model (5.4) describes a single observation at delay k and time n. In order to extend this model to the complete set of observations the vectors $\mathbf{k} = \begin{bmatrix} 0, \dots, K-1 \end{bmatrix}^{\mathrm{T}}$ and $\mathbf{n} = \begin{bmatrix} 0, \dots, N-1 \end{bmatrix}^{\mathrm{T}}$ of sampling points are introduced. Then, the matrix $\mathbf{X} \in \mathbb{C}^{K \times N}$ comprising all samples can be modelled.

$$\mathbf{X} = x(\mathbf{k}, \mathbf{n}) \approx \gamma_p \cdot \exp\left\{J\tau_p \mathbf{k}\right\} \cdot \exp\left\{Jv_p \mathbf{n}^{\mathrm{T}}\right\} \tag{5.5}$$

Commonly, the observations are stacked into a vector rather than stacking them into the columns of a matrix [Ric05]. Using the vec-operator the model of the vectorised matrix $\mathbf{X}$ is

$$\mathbf{x} = \mathrm{vec}\left\{\mathbf{X}\right\} \approx \gamma_p \cdot \left(\exp\left\{Jv_p \mathbf{n}\right\} \otimes \exp\left\{J\tau_p \mathbf{k}\right\}\right) \in \mathbb{C}^{K \cdot N} \quad , \tag{5.6}$$

which has a so called Kronecker structure due to the Kronecker operator $\otimes$. Kronecker structured models are very popular in signal processing, because their structure can be employed for numerical simplifications, see e.g. [Ric05].

Now consider the presence of P resolvable propagation paths. According to the findings in Section 3.2.2 these propagation paths are superimposed. Hence, the model for P paths is

$$\begin{aligned} \mathbf{x} &\approx \gamma_1 \cdot \left(\exp\left\{Jv_1 \mathbf{n}\right\} \otimes \exp\left\{J\tau_1 \mathbf{k}\right\}\right) + \dots + \gamma_P \cdot \left(\exp\left\{Jv_P \mathbf{n}\right\} \otimes \exp\left\{J\tau_P \mathbf{k}\right\}\right) \\ &= \left(\begin{bmatrix} \exp\left\{Jv_1 \mathbf{n}\right\}, \dots, \exp\left\{Jv_P \mathbf{n}\right\} \end{bmatrix} \diamond \begin{bmatrix} \exp\left\{J\tau_1 \mathbf{k}\right\}, \dots, \exp\left\{J\tau_P \mathbf{k}\right\} \end{bmatrix}\right) \cdot \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_P \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{B}_v(\boldsymbol{\nu}) \diamond \mathbf{B}_\tau(\boldsymbol{\tau}) \end{bmatrix} \cdot \boldsymbol{\gamma} \quad , \end{aligned} \tag{5.7}$$

with $\diamond$ denoting the Khatri-Rao product (column-wise Kronecker product). For sake of

notational convenience, the parameter vectors

$$\boldsymbol{\varphi} = \left[\varphi_1, \ldots, \varphi_P\right]^{\mathrm{T}} \in \mathbb{R}^P \tag{5.8a}$$

$$\boldsymbol{\vartheta} = \left[\vartheta_1, \ldots, \vartheta_P\right]^{\mathrm{T}} \in \mathbb{R}^P \tag{5.8b}$$

$$\boldsymbol{\tau} = \left[\tau_1, \ldots, \tau_P\right]^{\mathrm{T}} \in \mathbb{R}^P \tag{5.8c}$$

$$\boldsymbol{\nu} = \left[\nu_1, \ldots, \nu_P\right]^{\mathrm{T}} \in \mathbb{R}^P \tag{5.8d}$$

$$\boldsymbol{\gamma} = \left[\gamma_1, \ldots, \gamma_P\right]^{\mathrm{T}} \in \mathbb{C}^P \tag{5.8e}$$

and the matrices

$$\mathbf{B}_\tau \left(\boldsymbol{\tau}\right) = \exp\left\{\jmath \mathbf{k}\boldsymbol{\tau}^{\mathrm{T}}\right\} \in \mathbb{C}^{K \times P} \tag{5.9a}$$

$$\mathbf{B}_\nu \left(\boldsymbol{\nu}\right) = \exp\left\{\jmath \mathbf{n}\boldsymbol{\nu}^{\mathrm{T}}\right\} \in \mathbb{C}^{N \times P} \quad, \tag{5.9b}$$

have been introduced. The columns of the matrices comprise the respective delay and Doppler harmonic for each propagation path.

So far, the influence of the Tx and Rx antenna on the observation has been ignored. In order to account also for the influence of the antennas, equation (5.4) will be considered again. In Section 4.3 it was found, that a wideband model of the antenna array has to be considered to properly account for the antenna embedding in the measurements. Hence, the antenna influence varies over the delay dimension k. Furthermore, because dual-polarimetric calibration data of the antenna are not available, a single polarised antenna is considered only. Neglecting the polarimetric properties of antennas cause model errors, which influence the parameter estimation accuracy [23, LKT12]. However, these model errors cannot be circumvented. In order to account for the influence of the Tx and Rx antenna equation (5.4) is extended to

$$x\left(k, n\right) \approx \gamma_p \cdot \left[b_{\mathrm{Tx}}\left(k; \vartheta_{\mathrm{Tx}_p}, \varphi_{\mathrm{Tx}_p}\right) \cdot b_{\mathrm{Rx}}\left(k; \vartheta_{\mathrm{Rx}_p}, \varphi_{\mathrm{Rx}_p}\right)\right] \cdot \exp\left\{\jmath\tau_p k\right\} \cdot \exp\left\{\jmath\nu_p n\right\} \quad, \tag{5.10}$$

with $b_{\mathrm{Tx}}\left(k; \vartheta_{\mathrm{Tx}_p}, \varphi_{\mathrm{Tx}_p}\right)$ and $b_{\mathrm{Rx}}\left(k; \vartheta_{\mathrm{Rx}_p}, \varphi_{\mathrm{Rx}_p}\right)$ the response of the Tx and Rx antenna w.r.t. the directions of radiation $\left(\vartheta_{\mathrm{Tx}_p}, \varphi_{\mathrm{Tx}_p}\right)$ and impingement $\left(\vartheta_{\mathrm{Rx}_p}, \varphi_{\mathrm{Rx}_p}\right)$ of the p-th path, respectively. Next, a monostatic radar with co-located Tx and Rx antennas is considered. Hence, the angles at Tx and Rx side are equal: $\varphi = \varphi_{\mathrm{Rx}} = \varphi_{\mathrm{Tx}}$ and $\vartheta = \vartheta_{\mathrm{Rx}} = \vartheta_{\mathrm{Tx}}$. Thus, the model reduces

$$x\left(k, n\right) \approx \gamma_p \cdot b\left(k; \varphi_p, \vartheta_p\right) \cdot \exp\left\{\jmath\tau_p k\right\} \cdot \exp\left\{\jmath\nu_p n\right\} \quad, \tag{5.11}$$

with $b\left(k; \varphi_p, \vartheta_p\right) = b_{\mathrm{Tx}}\left(k; \varphi_p, \vartheta_p\right) \cdot b_{\mathrm{Rx}}\left(k; \varphi_p, \vartheta_p\right)$ the joint response of the Tx and Rx antenna w.r.t. the azimuth angle φ_p and elevation angle ϑ_p. Introduce matrix $\mathbf{B}_\varphi\left(\boldsymbol{\varphi}, \boldsymbol{\vartheta}\right)$, in which the responses of the antenna to all paths are column-wisely stacked.

$$\mathbf{B}_\varphi\left(\boldsymbol{\varphi}, \boldsymbol{\vartheta}\right) = \left[\mathbf{b}\left(\mathbf{k}; \varphi_1, \vartheta_1\right), \ldots, \mathbf{b}\left(\mathbf{k}; \varphi_P, \vartheta_P\right)\right] \in \mathbb{C}^{K \times P} \tag{5.12a}$$

$$\mathbf{b}\left(\mathbf{k}; \varphi_p, \vartheta_p\right) = \left[b\left(0; \varphi_p, \vartheta_p\right), \ldots, b\left(K-1; \varphi_p, \vartheta_p\right)\right]^{\mathrm{T}} \in \mathbb{C}^K \tag{5.12b}$$

Then, the model becomes

$$\mathbf{x} = \left[\mathbf{B}_\nu(\boldsymbol{\nu}) \diamond \left[\mathbf{B}_\varphi(\boldsymbol{\varphi}, \boldsymbol{\vartheta}) \odot \mathbf{B}_\tau(\boldsymbol{\tau})\right]\right] \cdot \boldsymbol{\gamma} \quad , \tag{5.13}$$

with $\odot$ the Schur product (element-wise matrix product).

Last, the embedding of the device response into the observations has to be clarified. In Section 4.2, two models for the response of a SISO FMCW radar have been derived. In the following, the model presented in Section 4.2.1 will be assumed as sufficient. Hence, bandwidth limitation is done as a pre-processing step before the estimation, in order to suppress the ghost targets. Also, the more complicated model presented in Section 4.2.2 has to be accounted for under very high SNR conditions, which will not be considered here. Accounting for the device response $G(k)$ the model becomes

$$x(k, n) \approx \gamma_p \cdot g(k) \cdot b\left(k; \varphi_p, \vartheta_p\right) \cdot \exp\left\{J\tau_p k\right\} \cdot \exp\left\{J\nu_p n\right\} \quad . \tag{5.14}$$

Plugging the sampled time-invariant radar response $\mathbf{g} = \left[G(0), \dots, G(K-1)\right]^{\mathrm{T}} \in \mathbb{C}^K$ into model (5.13) yields the model for a SISO observation in the noiseless case.

$$\mathbf{x} = \left[\mathbf{B}_\nu(\boldsymbol{\nu}) \diamond \left[\left(\mathrm{diag}\{\mathbf{g}\} \cdot \mathbf{B}_\varphi(\boldsymbol{\varphi}, \boldsymbol{\vartheta})\right) \odot \mathbf{B}_\tau(\boldsymbol{\tau})\right]\right] \cdot \boldsymbol{\gamma} \tag{5.15}$$

Model of a MIMO Observation

In case of MIMO observations not only multiple channels have to taken into account to derive a model. It has been pointed out in Section 4.5, that the MIMO accessing scheme, i.e. how all MIMO channels are accessed, has to be taken into account too. Here, a TDM scheme at the Tx and parallel receiving channels will be considered. Thus, the phase term ϕ_{Doppler} related to the Doppler shift ν becomes dependent on the actual Tx channel, see Section 4.5.

$$\frac{1}{2\pi}\phi_{\mathrm{Doppler}} \approx \nu' \cdot m_{\mathrm{Tx}}T_{\mathrm{P}} + \nu' \cdot nM_{\mathrm{Tx}}T_{\mathrm{P}} = \nu' \cdot T_{\mathrm{P}} \cdot M_{\mathrm{Tx}} \cdot \left(n + \frac{m_{\mathrm{Tx}}}{M_{\mathrm{Tx}}}\right) \tag{5.16}$$

Variable $m_{\mathrm{Tx}} = 0 \dots M_{\mathrm{Tx}} - 1$ denotes the actually active Tx channel. Note, that the normalised Doppler ν is now given by $\nu = 2\pi\nu' \cdot T_{\mathrm{P}} \cdot M_{\mathrm{Tx}}$, as the duration for a MIMO observation is $M_{\mathrm{Tx}} \cdot T_{\mathrm{P}}$. In order to account for the variation of the observed Doppler phase due to the TDM scheme, the matrix $\mathbf{B}_\nu(\boldsymbol{\nu})$, comprising the Doppler shift related harmonics, has to be modified. If the MIMO channel index m is a consecutive number from $1 \dots M_{\mathrm{Rx}} \dots M_{\mathrm{Tx}}M_{\mathrm{Rx}}$, the index m_{Tx} of the active Tx channel is given by

$$m_{\mathrm{Tx}} = (m - 1) \setminus M_{\mathrm{Tx}} \quad , \tag{5.17}$$

with operator $(\setminus)$ denoting the integer division. Accordingly, the matrix $\mathbf{B}_{\nu_m}(\boldsymbol{\nu}) \in \mathbb{C}^{N \times P}$ of Doppler harmonics for all paths and the m-th MIMO channel being active is

$$\mathbf{B}_{\nu_m}(\boldsymbol{\nu}) = \exp\left\{J\left(\mathbf{n} + \frac{(m-1) \setminus M_{\mathrm{Tx}}}{M_{\mathrm{Tx}}} \cdot \mathbf{1}_N\right)\boldsymbol{\nu}^{\mathrm{T}}\right\} = \mathbf{A}_m(\boldsymbol{\nu}) \odot \mathbf{B}_\nu(\boldsymbol{\nu}) \quad . \tag{5.18}$$

Vector $\mathbf{1}_N \in \mathbb{R}^N$ is the unity vector and

$$\mathbf{A}_m(\boldsymbol{\nu}) = \exp\left\{ J \frac{(m-1) \setminus M_{\text{Tx}}}{M_{\text{Tx}}} \cdot \mathbf{1}_N \cdot \boldsymbol{\nu}^{\text{T}} \right\} \quad . \tag{5.19}$$

The considered radar system is able to resolve the azimuth angle φ only, because the array at Tx and Rx is an ULA, and hence the virtual array too. The elevation angle is assumed to be $\vartheta = 0°$. Resolution of azimuth only causes the ambiguity cone [LKT12]. Hence, an unresolved or wrongly presumed elevation angle cause the estimation of a different azimuth angle than the actual one. Because the considered radar is equipped with a lens focusing the azimuth plane, the assumption $\vartheta = 0°$ is ensured by the radar hardware. The model for an observation of the m-th MIMO channel in the noiseless case has to be re-stated.

$$\mathbf{x}_m(\boldsymbol{\varpi}) = \left(\left[\mathbf{A}_m(\boldsymbol{\nu}) \odot \mathbf{B}_\nu(\boldsymbol{\nu}) \right] \diamond \left[\left(\text{diag}\{\mathbf{g}_m\} \cdot \mathbf{B}_{\varphi_m}(\boldsymbol{\varphi}) \right) \odot \mathbf{B}_\tau(\boldsymbol{\tau}) \right] \right) \cdot \boldsymbol{\gamma} = \mathbf{S}_m(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}) \cdot \boldsymbol{\gamma} \tag{5.20}$$

For sake of convenience the parameter vector $\boldsymbol{\varpi} \in \mathbb{R}^{5P}$ has been introduced, which comprises the parameters of all P propagation paths

$$\boldsymbol{\varpi} = \left[\boldsymbol{\varphi}^{\text{T}}, \boldsymbol{\tau}^{\text{T}}, \boldsymbol{\nu}^{\text{T}}, \Re\{\boldsymbol{\gamma}\}^{\text{T}}, \Im\{\boldsymbol{\gamma}\}^{\text{T}} \right]^{\text{T}} \quad . \tag{5.21}$$

Finally, the observations of all M MIMO channels are column-wisely stacked into a matrix. Afterwards, this matrix is vectorised, which gives the vector $\boldsymbol{\mu} \in \mathbb{C}^{M \cdot K \cdot N}$ of MIMO observations in the noiseless case.

$$\boldsymbol{\mu}(\boldsymbol{\varpi}) = \text{vec}\left\{ \left[\mathbf{x}_1(\boldsymbol{\varpi}), \dots, \mathbf{x}_M(\boldsymbol{\varpi}) \right] \right\} = \begin{bmatrix} \mathbf{S}_1(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}) \\ \vdots \\ \mathbf{S}_M(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}) \end{bmatrix} \cdot \boldsymbol{\gamma} = \mathbf{S}(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}) \cdot \boldsymbol{\gamma} \tag{5.22}$$

Unfortunately, a notational more convenient formula for $\boldsymbol{\mu}$ could not be found. Equation (5.22) will be denoted as *system model* of the co-located FMCW MIMO radar.

A note regarding the neglection of dual-polarimetric properties of the antennas, especially if arrays are employed and angles shall be estimated, is necessary. In case of a SISO system, neglection of polarimetry is not harmful to the parameter estimation performance, as the estimator compensates for the neglected polarimetry by tuning the path weights. Hence, only the estimate of the path weights are effected. However, in case of antenna arrays at Tx, Rx or both sides, neglection of polarimetry cause severe model errors, such that the estimation performance for all parameters is degraded [23, LKT12]. The reason is, that the polarimetric properties are different for each antenna of the array compound and they are direction dependent too. Hence, the model error by neglecting polarimetry cannot be compensated by tuning a path weight, which is fix for each antenna. Unfortunately, the resulting degradation of the estimation performance cannot be circumvented here, because of the considered radar hardware and available calibration data.

5.1.2 The Statistical Model – Model of Observations in the Noisy Case

In the noisy case the system model $\mathcal{M}^S$ and the noise model $\mathcal{M}^N$ (see Section 4.6) are used to describe the observations. Incorporation of the noise model in the system model results in the statistical model, which is actually the observation distribution or likelihood. It has been stated in Section 2.3, that the likelihood is required to solve the inverse problem.

In Section 4.6, thermal noise, phase noise and leakage noise has been introduced as noise sources of the radar device. Subsequently, consideration of each of this noise sources by a statistical model will be discussed, and the final statistical model will be derived last.

Thermal Noise

In Section 4.6.1 thermal noise has been introduced. Taking into account a thermal noise source in the observation $\mathbf{y}_m \in \mathbb{C}^{K \cdot N}$ of the m-th MIMO channel, the respective observation model is

$$\mathbf{y}_m = \mathbf{x}_m(\boldsymbol{\varpi}) + \mathbf{w}_m \quad . \tag{5.23}$$

The sampled thermal noise process $\mathbf{w}_m$ is a zero-mean, proper complex, normally distributed random process with PDF

$$\mathbf{w}_m \sim p_{\mathbf{w}_m} = \mathcal{CN}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{w}_m}\right) \quad . \tag{5.24}$$

$\boldsymbol{\Sigma}_{\mathbf{w}_m}$ is the respective covariance matrix. As the thermal noise is assumed as white, i.e. the process is uncorrelated in delay and time, the covariance matrix is a diagonal matrix.

$$\boldsymbol{\Sigma}_{\mathbf{w}_m} = \eta_m \cdot \mathbf{I}_{NK} \tag{5.25}$$

η_m the respective power of the thermal noise process.

Treating the noise parameters η_m and the propagation channel parameters $\boldsymbol{\varpi}$ as random quantities, the distribution of the observation $\mathbf{y}_m$ is [KS05]

$$p\left(\mathbf{y}_m \middle| \boldsymbol{\varpi}, \eta_m\right) = p_{\mathbf{w}}\left(\mathbf{y}_m - \mathbf{x}_m(\boldsymbol{\varpi}) \middle| \boldsymbol{\varpi}, \eta_m\right) \quad . \tag{5.26}$$

Assuming mutually independence of the generation process of the propagation parameters and noise parameters, the PDF reduces to [KS05]

$$p\left(\mathbf{y}_m \middle| \boldsymbol{\varpi}, \eta_m\right) = p_{\mathbf{w}}\left(\mathbf{y}_m - \mathbf{x}_m(\boldsymbol{\varpi})\right) \quad . \tag{5.27}$$

Accordingly, the PDF for the observation of the m-th MIMO channel is

$$\mathbf{y}_m \sim \mathcal{CN}\left(\mathbf{x}_m(\boldsymbol{\varpi}), \boldsymbol{\Sigma}_{\mathbf{w}_m}\right) \quad . \tag{5.28}$$

Extending the model of additive thermal noise to the observation of M MIMO channels, one get

$$\mathbf{y} = \text{vec}\left\{\left[\mathbf{x}_1, \ldots, \mathbf{x}_M\right]\right\} + \text{vec}\left\{\left[\mathbf{w}_1, \ldots, \mathbf{w}_M\right]\right\} = \boldsymbol{\mu} + \mathbf{w} \quad . \tag{5.29}$$

Assuming that the noise processes of the channels are uncorrelated to each other, therefore

$$\mathrm{E}\left\{\mathbf{w}_{m_1} \cdot \mathbf{w}_{m_2}^{\mathrm{H}}\right\} = \mathbf{0}_{NK \times NK} \qquad , \qquad \forall m_1 \neq m_2 \tag{5.30}$$

holds, the PDF of the MIMO thermal noise process $\mathbf{w}$ is

$$\mathbf{w} \sim p_{\mathrm{w}} = \mathcal{CN}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\mathrm{w}}(\boldsymbol{\eta})\right) \qquad . \tag{5.31}$$

Matrix $\boldsymbol{\Sigma}_{\mathrm{w}}(\boldsymbol{\eta}) = \mathrm{diag}\{\boldsymbol{\eta}\} \otimes \mathbf{I}_{NK}$ is the covariance matrix, and vector $\boldsymbol{\eta} = \left[\eta_1, \ldots, \eta_M\right]^{\mathrm{T}} \in \mathbb{R}_+^M$ comprises the powers of the thermal noise processes for all M MIMO channels.

Finally, the PDF of a MIMO radar observation (radar data cube) for the presence of thermal noise is

$$\mathrm{vec}\{\mathcal{Y}\} = \mathbf{y} \sim p(\mathbf{y} | \boldsymbol{\varpi}, \boldsymbol{\eta}) = \mathcal{CN}\left(\boldsymbol{\mu}(\boldsymbol{\varpi}), \boldsymbol{\Sigma}_{\mathrm{w}}(\boldsymbol{\eta})\right) \qquad . \tag{5.32}$$

Leakage Noise

In Section 4.6.2 leakage noise has been introduced. The corresponding model for an observation of the m-th MIMO channel is

$$\mathbf{y}_m = \mathbf{x}_m + \mathbf{l}_m \qquad . \tag{5.33}$$

The leakage noise $\mathbf{l}_m$ is assumed as a zero-mean, proper complex and coloured, normally distributed random process.

$$\mathbf{l}_m \sim \mathcal{CN}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\mathrm{l}_m}\right) \tag{5.34}$$

Covariance matrix $\boldsymbol{\Sigma}_{\mathrm{l}_m}$ is the sampled version of the correlation function (4.133).

$$\boldsymbol{\Sigma}_{\mathrm{l}_m} = \mathbf{I}_N \otimes \boldsymbol{\Sigma}_{\mathrm{l}_m}^{(\tau)} \tag{5.35}$$

Matrix $\boldsymbol{\Sigma}_{\mathrm{l}_m}^{(\tau)}$ is the covariance matrix of the leakage noise process in delay domain, given by sampling model (4.130). Because $\boldsymbol{\psi}_{\mathrm{l}}(-\tau) = \boldsymbol{\psi}_{\mathrm{l}}(\tau)^{\dagger}$ holds, the covariance matrix in delay domain is of hermitian Toeplitz structure.

$$\boldsymbol{\Sigma}_{\mathrm{l}_m}^{(\tau)} = \mathrm{T}\left\{\boldsymbol{\psi}_{\mathrm{l}}\left(\varsigma_m, \varepsilon_m, \varrho_m\right)\right\} \tag{5.36}$$

Operator T denotes the constructor of a hermitian Toeplitz matrix. Vector $\boldsymbol{\psi}_{\mathrm{l}} \in \mathbb{C}^K$ comprises the samples of model (4.130) of the ACF, taken at the delay points $\tau = 0, \ldots, (K-1)T_{\mathrm{S}}$.

$$\boldsymbol{\psi}_{\mathrm{l}}\left(\varsigma_m, \varepsilon_m, \varrho_m\right) = \varsigma_m \left[\frac{1}{\varepsilon_m}, \ldots, \frac{\exp\{j2\pi\varrho_m(K-1)\}}{\varepsilon_m - j2\pi(K-1)}\right]^{\mathrm{T}} \tag{5.37}$$

Note, that the parameters $\varsigma_m, \varepsilon_m, \varrho_m$ are normalised versions of the parameters of the ACF (4.130): $\varsigma_m = \frac{\varsigma}{T_{\mathrm{S}}}, \varepsilon_m = \frac{\varepsilon}{T_{\mathrm{S}}}, \varrho_m = \varrho \cdot T_{\mathrm{S}}$.

Next, leakage noise in the observation of all M MIMO channels is considered.

$$\mathbf{y} = \text{vec}\left\{\left[\mathbf{x}_1, \ldots, \mathbf{x}_M\right]\right\} + \text{vec}\left\{\left[\mathbf{l}_1, \ldots, \mathbf{l}_M\right]\right\} = \boldsymbol{\mu} + \mathbf{l} \tag{5.38}$$

The leakage noise processes are assumed as uncorrelated among the MIMO channels.

$$\text{E}\left\{\mathbf{l}_{m_1} \cdot \mathbf{l}_{m_2}^{\text{H}}\right\} = \mathbf{0} \quad , \quad \forall m_1 \neq m_2 \tag{5.39}$$

Accordingly, the overall covariance matrix $\boldsymbol{\Sigma}_{\mathbf{l}}$ of the leakage noise processes is a block diagonal matrix, whereas the m-th block entry corresponds to the leakage noise covariance matrix of the m-th MIMO channel.

$$\boldsymbol{\Sigma}_{\mathbf{l}}(\boldsymbol{\varsigma}, \boldsymbol{\varepsilon}, \boldsymbol{\varrho}) = \text{E}\left\{\mathbf{l} \cdot \mathbf{l}^{\text{H}}\right\} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{l}_1}\left(\varsigma_1, \varepsilon_1, \varrho_1\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma}_{\mathbf{l}_M}\left(\varsigma_M, \varepsilon_M, \varrho_M\right) \end{bmatrix} \tag{5.40}$$

The parameter vectors $\boldsymbol{\varsigma}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varrho}$ are introduced for convenience, which comprise the parameters of the leakage noise model for all M MIMO channels.

$$\boldsymbol{\varsigma} = \left[\varsigma_1, \ldots, \varsigma_M\right]^{\text{T}} \in \mathbb{R}_+^M \tag{5.41a}$$

$$\boldsymbol{\varepsilon} = \left[\varepsilon_1, \ldots, \varepsilon_M\right]^{\text{T}} \in \mathbb{R}_+^M \tag{5.41b}$$

$$\boldsymbol{\varrho} = \left[\varrho_1, \ldots, \varrho_M\right]^{\text{T}} \in \mathbb{R}^M \tag{5.41c}$$

The PDF of the MIMO leakage noise process $\mathbf{l}$ is

$$\mathbf{l} \sim p_{\mathbf{l}} = \mathcal{CN}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{l}}(\boldsymbol{\varsigma}, \boldsymbol{\varepsilon}, \boldsymbol{\varrho})\right) \quad . \tag{5.42}$$

Finally, the PDF of the MIMO radar observation (radar data cube) for the presence of leakage noise is

$$\text{vec}\left\{\boldsymbol{\mathcal{Y}}\right\} = \mathbf{y} \sim p\left(\mathbf{y}\mid \boldsymbol{\varpi}, \boldsymbol{\varsigma}, \boldsymbol{\varepsilon}\right) = \mathcal{CN}\left(\boldsymbol{\mu}\left(\boldsymbol{\varpi}\right), \boldsymbol{\Sigma}_{\mathbf{l}}(\boldsymbol{\varsigma}, \boldsymbol{\varepsilon}, \boldsymbol{\varrho})\right) \quad . \tag{5.43}$$

Phase Noise

In Section 4.6.3 a model for the occurrence of phase noise has been derived. Consider the noisy observation $\mathbf{y}_m$ of the m-th MIMO channel. The corresponding observation model is

$$\mathbf{y}_m = \mathbf{x}_m \odot \exp\left\{J\mathbf{e}_m\right\} \approx \mathbf{x}_m \odot \left[\mathbf{1} + J\mathbf{e}_m\right] = \mathbf{x}_m + \mathbf{x}_m \odot J\mathbf{e}_m = \mathbf{x}_m + \mathbf{p}_m \quad . \tag{5.44}$$

Mean vector and covariance matrix of the observation $\mathbf{y}_m$ have to be determined in order to describe the respective PDF. The random process $\mathbf{e}_m$ is assumed as zero mean circular normally distributed. Because the phase noise process $\mathbf{e}_m$ is assumed as zero-mean, the expectation (mean vector) of the observation is

$$\text{E}\left\{\mathbf{y}_m\right\} = \text{E}\left\{\mathbf{x}_m + \mathbf{x}_m \odot J\mathbf{e}_m\right\} = \mathbf{x}_m + J\mathbf{x}_m \odot \text{E}\left\{\mathbf{e}_m\right\} = \mathbf{x}_m \quad . \tag{5.45}$$

In order to derive the covariance matrix, the observation vector $\mathbf{y}_m$ of the m-th MIMO channel is split into its N snapshot vectors $\mathbf{y}_{m,n}$.

$$\mathbf{y}_m = \text{vec}\left\{\left[\mathbf{y}_{m,1}, \ldots, \mathbf{y}_{m,N}\right]\right\}$$

According to the MIMO phase noise model in equation (4.165), the covariance matrix $\boldsymbol{\Sigma}_{\mathrm{p}_{m,n}}^{(\tau)}$ for observation $\mathbf{y}_{m,n}$ (m-th MIMO channel and the n-th snapshot) is given by

$$\boldsymbol{\Sigma}_{\mathrm{p}_{m,n}}^{(\tau)} = \mathrm{E}\left\{\left[\mathbf{y}_{m,n} - \mathrm{E}\left\{\mathbf{y}_{m,n}\right\}\right]\cdot\left[\mathbf{y}_{m,n} - \mathrm{E}\left\{\mathbf{y}_{m,n}\right\}\right]^{\mathrm{H}}\right\} = \mathrm{E}\left\{\mathbf{p}_{m,n}\cdot\mathbf{p}_{m,n}^{\mathrm{H}}\right\}$$

$$= \mathrm{E}\left\{\left[\mathbf{x}_{m,n}\odot J\mathbf{e}_{m,n}\right]\cdot\left[\mathbf{x}_{m,n}\odot J\mathbf{e}_{m,n}\right]^{\mathrm{H}}\right\} \approx \boldsymbol{\Sigma}_{\mathrm{e}_m}^{(\tau)}\odot\sum_{p=1}^{P}\mathbf{x}_{m,n}\left(\boldsymbol{\varpi}_p\right)\cdot\mathbf{x}_{m,n}\left(\boldsymbol{\varpi}_p\right)^{\mathrm{H}} \quad . \quad (5.46)$$

Matrix $\boldsymbol{\Sigma}_{\mathrm{e}_m}^{(\tau)}$ is the covariance matrix of the phase noise process $\mathbf{e}_{m,n}$ in the delay domain. Note, that the property[1] of the Schur product has been utilised for derivation. Because the phase noise process is assumed as WSS in delay domain, the covariance matrix is of hermitian Toeplitz structure.

$$\boldsymbol{\Sigma}_{\mathrm{e}_{m,n}}^{(\tau)} = \mathrm{T}\left\{\boldsymbol{\psi}_{\mathrm{e}}\left(\alpha_m, \beta_m\right)\right\} \tag{5.47}$$

Vector $\boldsymbol{\psi}_{\mathrm{e}}\left(\alpha_m, \beta_m\right) \in \mathbb{C}^K$ comprises the samples of model (4.142) of the ACF, which are taken at the delay points $\tau = 0, \ldots, (K-1)T_{\mathrm{S}}$.

$$\boldsymbol{\psi}_{\mathrm{e}}\left(\alpha_m, \beta_m\right) = \alpha_m\cdot\left[1, \ldots, \exp\left\{-(K-1)\cdot\beta_m\right\}\right]^{\mathrm{T}} \tag{5.48}$$

Note, that the parameter β_m of the sampled ACF has been normalised to T_{S} compared to the parameter in equation (4.142): $\beta_m = \frac{T_{\mathrm{S}}}{\beta}$. The phase noise process is assumed as i.i.d. process. Hence, the covariance matrix $\boldsymbol{\Sigma}_{\mathrm{p}_m}\left(\boldsymbol{\varpi}, \alpha_m, \beta_m\right)$ in delay-time domain for the m-th MIMO channel is a block diagonal matrix.

$$\boldsymbol{\Sigma}_{\mathrm{p}_m}^{(\tau)}\left(\boldsymbol{\varpi}, \alpha_1, \beta_1\right) = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{p}_{m,1}}^{(\tau)}\left(\boldsymbol{\varpi}, \alpha_1, \beta_1\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma}_{\mathrm{p}_{m,N}}^{(\tau)}\left(\boldsymbol{\varpi}, \alpha_1, \beta_1\right) \end{bmatrix} = \left[\mathbf{I}_N\otimes\boldsymbol{\Sigma}_{\mathrm{e}_m}^{(\tau)}\right]\odot\sum_{p=1}^{P}\mathbf{x}_m\left(\boldsymbol{\varpi}_p\right)\cdot\mathbf{x}_m\left(\boldsymbol{\varpi}_p\right)^{\mathrm{H}}$$

$$\tag{5.49}$$

Due to the assumption of uncorrelated phase noise among the MIMO channels, the respective covariance matrix $\boldsymbol{\Sigma}_{\mathrm{p}}\left(\boldsymbol{\varpi}, \boldsymbol{\alpha}, \boldsymbol{\beta}\right)$ for all M MIMO channels is a block diagonal matrix.

$$\boldsymbol{\Sigma}_{\mathrm{p}}\left(\boldsymbol{\varpi}, \boldsymbol{\alpha}, \boldsymbol{\beta}\right) = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathrm{p}_1}\left(\boldsymbol{\varpi}, \alpha_1, \beta_1\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma}_{\mathrm{p}_M}\left(\boldsymbol{\varpi}, \alpha_M, \beta_M\right) \end{bmatrix} \tag{5.50}$$

---

[1] $(\mathbf{a}\odot\mathbf{b})\cdot(\mathbf{c}\odot\mathbf{d})^{\mathrm{H}} = \left(\mathbf{a}\cdot\mathbf{c}^{\mathrm{H}}\right)\odot\left(\mathbf{b}\cdot\mathbf{d}^{\mathrm{H}}\right)$

For sake of convenience the parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are introduced, comprising the phase noise parameters for all M channels.

$$\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_M\right]^{\mathrm{T}} \in \mathbb{R}_+^M \tag{5.51a}$$

$$\boldsymbol{\beta} = \left[\alpha_1, \ldots, \beta_M\right]^{\mathrm{T}} \in \mathbb{R}_+^M \tag{5.51b}$$

Finally, the PDF of the MIMO radar observation (radar data cube) for the presence of phase noise is

$$\mathrm{vec}\left\{\mathcal{Y}\right\} = \mathbf{y} \sim p\left(\mathbf{y}|\,\boldsymbol{\varpi}, \boldsymbol{\alpha}, \boldsymbol{\beta}\right) = \mathcal{CN}\left(\boldsymbol{\mu}\left(\boldsymbol{\varpi}\right), \boldsymbol{\Sigma}_{\mathrm{p}}\left(\boldsymbol{\varpi}, \boldsymbol{\alpha}, \boldsymbol{\beta}\right)\right) \quad . \tag{5.52}$$

The Statistical Model or Likelihood

Phase noise, thermal noise and leakage noise are modelled as additive processes. The corresponding model for the observations of the m-th MIMO channel is

$$\mathbf{y}_m = \mathbf{x}_m + \mathbf{w}_m + \mathbf{l}_m + \mathbf{p}_m \quad . \tag{5.53}$$

The individual noise processes are assumed as *uncorrelated* to each other. The noise processes are modelled as circular normally distributed, such that the measurements are circular normally distributed too. The mean vector and covariance matrix have to be defined in order to describe the PDF of the observations. The mean vector of the observation of the m-th MIMO channel is

$$\mathrm{E}\left\{\mathbf{y}_m\right\} = \mathrm{E}\left\{\mathbf{x}_m\right\} + \mathrm{E}\left\{\mathbf{x}_m{\odot}\!J\mathbf{p}_m\right\} + \mathrm{E}\left\{\mathbf{l}_m\right\} + \mathrm{E}\left\{\mathbf{w}_m\right\} = \mathbf{x}_m + \mathbf{x}_m{\odot}\!J\mathrm{E}\left\{\mathbf{p}_m\right\} + \mathbf{0} = \mathbf{x}_m \quad . \tag{5.54}$$

Taking into account, that the noise sources are uncorrelated, the covariance matrix is given by the sum of the covariance matrices of the noise sources.

$$\boldsymbol{\Sigma}_m = \mathrm{E}\left\{\left[\mathbf{y}_m - \mathrm{E}\left\{\mathbf{y}_m\right\}\right]\cdot\left[\mathbf{y}_m - \mathrm{E}\left\{\mathbf{y}_m\right\}\right]^{\mathrm{H}}\right\} = \mathrm{E}\left\{\left[\mathbf{p}_m + \mathbf{l}_m + \mathbf{w}_m\right]\cdot\left[\mathbf{p}_m + \mathbf{l}_m + \mathbf{w}_m\right]^{\mathrm{H}}\right\}$$
$$= \boldsymbol{\Sigma}_{\mathrm{l}_m} + \boldsymbol{\Sigma}_{\mathrm{w}_m} + \boldsymbol{\Sigma}_{\mathrm{p}_m} \tag{5.55}$$

A model for the ACF of all noise processes has been derived in Section 4.6.4. The covariance matrix is given by sampling this model (4.173).

$$\boldsymbol{\Sigma}_m\left(\boldsymbol{\sigma}_m\right) = \mathbf{I}_N{\otimes}\left[\boldsymbol{\Sigma}_{\mathrm{l}_m}^{(\tau)}\left(\varsigma_m, \varepsilon_m, \varrho_m\right) + \eta_m\mathbf{I}_K\right] + \left[\mathbf{I}_N{\otimes}\boldsymbol{\Sigma}_{\mathrm{e}_m}^{(\tau)}\left(\alpha_m, \beta_m\right)\right] \odot \sum_{p=1}^{P} \mathbf{x}_m\left(\boldsymbol{\varpi}_p\right)\mathbf{x}_m\left(\boldsymbol{\varpi}_p\right)^{\mathrm{H}} \tag{5.56}$$

Vector $\boldsymbol{\sigma}_m$ comprises the parameters of the covariance model for the m-th MIMO channel.

$$\boldsymbol{\sigma}_m = \left[\varsigma_m, \varepsilon_m, \varrho_m, \alpha_m, \beta_m, \eta_m\right]^{\mathrm{T}} \in \mathbb{R}^6 \tag{5.57}$$

Summarised, the PDF of an observation of the m-th MIMO channel is

$$\mathbf{y}_m \sim \mathcal{CN}\left(\boldsymbol{\mu}_m\left(\boldsymbol{\varpi}\right), \boldsymbol{\Sigma}_m\left(\boldsymbol{\varpi}, \boldsymbol{\sigma}_m\right)\right) \quad . \tag{5.58}$$

Because thermal noise, leakage noise and phase noise are assumed as uncorrelated among the MIMO channels, the covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\varpi}, \boldsymbol{\sigma})$ of the MIMO observation is block diagonal.

$$\boldsymbol{\Sigma}(\boldsymbol{\varpi}, \boldsymbol{\sigma}) = \mathrm{E}\left\{[\mathbf{y} - \mathrm{E}\{\mathbf{y}\}] \cdot [\mathbf{y} - \mathrm{E}\{\mathbf{y}\}]^{\mathrm{H}}\right\} = \begin{bmatrix} \boldsymbol{\Sigma}_1\left(\boldsymbol{\varpi}, \boldsymbol{\sigma}_1\right) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma}_M\left(\boldsymbol{\varpi}, \boldsymbol{\sigma}_M\right) \end{bmatrix} \quad (5.59)$$

Vector $\boldsymbol{\sigma}$ comprises the parameters of the whole covariance model.

$$\boldsymbol{\sigma} = \left[\boldsymbol{\varsigma}^{\mathrm{T}}, \boldsymbol{\varepsilon}^{\mathrm{T}}, \boldsymbol{\varrho}^{\mathrm{T}}, \boldsymbol{\alpha}^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\eta}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{6M} \quad (5.60)$$

Finally, the *statistical model or likelihood* of the MIMO radar observation (radar data cube) for the presence of thermal, leakage and phase noise is

$$\mathbf{y} \sim p(\mathbf{y} | \boldsymbol{\varpi}, \boldsymbol{\sigma}) = \mathcal{CN}(\boldsymbol{\mu}(\boldsymbol{\varpi}), \boldsymbol{\Sigma}(\boldsymbol{\varpi}, \boldsymbol{\sigma})) \quad . \quad (5.61)$$

For sake of convenience the parameters of the system model $\boldsymbol{\varpi}$ and the parameters of the noise model $\boldsymbol{\sigma}$ are concatenated in the vector of model parameters $\boldsymbol{\theta} = \left[\boldsymbol{\varpi}^{\mathrm{T}}, \boldsymbol{\sigma}^{\mathrm{T}}\right]^{\mathrm{T}} \in \Theta = \Pi \times \Sigma \subset \mathbb{R}^{5P+6M}$, which has to be estimated from the observations $\mathbf{y}$.

5.2 Derivation of the Parameter Estimator

5.2.1 Maximum-Likelihood Estimation

In Chapter 2 Bayesian inversion has been pointed out as a method to estimate parameter values from noisy observations according to a stated statistical model. The statistical model (or the likelihood, respectively) of the observations $p(\mathbf{y} | \boldsymbol{\theta})$ has been described in the previous section 5.1. Knowing also the prior distribution $p(\boldsymbol{\theta})$, which is so far subject to clarification, the point estimators maximum a posteriori (MAP) and minimum mean-square error (MMSE) can be derived (see Section 2.3). Here, the MAP estimator will be considered, because its calculation requires just the optimisation of an objective function and numerical exhaustive MCMC sampling techniques are not required as the MMSE estimator requires [GL06, 28]. The MAP estimator examines the mode of the posterior distribution as the estimation criterion.

$$\hat{\boldsymbol{\theta}}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\theta}} \frac{p(\mathbf{y} | \boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})}{p(\mathbf{y})} = \arg\max_{\boldsymbol{\theta}} p(\mathbf{y} | \boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}) \quad (5.62)$$

Hence, the probability of the parameters to be included in the observation is maximised. The denominator does not depend on the parameters and hence becomes negligible.

Assignment of the Prior

So far, a prior distribution of the parameters has not been assigned. As stated in Section 2.3 the prior must be properly selected, because the prior influences the posterior and hence the estimation result. Because no certain parameter realisation can be preferred over the other

(*principle of indifference*), a constant or uniform prior is considered for the parameters.

$$p(\boldsymbol{\theta}) = const \tag{5.63}$$

Furthermore, the prior can represent any information regarding the parameters, which is available before the observation. Focusing on the set space for the path and noise parameters, Π and Σ, it can be concluded that these sets are bounded due to the parameter's nature. Accounting for this bounds, the prior distribution becomes

$$p(\boldsymbol{\theta}) = const \cdot cu_{\Theta}(\boldsymbol{\theta}) \quad, \tag{5.64}$$

with $cu_{\Theta}(\bullet)$ the characteristic function of the cube, who is limited according to the parameter bounds. Due to the principle of indifference and the parameter bounds a uniform distribution can be assigned as prior distribution.

$$p(\boldsymbol{\theta}) = \mathcal{U}(\boldsymbol{\theta}) \tag{5.65}$$

Maximum-Likelihood Estimator with Bound Constraints

Considering a uniform prior distribution the MAP estimator reduces to

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \underset{\boldsymbol{\theta} \in \mathbb{R}^{5P+6M}}{\arg \max} \; p(\mathbf{y}|\boldsymbol{\theta}) \cdot \mathcal{U}(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta} \in \Theta}{\arg \max} \; p(\mathbf{y}|\boldsymbol{\theta}) \quad. \tag{5.66}$$

This is actually an ML estimator, i.e., the maximisation of the likelihood function only. The maximisation is conducted over the (bounded) space Θ, and not the entire space $\mathbb{R}^{5P+6M}$. Hence, the considered parameter estimator is a *maximum-likelihood estimator with bound constraints*.

According to the statistical model (5.61) the observation is multidimensional, proper complex normally distributed.

$$p(\mathbf{y}|\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})) = \frac{1}{|\pi\boldsymbol{\Sigma}(\boldsymbol{\theta})|} \cdot \exp\left\{ -\left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\theta})\right]^{\mathrm{H}} \cdot \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\theta})\right] \right\} \quad. \tag{5.67}$$

Variables $\boldsymbol{\mu}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ denote the parametric model of the mean vector and covariance matrix, respectively. Plugging into the ML estimator yields a maximisation problem for parameter estimation.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \Theta}{\arg \max} \; \frac{1}{|\pi\boldsymbol{\Sigma}(\boldsymbol{\theta})|} \cdot \exp\left\{ -\left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\theta})\right]^{\mathrm{H}} \cdot \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\theta})\right] \right\} \tag{5.68}$$

Commonly, the negative natural logarithm of the objective function is taken. Ignoring constant terms, because the do not influence the maximisation problem, the *negative log-likelihood function or objective function* $\mathcal{L}(\mathbf{y}|\boldsymbol{\theta})$ can be stated.

$$\mathcal{L}(\mathbf{y}|\boldsymbol{\theta}) = \ln\left\{|\boldsymbol{\Sigma}(\boldsymbol{\theta})|\right\} + \left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\theta})\right]^{\mathrm{H}} \cdot \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\theta})\right] \tag{5.69}$$

Minimising the objective function over the parameter space Θ yields the ML estimate of the

parameter vector $\hat{\boldsymbol{\theta}}$.

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \Theta} \ \{\mathcal{L}(\mathbf{y}|\boldsymbol{\theta})\}$$

$$\Downarrow$$

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} = \left[\boldsymbol{\varpi}^{\mathrm{T}}, \boldsymbol{\sigma}^{\mathrm{T}}\right]^{\mathrm{T}} \in \Theta} \ \left\{ \ln\left\{|\boldsymbol{\Sigma}(\boldsymbol{\varpi}, \boldsymbol{\sigma})|\right\} + \left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\varpi})\right]^{\mathrm{H}} \cdot \boldsymbol{\Sigma}(\boldsymbol{\varpi}, \boldsymbol{\sigma})^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\varpi})\right] \right\} \qquad (5.70)$$

Because the parameter space Θ is not an open space, i.e. bounds on the parameters exist, the optimisation problem (5.70) is a *box-constrained problem*.

A necessary but not sufficient condition for the ML estimate $\hat{\boldsymbol{\theta}}$ is, that for the first-order derivative (gradient vector) of the objective

$$\frac{\partial}{\partial \boldsymbol{\theta}} \mathcal{L}(\mathbf{y}|\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \mathbf{0} \qquad (5.71)$$

holds. Hence, the objective function must exhibit an optimum at the point $\hat{\boldsymbol{\theta}}$ of the ML estimate. This criterion ensures an optimum only, whereas the ML estimate requires a minimum. In order to ensure the optimum to be a minimum, the second-order derivative (Hessian matrix) must be positive definite.

$$\frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\mathrm{T}}} \mathcal{L}(\mathbf{y}|\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \succ \mathbf{0} \qquad (5.72)$$

In order to have an identifiable problem, i.e. the inverse problem can be solved assuming the Hadamard conditions as fulfilled, the number of parameters $5P + 6M$ must be less than the number of real-valued measurement samples $2MNK$. As only the number of paths P can vary for given observations, the condition

$$P \leq \frac{2MNK - 6M}{5} \qquad (5.73)$$

must be fulfilled in order to have an identifiable problem.

5.2.2 Some Asymptotic Properties of Maximum-Likelihood Estimators

The maximum-likelihood estimator exhibits some asymptotic properties, why it becomes one of the most popular parameter estimators [Van04].

The *estimation error* $\boldsymbol{\theta}_{\epsilon}$ for any parameter estimator can be expressed as [TB07]

$$\boldsymbol{\theta}_{\epsilon} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \quad . \qquad (5.74)$$

The *bias* $\mathbf{b}(\hat{\boldsymbol{\theta}})$ is quite often considered as performance metric of an estimator. Having

defined the estimation error, the bias is given as its respective expectation.

$$\mathbf{b}\left(\hat{\boldsymbol{\theta}}\right) = \mathrm{E}\left\{\boldsymbol{\theta}_{\epsilon}\right\} = \mathrm{E}\left\{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right\} \tag{5.75}$$

Under some regularity conditions, which are often encountered in practise, the ML estimator is *unbiased*. However, unbiasedness of the estimator also depends on the data model and whether the inverse problem fulfils the Hadamard condition of uniqueness. For an unbiased estimator, the expectation of the estimated parameters $\hat{\boldsymbol{\theta}}$ is equal to the true parameters.

$$\mathrm{E}\left\{\hat{\boldsymbol{\theta}}\right\} = \boldsymbol{\theta} \quad \Longleftrightarrow \quad \mathbf{b}\left(\hat{\boldsymbol{\theta}}\right) = \mathbf{0} \tag{5.76}$$

The *covariance matrix* $\mathbf{C}\left(\hat{\boldsymbol{\theta}}\right)$ of the parameter estimates is a further performance metric of any parameter estimator.

$$\mathbf{C}\left(\hat{\boldsymbol{\theta}}\right) = \mathrm{E}\left\{\left(\hat{\boldsymbol{\theta}} - \mathrm{E}\left\{\hat{\boldsymbol{\theta}}\right\}\right)^{\mathrm{T}} \cdot \left(\hat{\boldsymbol{\theta}} - \mathrm{E}\left\{\hat{\boldsymbol{\theta}}\right\}\right)\right\} \quad . \tag{5.77}$$

Assuming an unbiased estimator, the respective covariance matrix of the parameter estimates becomes

$$\mathbf{C}\left(\hat{\boldsymbol{\theta}}\right) = \mathrm{E}\left\{\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^{\mathrm{T}}\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)\right\} \quad . \tag{5.78}$$

Having defined the bias and the covariance matrix, the mean-squared error matrix $\mathbf{E}\left(\hat{\boldsymbol{\theta}}\right)$ of the parameter estimation error can given [TB07].

$$\mathbf{E}\left(\hat{\boldsymbol{\theta}}\right) = \mathrm{E}\left\{\boldsymbol{\theta}_{\epsilon}\boldsymbol{\theta}_{\epsilon}^{\mathrm{T}}\right\} = \mathbf{C}\left(\hat{\boldsymbol{\theta}}\right) + \mathbf{b}\left(\hat{\boldsymbol{\theta}}\right)\mathbf{b}\left(\hat{\boldsymbol{\theta}}\right)^{\mathrm{T}} \tag{5.79}$$

Note, that the covariance matrix $\mathbf{C}\left(\hat{\boldsymbol{\theta}}\right)$ and the mean-squared error matrix $\mathbf{E}\left(\hat{\boldsymbol{\theta}}\right)$ are equal for an unbiased estimator.

It is known from theory [TB07], that the variance of the ML estimates (diagonal entries of $\mathbf{C}$) asymptotically attain the Cramér-Rao lower bound $\mathbf{c}\left(\boldsymbol{\theta}\right)$.

$$\lim_{MNK\to\infty} \mathrm{diag}\left\{\mathrm{E}\left\{\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)^{\mathrm{T}} \cdot \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right)\right\}\right\} = \mathbf{c}\left(\boldsymbol{\theta}\right) \tag{5.80}$$

Thus, an ML estimator is an asymptotically *efficient estimator*. The Cramér-Rao lower bound is given by the diagonal entries of the Cramér-Rao matrix, which is given by the inverse of the FIM $\mathbf{F}$. Calculation of the FIM will be elaborated in the next section.

Last, ML estimates are asymptotically normal distributed [Van04].

$$\hat{\boldsymbol{\theta}} \sim \mathcal{N}\left(\boldsymbol{\theta}, \mathbf{C}\right) \tag{5.81}$$

5.2.3 The Bayesian Cramér-Rao Bound

Considering Bayesian estimators as e.g. the MAP or MMSE (see Section 2.3), a likelihood $p\left(\mathbf{y}\,|\,\boldsymbol{\theta}\right)$ and a prior $p\left(\boldsymbol{\theta}\right)$ have been stated. The lowest achievable variance of Bayesian parameter estimates is given by the Bayesian Cramér-Rao bound (BCRB), if the parameter estimator is unbiased. In this case, the expectation of the mean-squared error matrix $\mathbf{E}\left(\boldsymbol{\theta}\right)$ (expectation

w.r.t. the prior) is lower bounded by the inverse of the Bayesian information matrix (BIM) **B** [TB07].

$$\mathrm{E}\left\{\mathbf{E}\left(\boldsymbol{\theta}\right)\right\} \succeq \mathbf{B}\left(\boldsymbol{\theta}\right)^{-1} \tag{5.82}$$

Under some regularity conditions, the entry for the i-th column and j-th row of the BIM is given by [TB07]

$$[\mathbf{B}]_{(i,j)} = -\mathrm{E}\left\{ \frac{\partial^2 \ln\left\{p\left(\mathbf{y}\mid\boldsymbol{\theta}\right)\right\}}{\partial\theta_i\partial\theta_j} \right\} - \mathrm{E}\left\{ \frac{\partial^2 \ln\left\{p\left(\boldsymbol{\theta}\right)\right\}}{\partial\theta_i\partial\theta_j} \right\} \quad . \tag{5.83}$$

The first summand in equation (5.83) represents the contribution of the observations (i.e. the likelihood), whereas the second summand represents contributions of the prior to the BIM.

The BCRB requires some regularity conditions as e.g. the continuity conditions of a twice differentiable prior [TB07]. This condition is not satisfied, if a uniform prior is assigned to $\boldsymbol{\theta}$. A uniform prior is a cube (a rectangular function in the one-dimensional space), which is not continuous and hence not differentiable. However, asymptotically, i.e. for large sample sizes $MNK \to \infty$ or equivalently high SNR, the contribution of the observation dominates the contribution of the prior to the BIM. Hence, the inverse of the BIM **B** reduces to the inverse of the expectation of the FIM **F** [TB07].

$$\lim_{MNK\to\infty} \mathbf{B}^{-1} = \mathrm{E}\left\{\mathbf{F}\right\}^{-1} \tag{5.84}$$

This presupposes, that the MAP estimator asymptotically converges to the ML estimator [TB07].

$$\lim_{MNK\to\infty} \hat{\boldsymbol{\theta}}_{\mathrm{MAP}} = \hat{\boldsymbol{\theta}}_{\mathrm{ML}} \tag{5.85}$$

The entries of the FIM **F** are given by partial derivatives of the likelihood $p\left(\mathbf{y}\mid\boldsymbol{\theta}\right)$.

$$[\mathbf{F}]_{(i,j)} = -\mathrm{E}\left\{ \frac{\partial^2 \ln\left\{p\left(\mathbf{y}\mid\boldsymbol{\theta}\right)\right\}}{\partial\theta_i\partial\theta_j} \right\} = \mathrm{E}\left\{ \frac{\partial \ln\left\{p\left(\mathbf{y}\mid\boldsymbol{\theta}\right)\right\}}{\partial\theta_i} \cdot \frac{\partial \ln\left\{p\left(\mathbf{y}\mid\boldsymbol{\theta}\right)\right\}}{\partial\theta_j} \right\} \tag{5.86}$$

In order to calculate the FIM lets consider a multivariate, complex normal distributed random variable with parametric mean $\boldsymbol{\mu}\left(\boldsymbol{\theta}\right)$ and structural covariance $\boldsymbol{\Sigma}\left(\boldsymbol{\theta}\right)$. The FIM entries are given by the summation of a stochastic part (derivative of the covariance matrix w.r.t. the parameters) and a deterministic part (derivative of the mean vector w.r.t. the parameters).

$$[\mathbf{F}]_{(i,j)} = \mathrm{trace}\left\{ \frac{\partial\boldsymbol{\Sigma}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1} \cdot \frac{\partial\boldsymbol{\Sigma}}{\partial\theta_j}\boldsymbol{\Sigma}^{-1} \right\} + 2\cdot\Re\left\{ \frac{\partial\boldsymbol{\mu}^{\mathrm{H}}}{\partial\theta_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\mu}}{\partial\theta_j} \right\} \tag{5.87}$$

For convenience lets consider the Cholesky factorisation [PP12] of the covariance matrix.

$$\boldsymbol{\Sigma} = \mathbf{L}\cdot\mathbf{L}^{\mathrm{H}} \quad\Longleftrightarrow\quad \boldsymbol{\Sigma}^{-1} = \mathbf{L}^{-\mathrm{H}}\cdot\mathbf{L}^{-1} \tag{5.88}$$

Furthermore lets introduce the variables $\mathbf{J}_i$ and $\mathbf{j}_i$, which comprise the partial derivatives of the covariance matrix and the mean vector w.r.t. the i-th parameter in $\boldsymbol{\theta}$, multiplied with the

Cholesky factor of the covariance matrix.

$$\mathbf{J}_i = \mathbf{L}^{-1} \frac{\partial \mathbf{\Sigma}}{\partial \theta_i} \mathbf{L}^{-H} \tag{5.89a}$$

$$\mathbf{j}_i = \mathbf{L}^{-1} \frac{\partial \mathbf{\mu}}{\partial \theta_i} \tag{5.89b}$$

The partial derivatives of the mean vector and the structural covariance matrix w.r.t. the parameters are given in Section 5.3. Plugging into equation (5.87) the i, j-th entry of the FIM can be given, using mathematical identity [2] and [3].

$$[\mathbf{F}]_{(i,j)} = \text{trace}\left\{\mathbf{J}_i \cdot \mathbf{J}_j\right\} + 2 \cdot \Re\left\{\mathbf{j}_i^H \cdot \mathbf{j}_j\right\} = \text{vec}\left\{\mathbf{J}_i^T\right\}^T \cdot \text{vec}\left\{\mathbf{J}_j\right\} + 2 \cdot \Re\left\{\mathbf{j}_i^H \cdot \mathbf{j}_j\right\} \tag{5.90}$$

Now, the complete FIM can be calculated.

$$\mathbf{F} = \mathbf{J}''^T \cdot \mathbf{J}''' + 2 \cdot \Re\left\{\mathbf{J}'^H \cdot \mathbf{J}'\right\} \tag{5.91}$$

The matrices $\mathbf{J}'$, $\mathbf{J}''$ and $\mathbf{J}'''$ comprise the partial derivatives all variables in $\mathbf{\theta}$.

$$\mathbf{J}' = \left[\mathbf{j}_1, \ldots, \mathbf{j}_J\right] \tag{5.92a}$$

$$\mathbf{J}'' = \left[\text{vec}\left\{\mathbf{J}_1^T\right\}, \ldots, \text{vec}\left\{\mathbf{J}_J^T\right\}\right] \tag{5.92b}$$

$$\mathbf{J}''' = \left[\text{vec}\left\{\mathbf{J}_1\right\}, \ldots, \text{vec}\left\{\mathbf{J}_J\right\}\right] \tag{5.92c}$$

The true parameter values are unknown in practical situations, such that the FIM cannot be calculated. However, the covariance of the estimated parameter values can be necessary in practical situations. Hence, a consistent estimate of the FIM is required in practical situations. A consistent estimate of the FIM can be get by evaluating the FIM at the actual ML parameter estimates [Van04].

$$\hat{\mathbf{F}} = \mathbf{F}\left(\hat{\mathbf{\theta}}\right) \tag{5.93}$$

$\hat{\mathbf{F}}$ will be denoted as *observed Fisher information matrix*. Consequently, the Hessian at the ML estimates is the observed FIM, which is asymptotically the FIM.

Matrix Block Partitioning

The covariance matrix of the parameter estimates is required for e.g. model order estimation, see Chapter 6. Hence, the FIM has to be inverted. If the covariance of a subset of the parameters in $\mathbf{\theta}$ is required only, e.g. the covariance matrix for the parameter vector $\mathbf{\theta}$ is of interest only, inversion of the complete FIM can be circumvented by matrix block partitioning.

Partition the parameter vector $\mathbf{\theta}$ in the vector of system parameters $\mathbf{\varpi}$ and the vector of noise parameters $\mathbf{\sigma}$. Then, the covariance matrix of the parameter estimates can be represented as a

---

[2]$\text{trace}\left\{\mathbf{A} \cdot \mathbf{B}\right\} = \text{vec}\left\{\mathbf{A}^T\right\}^T \cdot \text{vec}\left\{\mathbf{B}\right\}$
[3]$\text{trace}\left\{\mathbf{A} \cdot \mathbf{B} \cdot \mathbf{C}\right\} = \text{trace}\left\{\mathbf{B} \cdot \mathbf{C} \cdot \mathbf{A}\right\}$

block partitioned matrix [Van04].

$$\mathbf{F}\left(\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\varpi}^{\mathrm{T}} & \boldsymbol{\sigma}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}\right)^{-1} = \mathbf{C}\left(\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\varpi}^{\mathrm{T}} & \boldsymbol{\sigma}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}\right) = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\varpi}} & \mathbf{C}_{\boldsymbol{\varpi},\boldsymbol{\sigma}} \\ \mathbf{C}_{\boldsymbol{\varpi},\boldsymbol{\sigma}}^{\mathrm{T}} & \mathbf{C}_{\boldsymbol{\sigma}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\boldsymbol{\varpi}} & \mathbf{F}_{\boldsymbol{\varpi},\boldsymbol{\sigma}} \\ \mathbf{F}_{\boldsymbol{\varpi},\boldsymbol{\sigma}}^{\mathrm{T}} & \mathbf{F}_{\boldsymbol{\sigma}} \end{bmatrix}^{-1} \quad (5.94)$$

Calculating the inverse of the FIM can be costly, because the size of the FIM scales with the number of paths: $5P + 6M \times 5P + 6M$. Exploiting the theorem of inverting a block partitioned matrix [MS00], only the inversion of smaller matrices is necessary. Knowing the block entries of the FIM, the covariance $\mathbf{C}_{\boldsymbol{\varpi}}$ for the system parameters $\boldsymbol{\varpi}$ can be stated.

$$\mathbf{C}_{\boldsymbol{\varpi}} = \left(\mathbf{F}_{\boldsymbol{\varpi}} - \mathbf{F}_{\boldsymbol{\varpi},\boldsymbol{\sigma}} \cdot \mathbf{F}_{\boldsymbol{\sigma}}^{-1} \cdot \mathbf{F}_{\boldsymbol{\varpi},\boldsymbol{\sigma}}^{\mathrm{T}}\right)^{-1} \quad (5.95)$$

Hence, inversion of the complete FIM is not necessary.

5.3 Global Optimisation of the Bounded Non-Convex Objective

According to the ML estimator derived in Section 5.2 the parameters are estimated from the observations by finding the *global minimum* of objective function (5.69) (the negative log-likelihood function) subject to bounds. A global minimum $\hat{\boldsymbol{\theta}} \in \Theta$ is defined as [Wei09]

$$\mathcal{L}\left(\mathbf{y} | \hat{\boldsymbol{\theta}}\right) < \mathcal{L}\left(\mathbf{y} | \boldsymbol{\theta}\right) \quad , \forall \boldsymbol{\theta} \in \Theta \quad . \quad (5.96)$$

Hence, a global minimum is a *unique* point of the multidimensional objective function in the parameter space Θ. An *optimisation method* is used to detect this global minimum. An optimisation method shall converge to the global minimum, while maintaining numerical efficiency.

The demand for a global minimum somehow coincides with the Hadamard criteria for a well-posed inverse problem. Hadamard criteria are, that a solution should exist and be unique. Hence, this criteria demands an unique global minimum. However, if the posterior has multiple modes, multiple global optima exist and the inverse problem is ill-posed in the sense of the Hadamard criteria.

Parameter Bounds

Uniqueness of a global minimum is determined by the the the statistical model, which determines the "shape" of the objective function (the posterior), and also by the parameter bounds. For example, consider the estimation of AoA with an ULA. ULAs suffer, apart from the ambiguity cone for unknown elevation angles [LKT12], from a forward-backward ambiguity in their array response, i.e. the array cannot distinguish between the incidence of a wave from the front or the rear. If the set of possible AoA angles is bounded, e.g. angles from the front are taken into account only, a unique global minimum exists. The parameter bounds have to be chosen according to physical reasoning. For example, capturing of impinging waves from the rear can be excluded, if directive antennas are employed in the ULA. Summarised, a proper selection of the parameter space (i.e. the parameter bounds) can be advantageous the ensure the

Table 5.1: *Bounds for the parameters of the mean and covariance model*

Mean Parameters	Bounds	Covariance Parameters	Bounds
φ	$-\frac{2\pi}{9} \dots \frac{2\pi}{9}$	ς	$0 \dots \infty$
τ	$0 \dots 2\pi$	ε	$0 \dots \infty$
ν	$-\pi \dots \pi$	ϱ	$0 \dots 1$
$\Re\{\gamma\}$	$-\infty \dots \infty$	α	$0 \dots \infty$
$\Im\{\gamma\}$	$-\infty \dots \infty$	β	$0 \dots \infty$
		η	$0 \dots \infty$

existence of a unique global minimum. The parameter bounds must be carefully chosen under hardware and environment considerations.

The minimisation problem (5.70) takes place in the parameter space Θ. Necessary or at least sufficient bounds exist for all of the considered parameters. Bounds on the mean model parameters are given due to the limited observation space; and bounds on the covariance parameters are given due to model restrictions. Table 5.1 summarises the parameter bounds. Note, that the bounds for the azimuth angle φ have been chosen in accordance to the 3 dB opening angle of the array antennas of the considered radar.

Non-Convex Objective

Presume the existence and uniqueness of a global minimum. Then, the optimisation problem can still be problematic, if local minima exist. A local minimum is a point of the objective function, which is minimal compared to all other feasible points in its vicinity [NW06]. If an objective function has a unique minimum [Sor93], the function is stated as *convex*. Optimisation problems featuring convex objectives can be efficiently solved [BV04]. An objective function exhibiting multiple minima, which can be local or global, is stated as *non-convex*, see APPENDIX B. It cannot be easily decided, whether a found minimum is of global or local nature. Hence, optimisation methods can fail to found the global minimum, even if there is one. In order to circumvent this problem, multiple optimisation trials can be run and the values of the objective function at the detected minima are compared. However, it cannot be ensured that one of the detected minima is the global minimum. Furthermore, the computational complexity is high. Summarised, detection of the global minimum of a non-convex objective cannot be assured in general and under practical considerations. Therefore, optimisation problems including non-convex objective functions are quite difficult to solve and no general solvers exist [BV04]. Unfortunately, the objective function for the considered parameter estimation problem is non-convex.

Many computational optimisation methods for non-convex objectives are known from literature, see e.g. [Koz11, NW06, Wei11, Yan17]. Optimisation methods can be roughly classified into global and local optimisation, i.e. the methods converge surely to a local or global minimum, and heuristic and deterministic methods, i.e. the methods exhibit some heuristics or randomness in finding the optima or not. Some optimisation frameworks featuring global convergence are e.g. simulated annealing, particle swarm, genetic algorithm or cuckoo search [Wei09]. Advantages of these methods are, in addition to global convergence, that they

require the evaluation of the objective function only and they are numerically stable, e.g. do not require the calculation of derivatives. On the other hand, these methods feature a low convergence speed. Gradient-based methods as gradient descent, Gauß–Newton or Levenberg-Marquardt [NW06] feature a fast convergence speed, but only convergence to a local optimum is ensured. Therefore, an initial guess close to the global minimum is necessary, which acts as the starting point for the local optimisation method. Furthermore, gradient methods can become numerically unstable as the calculation of derivatives is required. All of the mentioned methods are iterative or feature multiple estimation trials, such that they are computationally costly. Because online estimation fulfilling real-time requirements will not be considered here, there is no demand for high computational efficiency. A simplified version of the RIMAX framework [RT05, Ric05] will be employed to minimise the objective function.

Reduction of the Dimensionality

The objective function (5.69) has a high-dimensional parameter space.

$$\mathcal{L}\left(\mathbf{y}\,|\,\boldsymbol{\theta} = \left[\boldsymbol{\varpi}^{\mathrm{T}}, \boldsymbol{\sigma}^{\mathrm{T}}\right]^{\mathrm{T}}\right) : \mathbb{R}^{5P} \times \mathbb{R}^{6M} \mapsto \mathbb{R} \tag{5.97}$$

Furthermore, the size of the parameter space scales with the number of propagation paths *P*. Hence, exhaustive search for a global minimum is highly computationally costly and hence not conceivable. Hence, a compromise must be drawn between reduced estimation accuracy and computational effort. Simplifications or dimensionality reductions can be employed to reduce the numerical effort to the risk of a lower estimation accuracy. The dimensionality of the optimisation problem can be reduced by optimising the objective w.r.t. a subset of the parameters only, whereas all other parameters are kept fix. Hence, the parameter search space and therefore the numerical complexity for a single optimisation step is reduced. In other words, the optimisation problem is split into multiple subproblems with disjunct parameter sets, and these subproblems are solved in an alternating manner. This optimisation strategy is known as the space-alternating generalized expectation-maximization (SAGE) algorithm [FH94]. A proper choice for the optimisation problem under consideration is the separate optimisation w.r.t. the parameters of the mean and covariance model. Hence, the vector of unknown parameters $\boldsymbol{\theta}$ is split into the vectors $\boldsymbol{\varpi}$ and $\boldsymbol{\sigma}$, which yields two subproblems for optimisation.

$$\hat{\boldsymbol{\varpi}} = \arg\min_{\boldsymbol{\varpi} \in \Pi} \mathcal{L}\left(\mathbf{y}\,|\,\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}}\right) \tag{5.98a}$$

$$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma} \in \Sigma} \mathcal{L}\left(\mathbf{y}\,|\,\boldsymbol{\sigma}, \hat{\boldsymbol{\varpi}}\right) \tag{5.98b}$$

Both optimisation problems are considered in an alternating manner until convergence. However, if the parameters are coupled, i.e. variation of one parameter causes variations of the objective in the dimension of another parameter, this approach is suboptimal and many iterations are necessary. Also, separation of the optimisation problem into an optimisation w.r.t. the mean parameters and an optimisation w.r.t. the covariance parameters necessitates, that the mean and covariance model are driven by disjunct parameters. This, however, is not ensured as the covariance matrix depends on the propagation parameters due to the phase noise model (see Section 4.6). As a consequence, more iterations are necessary. However, this is still computationally more efficient than optimising w.r.t. all unknown parameters jointly.

5.3.1 Optimisation w.r.t. the Parameters of the Mean Model

Estimation of the parameters of the mean model requires the minimisation of the objective $\mathcal{L}(\mathbf{y}|\,\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}})$.

$$\mathcal{L}(\mathbf{y}|\,\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}}) = \ln\{|\boldsymbol{\Sigma}(\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}})|\} + [\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\varpi})]^{\mathrm{H}} \cdot \boldsymbol{\Sigma}(\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}})^{-1} \cdot [\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\varpi})] \qquad (5.99)$$

This objective is non-convex, such that an iterative approach is required. An iterative approach requires the evaluation of the objective for various $\boldsymbol{\varpi}$. Thus, the determinant and the inverse of the covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}})$ has to be evaluated multiple times, which is numerically very expensive.

Because an iterative optimisation scheme will be employed and based on numerical experiences, a simplification will be drawn: the covariance matrix will be kept fix $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\hat{\boldsymbol{\varpi}}, \hat{\boldsymbol{\sigma}})$ while optimising the objective w.r.t. the parameters $\boldsymbol{\varpi}$. This simplification changes the shape of the objective, but does not alter the location of the global optimum. The varied shape will tend to more iterations of the optimisation algorithm, but each iteration is numerically much cheaper. Also, the estimator does not become biased due to the simplification. The simplification might not be optimal in the sense of the parameter estimates, but the estimator still converges. In order to draw the above simplification, the parameters $\boldsymbol{\sigma}$ have to be known prior to the mean parameters $\boldsymbol{\varpi}$. Hence, the parameters of the covariance model have to be estimated before the parameters of the mean model, which is especially important for the beginning of the iteration of the optimisation algorithm. Assuming the covariance matrix as fix w.r.t. the parameters $\boldsymbol{\varpi}$, the objective function can be simplified.

$$\hat{\boldsymbol{\varpi}} = \arg\min_{\boldsymbol{\varpi} \in \Pi} \mathcal{L}(\mathbf{y}|\,\boldsymbol{\varpi}, \hat{\boldsymbol{\sigma}}) \approx \arg\min_{\boldsymbol{\varpi} \in \Pi} \ [\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\varpi})]^{\mathrm{H}} \cdot \hat{\boldsymbol{\Sigma}}^{-1} \cdot [\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\varpi})] \qquad (5.100)$$

The estimation of the parameter vector $\boldsymbol{\varpi}$ simplifies to a *box-constrained weighted non-linear least-squares problem*. The computational burden is much lower, because the inverse of the covariance matrix has to be calculated only once.

Many algorithms are known in literature to solve a weighted non-linear least-squares optimisation problem (5.100): SAGE-based algorithm with line-search [Fle+99], alternating projection [ZW88], RELAX algorithm [LS96, LL98] or CLEAN algorithms [CSW02, MSE08, TS88]. The subsequently described optimisation procedure is related to the RIMAX framework [RLT03, RLT02], which among the other algorithms features a good convergence speed and convergence to a global minimum is not guaranteed but highly probable. The algorithm is twofold. First, a coarse initial estimate of the non-linear parameters delay, angle and Doppler is conducted using an RELAX-like approach. The linear parameters are jointly estimated using the BLUE. Second, the coarse parameter estimates are refined by joint optimisation of all parameters by a gradient-based optimisation approach to ensure the high-resolution capability of the parameter estimator.

Because the covariance matrix is hermitian and positive semi-definite, their Cholesky factorisation [GV13] can be calculated.

$$\hat{\boldsymbol{\Sigma}} = \mathbf{L} \cdot \mathbf{L}^{\mathrm{H}} \qquad (5.101)$$

Plugging into (5.100), the weighted non-linear least-squares problem becomes a *non-linear least-squares problem* [NW06].

$$\left[\mathbf{y} - \boldsymbol{\mu}\left(\boldsymbol{\varpi}\right)\right]^{\mathrm{H}} \cdot \hat{\boldsymbol{\Sigma}}^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}\left(\boldsymbol{\varpi}\right)\right] = \left\|\mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}}\left(\boldsymbol{\varpi}\right)\right\|^{2} \tag{5.102}$$

Variables $\mathbf{y}_{\mathrm{L}} = \mathbf{L}^{-1}\mathbf{y}$ and $\boldsymbol{\mu}_{\mathrm{L}}\left(\boldsymbol{\varpi}\right) = \mathbf{L}^{-1}\boldsymbol{\mu}\left(\boldsymbol{\varpi}\right)$ are introduced for short hand notation.

Initial Estimation of Parameter Values

Finding initial and coarse estimates for the parameter vector $\boldsymbol{\varpi}$ is possible by a multidimensional exhaustive search of the minimum of the objective (5.102). However, an exhaustive search is intractable under practical considerations of numerical complexity, because the dimensionality of (5.102) is too high and also scales with the number of paths. Hence, one may seek for methods, which are possibly iterative but numerically less complex in each iteration.

Variable Projection Method   Consider the matrix-vector representation of model (5.22) of the mean vector.

$$\boldsymbol{\mu}\left(\boldsymbol{\varpi}\right) = \mathbf{S}\left(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}\right) \cdot \boldsymbol{\gamma} = \left[\mathbf{s}\left(\varphi_{1}, \tau_{1}, \nu_{1}\right), \ldots, \mathbf{s}\left(\varphi_{P}, \tau_{P}, \nu_{P}\right)\right] \cdot \boldsymbol{\gamma} = \sum_{p=1}^{P} \gamma_{p} \cdot \mathbf{s}\left(\varphi_{p}, \tau_{p}, \nu_{p}\right)$$

Note, that $\mathbf{s}\left(\varphi_{p}, \tau_{p}, \nu_{p}\right)$ describes the model for a single path. Plugging into the least-squares objective (5.102) yields

$$\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}, \boldsymbol{\gamma}}{\arg \min} \left\|\mathbf{y}_{\mathrm{L}} - \mathbf{S}_{\mathrm{L}}\left(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}\right) \cdot \boldsymbol{\gamma}\right\|^{2} \quad , \tag{5.103}$$

with $\mathbf{S}_{\mathrm{L}} = \mathbf{L}^{-1}\mathbf{S}$ introduced for short hand notation. This optimisation problem is a *separable non-linear least-squares problem* [GP73, GP03], as the non-linear functions $\mathbf{s}\left(\varphi_{p}, \tau_{p}, \nu_{p}\right)$ are weighted add up in the model. Hence, the *variable projection algorithm* [GP73, GP03] can be used. Once estimates of the non-linear parameters $\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}$ are known, an estimate of the vector of linear parameters $\boldsymbol{\gamma}$ is given by the BLUE [MS00].

$$\hat{\boldsymbol{\gamma}} = \mathbf{S}_{\mathrm{L}}\left(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}\right)^{+} \cdot \mathbf{y}_{\mathrm{L}} = \left[\mathbf{S}\left(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}\right)^{\mathrm{H}} \cdot \hat{\boldsymbol{\Sigma}}^{-1} \cdot \mathbf{S}\left(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}\right)\right]^{-1} \cdot \mathbf{S}\left(\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}\right)^{\mathrm{H}} \cdot \hat{\boldsymbol{\Sigma}}^{-1} \cdot \mathbf{y} \tag{5.104}$$

Plugging the BLUE into equation (5.103) yields the variable projection functional, which is the profiled (concentrated) objective (5.103) w.r.t. the linear parameters $\boldsymbol{\gamma}$.

$$\hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}} = \underset{\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}}{\arg \min} \left\|\mathbf{y}_{\mathrm{L}} - \mathbf{S}_{\mathrm{L}}\left(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}\right) \cdot \mathbf{S}_{\mathrm{L}}\left(\boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\nu}\right)^{+} \cdot \mathbf{y}_{\mathrm{L}}\right\|^{2} \tag{5.105}$$

Instead of optimising the $5P$-dimensional objective (5.103), the variable projection method splits the optimisation problem into a non-linear least squares problem of dimensionality $3P$ and a linear optimisation problem, which has a closed-form solution. Hence, the optimisation problem is relaxed.

SAGE Algorithm with Beamforming Estimation   From the variable projection method a $3P$-dimensional non-linear least squares optimisation problem remains. However, an exhaustive multidimensional search is still not conceivable from the viewpoint of numerical complexity.

A further relaxation of the search space dimensionality can be achieved by exploiting the concept of the SAGE algorithm [FH94]. The single search in the $3P$-dimensional parameter space is decoupled into P successive searches in a 3-dimensional parameter space, and the P searches are considered in an alternating manner. Presume that estimates of $P - 1$ paths, i.e. the respective vector of parameter estimates, is known.

$$\hat{\boldsymbol{\varpi}}_{1:P-1} = \left[\hat{\boldsymbol{\phi}}_{1:P-1}^{\mathrm{T}}, \hat{\boldsymbol{\tau}}_{1:P-1}^{\mathrm{T}}, \hat{\boldsymbol{\nu}}_{1:P-1}^{\mathrm{T}}, \Re\left\{\hat{\boldsymbol{\gamma}}_{1:P-1}\right\}^{\mathrm{T}}, \Im\left\{\hat{\boldsymbol{\gamma}}_{1:P-1}\right\}^{\mathrm{T}}\right]^{\mathrm{T}}$$

By removing the contributions of all $P - 1$ paths from the observation (*expectation step*), only the contribution from the P-th path remains.

$$\mathbf{r} = \mathbf{y} - \boldsymbol{\mu}\left(\hat{\boldsymbol{\varpi}}_{1:P-1}\right) = \mathbf{y} - \mathbf{S}\left(\hat{\boldsymbol{\phi}}_{1:P-1}, \hat{\boldsymbol{\tau}}_{1:P-1}, \hat{\boldsymbol{\nu}}_{1:P-1}\right) \cdot \boldsymbol{\gamma}_{1:P-1} \qquad (5.106)$$

Vector $\mathbf{r}$ denotes the *remaining signal*. The parameter estimation problem is reduced to the estimation of the parameters of the P-th path only, which is a 3-dimensional problem. Restating the variable projection functional (5.105) w.r.t. a single path yields

$$\hat{\varphi}, \hat{\tau}, \hat{\nu} = \arg\min_{\varphi, \tau, \nu} \left\| \mathbf{r}_{\mathrm{L}} - \mathbf{s}_{\mathrm{L}}\left(\varphi, \tau, \nu\right) \cdot \mathbf{s}_{\mathrm{L}}\left(\varphi, \tau, \nu\right)^{+} \cdot \mathbf{r}_{\mathrm{L}} \right\|^{2} \qquad . \qquad (5.107)$$

After some numerical manipulations and ignoring constant terms, the minimisation problem can be formulated as an equivalent maximisation problem (*maximisation step*).

$$\hat{\varphi}, \hat{\tau}, \hat{\nu} = \arg\max_{\varphi, \tau, \nu} \frac{\left| \mathbf{r}_{\mathrm{L}}^{\mathrm{H}} \cdot \mathbf{s}_{\mathrm{L}}\left(\varphi, \tau, \nu\right) \right|^{2}}{\left\| \mathbf{s}_{\mathrm{L}}\left(\varphi, \tau, \nu\right) \right\|^{2}} \qquad (5.108)$$

Objective equation (5.108) can be considered as correlation function or beamforming, whereas the "beam" is adjusted by tuning the non-linear parameters of interest.

For maximisation of the 3-dimensional correlation function a grid-based search is conducted. A grid for each parameter dimension is defined beforehand, and the grid bounds are chosen in accordance to the respective parameter bounds in Tab. 5.1.

- Grid size in delay domain: variable resolution of $\frac{360°}{4 \cdot K}$

- Grid size in azimuth domain: 1° resolution

- Grid size in Doppler domain: variable resolution of $\frac{360°}{10 \cdot N}$

The search grid is very densely chosen, because the considered radar has a large aperture in spatial and delay domain. Due to the large aperture, the main lobe of the beamforming function (5.108) is very narrow and a dense grid becomes necessary for proper detection of it. Remember, that a proper detection is required to ensure the initial estimates being close to the optimum. For each 3-dimensional grid point $\left(\varphi_{h}, \tau_{i}, \nu_{j}\right)$ a pattern $\mathbf{s}\left(\varphi_{h}, \tau_{i}, \nu_{j}\right)$ is calculated beforehand. In the maximisation step these patterns are compared by calculating the respective correlation function equation (5.108), and the parameters which correspond to the pattern with

---

Algorithm 1 Path Search Algorithm

Define grid: $\quad \varphi_1, \dots, \varphi_H; \tau_1, \dots, \tau_I; v_1, \dots, v_J$

Calculate model patterns: $\quad \mathbf{s}\left(\varphi_1, \tau_1, v_1\right), \dots, \mathbf{s}\left(\varphi_H, \tau_I, v_J\right)$

Require: $\mathbf{y}, \mathbf{L}^{-1}, \hat{\boldsymbol{\omega}}_{1:P}$

Expectation step: $\quad \mathbf{r}_{\mathrm{L}} \leftarrow \mathbf{L}^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}\left(\hat{\boldsymbol{\omega}}_{1:P}\right)\right]$

Maximisation step: $\quad \hat{\varphi}, \hat{\tau}, \hat{v} \leftarrow \underset{\substack{\varphi_h \in \{\varphi_1, \dots, \varphi_H\} \\ \tau_i \in \{\tau_1, \dots, \tau_I\} \\ v_j \in \{v_1, \dots, v_J\}}}{\arg\max} \left|\mathbf{r}_{\mathrm{L}}^{\mathrm{H}} \cdot \mathbf{s}\left(\varphi_h, \tau_i, v_j\right)\right|^2 \Big/ \left\|\mathbf{s}\left(\varphi_h, \tau_i, v_j\right)\right\|^2$

$\hat{\boldsymbol{\varphi}}_{1:P+1} \leftarrow \left[\hat{\boldsymbol{\varphi}}_{1:P}^{\mathrm{T}}, \hat{\varphi}\right]^{\mathrm{T}}$

$\hat{\boldsymbol{\tau}}_{1:P+1} \leftarrow \left[\hat{\boldsymbol{\tau}}_{1:P}^{\mathrm{T}}, \hat{\tau}\right]^{\mathrm{T}}$

$\hat{\boldsymbol{v}}_{1:P+1} \leftarrow \left[\hat{\boldsymbol{v}}_{1:P}^{\mathrm{T}}, \hat{v}\right]^{\mathrm{T}}$

---

the highest correlation are considered as results of the maximisation step, i.e are the estimates of the non-linear parameters.

The maximisation (5.108) can be costly, because a 3-dimensional objective is considered. In [Ric05] the domains of the correlation function (i.e. time, frequency or spatial domain) are individually processed in the maximisation step, i.e. three 1-dimensional maximisation are conducted. All other dimensions are collapsed by either incoherent averaging or application of beamforming using the previously gained estimates. This kind of optimisation coincides with the optimisation approach in [Fle+99]. The individual processing of each dimension is possible due to the Kronecker structured data model, which causes the data dimensions to be decoupled in the model. The individual processing causes loss of correlation gain, such that the results of the maximisation step are just coarse estimates. Advantageously, only a 1-dimensional maximisation has to be conducted, which has advantages in terms of computational complexity. However, the data model under consideration has no Kronecker structure due to the angle-Doppler and angle-delay (wideband array model) coupling. Hence, it is not possible to collapse the dimensions and a complete 3-dimensional search becomes necessary.

Path Search Algorithm Assume, that the parameters of *P*-paths (*P* can be zero as well!) have been estimated, or are known due to e.g. tracking [28] or a previous run of the estimator: $\hat{\boldsymbol{\omega}}_{1:P}$. Then, an approach to estimate a single, additional path is given by using the previously described SAGE algorithm with the beamforming estimation. Here, the algorithm is used to estimate the parameters of a single new path from the remaining signal defined in (5.106), rather than alternating between the already known paths to improve their estimates. The path search algorithm is given in Algorithm 1.

Having a new set of non-linear parameters, the corresponding linear parameters can be estimated using the BLUE.

$$\hat{\boldsymbol{\gamma}}_{1:P+1} = \mathbf{S}_{\mathrm{L}}\left(\hat{\boldsymbol{\varphi}}_{1:P+1}, \hat{\boldsymbol{\tau}}_{1:P+1}, \hat{\boldsymbol{v}}_{1:P+1}\right)^{+} \cdot \mathbf{y}_{\mathrm{L}}$$

Note, that the vector of linear parameters $\boldsymbol{\gamma}$ is estimated in each iteration step, i.e., the path weights are jointly estimated after each newly detected path. This is a difference compared to the path search approach in [Ric05], where the path weight of each newly detected path

is individually estimated. By jointly estimating the path weights w.r.t. all detected paths the estimation accuracy is improved, such that the expectation step becomes more accurate. Last, the new set of parameter estimates is given by concatenating all latest estimates.

$$\hat{\boldsymbol{\varpi}}_{1:P+1} = \left[\hat{\boldsymbol{\varphi}}_{1:P+1}^{\mathrm{T}}, \hat{\boldsymbol{\tau}}_{1:P+1}^{\mathrm{T}}, \hat{\boldsymbol{\nu}}_{1:P+1}^{\mathrm{T}}, \Re\left\{ \hat{\boldsymbol{\gamma}}_{1:P+1} \right\}^{\mathrm{T}}, \Im\left\{ \hat{\boldsymbol{\gamma}}_{1:P+1} \right\}^{\mathrm{T}} \right]^{\mathrm{T}}$$

This path search scheme is comparable to the RELAX algorithm [LS96, LL98]. On the contrary, no re-estimation of the non-linear parameters of the previously detected paths is considered, and the search grid is fix and not re-adjusted/zoomed to improve the resolution of the estimates. Hence, the presented path search scheme may result in worse estimates, if multiple propagation paths, which are close to each other [Ric05], occur. Therefore, an additional optimisation step is necessary to improve the parameter estimates.

Joint Estimation by Gradient Method

The initial parameter estimates are probably imprecise, such that a refinement becomes necessary. Furthermore, optimisation w.r.t. all estimated paths is necessary to ensure high-resolution. A gradient-based optimisation method will be used, because it features a better convergence speed than e.g. axis parallel optimisation as done by the SAGE algorithm [FJS02, Fle+99].

Consider a point $\boldsymbol{\varpi}$, which is $\varDelta\boldsymbol{\varpi}$ apart from a (local) minimum $\hat{\boldsymbol{\varpi}}$ of the objective (5.102). By linearisation (first-order Taylor series expansion) of the mean model $\boldsymbol{\mu}_{\mathrm{L}}(\boldsymbol{\varpi})$ at point $\boldsymbol{\varpi}+\varDelta\boldsymbol{\varpi}$, the descent direction $\varDelta\boldsymbol{\varpi}$ to the minimum can be calculated.

$$\boldsymbol{\mu}_{\mathrm{L}}(\boldsymbol{\varpi} + \varDelta\boldsymbol{\varpi}) \approx \boldsymbol{\mu}_{\mathrm{L}}(\boldsymbol{\varpi}) + \mathbf{J}_{\mathrm{L}}(\boldsymbol{\varpi}) \cdot \varDelta\boldsymbol{\varpi} \tag{5.109}$$

Matrix $\mathbf{J}_{\mathrm{L}}(\boldsymbol{\varpi})$ is the *Jacobian matrix*.

$$\mathbf{J}_{\mathrm{L}}(\boldsymbol{\varpi}) = \frac{\partial\boldsymbol{\mu}_{\mathrm{L}}(\boldsymbol{\varpi})}{\partial\boldsymbol{\varpi}^{\mathrm{T}}} \tag{5.110}$$

Plugging the Taylor series expansion into the non-linear least-squares objective yields the *subproblem of gradient-based optimisation* (arguments are dropped for notational convenience).

$$\underset{\varDelta\boldsymbol{\varpi}\in\mathbb{R}^{5P}}{\arg\min} \left\| (\mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}}) - \mathbf{J}_{\mathrm{L}} \cdot \varDelta\boldsymbol{\varpi} \right\|^2 \iff \mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}} = \mathbf{J}_{\mathrm{L}} \cdot \varDelta\boldsymbol{\varpi} \tag{5.111}$$

The solution to this optimisation problem (left-hand side) is given by solving a linear system of equation (right-hand side). Because the descent direction $\varDelta\boldsymbol{\varpi}$ is real-valued and $\mathbf{y}_{\mathrm{L}}$, $\boldsymbol{\mu}_{\mathrm{L}}$ and $\mathbf{J}_{\mathrm{L}}$ are complex-valued, the system of equations is split into its real and imaginary part.

$$\begin{bmatrix} \Re\left\{ \mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}} \right\} \\ \Im\left\{ \mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}} \right\} \end{bmatrix} = \begin{bmatrix} \Re\left\{ \mathbf{J}_{\mathrm{L}} \right\} \\ \Im\left\{ \mathbf{J}_{\mathrm{L}} \right\} \end{bmatrix} \cdot \varDelta\boldsymbol{\varpi}$$

The solution of the subproblem is given by the Moore–Penrose pseudo inverse [GV13].

$$\Delta\boldsymbol{\varpi} = \left(\begin{bmatrix} \Re\left\{\mathbf{J}_{\mathrm{L}}\right\} \\ \Im\left\{\mathbf{J}_{\mathrm{L}}\right\} \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} \Re\left\{\mathbf{J}_{\mathrm{L}}\right\} \\ \Im\left\{\mathbf{J}_{\mathrm{L}}\right\} \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} \Re\left\{\mathbf{J}_{\mathrm{L}}\right\} \\ \Im\left\{\mathbf{J}_{\mathrm{L}}\right\} \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} \Re\left\{\mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}}\right\} \\ \Im\left\{\mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}}\right\} \end{bmatrix}$$

By exploiting[4] it can be re-written in more compact form.

$$\Delta\boldsymbol{\varpi} = \left[\Re\left\{ \mathbf{J}_{\mathrm{L}}\left(\boldsymbol{\varpi}\right)^{\mathrm{H}} \cdot \mathbf{J}_{\mathrm{L}}\left(\boldsymbol{\varpi}\right) \right\} \right]^{-1} \cdot \Re\left\{ \mathbf{J}_{\mathrm{L}}\left(\boldsymbol{\varpi}\right)^{\mathrm{H}} \cdot \left[\mathbf{y}_{\mathrm{L}} - \boldsymbol{\mu}_{\mathrm{L}}\left(\boldsymbol{\varpi}\right) \right] \right\}$$

The vector $\mathbf{g}\left(\boldsymbol{\varpi}\right)$ and the matrix $\mathbf{H}\left(\boldsymbol{\varpi}\right)$ will be introduced for short hand notation.

$$\mathbf{g}\left(\boldsymbol{\varpi}\right) = \Re\left\{ \mathbf{J}\left(\boldsymbol{\varpi}\right)^{\mathrm{H}} \cdot \boldsymbol{\Sigma}^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu}\left(\boldsymbol{\varpi}\right) \right] \right\} \tag{5.112a}$$

$$\mathbf{H}\left(\boldsymbol{\varpi}\right) = \Re\left\{ \mathbf{J}\left(\boldsymbol{\varpi}\right)^{\mathrm{H}} \cdot \boldsymbol{\Sigma}^{-1} \cdot \mathbf{J}\left(\boldsymbol{\varpi}\right) \right\} \tag{5.112b}$$

Last, the location of the optimum (respectively the parameter estimate) can be calculated by

$$\hat{\boldsymbol{\varpi}} = \boldsymbol{\varpi} + \Delta\boldsymbol{\varpi} \quad . \tag{5.113}$$

The linearisation (5.109) of the mean model is just an approximation, such that the descent direction $\Delta\boldsymbol{\varpi}$ towards the minimum is not optimal. Hence, an iterative minimisation of (5.100) is commonly applied. For each iteration i an update of the parameter estimate is given by the previous parameter estimate $\boldsymbol{\varpi}_{i-1}$ and an estimated update step $\Delta\boldsymbol{\varpi}_i$.

$$\boldsymbol{\varpi}_i = \boldsymbol{\varpi}_{i-1} + \Delta\boldsymbol{\varpi}_i \tag{5.114}$$

The initial guess $\boldsymbol{\varpi}_0$ is known from the previously described path search approach. Furthermore, because the first-order Taylor approximation of the objective is minimised rather than the true objective, usage of the raw descent direction as the update step is probably not optimal too. The objective is possibly not decreased or not decreased in the best manner. Hence, the update step has to be adjusted. Two methods are known from literature to adjust the update step: the *line search* and the *trust region* approach, resulting in the Gauß-Newton and Levenberg-Marquardt algorithm, respectively.

Line Search Approach   The line search approach attempts to adjust the length of the descent direction $\Delta\boldsymbol{\varpi}_i$ in order to find a proper update step. Therefore, a step length factor δ_i is introduced [NW06].

$$\Delta\boldsymbol{\varpi}_i = \delta_i \cdot \mathbf{H}\left(\boldsymbol{\varpi}_i\right)^{-1} \cdot \mathbf{g}\left(\boldsymbol{\varpi}_i\right) \tag{5.115}$$

The factor δ_i is chosen by a line search algorithm, which attempts to satisfy the Wolfe conditions [NW06] in order to terminate the line search. The parameter update in each iteration is now given by

$$\boldsymbol{\varpi}_{i+1} = \boldsymbol{\varpi}_i + \delta_i \cdot \mathbf{H}\left(\boldsymbol{\varpi}_i\right)^{-1} \cdot \mathbf{g}\left(\boldsymbol{\varpi}_i\right) \quad . \tag{5.116}$$

---

[4] $\begin{bmatrix} \Re\left\{\mathbf{A}\right\} \\ \Im\left\{\mathbf{A}\right\} \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} \Re\left\{\mathbf{B}\right\} \\ \Im\left\{\mathbf{B}\right\} \end{bmatrix} = \Re\left\{\mathbf{A}\right\}^{\mathrm{T}} \cdot \Re\left\{\mathbf{B}\right\} + \Im\left\{\mathbf{A}\right\}^{\mathrm{T}} \cdot \Im\left\{\mathbf{B}\right\} = \Re\left\{\mathbf{A}^{\mathrm{H}} \cdot \mathbf{B}\right\}$

The above parameter update is known as Gauß-Newton update.

The line-search approach is able to minimise the objective towards a local minimum. However, the Gauß-Newton update may face numerical issues, if matrix $\mathbf{H}\left(\boldsymbol{\varpi}_i\right)$ becomes singular or nearly singular, i.e. the Jacobian becomes rank-deficient. This may happen, if two propagation paths are close and the parameter estimator attempts to resolve both paths, i.e. high-resolution shall be achieved. Hence, the Gauß-Newton update is possibly not the appropriate method to calculate the update step for the problem under consideration.

Trust Region Approach   The trust region approach attempts to determine the direction of a proper update step, while the length $\left\|\Delta\boldsymbol{\varpi}_i\right\|$ of the update step is pre-determined. Because the first-order Taylor approximation of the objective is only valid in a small region (trust region) around the actual iteration point $\boldsymbol{\varpi}_{i-1}$, one may restrict the length of the update step $\Delta\boldsymbol{\varpi}_i$, such that the new iteration point $\boldsymbol{\varpi}_{i+1}$ lies inside of the trust region. The respective descent direction is found by solving the modified subproblem.

$$\underset{\Delta\boldsymbol{\varpi}_i\in\mathbb{R}^{5P}}{\arg\min}\ \left\|\left[\mathbf{y}_{\mathrm{L}}-\boldsymbol{\mu}_{\mathrm{L}}\left(\boldsymbol{\varpi}_i\right)\right]-\mathbf{J}_{\mathrm{L}}\left(\boldsymbol{\varpi}_i\right)\cdot\Delta\boldsymbol{\varpi}_i\right\|^2\quad\text{s.t.}\quad\left\|\Delta\boldsymbol{\varpi}_i\right\|<\epsilon_i\qquad(5.117)$$

Length ϵ_i the radius of the trust region or length of the update step. Hence, the descent direction is calculated for a predefined step length. It is proven, that a solution for the trust region subproblem is given if and only if a parameter $\delta_i>0$ exists [NW06].

$$\boldsymbol{\varpi}_{i+1}=\boldsymbol{\varpi}_i+\left[\mathbf{H}\left(\boldsymbol{\varpi}_i\right)+\delta_i\cdot\mathbf{I}_{5P}\right]^{-1}\cdot\mathbf{g}\left(\boldsymbol{\varpi}_i\right)\qquad(5.118)$$

The parameter δ_i is tuned in order to adjust the trust region radius (i.e. tuning the step length) and also the update direction. The above update step is known as Levenberg's update step. An improvement of Levenberg's update step, which improves the convergence rate in the direction of a small gradient, is the Marquardt update step.

$$\boldsymbol{\varpi}_{i+1}=\boldsymbol{\varpi}_i+\left[\mathbf{H}\left(\boldsymbol{\varpi}_i\right)+\delta_i\cdot\mathbf{I}_{5P}\odot\mathbf{H}\left(\boldsymbol{\varpi}_i\right)\right]^{-1}\cdot\mathbf{g}\left(\boldsymbol{\varpi}_i\right)\qquad(5.119)$$

The trust region approach has a clear advantage compared to the line search approach and will be considered therefore. Instead of inverting the (possible) singular matrix $\mathbf{H}\left(\boldsymbol{\varpi}_i\right)$, the regularised matrix $\left[\mathbf{H}\left(\boldsymbol{\varpi}_i\right)+\delta_i\cdot\mathbf{I}_{5P}\odot\mathbf{H}\left(\boldsymbol{\varpi}_i\right)\right]$ is inverted, which results in numerically more stable solutions. Therefore, the Levenberg-Marquardt update step seems to be more appropriate to achieve high-resolution.

The complete Levenberg-Marquardt optimisation algorithm is shown in Algorithm 2.

Calculation of the Partial Derivatives

The Jacobian matrix $\mathbf{J}$ comprises the partial derivatives of the mean model $\boldsymbol{\mu}\left(\boldsymbol{\varpi}\right)$ w.r.t. its parameters $\boldsymbol{\varpi}$.

$$\mathbf{J}=\frac{\partial}{\partial\boldsymbol{\varpi}^{\mathrm{T}}}\boldsymbol{\mu}\left(\boldsymbol{\varpi}\right)\qquad(5.120)$$

Algorithm 2 Levenberg-Marquardt Algorithm

Require: $\boldsymbol{\varpi}_0, \mathbf{L}^{-1}$
  $i \leftarrow 0$
  $\delta_i \leftarrow 0.01$
  repeat
    $\mathbf{g}_i, \mathbf{H}_i \leftarrow$ *equation* (5.112)
    $\boldsymbol{\varpi}_{i+1} \leftarrow \boldsymbol{\varpi}_i + \left[\mathbf{H}_i + \delta_i \cdot \mathrm{diag} \left\{ \mathrm{diag} \left\{ \mathbf{H}_i \right\} \right\} \right]^{-1} \cdot \mathbf{g}_i$
    if $\left\| \mathbf{L}^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu} \left(\boldsymbol{\varpi}_{i+1} \right) \right] \right\|^2 < \left\| \mathbf{L}^{-1} \cdot \left[\mathbf{y} - \boldsymbol{\mu} \left(\boldsymbol{\varpi}_i \right) \right] \right\|^2$ **then**
      $\delta_{i+1} \leftarrow \delta_i / 10$
    else
      $\boldsymbol{\varpi}_{i+1} \leftarrow \boldsymbol{\varpi}_i$
      $\delta_{i+1} \leftarrow 10 \cdot \delta_i$
    end if
    $i \leftarrow i + 1$
  until *Convergence*

The mean model is given by equation (5.22). Therefore, the partial derivative of vector $\boldsymbol{\mu}\left(\boldsymbol{\varpi} \right)$ w.r.t. an arbitrary parameter ϖ is given by the partial derivatives of each vector entry w.r.t. this variable.

$$\frac{\partial}{\partial \varpi} \boldsymbol{\mu}\left(\boldsymbol{\varpi} \right) = \mathrm{vec} \left\{ \left[\frac{\partial}{\partial \varpi} \boldsymbol{\mu}_1\left(\boldsymbol{\varpi} \right), \dots, \frac{\partial}{\partial \varpi} \boldsymbol{\mu}_M\left(\boldsymbol{\varpi} \right) \right] \right\} \tag{5.121}$$

The vectors $\boldsymbol{\mu}_m\left(\boldsymbol{\varpi} \right)$ are given by equation (5.20). Applying the product rule for derivatives [AS72], the partial derivative becomes (arguments are dropped for convenience)

$$\begin{aligned}
\frac{\partial}{\partial \varpi} \boldsymbol{\mu}_m = & \left(\frac{\partial \mathbf{A}_m \odot \mathbf{B}_\nu}{\partial \varpi} \diamond \left[\left(\mathrm{diag}\left\{ \mathbf{g}_m \right\} \cdot \mathbf{B}_{\varphi_m} \right) \odot \mathbf{B}_\tau \right] \right) \cdot \boldsymbol{\gamma} \\
& + \left(\left[\mathbf{A}_m \odot \mathbf{B}_\nu \right] \diamond \left[\frac{\partial \mathrm{diag}\left\{ \mathbf{g}_m \right\} \cdot \mathbf{B}_{\varphi_m}}{\partial \varpi} \odot \mathbf{B}_\tau \right] \right) \cdot \boldsymbol{\gamma} \\
& + \left(\left[\mathbf{A}_m \odot \mathbf{B}_\nu \right] \diamond \left[\left(\mathrm{diag}\left\{ \mathbf{g}_m \right\} \cdot \mathbf{B}_{\varphi_m} \right) \odot \frac{\partial \mathbf{B}_\tau}{\partial \varpi} \right] \right) \cdot \boldsymbol{\gamma} \\
& + \left(\left[\mathbf{A}_m \odot \mathbf{B}_\nu \right] \diamond \left[\left(\mathrm{diag}\left\{ \mathbf{g}_m \right\} \cdot \mathbf{B}_{\varphi_m} \right) \odot \mathbf{B}_\tau \right] \right) \cdot \frac{\partial \boldsymbol{\gamma}}{\partial \varpi} \quad .
\end{aligned} \tag{5.122}$$

Derivation w.r.t. the Doppler Shift The partial derivatives w.r.t. the Doppler shift are the derivatives of the matrix $\left[\mathbf{A}_m\left(\boldsymbol{\nu} \right) \odot \mathbf{B}_\nu\left(\boldsymbol{\nu} \right) \right]$. According to the product rule of derivatives, the partial derivative is

$$\frac{\partial}{\partial \varpi} \left[\mathbf{A}_m\left(\boldsymbol{\nu} \right) \odot \mathbf{B}_\nu\left(\boldsymbol{\nu} \right) \right] = \left[\frac{\partial}{\partial \varpi} \mathbf{A}_m\left(\boldsymbol{\nu} \right) \right] \odot \mathbf{B}_\nu\left(\boldsymbol{\nu} \right) + \mathbf{A}_m\left(\boldsymbol{\nu} \right) \odot \left[\frac{\partial}{\partial \varpi} \mathbf{B}_\nu\left(\boldsymbol{\nu} \right) \right] \quad . \tag{5.123}$$

Matrix $\mathbf{B}_\nu(\nu)$ and $\mathbf{A}_m(\nu)$ are given by (5.9b) and (5.19), respectively, and their derivatives are

$$\frac{\partial}{\partial \varpi} \mathbf{B}_\nu(\nu) = \exp\left\{j\mathbf{k}\nu^{\mathrm{T}}\right\} \odot \left[j\mathbf{k}\left(\frac{\partial}{\partial \varpi}\nu\right)^{\mathrm{T}}\right] \tag{5.124a}$$

$$\frac{\partial}{\partial \varpi} \mathbf{A}_m(\nu) = \exp\left\{j\frac{(m-1)\setminus M_{\mathrm{Tx}}}{M_{\mathrm{Tx}}} \cdot \mathbf{1}_N \cdot \nu^{\mathrm{T}}\right\} \odot \left[j\frac{(m-1)\setminus M_{\mathrm{Tx}}}{M_{\mathrm{Tx}}} \cdot \mathbf{1}_N \cdot \left(\frac{\partial}{\partial \varpi}\nu\right)^{\mathrm{T}}\right] \quad . \tag{5.124b}$$

Derivation w.r.t. the Azimuth Angle   The partial derivatives w.r.t. the azimuth angle are the derivatives of matrix $\mathrm{diag}\left\{\mathbf{g}_m\right\} \cdot \mathbf{B}_{\varphi_m}(\varphi)$. Applying the product rule, the partial derivative is given by

$$\frac{\partial}{\partial \varpi} \left[\mathrm{diag}\left\{\mathbf{g}_m\right\} \cdot \mathbf{B}_{\varphi_m}(\varphi)\right] = \mathrm{diag}\left\{\mathbf{g}_m\right\} \cdot \frac{\partial}{\partial \varpi} \mathbf{B}_{\varphi_m}(\varphi) \quad . \tag{5.125}$$

Hence, the partial derivative of the antenna response matrix $\mathbf{B}_{\varphi_m}(\varphi)$ of the m-th MIMO channel has to be calculated. Using the concept of array manifold decomposition [CRK10] as e.g. exploited by the ETADF [24], the antenna response can be represented by

$$\mathbf{B}_{\varphi_m}(\varphi) = \mathbf{G}_m \cdot \exp\left\{j\mathbf{u}\varphi^{\mathrm{T}}\right\} \quad , \tag{5.126}$$

with $\mathbf{u} = \frac{1}{U} \cdot \left[-\frac{U}{2}, \dots, \frac{U}{2} - 1\right]^{\mathrm{T}} \in \mathbb{R}^U$. Matrix $\mathbf{G}_m \in \mathbb{C}^{M \times U}$ is a specific matrix for the m-th MIMO channel and can be derived from calibration measurements (see Section 4.4). Note, that this matrix is *independent* on the azimuth angle. Using the array manifold decomposition, the partial derivative becomes

$$\frac{\partial}{\partial \varpi} \mathbf{B}_{\varphi_m}(\varphi) = \mathbf{G}_m \cdot \frac{\partial}{\partial \varpi} \exp\left\{j\mathbf{u}\varphi^{\mathrm{T}}\right\} = \mathbf{G}_m \cdot \left[\left(j\mathbf{u}\frac{\partial}{\partial \varpi}\varphi^{\mathrm{T}}\right) \odot \exp\left\{j\mathbf{u}\varphi^{\mathrm{T}}\right\}\right] \quad . \tag{5.127}$$

Derivation w.r.t. the Delay   The partial derivatives w.r.t. the delay are the derivatives of matrix $\mathbf{B}_\tau(\tau)$. Matrix $\mathbf{B}_\tau(\tau)$ is given by equation (5.9a). Its partial derivative is

$$\frac{\partial}{\partial \varpi} \mathbf{B}_\tau(\tau) = \exp\left\{j\mathbf{k}\tau^{\mathrm{T}}\right\} \odot \left[j\mathbf{k}\left(\frac{\partial}{\partial \varpi}\tau\right)^{\mathrm{T}}\right] \quad . \tag{5.128}$$

Derivation w.r.t. the Path Weights   The partial derivatives w.r.t. the path weights are the derivatives of the path weight vector γ itself. The path weight vector is split into his real and imaginary part. Using the sum rule of derivatives [AS72], the partial derivative of the path weight vector is

$$\frac{\partial}{\partial \varpi} \gamma = \frac{\partial}{\partial \varpi} \left[\mathfrak{R}\left\{\gamma\right\} + j\mathfrak{I}\left\{\gamma\right\}\right] = \frac{\partial}{\partial \varpi} \mathfrak{R}\left\{\gamma\right\} + j\frac{\partial}{\partial \varpi} \mathfrak{I}\left\{\gamma\right\} \quad . \tag{5.129}$$

5.3.2 Optimisation w.r.t. the Parameters of the Covariance Model

In order to estimate the parameters $\boldsymbol{\sigma}$ of the covariance model the objective $\mathcal{L}\left(\mathbf{y}\,|\,\boldsymbol{\sigma},\hat{\boldsymbol{\omega}}\right)$ has to be minimised w.r.t. these parameters.

$$\mathcal{L}\left(\mathbf{y}\,|\,\boldsymbol{\sigma},\hat{\boldsymbol{\omega}}\right) = \ln\left\{|\boldsymbol{\Sigma}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}\right)|\right\} + \left[\mathbf{y}-\boldsymbol{\mu}\left(\hat{\boldsymbol{\omega}}\right)\right]^{\mathrm{H}}\cdot\boldsymbol{\Sigma}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}\right)^{-1}\cdot\left[\mathbf{y}-\boldsymbol{\mu}\left(\hat{\boldsymbol{\omega}}\right)\right] \tag{5.130}$$

Model (5.59) of the covariance matrix is block diagonal w.r.t. the MIMO channels and the model parameters of the blocks (of the MIMO channels, respectively) are disjunct. Hence, the objective can be restated using the properties of block diagonal matrices [GV13]. The determinant of a block diagonal matrix is given by the product of the determinants of the blocks.

$$|\boldsymbol{\Sigma}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}\right)| = \prod_{m=1}^{M}\left|\boldsymbol{\Sigma}_{m}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}_{m}\right)\right| \tag{5.131}$$

The inverse of a block diagonal matrix is given by the inverse of the blocks.

$$\boldsymbol{\Sigma}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}\right)^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_{1}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}_{1}\right)^{-1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma}_{M}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}_{M}\right)^{-1} \end{bmatrix}$$

The vector $\mathbf{y}$ of MIMO observations will be split into the observations at each MIMO channel, with $\mathbf{y}_{m}$ the observation at the m-th MIMO channel.

$$\mathbf{y} = \left[\mathbf{y}_{1}^{\mathrm{T}},\ldots,\mathbf{y}_{M}^{\mathrm{T}}\right]^{\mathrm{T}}$$

Now, objective (5.130) can be re-stated.

$$\mathcal{L}\left(\mathbf{y}\,|\,\boldsymbol{\sigma},\hat{\boldsymbol{\omega}}\right) = \sum_{m=1}^{M}\mathcal{L}_{m}\left(\mathbf{y}_{m}\,|\,\boldsymbol{\sigma}_{m},\hat{\boldsymbol{\omega}}\right)$$

The respective objective $\mathcal{L}_{m}\left(\mathbf{y}_{m}\,|\,\boldsymbol{\sigma}_{m},\hat{\boldsymbol{\omega}}\right)$ for each MIMO channel m is

$$\mathcal{L}_{m}\left(\mathbf{y}_{m}\,|\,\boldsymbol{\sigma}_{m},\hat{\boldsymbol{\omega}}\right) = \ln\left\{\left|\boldsymbol{\Sigma}_{m}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}_{m}\right)\right|\right\} + \mathbf{r}_{m}^{\mathrm{H}}\cdot\boldsymbol{\Sigma}_{m}\left(\hat{\boldsymbol{\omega}},\boldsymbol{\sigma}_{m}\right)^{-1}\cdot\mathbf{r}_{m} \quad. \tag{5.132}$$

Vector $\mathbf{r}_{m} = \mathbf{y}_{m}-\mathbf{x}_{m}\left(\hat{\boldsymbol{\omega}}\right)$ denotes the residuals of the m-th MIMO channel after coherently subtracting the estimated paths using equation (5.20). As the parameters of the noise model are disjunct w.r.t. the MIMO channels, the minimum of the sum is given by the minimum of each summand. Hence, the parameters can be estimated for each MIMO channel separately.

$$\hat{\boldsymbol{\sigma}}_{m} = \operatorname*{arg\,min}_{\boldsymbol{\sigma}_{m}\in\Sigma_{m}}\left\{\mathcal{L}_{m}\left(\mathbf{y}_{m}\,|\,\boldsymbol{\sigma}_{m},\hat{\boldsymbol{\omega}}\right)\right\} \tag{5.133}$$

Summarised, because the MIMO channels are independent regarding their noise model parameters (disjunct parameter sets) and the noise covariance matrix is block diagonal w.r.t. the MIMO channels, the parameter estimation for all MIMO channels is independent.

Hence, instead of solving the $6M$-dimensional minimisation problem (5.130), M minimisation problems of dimensionality 6 are solved, which is numerically more efficient. Also, the M minimisations can be *run in parallel*, which reduces computational time. Subsequently, the minimisation w.r.t. a single MIMO channel will be discussed. Note, that the index m, indicating the considered MIMO channel, will be dropped in the following for notational convenience.

A closed-form solution of problem (5.133) is not possible, because the objective is non-convex. Hence, a two stage and iterative optimisation is employed. The minimisation method consists of an approach to estimate initial parameter values, which are refined by a gradient-based method. The presented minimisation method has to be applied to all M minimisation problems (all M MIMO channels).

Minimisation by Gradient Method

The objective (5.133) will be replaced by an approximation, whose minimisation is denoted as subproblem. Quite often the Taylor series expansion of the objective is considered as approximation. Since the approximation results only in suboptimal solution of (5.133), an iterative scheme is employed, which alternates between the approximation of the objective and the solution of the respective subproblem. Indicating the parameters in the i-th iteration as $\boldsymbol{\sigma}_i$, which is $\Delta\boldsymbol{\sigma}_i$ apart from the minimum, the second-order Taylor series expansion of the objective $\mathcal{L}_m\left(\mathbf{y}_m\middle|\boldsymbol{\sigma}_m,\hat{\boldsymbol{\omega}}\right)$ is [PP12] is

$$\mathcal{L}\left(\boldsymbol{\sigma}_i + \Delta\boldsymbol{\sigma}_i\right) \approx \mathcal{L}\left(\boldsymbol{\sigma}_i\right) + \mathbf{g}_i\left(\boldsymbol{\sigma}_i\right)^{\mathrm{T}} \cdot \Delta\boldsymbol{\sigma}_i + \frac{1}{2} \cdot \Delta\boldsymbol{\sigma}_i^{\mathrm{T}} \cdot \mathbf{H}_i\left(\boldsymbol{\sigma}_i\right) \cdot \Delta\boldsymbol{\sigma}_i \quad . \tag{5.134}$$

Vector $\mathbf{g}_i\left(\boldsymbol{\sigma}_i\right)$ is the *gradient vector* and matrix $\mathbf{H}_i\left(\boldsymbol{\sigma}_i\right)$ is the *Hessian matrix*, evaluated at the iteration point $\boldsymbol{\sigma}_i$. An estimate of $\Delta\boldsymbol{\sigma}_i$, i.e. the solution of the subproblem, is given by minimising function (5.134) w.r.t. $\Delta\boldsymbol{\sigma}_i$. This is a *quadratic programming problem*, which has has a minimum, if the Hessian matrix is positive definite. Subsequently, the iteration index i and the arguments will be dropped for convenience.

The entry g_x of the gradient vector $\mathbf{g}$ for the arbitrary parameter x is given by

$$g_x = \frac{\partial}{\partial x}\mathcal{L} = \mathrm{trace}\left\{\boldsymbol{\Sigma}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial x}\right\} - \mathbf{r}^{\mathrm{H}} \cdot \boldsymbol{\Sigma}^{\text{-}1} \cdot \frac{\partial\boldsymbol{\Sigma}}{\partial x} \cdot \boldsymbol{\Sigma}^{\text{-}1} \cdot \mathbf{r} \quad . \tag{5.135}$$

The entry $H_{x,y}$ of the Hessian matrix for the arbitrary parameters x and y is given by

$$H_{x,y} = \frac{\partial^2}{\partial x \partial y}\mathcal{L} = \mathrm{trace}\left\{\frac{\partial\boldsymbol{\Sigma}}{\partial x}\boldsymbol{\Sigma}^{\text{-}1} \cdot \frac{\partial\boldsymbol{\Sigma}}{\partial y}\boldsymbol{\Sigma}^{\text{-}1}\right\} = \mathrm{trace}\left\{\boldsymbol{\Sigma}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial y} \cdot \boldsymbol{\Sigma}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial x}\right\} \quad . \tag{5.136}$$

Consider the Cholesky factorisation (5.101) of the covariance matrix and plugging into (5.135) and (5.136) yields

$$g_x = \frac{\partial}{\partial x}\mathcal{L} = \mathrm{trace}\left\{\mathbf{L}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial x}\mathbf{L}^{\text{-H}}\right\} - \left(\mathbf{L}^{\text{-}1}\mathbf{r}\right)^{\mathrm{H}} \cdot \mathbf{L}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial x}\mathbf{L}^{\text{-H}} \cdot \left(\mathbf{L}^{\text{-}1}\mathbf{r}\right) \tag{5.137a}$$

$$H_{x,y} = \frac{\partial^2}{\partial x \partial y}\mathcal{L} = \mathrm{trace}\left\{\mathbf{L}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial y}\mathbf{L}^{\text{-H}} \cdot \mathbf{L}^{\text{-}1}\frac{\partial\boldsymbol{\Sigma}}{\partial x}\mathbf{L}^{\text{-H}}\right\} \quad . \tag{5.137b}$$

Note, that property[5] of the trace operator has been used for simplifications. Furthermore, the Cholesky factor $\mathbf{L}$ can be used to calculate the entropy, i.e., log-determinant of the covariance matrix.

$$\ln\left\{|\boldsymbol{\Sigma}|\right\} = 2 \cdot \sum_{k=1}^{K} \ln\left\{\left[\operatorname{diag}\left\{\mathbf{L}\right\}\right]_{(k)}\right\} \tag{5.138}$$

Plugging the Cholesky factorisation into objective (5.132) yields

$$\mathcal{L} = 2 \cdot \sum_{k=1}^{K} \ln\left\{\left[\operatorname{diag}\left\{\mathbf{L}\right\}\right]_{(k)}\right\} + \left(\mathbf{L}^{-1} \cdot \mathbf{r}\right)^{\mathrm{H}} \cdot \left(\mathbf{L}^{-1} \cdot \mathbf{r}\right) \quad . \tag{5.139}$$

Hence, knowing $\mathbf{L}^{-1}$, $\mathbf{L}^{-1} \cdot \frac{\partial \boldsymbol{\Sigma}}{\partial x} \cdot \mathbf{L}^{-\mathrm{H}}$ and $\mathbf{L}^{-1} \cdot \mathbf{r}$ is sufficient to evaluate the objective function, and to calculate the entries of the gradient vector and the Hessian matrix. Introduce

$$\mathbf{J}_x = \mathbf{L}^{-1} \cdot \frac{\partial \boldsymbol{\Sigma}}{\partial x} \cdot \mathbf{L}^{-\mathrm{H}} \tag{5.140a}$$

$$\mathbf{r}_{\mathrm{L}} = \mathbf{L}^{-1} \cdot \mathbf{r} \tag{5.140b}$$

for short hand notation, an entry of the Hessian matrix is given by

$$H_{x,y} = \operatorname{trace}\left\{\mathbf{J}_x \cdot \mathbf{J}_y\right\} = \operatorname{vec}\left\{\mathbf{J}_x^{\mathrm{T}}\right\}^{\mathrm{T}} \cdot \operatorname{vec}\left\{\mathbf{J}_y\right\} \quad . \tag{5.141}$$

Identity[6] has been used for simplification. The complete Hessian matrix can be calculated by

$$\mathbf{H} = \mathbf{J}'^{\mathrm{T}} \cdot \mathbf{J}'' \tag{5.142}$$

with

$$\mathbf{J}'\acute{\mathbf{J}}\grave{\mathbf{J}} = \left[\operatorname{vec}\left\{\mathbf{J}_{\varsigma}^{\mathrm{T}}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\varepsilon}^{\mathrm{T}}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\varrho}^{\mathrm{T}}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\alpha}^{\mathrm{T}}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\beta}^{\mathrm{T}}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\eta}^{\mathrm{T}}\right\}\right] \tag{5.143a}$$

$$\mathbf{J}'' = \left[\operatorname{vec}\left\{\mathbf{J}_{\varsigma}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\varepsilon}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\varrho}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\alpha}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\beta}\right\}, \operatorname{vec}\left\{\mathbf{J}_{\eta}\right\}\right] \quad . \tag{5.143b}$$

Similarly, the gradient vector $\mathbf{g}$ can be calculated.

$$\mathbf{g} = \mathbf{J}'^{\mathrm{T}} \cdot \operatorname{vec}\left\{\mathbf{I}_{N \cdot K} - \mathbf{r}_{\mathrm{L}} \cdot \mathbf{r}_{\mathrm{L}}^{\mathrm{H}}\right\} \tag{5.144}$$

Trust Region Algorithm for Bound Constraints The minimisation problem (5.133) is subject to the parameters $\boldsymbol{\sigma}_m$ being in the set $\Sigma_m \subset \mathbb{R}_+^6$. This set is bounded according to Tab. 5.1. These bounds have to be fulfilled in the gradient-based method. According to the parameter bounds, minimisation problem (5.133) can be restated as a box constrained minimisation problem (channel index m dropped for convenience).

$$\hat{\boldsymbol{\sigma}} = \arg\min_{\boldsymbol{\sigma}} \mathcal{L}\left(\boldsymbol{\sigma}\right) \quad \text{s.t.} \quad \mathbf{l} \leq \boldsymbol{\sigma} \leq \mathbf{u} \tag{5.145}$$

---

[5] $\operatorname{trace}\left\{\mathbf{A} \cdot \mathbf{B} \cdot \mathbf{C}\right\} = \operatorname{trace}\left\{\mathbf{B} \cdot \mathbf{C} \cdot \mathbf{A}\right\}$
[6] $\operatorname{trace}\left\{\mathbf{A} \cdot \mathbf{B}\right\} = \operatorname{vec}\left\{\mathbf{A}^{\mathrm{T}}\right\}^{\mathrm{T}} \cdot \operatorname{vec}\left\{\mathbf{B}\right\}$

Vector $\mathbf{l} \in \mathbb{R}^6_+$ and $\mathbf{u} \in \mathbb{R}^6_+$ are the vector of the lower and upper parameter bounds, respectively. The algorithm in [CL96] will be used to solve the box constrained minimisation problem, as it is numerically less costly than other competitive methods [CL96].

Be $\boldsymbol{\sigma}^\star$ a (local) minimum of $\mathcal{L}$. Then, the first-order necessary (but not sufficient) optimality conditions (also denoted as Karush–Kuhn–Tucker (KKT) conditions) for $\boldsymbol{\sigma}^\star$ to be a minimum have to be fulfilled. If no constraints are set, the first-order necessary conditions are

$$\left. \frac{\partial \mathcal{L}}{\partial \boldsymbol{\sigma}} \right|_{\boldsymbol{\sigma} = \boldsymbol{\sigma}^\star} = \mathbf{0} \quad .$$

Because box constraints are employed, the KKT conditions have to be modified. Be g_j the j-th entry of the gradient vector $\mathbf{g}$, σ_j the j-th entry of the parameter vector $\boldsymbol{\sigma}$, and u_j and l_j the j-th entries of the boundary vectors. Then, the KKT conditions are [CL94]

$$\left. \begin{aligned} &\text{if } l_j < \hat{\sigma}_j < u_j \text{ then } g_j\left(\sigma_j^\star\right) = 0 \\ &\text{if } \sigma_j^\star = u_j \text{ then } g_j\left(\sigma_j^\star\right) \leq 0 \\ &\text{if } \sigma_j^\star = l_j \text{ then } g_j\left(\sigma_j^\star\right) \geq 0 \end{aligned} \right\} \underset{j=1,\dots,6}{\Longrightarrow} \quad \mathbf{D}\left(\boldsymbol{\sigma}^\star\right) \cdot \mathbf{D}\left(\boldsymbol{\sigma}^\star\right) \cdot \mathbf{g}\left(\boldsymbol{\sigma}^\star\right) = \mathbf{0} \quad .$$

Considering the KKT conditions as additional constraints for the optimisation problem, a trust region update step $\Delta\boldsymbol{\sigma}_i$ for the i-th iteration can be defined. The respective subproblem is solved in a transformed variable space by a affine scaling transformation. The affine scaling transformation accounts for the additional constraints due to the KKT conditions.

$$\Delta\tilde{\boldsymbol{\sigma}}_i = \mathbf{D}\left(\boldsymbol{\sigma}_i\right)^{-1} \cdot \Delta\boldsymbol{\sigma}_i \tag{5.146}$$

The trust region update step in the transformed variable space is given by the following minimisation problem (arguments are dropped) [CL96].

$$\Delta\tilde{\boldsymbol{\sigma}}_i = \underset{\Delta\tilde{\boldsymbol{\sigma}}_i}{\arg\min} \left\{ \left[\mathbf{D}_i\mathbf{g}_i\right]^{\mathrm{T}} \Delta\tilde{\boldsymbol{\sigma}}_i + \frac{1}{2}\Delta\tilde{\boldsymbol{\sigma}}_i^{\mathrm{T}} \left[\mathbf{D}_i\mathbf{H}_i\mathbf{D}_i + \mathrm{diag}\left\{\mathbf{g}_i\right\}\mathbf{V}_i\right] \cdot \Delta\tilde{\boldsymbol{\sigma}}_i \right\} \text{ s.t. } \left\|\Delta\tilde{\boldsymbol{\sigma}}_i\right\| < \epsilon_i \tag{5.147}$$

The diagonal scaling matrix $\mathbf{D}_i = \mathbf{D}\left(\boldsymbol{\sigma}_i\right) = \mathrm{diag}\left\{\mathbf{d}\left(\boldsymbol{\sigma}_i\right)\right\}^{\frac{1}{2}}$ is related to the distance to the boundaries. Diagonal matrix $\mathbf{V}_i = \mathrm{diag}\left\{\mathbf{v}\left(\boldsymbol{\sigma}_i\right)\right\}$ is the Jacobian matrix of the constraint $\mathbf{D}_i \cdot \mathbf{D}_i \cdot \mathbf{g}_i = \mathbf{0}$, i.e. the Jacobian of vector $\mathbf{d}\left(\boldsymbol{\sigma}_i\right)$. The entries d_j and v_j of vector $\mathbf{d}\left(\boldsymbol{\sigma}\right)$ and $\mathbf{v}\left(\boldsymbol{\sigma}\right)$ are given by [CL96]

$$\text{if } g_j < 0 \text{ and } u_j < \infty \text{ then } d_j = \left|\sigma_j - u_j\right| \text{ and } v_j = 1 \tag{5.148a}$$

$$\text{if } g_j \geq 0 \text{ and } l_j > -\infty \text{ then } d_j = \left|\sigma_j - l_j\right| \text{ and } v_j = 1 \tag{5.148b}$$

$$\text{if } g_j < 0 \text{ and } u_j = \infty \text{ then } d_j = \left|-1\right| \text{ and } v_j = 0 \tag{5.148c}$$

$$\text{if } g_j \geq 0 \text{ and } l_j = -\infty \text{ then } d_j = \left|1\right| \text{ and } v_j = 0 \quad . \tag{5.148d}$$

Note, that $\mathbf{D}\left(\boldsymbol{\sigma}\right) \cdot \mathbf{D}\left(\boldsymbol{\sigma}\right) \cdot \mathbf{g}\left(\boldsymbol{\sigma}\right)$ is not differentiable if $d_j = 0$, i.e. the parameter attains the upper or lower bound. Non-differentiability can be avoided by ensuring strict feasibility, e.g. $\mathbf{l} < \boldsymbol{\sigma} < \mathbf{u}$.

Algorithm 3 Trust Region Algorithm with Box Constraints and Step Size Control by Levenberg Method

---

Require: $\boldsymbol{\sigma}_0, \mathbf{r}$
  $i \leftarrow 0$
  $\delta_i \leftarrow 0.1$
  $\mathcal{L}_i \leftarrow$ equation (5.139)
  repeat
    $\boldsymbol{\Sigma}_i \leftarrow$ equation (5.56)
    $\mathbf{L}_i \leftarrow$ equation (5.101)
    $\mathbf{g}_i \leftarrow$ equation (5.144)
    $\mathbf{H}_i \leftarrow$ equation (5.142)
    $\mathbf{D}_i, \mathbf{V}_i \leftarrow$ equation (5.148)
    $\boldsymbol{\sigma}_{i+1} \leftarrow \boldsymbol{\sigma}_i - \mathbf{D}_i \cdot \left[\mathbf{D}_i \cdot \mathbf{H}_i \cdot \mathbf{D}_i + \mathrm{diag}\left\{ \mathbf{g}_i \right\} \cdot \mathbf{V}_i + \delta_i \cdot \mathbf{I} \right]^{-1} \cdot \mathbf{D}_i \cdot \mathbf{g}_i$
    $\mathbf{L}_{i+1} \leftarrow$ equation (5.101)
    $\mathcal{L}_{i+1}$ equation (5.139)
    if $\mathcal{L}_{i+1} < \mathcal{L}_i$ **then**
      $\delta_{i+1} \leftarrow \delta_i / 10$
    else
      $\boldsymbol{\sigma}_{i+1} \leftarrow \boldsymbol{\sigma}_i$
      $\delta_{i+1} \leftarrow 10 \cdot \delta_i$
    end if
    $i \leftarrow i + 1$
  until *Convergence*

A solution of the trust region subproblem (5.147) is given by the Levenberg update step.

$$\Delta\tilde{\boldsymbol{\sigma}}_i = -\left[\mathbf{D}_i \mathbf{H}_i \mathbf{D}_i + \mathrm{diag}\left\{ \mathbf{g}_i \right\} \mathbf{V}_i + \delta_i \mathbf{I} \right]^{-1} \cdot \mathbf{D}_i \cdot \mathbf{g}_i \tag{5.149}$$

Parameter $\delta_i > 0$ is tuned to adjust the trust region size and hence the step length, and to adjust the direction [NW06]. After calculating the update step the affine scaling transformation has to be reverted to get the update step in the original variable space: $\Delta\boldsymbol{\sigma}_i = \mathbf{D}\left(\boldsymbol{\sigma}_i \right) \cdot \Delta\tilde{\boldsymbol{\sigma}}_i$.

The complete trust region algorithm for bound constraints is shown in Algorithm 3.

Calculation of the Partial Derivatives

Lets recap the model of the covariance matrix of the *m*-th MIMO channel from equation (5.56).

$$\boldsymbol{\Sigma}_m\left(\boldsymbol{\sigma}_m \right) = \mathbf{I}_N \otimes \left[\boldsymbol{\Sigma}_{1_m}^{(\tau)}\left(\varsigma_m, \varepsilon_m, \varrho_m \right) + \eta_m \mathbf{I}_K \right] + \left[\mathbf{I}_N \otimes \boldsymbol{\Sigma}_{\mathrm{e}_m}^{(\tau)}\left(\alpha_m, \beta_m \right) \right] \odot \sum_{p=1}^{P} \mathbf{x}_m\left(\boldsymbol{\varpi}_p \right) \mathbf{x}_m\left(\boldsymbol{\varpi}_p \right)^{\mathrm{H}}$$

For the calculation of the gradient vector and the Hessian matrix, the partial derivatives of the covariance model w.r.t. the parameter vector $\boldsymbol{\sigma}_m$ are required.

Derivative w.r.t. the Thermal Noise Parameters According to the sum and product rule for derivatives, the partial derivative w.r.t. the thermal noise power is

$$\frac{\partial}{\partial \eta_m} \boldsymbol{\Sigma}_m \left(\boldsymbol{\sigma}_m \right) = \mathbf{I}_N \otimes \mathbf{I}_K \quad . \tag{5.150}$$

Derivative w.r.t. the Leakage Noise Parameters According to the sum and product rule for derivatives, the partial derivative w.r.t. a parameter of the leakage noise is

$$\frac{\partial}{\partial x} \boldsymbol{\Sigma}_m \left(\boldsymbol{\sigma}_m \right) = \mathbf{I}_N \otimes \frac{\partial}{\partial x} \boldsymbol{\Sigma}_{l_m}^{(\tau)} \left(\varsigma_m, \varepsilon_m, \varrho_m \right) \quad , \tag{5.151}$$

with $x \in \{\varsigma_m, \varepsilon_m, \varrho_m\}$. Matrix $\boldsymbol{\Sigma}_{l_m}^{(\tau)}$ is constructed by the Toeplitz operator and the vector $\boldsymbol{\psi}_l$ of the sampled leakage noise ACF, see equation (5.37). Considering the product rule the partial derivate is

$$\frac{\partial}{\partial x} \boldsymbol{\Sigma}_{l_m}^{(\tau)} = \text{diag} \left\{ \mathbf{g}_m \right\} \cdot \mathrm{T} \left\{ \frac{\partial}{\partial x} \boldsymbol{\psi}_l \left(\varsigma_m, \varepsilon_m, \varrho_m \right) \right\} \cdot \text{diag} \left\{ \mathbf{g}_m \right\}^{\mathrm{H}} \quad . \tag{5.152}$$

The derivative of vector $\boldsymbol{\psi}_l$ is given by sampling the respective derivative of the ACF at delay points $\tau = 0, \ldots, (K-1) \cdot T_{\mathrm{S}}$.

$$\frac{\partial}{\partial \varsigma_m} \boldsymbol{\psi}_l \left(\varsigma_m, \varepsilon_m, \varrho_m \right) = \left[\frac{1}{\varepsilon_m}, \ldots, \frac{\exp\{j2\pi\varrho_m(K-1)\}}{\varepsilon_m - j2\pi(K-1)} \right]^{\mathrm{T}} \tag{5.153a}$$

$$\frac{\partial}{\partial \varepsilon_m} \boldsymbol{\psi}_l \left(\varsigma_m, \varepsilon_m, \varrho_m \right) = \varsigma_m \cdot \left[-\frac{1}{\varepsilon_m^2}, \ldots, -\frac{\exp\{j2\pi\varrho_m(K-1)\}}{(\varepsilon_m - j2\pi(K-1))^2} \right]^{\mathrm{T}} \tag{5.153b}$$

$$\frac{\partial}{\partial \varrho_m} \boldsymbol{\psi}_l \left(\varsigma_m, \varepsilon_m, \varrho_m \right) = \varsigma_m \cdot \left[0, \ldots, j2\pi(K-1) \cdot \frac{\exp\{j2\pi\varrho_m(K-1)\}}{\varepsilon_m - j2\pi(K-1)} \right]^{\mathrm{T}} \tag{5.153c}$$

Derivative w.r.t. the Phase Noise Parameters Considering the sum and product rule, the partial derivatives w.r.t. the parameters of the phase noise model are

$$\frac{\partial}{\partial x} \boldsymbol{\Sigma}_m \left(\boldsymbol{\sigma}_m \right) = \left(\mathbf{I}_N \otimes \frac{\partial}{\partial x} \boldsymbol{\Sigma}_{e_m}^{(\tau)} \left(\alpha_m, \beta_m \right) \right) \odot \sum_{p=1}^{P} \mathbf{x}_m \left(\boldsymbol{\varpi}_p \right) \cdot \mathbf{x}_m \left(\boldsymbol{\varpi}_p \right)^{\mathrm{H}} \quad , \tag{5.154}$$

with $x \in \{\alpha_m, \beta_m\}$. The partial derivative of the Toeplitz matrix $\boldsymbol{\Sigma}_{e_{m,n}}^{(\tau)}$ is

$$\frac{\partial}{\partial x} \boldsymbol{\Sigma}_{e_{m,n}}^{(\tau)} = \mathrm{T} \left\{ \frac{\partial}{\partial x} \boldsymbol{\psi}_e \left(\alpha_m, \beta_m \right) \right\} \quad . \tag{5.155}$$

The derivative of vector $\boldsymbol{\psi}_e$ is given by sampling the respective derivative of the ACF (4.142) at the delay points $\tau = 0, \ldots, (K-1) \cdot T_{\mathrm{S}}$.

$$\frac{\partial}{\partial \alpha_m} \boldsymbol{\psi}_e \left(\alpha_m, \beta_m \right) = \left[1, \ldots, \exp\left\{ -(K-1) \cdot \beta_m \right\} \right]^{\mathrm{T}} \tag{5.156a}$$

$$\frac{\partial}{\partial \beta_m} \boldsymbol{\psi}_e \left(\alpha_m, \beta_m \right) = \alpha_m \cdot \left[0, \ldots, -(K-1) \cdot \exp\left\{ -(K-1) \cdot \beta_m \right\} \right]^{\mathrm{T}} \tag{5.156b}$$

Estimation of Initial Parameter Values

The gradient-based optimisation requires initial estimates $\boldsymbol{\sigma}_0$. These estimates have to be close to the global minimum, such that the gradient-based optimisation does not end up in a local minimum.

Leakage Noise and Thermal Noise Parameters In order to estimate initial values of the parameters of the leakage and thermal noise model, a suboptimal but numerically efficient estimator will be employed.

First, the vector $\hat{\mathbf{z}}_m \in \mathbb{R}^K$ of an estimate of the PSD of the m-th MIMO channel is introduced. As thermal noise and leakage noise are assumed as i.i.d., the PSD can be averaged over the captured signal periods N.

$$\hat{\mathbf{z}}_m = \frac{1}{N} \sum_{n=1}^{N} \left(\mathbf{E} \cdot \mathbf{r}_{m,n} \right)^{\dagger} \odot \left(\mathbf{E} \cdot \mathbf{r}_{m,n} \right) \tag{5.157}$$

Matrix $\mathbf{E} \in \mathbb{C}^{K \times K}$ denotes the discrete Fourier transform (DFT) matrix, and $\mathbf{r}_{m,n} \in \mathbb{C}^K$ are the residuals at the m-th MIMO channel for the n-th signal period. If no previous estimates $\hat{\boldsymbol{\omega}}$ of the path parameters are available, then $\mathbf{r}_{m,n} = \mathbf{y}_{m,n}$. Note, that $\hat{\mathbf{z}}_m$ is an estimate of the PSD of the noise processes, if all paths have been coherently subtracted. Now recall, that both noise processes are assumed as WSS. Hence, the covariance matrix of the leakage noise $\boldsymbol{\Sigma}_{\mathrm{l}}^{(\tau)}$ and the covariance matrix of the thermal noise $\boldsymbol{\Sigma}_{\mathrm{w}}^{(\tau)}$ are diagonal matrices in the frequency domain. However, due to the limited observation bandwidth of the measurements, this is only approximately true. Furthermore, thermal noise and leakage noise are assumed as uncorrelated with each other. Hence, the sum of the diagonal entries are a model $\mathbf{z}_m$ of the PSD of the noise processes.

$$\mathbf{z}_m = \mathrm{diag}\left\{ \mathbf{E} \cdot \boldsymbol{\Sigma}_{\mathrm{l}} \left(\varsigma_m, \varepsilon_m, \varrho_m \right) \cdot \mathbf{E}^{\mathrm{H}} \right\} + \mathrm{diag}\left\{ \mathbf{E} \cdot \boldsymbol{\Sigma}_{\mathrm{w}} \left(\eta_m \right) \cdot \mathbf{E}^{\mathrm{H}} \right\} \tag{5.158}$$

Assume, that the frequency shift parameter is approx. $\varrho_m = 0$, because the leakage noise is commonly not significantly shifted. Then, a model for the PSD of the noise processes can be given.

$$\mathbf{z}_m = \mathbf{W} \left(\varepsilon_m \right) \cdot \left[\varsigma_m, \eta_m \right]^{\mathrm{T}} \tag{5.159}$$

Matrix $\mathbf{W} \left(\varepsilon_m \right)$ contains the normalised PSD of the leakage noise and thermal noise process.

$$\mathbf{W} \left(\varepsilon_m \right) = \left[\mathrm{diag}\left\{ \mathbf{E} \, \mathrm{diag}\left\{ \mathbf{g}_m \right\} \cdot \mathrm{T} \left\{ \exp\left\{ -\ln\left\{ \varepsilon_m \mathbf{1}_K - \jmath 2\pi \mathbf{k} \right\} \right\} \right\} \cdot \mathrm{diag}\left\{ \mathbf{g}_m \right\}^{\mathrm{H}} \mathbf{E}^{\mathrm{H}} \right\}, \mathbf{1}_K \right] \tag{5.160}$$

Now, a suboptimal estimator, i.e. not the ML estimator, of the model parameters can be given by the least-squares of the estimated PSD $\hat{\mathbf{z}}_m$ and its respective model (5.159).

$$\hat{\varsigma}_m, \hat{\varepsilon}_m, \hat{\eta}_m = \underset{\varsigma_m, \varepsilon_m, \eta_m}{\arg\min} \left\{ \left\| \hat{\mathbf{z}}_m - \mathbf{W} \left(\varepsilon_m \right) \cdot \left[\varsigma_m, \eta_m \right]^{\mathrm{T}} \right\|^2 \right\} \tag{5.161}$$
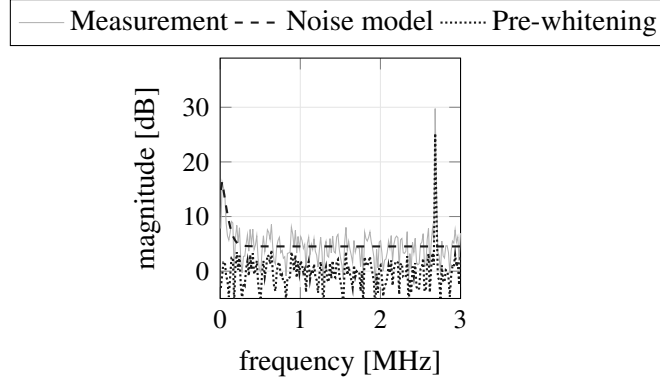
Figure 5.2: *Average range spectrum, the initially estimated PSD of the leakage and thermal noise, and the range spectrum after whitening with the estimated noise model.*

Because the parameters ς_m and η_m are linearly included in the objective (5.161), the variable projection method can be employed. An estimator of the parameters ς_m and η_m is the BLUE.

$$\left[\hat{\varsigma}_m, \hat{\eta}_m\right]^{\mathrm{T}} = \mathbf{W}\left(\varepsilon_m\right)^+ \cdot \hat{\mathbf{z}}_m \tag{5.162}$$

Plugging equation (5.162) into equation (5.161) and applying some mathematical manipulations, the profiled objective function, which depends on the parameter ε_m only, can be stated.

$$\hat{\varepsilon}_m = \arg\max_{\varepsilon_m} \hat{\mathbf{z}}_m^{\mathrm{T}} \cdot \mathbf{W}\left(\varepsilon_m\right) \cdot \mathbf{W}\left(\varepsilon_m\right)^+ \cdot \hat{\mathbf{z}}_m \tag{5.163}$$

Maximisation of objective (5.163) is done by a 1-dimensional search over a finite grid. Afterwards, the noise level η_m and the amplitude ς_m is estimated using equation (5.162).

Figure 5.2 shows an example measurement and the initially estimated model of the leakage and thermal noise using the stated method. The estimated noise model has been used to pre-whiten the measurement. As the noise model is quite properly estimated, the pre-whitening result shows a uniform noise floor.

Phase Noise Parameters Because the covariance model of the phase noise depends on parameters of the mean model (path parameters), initial estimates of the parameters of the phase noise model cannot be drawn until one path has been detected. Hence, the parameters α_m and β_m of the phase noise model are initialised by heuristic experiences. From multiple estimation trials it turned out, that the initial values $\alpha_m = 0$ and $\beta_m = 0$ are proper choices.

The complete algorithm to initialise the parameters of the covariance model is given in Algorithm 4.

Algorithm 4 Initialisation of Covariance Model Parameters

Require: $\mathcal{Y}, \hat{\varpi}$

  Define grid: $\quad \varepsilon_1, \ldots, \varepsilon_H$

  for $m = 1 \ldots M$ **do**

    $\hat{\varrho}_m \leftarrow 0$

    $\hat{\alpha}_m \leftarrow 0$

    $\hat{\beta}_m \leftarrow 0$

    $\mathbf{r}_{m,n} \leftarrow \mathbf{y}_{m,n} - \mathbf{x}_{m,n}(\hat{\varpi})$

    $\hat{\mathbf{z}}_m = \frac{1}{N} \sum_{n=1}^{N} \left(\mathbf{E} \cdot \mathbf{r}_{m,n}\right)^{\dagger} \odot \left(\mathbf{E} \cdot \mathbf{r}_{m,n}\right)$

    $\hat{\varepsilon}_m \leftarrow \underset{\varepsilon_h = \varepsilon_1, \ldots, \varepsilon_H}{\arg\max} \ \hat{\mathbf{z}}_m^{\mathrm{T}} \cdot \mathbf{W}(\varepsilon_h) \cdot \mathbf{W}(\varepsilon_h)^{+} \cdot \hat{\mathbf{z}}_m$

    $\left[\hat{\varsigma}_m, \hat{\eta}_m\right]^{\mathrm{T}} \leftarrow$ equation (5.162)

    $\hat{\boldsymbol{\sigma}}_m \leftarrow \left[\hat{\varsigma}_m, \hat{\varepsilon}_m, \hat{\varrho}_m, \hat{\alpha}_m, \hat{\beta}_m, \hat{\eta}_m\right]^{\mathrm{T}}$

  end for

5.4 Outline of the Algorithm for Parameter Estimation

The state diagram of the proposed optimisation and parameter estimator algorithm is shown in Fig. 5.3. As the simplification in equation (5.100) requires a proper guess of the noise parameters, an initial estimate of the noise parameters is deduced first. Afterwards, the iterative path detection and parameter estimates refinement is run. In each run a new propagation path is estimated first by the path search and path weight estimation method. Subsequently, the parameters of the yet detected paths (i.e. the mean model parameters) are jointly refined by the gradient-based optimisation. After refining the estimates of the mean model parameters ϖ, the estimates of the covariance model parameters σ are refined by the respective gradient-based approach. The algorithm continues as long as the selected number of propagation paths $\hat{P}$ has not been estimated.

5.5 Summary

The developed parameter estimator for a broadband FMCW MIMO radar has been presented in this chapter. The parameter estimator is a maximum-likelihood estimator with box constraints, resulting from the consideration of the MAP as point estimator and a prior derived from the principle of indifference.

    The ML estimator exploits a statistical model of the noisy observations, which is a multidimensional normal distribution featuring a parametric model of its mean and covariance. The mean model is given by the system model including the parameters of interest. The system model takes into account angle-Doppler and delay-Doppler coupling. By considering this system model in the parameter estimator, the angle-Doppler and delay-Doppler coupling are coped. The statistical model includes noise sources as leakage noise and phase noise, which have not been considered in signal processing for FMCW MIMO radars so far. As a result, the influence of these noises is tackled, which improves detection robustness against these noises
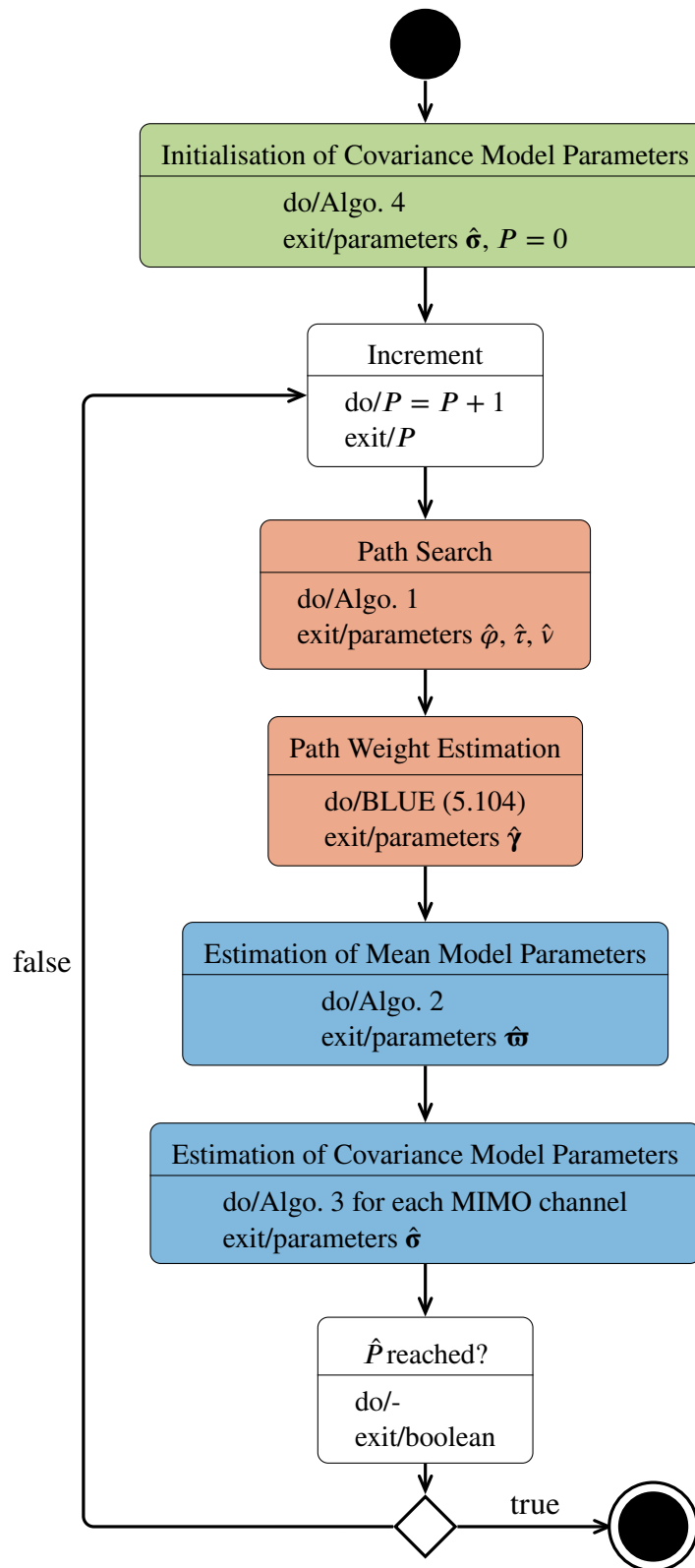

Figure 5.3: *State diagram of the optimisation algorithm to minimise the objective and to estimate the model parameters. The states of the initialisation approaches for the covariance model parameters and the mean model parameters are coloured in green and red, respectively, and the states of the joint optimisation approaches are coloured in blue.*

and prevents an overestimation of the number of targets.

The ML estimation requires the global optimisation of an objective, which turned out to be non-convex. A two-stage and iterative optimisation scheme has been proposed to detect the global optimum and hence get the parameter estimates. In the first stage a new path is detected and the parameters of the parameters of the mean model are refined by a gradient-based optimisation. In the second stage the parameters of the model of the covariance are refined by a gradient-based optimisation.

Because an iterative approach is used, whereas the number of iterations to attain a termination criterion are unknown and may vary, the optimisation algorithm is computational cumbersome. For example, in each iteration of the gradient optimisation, the respective model, being the mean or covariance model, has to be evaluated. Furthermore, the respective Jacobian matrix has to be evaluated in each iteration. These function evaluations are computational costly, because the respective functions are non-linear and multidimensional. It has to be pointed out, that the gradient-based optimisation of the covariance model parameters are the numerically most costly part of the optimisation algorithm. First, an optimisation has to be run for each MIMO channel, which are 64 for the considered radar. Second, calculation of the required Cholesky factorisation of the covariance matrix and its inverse (see Algorithm 3) are the numerically very costly. A numerical more efficient method, which circumvents the Cholesky factorisation or calculation of its inverse, could not be found. The basic problem is, that the covariance matrix does not feature structural properties as e.g. a Toeplitz structure, because of the phase noise covariance matrix. Hence, properties of Toeplitz matrices, as applied in [Ric05] to reduce the numerical effort, cannot be exploited. Consequently, calculation of the Cholesky factorisation and its inverse cannot be circumvented. Hence, the optimisation is computationally very complex and time consuming. As a result of the high computational burden a real-time application of the estimator is not possible.

MODEL ORDER ESTIMATION – DETERMINATION OF THE NUMBER OF PROPAGATION PATHS

SO far, the statistical model has been assumed as specified for the estimation of the respective model parameters $\boldsymbol{\theta}$. However, the model is not completely specified in advance, because the number of resolvable point targets/propagation paths P is unknown in advance. Hence, this number has to be determined too, taking into account the actual observations. Selection of the model has a severe influence on the parameter estimates, e.g. on their variance or bias [BA02, SMS02]. In this chapter, the estimation of the number of point targets/propagation paths is discussed.

Related Work and Own Contributions

In statistical inference, specification of the statistical model is a crucial before to the model parameter estimation and assignment of precisions [BA02]. The model specification can be split into the formulation of a *set of candidate models* $\{\mathcal{M}_1, \mathcal{M}_2, \dots\}$ and the *selection of the model* which most properly describes/predicts the underlying observations [BA02]. For the problem under consideration the model structure (i.e. the model family) is predetermined. Candidate models only differ in their *model order*, such that the set of candidate models composes of models of different order, i.e. $0, 1, \dots, P_{\max}$ and $\left\{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_{P_{\max}}\right\}$. The model order is the a-priori unknown number of resolvable point targets/propagation paths P. Hence, by selecting the best suited model $\mathcal{M}_{\hat{P}}$ from the set of candidate models, the best suited model order $\hat{P}$ (number of resolvable point targets) is inherently determined. As model $\mathcal{M}_P$ is a subset of model $\mathcal{M}_{P+1}$ in terms of the model parameters, the candidate models are so called *nested hierarchical models*.

$$\mathcal{M}_0 \subset \mathcal{M}_1 \subset \dots \subset \mathcal{M}_{P_{\max}}$$

Hence, the subsequently considered task is to select the best suited model from a set of nested, hierarchical candidate models. Methods for model order selection are widely dealt with in literature, see e.g. for an overview of some methods [Cos+09, SS04, Wal+05].

Some model order estimation methods determine the model order *before* the parameter

estimation step. Quite often, the eigenvalues of the data covariance matrix are exploited therefore [Cos+09, Qui+06]. This is advantageously, because the costly parameter estimation step is conducted only once. However, these methods need a certain sample size in order to properly estimate the data covariance matrix. Hence, these methods are not applicable in the single snapshot case. Also, they become less powerful in case of correlated sources [SN89]. Furthermore, these methods do not take into account the model structure or results of the parameter estimation, which can be beneficial to properly select the model order.

Other model selection schemes determine the best suited model by comparing the estimation results of all models in a set of candidate models. Such methods are in contrast to the previously stated approaches, because the best suited model is determined *after* the parameter estimation. A well known representative of such a model selection scheme is the stepwise regression search procedure. This procedure composes of

- A strategy to search for candidate models

- A criterion to compare candidate models.

After a candidate model has been discovered the respective model parameters are estimated. Afterwards, the model is compared to all yet discovered candidate models (whose parameters have been previously estimated). It is obvious, that this approach requires multiple estimation runs, such that this kind of model selection is computationally intensive. Therefore, a proper search strategy is required, in order to keep the number of candidate models and hence the number of estimation runs low. A proper search strategy for candidate models is to change the model in each search step based on a specified criterion. For the problem at hand only the model order has to be adjusted to get a new candidate model, such that two "search directions" can be distinguished: *step up* and *step down*. In the step up case the order is gradually increased ($\mathcal{M}_P \rightarrow \mathcal{M}_{P+1}$), whereas in the step down case the order is gradually decreased ($\mathcal{M}_P \rightarrow \mathcal{M}_{P+1}$). Step up requires the predefinition of a lower bound and step down requires the predefinition of an upper bound of the model order. A lower bound can be easily given by 0, i.e. no target/path is present, whereas an upper bound is hard to define. Hence, the step up strategy will be considered. The gradually increase of the model order continues, unless a stopping rule, which is defined from the comparison of candidate models, is fulfilled [Chu+07, Ric05]. As a consequence of this approach, the model with the lowest order is preferred against all models of higher order. In statistics, preferring a model of lower complexity (lower order) against all models of higher complexity (higher order) is known as principle of *Occam's razor*, which will be applied here.

As a criterion to compare candidate models, sometimes denoted as *goodness-of-fit criterion*, information theoretic criteria like Akaike information criterion [BA02], Takeuchi's information criterion [BA02], Bayesian information criterion [SS04] or minimum description length [WK85] are mostly employed. These criteria provide a figure-of-merit of each estimated candidate model. Hence, the best suited model is given by maximising/minimising this criteria. Information theoretic criteria tend to over- or under-estimate the model order in low or even high SNR scenarios [Qui+06, Won+90]. Also, they do not take into account knowledge on the model structure.

Another class of approaches to compare candidate models is hypothesis testing [BA02, Chu+07]. Recap, that candidate models are nested hierarchical for the problem under consideration. Hence, one can get $\mathcal{M}_P$ out of the models $\mathcal{M}_{P+1}, \mathcal{M}_{P+2}, \ldots$ by stating

restrictions on the parameters of the models of higher order. By validating these restrictions the models are compared and the model structure is inherently exploited for model order estimation. Parametric hypothesis tests on the estimated model parameters can be used for the validation of the restrictions. The authors in [WHG+02] propose a hypothesis test of the estimated path weights. Similar approaches can be found in [Ric05] and [SRK09]. These authors do not discuss the problem of having a non-linear statistical model, which causes the test statistic to not attain a theoretically assumed distribution. Nor they account for the problem of multi-hypothesis testing, which occurs in stepwise regression search. In [Chu+07] a solution for the multi-hypothesis testing problem is presented. Also, an approach to handle the problem of an unknown distribution of the test statistic is presented. However, this approach does not cover for the problem of having a non-linear statistical model.

In this chapter, a novel approach to estimate the number of point targets/propagation paths using step up regression search in conjunction with parametric hypothesis testing is presented. First, fundamentals of parametric hypothesis tests will be recapped in Section 6.1. Afterwards, the recapped hypothesis tests are adapted for the considered model order selection in Section 6.2. Their performance will be verified by Monte Carlo simulations. Because theoretical assumptions necessary to apply the hypothesis tests are violated, the test performance and hence the model order selection performance is low. Improvements to circumvent the problem of violated theoretical assumptions are proposed in Section 6.2, and Monte Carlo simulations are carried out to verify the improved tests. The model order estimation algorithm is outlined in Section 6.4, and the incorporation of the model order determination in the parameter estimation framework from Section 5.4 is presented. Section 6.5 concludes the chapter.

Some aspects of the presented model order estimation scheme have been used in the own publications [22, 25].

6.1 Parametric Hypothesis Testing

In parametric hypothesis testing statements on the parameters of the statistical model are drawn and validated. For example, a hypothesis on the ML parameter estimates $\hat{\boldsymbol{\theta}}$ is drawn and checked concerning their reasonableness. This hypothesis is called *null hypothesis*, symbolised by $\mathcal{H}_0$. Also, an *alternative hypothesis* $\mathcal{H}_1$ is stated, which opposites the null hypothesis. Say, that the ML parameter estimates under the null hypothesis $\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}$ attain certain values $\tilde{\boldsymbol{\theta}}$. The opposite statement would be, that the ML parameter estimates under the alternative hypothesis $\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}$ do not attain these values. The hypothesis test now verifies, whether $\mathcal{H}_1$ holds or whether $\mathcal{H}_0$ cannot be rejected.

$$\mathcal{H}_0 : \hat{\boldsymbol{\theta}}_{\mathcal{H}_0} = \tilde{\boldsymbol{\theta}} \tag{6.1a}$$

$$\mathcal{H}_1 : \hat{\boldsymbol{\theta}}_{\mathcal{H}_1} \neq \tilde{\boldsymbol{\theta}} \quad , \tag{6.1b}$$

As the null hypothesis states restrictions on the parameters of the statistical model, the statistical model under the null hypothesis is referred to as *restricted model*. Accordingly, the model under the alternative hypothesis is stated as *unrestricted model*. Prominent restrictions on the parameter estimates $\hat{\boldsymbol{\theta}}$ are linear restrictions or linear hypothesises. Introducing matrix $\mathbf{R}$, which

describes the superposition or selection of parameter values, the $\mathcal{H}_0$ and $\mathcal{H}_1$ hypothesis are

$$\mathcal{H}_0 \,:\, \mathbf{R} \cdot \hat{\boldsymbol{\theta}} = \mathbf{h} \tag{6.2a}$$

$$\mathcal{H}_1 \,:\, \mathbf{R} \cdot \hat{\boldsymbol{\theta}} \neq \mathbf{h} \quad . \tag{6.2b}$$

Based on the null hypothesis a test statistic or T-value t_{val} is calculated, which follows a certain distribution under the null hypothesis (denoted as *null distribution*). By calculating the p-value p_{val} or the critical value c_{val} the significance of the null hypothesis is verified w.r.t. the *significance level* α_{s}. The value p_{val} (also termed as *observed significance*) is the likelihood of the test statistic under the null distribution.

$$p_{\text{val}} = 1 - CDF_{\mathcal{H}_0}\left(t_{\text{val}}\right) \tag{6.3}$$

The value c_{val} is the distribution value for a probability given by the significance level.

$$c_{\text{val}} = iCDF_{\mathcal{H}_0}\left(1 - \alpha_{\text{s}}\right) \tag{6.4}$$

Here, $CDF_{\mathcal{H}_0}$ and $iCDF_{\mathcal{H}_0}$ denote the cumulative distribution function of the null distribution, and their respective inverse. Based on the p_{val} or the c_{val} it is decided, whether the null hypothesis can be rejected or not.

$$\alpha_{\text{s}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} p_{\text{val}} \quad \text{or} \quad t_{\text{val}} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} c_{\text{val}} \tag{6.5}$$

6.1.1 Popular Test Statistics

Three well known tests, which measure the discrepancy of the parameter estimates under the null and alternative hypothesis, are the Wald test, the Rao's score test (or Lagrangian multiplier test) and the likelihood-ratio test [Eng84, Kay98]. Figure 6.1 shows the measures, which are used by these tests to compare the parameter estimates under the null hypothesis and the alternative hypothesis. Comparing the Wald and Rao test w.r.t. their test statistic, both tests focus a different question in view of the problem under consideration. Basically, the Rao test calculates the test statistic using the parameter estimates under the null hypothesis and the Wald test calculates the test statistic using the parameters under the alternative hypothesis [Eng84]. Besides the Wald and Rao test, the likelihood-ratio test incorporates parameter estimates under the null and alternative-hypothesis. Other conceivable tests are the gradient test [Ter02] or the residual-ratio test (modified likelihood-ratio test) [Chu+07]. Furthermore, the F-test [Kay98] is well known for comparing nested models, whose parameters have been inferred by an ML estimator. Subsequently, only the mentioned Wald, Rao and likelihood-ratio test will be considered.

Wald Test

The Wald test verifies, how significant the distance between parameter values under the null hypothesis and parameter values under the alternative hypothesis is. Hence, the parameters $\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}$
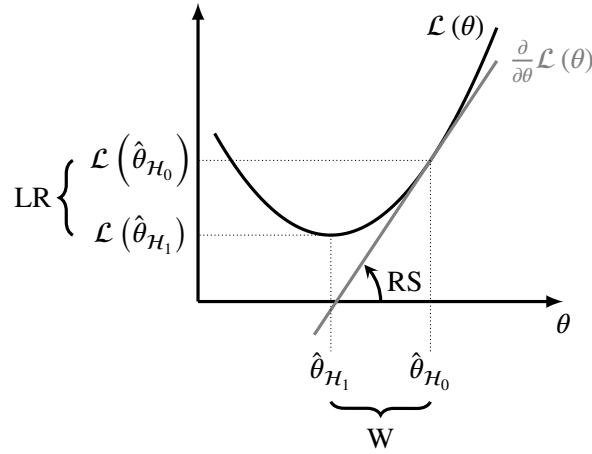
Figure 6.1: *Log-likelihood function $\mathcal{L}(\theta)$ and the measures of the Wald test (W), Rao's score test (RS) and the likelihood-ratio test (LR) to compare the ML parameter estimates under the null hypothesis $\hat{\theta}_{\mathcal{H}_0}$ (estimates for the restricted model) and the alternative hypothesis $\hat{\theta}_{\mathcal{H}_1}$ (estimates for the unrestricted model).*

of the unrestricted model has to be estimated from the observations. The test statistic is [Kay98]

$$t_{\mathrm{W}} = \left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_1} - \tilde{\boldsymbol{\theta}} \right)^{\mathrm{T}} \cdot \mathbf{F}\left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_1} \right) \cdot \left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_1} - \tilde{\boldsymbol{\theta}} \right) \quad , \tag{6.6}$$

with $\mathbf{F}\left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_1} \right)$ denoting the evaluated FIM (see Section 5.2.2) of the unrestricted model. If linear restrictions on the ML estimates shall be tested, the Wald test statistic is

$$t_{\mathrm{W}} = \left[\mathbf{R}\hat{\boldsymbol{\theta}} - \mathbf{h} \right]^{\mathrm{T}} \cdot \left[\mathbf{R}\mathbf{F}\left(\hat{\boldsymbol{\theta}} \right)^{-1} \mathbf{R}^{\mathrm{T}} \right]^{-1} \cdot \left[\mathbf{R}\hat{\boldsymbol{\theta}} - \mathbf{h} \right] \quad . \tag{6.7}$$

Rao's Score Test

The Rao test assumes, that the score function $\mathbf{g}(\boldsymbol{\theta})$ (gradient of the log-likelihood function) evaluated at the parameter estimates is zero, because the estimates are (at least) at a local minimum of the log-likelihood function (first-order optimality condition). The Rao test verifies, how significant the score function of the unrestricted model, evaluated by the parameter estimates $\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}$ of the restricted model, differs from zero. Hence, the restricted model has to be fitted to the data. The test statistic is [Kay98]

$$t_{\mathrm{RS}} = \mathbf{g}\left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_0} \right)^{\mathrm{T}} \cdot \mathbf{F}\left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_0} \right)^{-1} \cdot \mathbf{g}\left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_0} \right) \tag{6.8}$$

with $\mathbf{F}\left(\hat{\boldsymbol{\theta}}_{\mathcal{H}_0} \right)$ denoting the FIM evaluated for the parameter estimates $\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}$ under the restricted model.

Likelihood-Ratio Test

The likelihood-ratio test verifies, how significant the log-likelihood function under the null hypothesis $\mathcal{L}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}\right)$ differs from the log-likelihood under the alternative hypothesis $\mathcal{L}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}\right)$. Therefore, the restricted and the unrestricted model has to be fitted to the data. The test statistic is [Kay98]

$$t_{\mathrm{LR}} = 2\left[\mathcal{L}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}\right) - \mathcal{L}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}\right)\right] \quad . \tag{6.9}$$

Hence it is tested, whether a significant improvement of the log-likelihood is achieved by considering the unrestricted model.

6.1.2 Distribution of the T-Value Under the Null Hypothesis

The test statistics of the presented Wald, Rao and likelihood-ratio test are assumed to be asymptotic χ_L^2(Chi-quadrate)-distributed under $\mathcal{H}_0$, with L degrees of freedom. The degrees of freedom correspond to the number of restrictions under the null hypothesis. The test statistic distribution under the null hypothesis (null distribution) is commonly derived from first order approximations under asymptotic (large sample size) conditions [CC96]. Hence, the assumed null distribution is only an approximation of the true one. For example, the considered Wald and likelihood-ratio test are asymptotically similar w.r.t. their test statistic distribution [LF12, Ter02], if the log-likelihood function is well approximated by a quadratic function [Bus82]. However, different decisions are drawn by these tests in the finite-sample (low SNR) case, because the null distribution varies from the asymptotically one [Par17].

6.1.3 Multiple Hypothesis Testing

Because multiple candidate models are compared by hypothesis testing, multiple hypothesis tests are conducted. Multiple hypotheses tests on the same data set increase the type I error or false positive decision, which is the likelihood of rejecting a true null hypothesis [Kay98]. On the other hand, a false negative or type II error is committed, if a false null hypothesis is not rejected [Kay98].

Common metrics of describing the rate of type I errors are the *family-wise error rate (FWER)* and the *false discovery rate (FDR)*. The FWER is the probability of committing at least one type I error, whereas the FDR is the ratio of falsely rejected null hypothesises among the total number of rejected null hypothesises [Chu+07, Yoa10]. The FDR controlling methods have greater statistical power but an increased number of type I errors. The FWER is given by the significance level α_{s}, whereas the family-wise error rate for P tests is: $FWER = 1 - (1 - \alpha_{\mathrm{s}})^P$ for independent tests and $FWER \leq P \cdot \alpha_{\mathrm{s}}$ for dependent tests. Methods to control the FWER can be divided in the group of single step procedures and step down procedures. Single step procedures use the same criterion for all tests. Step down procedure apply different criteria for each test by priorly sorting the calculated p_{val}.

Several methods and concepts are known to control the FWER and the FDR. To control the FWER the Bonferroni correction (dependent tests) or the Šidak correction (independent tests) or the Holm–Bonferroni method can be applied [ADA14]. The Bonferroni and Šidak correction are single step procedures, whereas the Holm–Bonferroni method is a step down procedure. To control the FDR, methods like Benjamini-Hochberg (independent tests) or Benjamini-Yekutieli (dependent tests) procedure are known [ADA14]. The Benjamini-Hochberg procedure is uniformly more powerful than the Holm-Bonferroni procedure. However, the Benjamini-Hochberg procedure requires the hypotheses to be independent or under certain forms of positive dependence, whereas Holm–Bonferroni can be applied without such assumptions [Chu+07].

6.2 Considered Hypothesis Testing Approaches

6.2.1 Considered Test Statistics

Likelihood-Ratio Test

The likelihood-ratio test can be used to statistically compare the nested models of order $P - 1$ and P. The model of order $P - 1$ is considered as the restricted model (model under $\mathcal{H}_0$), and model of order P is considered as the unrestricted model (model under $\mathcal{H}_1$). Denoting the respective parameter estimates for both model orders as $\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}$ and $\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}$, respectively. Then, the null and alternative hypothesis is stated regarding the true parameter values $\boldsymbol{\theta}$ of the distribution of the observations as

$$\mathcal{H}_0 : \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{\mathcal{H}_0} \tag{6.10a}$$

$$\mathcal{H}_1 : \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{\mathcal{H}_1} \quad . \tag{6.10b}$$

The likelihood-ratio test verifies, if the improvement of the log-likelihood by considering a higher model order is statistically justifiable. In other words, is a higher model order (corresponding to parameter estimates $\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}$) to be favoured over a lower model order (corresponding to parameter estimates $\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}$). The corresponding test statistic is

$$t_{\mathrm{LR}} = 2 \cdot \left[\mathcal{L}\left(\mathbf{y} | \hat{\boldsymbol{\theta}}_{\mathcal{H}_1} \right) - \mathcal{L}\left(\mathbf{y} | \hat{\boldsymbol{\theta}}_{\mathcal{H}_0} \right) \right] \sim \chi_5^2 \quad . \tag{6.11}$$

Wald Amplitude Test

Instead of comparing models of adjacent order, a criterion to stop the stepwise regression search is given by testing the statistical significance of the current model order P. Be the vector of correspondingly estimated parameters $\hat{\boldsymbol{\omega}}_{1:P} = \left[\hat{\boldsymbol{\omega}}_1^{\mathrm{T}}, \dots, \hat{\boldsymbol{\omega}}_P^{\mathrm{T}} \right]^{\mathrm{T}}$. In order to determine, whether the model order P is statistically justifiable, i.e. the stepwise regression search must be continued or stops, the significance of all estimated propagation paths has to be jointly verified. As these propagation paths superimpose in the model, the joint significance test can be restated: Justifying the significance of all estimated paths is equivalent to justifying the significance of each estimated path separately. Testing the significance of one path only, e.g. the significance of the newly estimated path after increasing the model order, is not sufficient. In this case path

splitting, i.e. the approximation of a path in the observations by two paths in the data model, will not be detectable. Path splitting can only be detected by testing all paths, i.e. testing the *complete* data model, whether jointly (likelihood-ratio test) or separately (Wald test). The model order is statistically justifiable and hence the model order can be increased, if *all* detected paths are significant. Accordingly, the following null hypothesis and alternative hypothesis regarding the *p*-th path can be stated, where the *p*-th path is excluded under to null hypothesis.

$$\mathcal{H}_0 : \boldsymbol{\varpi}_{\mathcal{H}_0} = \hat{\boldsymbol{\varpi}}_{\{1:P\}\setminus\{p\}}, \ \hat{\boldsymbol{\sigma}}_{\mathcal{H}_1} \tag{6.12a}$$

$$\mathcal{H}_1 : \boldsymbol{\varpi}_{\mathcal{H}_1} = \hat{\boldsymbol{\varpi}}_{1:P}, \ \hat{\boldsymbol{\sigma}}_{\mathcal{H}_1} \tag{6.12b}$$

Noise parameters $\hat{\boldsymbol{\sigma}}_{\mathcal{H}_1}$ are the estimates under the alternative hypothesis, respectively. Hence, if the null hypothesis cannot be rejected, the *p*-th path is not significant and therefore negligible/not justifiable. As a consequence, the model order P is too largely chosen and the stepwise regression search procedure stops. Because the path parameters are jointly estimated, the path parameters may change for each model order incremental $P - 1 \rightarrow P$. Hence, validation of each path has to be conducted for every new model order.

Lets assume, that the *p*-th estimated path is not present, i.e. the path has been wrongly estimated due to a too largely chosen model order. Then, the estimated path weight $\hat{\gamma}_p$ should approach zero in expectation. In order to justify this assumption, the estimated path weight is compared to 0 under $\mathcal{H}_0$. The Wald test is used therefore, which becomes in this case the *Wald amplitude test* [WHG+02]. As the path weight is complex-valued the null hypothesis for the path *p*-th is

$$\mathcal{H}_0 : \left[\Re\left\{\hat{\gamma}_p\right\}, \Im\left\{\hat{\gamma}_p\right\}\right]^{\mathrm{T}} = \mathbf{0}, \ \hat{\boldsymbol{\sigma}}_{\mathcal{H}_1} \tag{6.13a}$$

$$\mathcal{H}_1 : \left[\Re\left\{\hat{\gamma}_p\right\}, \Im\left\{\hat{\gamma}_p\right\}\right]^{\mathrm{T}} \neq \mathbf{0}, \ \hat{\boldsymbol{\sigma}}_{\mathcal{H}_1} \quad . \tag{6.13b}$$

If the null hypothesis is not rejectable, the estimated path can be stated as unreliable and the model order can be decreased or the stepwise regression search procedure stops. The Wald test will be used to verify the null hypothesis regarding the path weight value.

$$t_{\mathrm{W}} = \left[\Re\left\{\hat{\gamma}_p\right\}, \Im\left\{\hat{\gamma}_p\right\}\right] \cdot \left[\mathbf{F}\left(\hat{\boldsymbol{\varpi}}_{\mathcal{H}_1}, \hat{\boldsymbol{\sigma}}_{\mathcal{H}_1}\right)^{-1}\right]_{\gamma_p}^{-1} \cdot \left[\Re\left\{\hat{\gamma}_p\right\}, \Im\left\{\hat{\gamma}_p\right\}\right]^{\mathrm{T}} \sim \chi_2^2 \tag{6.14}$$

Expression $\left[\mathbf{F}\left(\hat{\boldsymbol{\varpi}}_{\mathcal{H}_1}, \hat{\boldsymbol{\sigma}}_{\mathcal{H}_1}\right)^{-1}\right]_{\gamma_p}$ picks the entries from the inverse FIM, which correspond to the real and imaginary part of the *p*-th path weight, see Section 5.2.3.

6.2.2 Control of the Type I Error

As the model order is gradually increased and each increase is justified by statistical test(s), multiple tests are conducted. Denoting $\mathcal{H} : \hat{P} = P$ as the hypothesis that model order P is valid and $\hat{P}$ be the estimated model order. Then, the following null hypothesises and alternative

Table 6.1: *Simulation settings of the monostatic FMCW MIMO radar*

Parameter	Setting	Parameter	Setting
MIMO channels	8×8	T_S	500 ns
f_c	155 GHz	W	10 GHz
T_M	100 µs	T_P	150 µs

hypothesises are successively tested by the stepwise regression search [Chu+07, Eri01].

$$\mathcal{H}_0 : \hat{P} = 0 \qquad \text{vs.} \qquad \mathcal{H}_1 : \hat{P} = 1$$
$$\mathcal{H}_0 : \hat{P} = 1 \qquad \text{vs.} \qquad \mathcal{H}_1 : \hat{P} = 2$$
$$\vdots$$
$$\mathcal{H}_0 : \hat{P} = P_{\max} - 1 \qquad \text{vs.} \qquad \mathcal{H}_1 : \hat{P} = P_{\max}$$

Therefore, multiple hypothesis are tested, which increases the type I error. Note, that for the Wald amplitude test P tests are conducted to compare model order $P - 1$ vs. model order P. Hence, the problem of increased type I error becomes even more severe.

The Bonferroni correction of the significance level will be used in order to control the correspondingly increase of the FWER: $\alpha_s = \alpha_s / P$. Bonferroni control does not require any assumptions about dependence among p_{val} or about how many of the null hypotheses are true.

6.2.3 Experiments

Monte Carlo simulations are carried out to experimentally investigate the proposed estimator for the number of propagation paths (number of point targets, respectively). Calibration data of the 8×8 monostatic FMCW MIMO radar presented in [2] will be considered. A single point target will be considered, whose parameters are generated as uniformly distributed, random numbers: $\varphi \sim \mathcal{U}(-35°, 35°)$, $R \sim \mathcal{U}(0.1\,\text{m}, 1.3\,\text{m})$ and $v \sim \mathcal{U}(0\,\text{m s}^{-1}, 0.5\,\text{m s}^{-1})$. The path weight is set to $\gamma = 1$. The simulations are carried for various SNRs, see equation (7.1) for the considered SNR definition. The number of samples per MIMO channel, the number of MIMO channels and the number of signal periods is $K = 200$, $M = 64$ and $N = 8$, respectively. Thermal noise will be considered as the only noise source, and leakage and phase noise will be neglected for convenience. In total 1000 Monte Carlo trials are ran for each SNR value, and a new set of target parameters is randomly generated for each Monte Carlo trial.

T-Value and Null Distribution

In case of a single propagation path (single point target, respectively) the critical decision for the model order estimator is to distinguish between a single- and two-path scenario in the data model. Assuming a two-path model for the observations $\mathbf{y}$ one gets

$$\mathbf{y} = \mathbf{s}\left(\varphi_1, \tau_1, v_1\right) \cdot \gamma_1 + \mathbf{s}\left(\varphi_2, \tau_2, v_2\right) \cdot \gamma_2 + \mathbf{w} \quad .$$
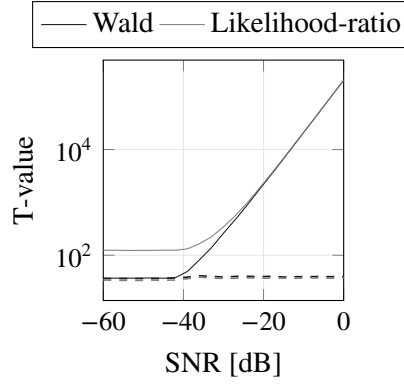
Figure 6.2: *95%-percentile value of the T-value for a single path scenario. The solid line corresponds to the T-value of the firstly estimated path (valid path), and the dashed line corresponds to the T-value of the secondly estimated path (invalid path).*

Without loss of generality one can say, that distinguishing between a single- and two-path model is equivalently to decide, whether the null hypothesis $\mathcal{H}_0 : \gamma_2 = 0$ holds. Again, in order to detect path splitting, the null hypothesis $\mathcal{H}_0 : \gamma_1 = 0$ has to be tested by the Wald test too. The proposed Wald amplitude and likelihood-ratio test both require the fitting of the two-path model. Hence, the number of paths to be inferred by the parameter estimator is set to two.

First, the 95%-percentile value of the test statistic value calculated over the Monte Carlo runs is investigated for both tests. The results are depicted in Figure 6.2 for various SNRs. The solid line represents the T-value for the firstly estimated path (valid path) and the dashed line corresponds to the T-value for the secondly estimated path (invalid path). It can be observed, that the T-value of the valid path for the Wald and likelihood-ratio test are equal for high SNRs. This fits to the theory, where the test performance of the Wald and likelihood-ratio test are considered as asymptotically (large sample size or high SNR, respectively) equal. The T-value of the invalid path are approx. independent of the SNR, and is always below the T-value of the valid path for a SNR above -40 dB. Hence, a simple threshold as e.g. c_{val} can be applied to separate the valid and invalid path by their T-value. For a SNRs below -40 dB, the T-value of the valid path converges to the T-value of the invalid path for the Wald test, such that a simple threshold for separation becomes hardly applicable. However, the T-values for the likelihood-ratio test are still separable by thresholding.

Second, the empirical distribution of the T-value under $\mathcal{H}_0$ will be investigated. According to theory, the distribution of this T-value should follow the null distribution, which is χ_2^2 and χ_5^2 for the Wald and likelihood-ratio test, respectively. The empirical null distribution must be well approximated by the theoretical one in the region of $1 - \alpha_{\text{s}}$, because the critical value is calculated there. The empirical cummulative distribution function (CDF) of the T-value for the Wald and likelihood-ratio test, as well as the CDFs of the assumed null distributions, are shown in Fig. 6.3. It becomes obvious, that the theoretical null distributions do not fit to the empirical ones. An explanation for this mismatch is, that the theoretic null distributions are derived under certain assumptions, as e.g. unbiasedness and normal distribution of the estimates or the proper quadratic approximability of the log-likelihood. If the data model in the likelihood is a linear model, i.e. the parameters are linearly included in the model, this is ensured. However, the considered data model (6.15) has linear (γ_1, γ_2) and non-linear
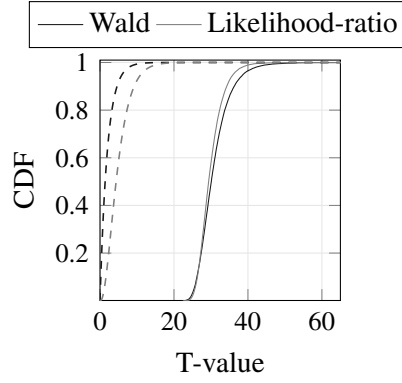
Figure 6.3: *Cummulative distribution function (CDF) of the T-value of the invalid path for the Wald and likelihood-ratio test. Solid lines are the empirical CDFs and dashed lines are the respectively theoretical CDFs.*

$(\varphi_1, \tau_1, \nu_1, \varphi_2, \tau_2, \nu_2)$ parameters, which all have to be estimated from the observations. If $\gamma_2 = 0$ is assumed under $\mathcal{H}_0$, the corresponding non-linear parameters φ_2, τ_2, ν_2 are in fact not observable [Eri01]. Consequently, these parameters are not estimable as the objective is nearly flat w.r.t. this parameters [Gal77], and their values are driven by noise and numerical accuracy as e.g. the convergence behaviour of the Levenberg-Marquardt algorithm. Hence, the parameter estimates $\hat{\varphi}_2, \hat{\tau}_2, \hat{\nu}_2$ do not converge in probability [Gal77], such that the estimator becomes biased. In principle, model of order P is nested in model of order $P + 1$ and can be obtained by setting the path weight of the corresponding path to zero. As a consequence, the regularity conditions for the standard asymptotic null distribution fail and furthermore the information matrix becomes singular. Therefore, the standard asymptotic null distribution cannot be observed. Vuong [Vuo89] has shown, that in general and especially under model misspecification the null distribution asymptotically tends to a weighted sum of independent chi-square distributions. However, determination of the weights and of the degrees-of-freedom of the respective chi-square distributions is hard to derive analytically. Summarised, the theory to analytically derive the null distribution breaks down, because assumptions are violated by forcing the estimation of an additional, invalid path. As a consequence, the critical value c_{val} cannot be calculated, as the null distribution is unknown.

In order to further investigate the breakdown of the theory, another Monte Carlo simulation is run. A single path with fix parameter settings $\varphi = 0°$, $R = 1.3\,\mathrm{m}$, $\upsilon = 0.5\,\mathrm{m\,s^{-1}}$ and $\gamma = 1$ is considered. The SNR is chosen to be $-9\,\mathrm{dB}$ and $I = 1000$ Monte Carlo trials are ran. The parameter estimator is forced to estimate two paths. The normalised histogram (empirical PDF) of the parameter estimates is shown in Fig. 6.4. Note, that the histograms have been centralised by their respective mean value. The distributions of the estimates of the first path (valid path) follow a normal distribution as forecast by theory, see Section 5.2. However, the distribution of the estimates of the second path (invalid path) do not indicate a normal distribution. In order to apply the proposed tests, at least the linear parameters should follow a normal distribution. The non-linear parameters (azimuth, delay, Doppler) are approx. uniformly distributed in their parameter space, and the distribution of the linear parameters is comparable to a Laplacian distribution. Summarised, the assumptions regarding the parameter's distribution are violated and therefore the theoretical null distribution is not observed.
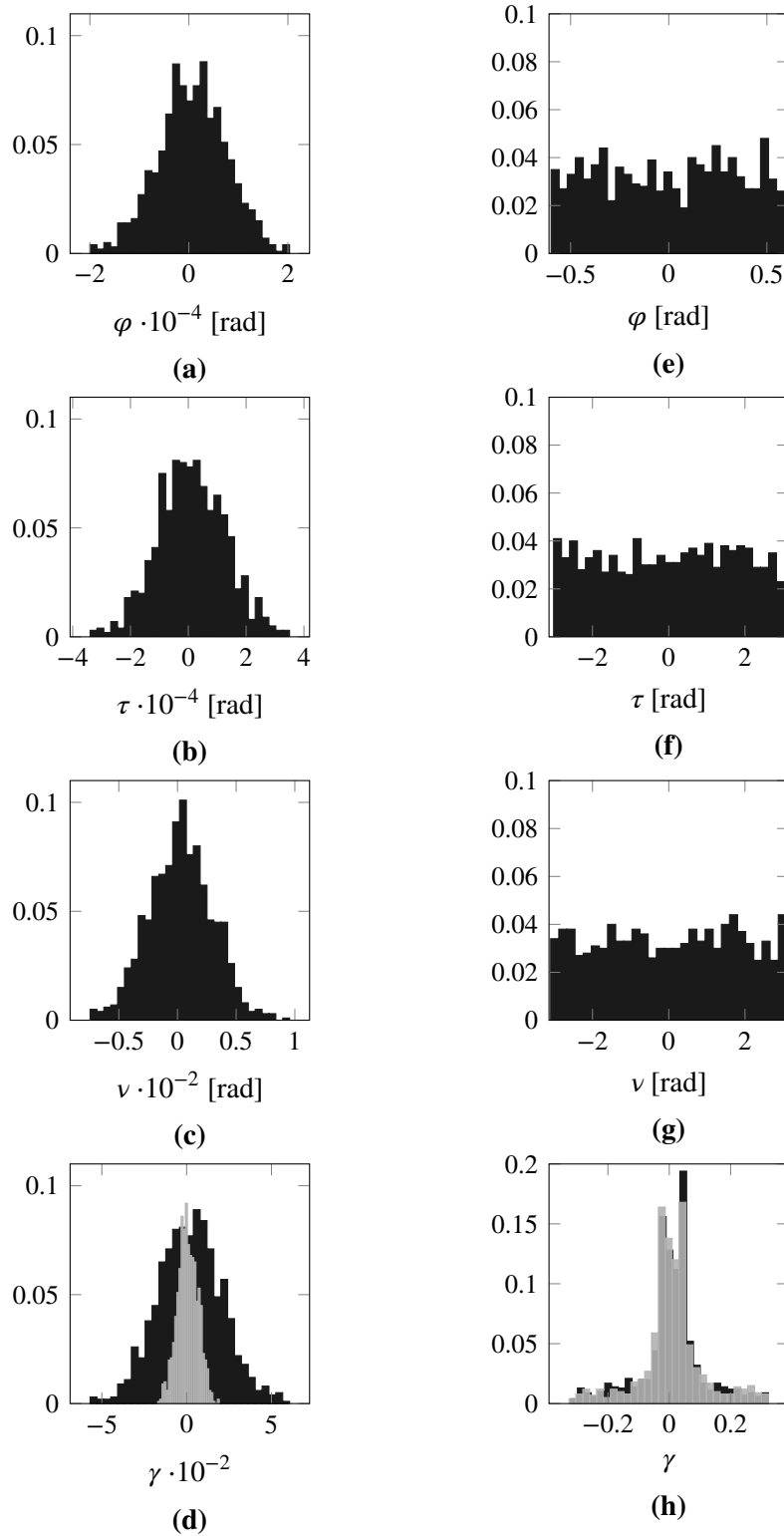
Figure 6.4: *Centralised and normalised histograms of the estimated parameters of the first path ((a) to (d)) and the second path ((e) to (h)) for SNR of −9 dB. Plot (d) and (h) shows the histograms of the real (dark) and imaginary (bright) part of the path weight estimates.*

6.3 Hypothesis Testing by Bootstrapping and Monte Carlo Simulation

In order to circumvent the problem of an unknown null distribution one can choose another distribution, which is a better approximation of true null distribution; or one can improve the test statistic such that the assumed null distribution becomes more valid. Approaches of the first kind are re-sampling techniques as e.g. bootstrapping [Efr79], where the null hypothesis' distribution is determined by an empirical sampling distribution. Methods of the second kind are e.g. Bartlett and Bartlett-type corrections [CC96]. The number of degrees of freedom can be adjusted too, in order to better approximate the true null distribution. Also, the empirical null distributions can be considered to calculate $c_{\text{val}} = 37$. For a significance level $\alpha_{\text{s}} = 0.05$ the empirically derived critical values are approx. $c_{\text{val}} = 37$ and $c_{\text{val}} = 34$ for the Wald test and likelihood-ratio test, respectively. Hence, with a probability of 5 % an invalid path will be considered as valid one and the model order is overestimated. This is not practical, as these critical values values depend on the number of samples or MIMO snapshots.

Other methods to derive the null distribution are *resampling methods* as *bootstrapping* or *Monte Carlo simulations*. Resampling methods are powerful tools to substitute theoretical derivations by computational calculations. For example, if the statistic of the data are unknown or only a few samples are available, asymptotic theory may not apply. In this case, resampling methods become conceivable. Resampling techniques will be considered for the likelihood-ratio and Wald test in order to tackle the problem of the unknown null distribution.

6.3.1 Likelihood-Ratio t-Test with Bootstrapping

An introduction to bootstrapping and its application for signal processing can be found in [ZB98] and [Pol98]. The problem, which is addressed by bootstrapping, can be summarised as follows: having the set $\mathsf{X} = \{X_1, \ldots, X_P\}$ of P observed i.i.d. random variables X, which have an unknown distribution F. Say that θ is an unknown characteristic of F as e.g. a location parameter like the mean or median. The question now is, how the estimator $\hat{\theta}$ is distributed, derived from the set of observations X. A possible way would be to repeat the experiment multiple times to get multiple sets of observations, and to estimate θ from each observed set. If a sufficient number of repeated experiments is conducted, the distribution of $\hat{\theta}$ can be approximated by empirical distribution $\hat{F}$ of the multiple estimates. However, repeating the same experiment multiple times may be inapplicable under practical considerations.

Assume, that X features the underlying distribution and the experiment can be repeated by random sampling with replacement from X. Draw C bootstrap samples $X_c^{\star}$ from X by sampling with replacement, and build the set $\mathsf{X}^{\star} = \{X_1^{\star}, \ldots, X_C^{\star}\}$. Repeat the bootstrap sampling B times to get $\mathsf{X}_1^{\star}, \ldots, \mathsf{X}_B^{\star}$ sets, and calculate the estimate $\hat{\theta}_b^{\star}$ of the unknown characteristic θ from each set $\mathsf{X}_b^{\star}$. Last, approximate the distribution of the estimator $\hat{\theta}$ by the empirical distribution of the $\hat{\theta}_1^{\star}, \ldots, \hat{\theta}_B^{\star}$. This kind of bootstrapping is known as *non-parametric bootstrapping*, because the observations are resampled.

In order to apply bootstrapping for the likelihood-ratio test, the respective test statistic in equation (6.9) will be restated. As the covariance matrix in equation (5.56) is a block-diagonal matrix in the dimension m of the MIMO channels and the dimension n of the MIMO

snapshots, the log-likelihood can be separately calculated for each MIMO channel and each MIMO snapshot: $\mathcal{L}_{m,n}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}\right)$. The overall log-likelihood is now given by summation over all M MIMO channels and N snapshots. Hence, the likelihood-ratio test statistic becomes

$$t_{\text{LR}} = \sum_{m=1}^{M}\sum_{n=1}^{N} 2 \cdot \left[\mathcal{L}_{m,n}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}_{\mathcal{H}_1}\right) - \mathcal{L}_{m,n}\left(\mathbf{y}|\hat{\boldsymbol{\theta}}_{\mathcal{H}_0}\right)\right] = \sum_{m=1}^{M}\sum_{n=1}^{N} t_{\text{LR}_{m,n}} \quad . \tag{6.15}$$

According to theory, $t_{\text{LR}_{m,n}} \sim \chi_5^2$ should hold under the null hypothesis. Hence, under asymptotic conditions $MN \to \infty$

$$\mu = \lim_{MN\to\infty} \frac{t_{\text{LR}}}{MN} = \lim_{MN\to\infty} \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} t_{\text{LR}_{m,n}} = 5 \tag{6.16}$$

should hold, as the expectation of a χ^2 random variable are the respective degrees-of-freedom. Hence, if the mean value of the likelihood-ratio test statistic is smaller than 5 it will be concluded, that the respective null hypothesis cannot be rejected. However, in the finite sample case a mean of 5 may not be observed, as only an estimate $\hat{\mu}$ is available. The check the estimated mean regarding the asymptotical value, a statistical test will be applied, whose null and alternative hypothesis are

$$\mathcal{H}_0 : \hat{\mu} \leq \mu_0 = 5 \tag{6.17a}$$
$$\mathcal{H}_1 : \hat{\mu} > \mu_0 = 5 \quad . \tag{6.17b}$$

The considered test for the null hypothesis is the Student's t-Test, whose test statistic t_{S} is

$$t_{\text{S}} = \frac{\hat{\mu} - \mu_0}{\hat{\sigma}_\mu} \quad . \tag{6.18}$$

The statistic is "studentised" by an estimate of the standard deviation in order to ensure pivoting.

$$\hat{\mu} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} t_{\text{LR}_{m,n}} \tag{6.19a}$$

$$\hat{\sigma} = \sqrt{\frac{1}{MN-1}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(t_{\text{LR}_{m,n}} - \hat{\mu}\right)^2} \tag{6.19b}$$

As the distribution of $t_{\text{LR}_{m,n}}$ is unknown, the statistic of t_{S} is unknown too. Hence, a t-Test with bootstrapping will be applied to justify the null hypothesis [Chu+07, ZB98]. Consider the likelihood-ratio values t_{LR} as the set of observations X for the bootstrapping.

$$\mathsf{X} = \left\{ t_{\text{LR}_{1,1}}, \ldots, t_{\text{LR}_{1,N}}, \ldots, t_{\text{LR}_{M,N}} \right\}$$

Draw C bootstrap samples by random sampling with replacement from X, and repeat this B times. As a result, one gets the sets $\mathsf{X}_1^\star, \ldots, \mathsf{X}_B^\star$ of bootstrap samples. For each set $\mathsf{X}_b^\star$ calculate

---

Algorithm 5 Likelihood-Ratio t-Test with Bootstrapping

Require: B, C, μ_0

Require: $\mathsf{X} = \left\{ t_{\mathrm{LR}_{1,1}}, \ldots, t_{\mathrm{LR}_{1,N}}, \ldots, t_{\mathrm{LR}_{M,N}} \right\}$

   Compute: $t_{\mathrm{S}_b} = \frac{\hat{\mu} - \mu_0}{\hat{\sigma}}$

   for $b = 1, \ldots, B$ **do**

      Resampling: Get $\mathsf{X}_b^\star$ by drawing C random samples with replacement from X

      Compute: $t_{\mathrm{S}_b}^\star = \frac{\hat{\mu}_b^\star - \hat{\mu}}{\hat{\sigma}_b^\star}$

   end for

   Ranking: $t_{\mathrm{S}_{(1)}}^\star < t_{\mathrm{S}_{(2)}}^\star < \ldots < t_{\mathrm{S}_{(B)}}^\star$

   Get L: $\cdots t_{\mathrm{S}_{(L-1)}}^\star < t_{\mathrm{S}} < t_{\mathrm{S}_{(L)}}^\star \cdots$

   Compute: $p_{\mathrm{val}} = 1 - L/B$

---

the respective Student's test statistic $t_{\mathrm{S}}^\star$.

$$t_{\mathrm{S}}^\star = \frac{\hat{\mu}^\star - \hat{\mu}}{\hat{\sigma}^\star} \tag{6.20}$$

The empirical distribution of the bootstrap statistics $t_{\mathrm{S}_1^\star}, \ldots, t_{\mathrm{S}_B^\star}$ are now an approximation of the null distribution. The p_{val} is given by

$$p_{\mathrm{val}} = 1 - P\left(t_{\mathrm{S}}^\star < t_{\mathrm{S}} \right) \hat{=} \frac{B - L}{B} \quad , \tag{6.21}$$

with L the number of bootstrap statistics $t_{\mathrm{S}}^\star$ smaller than t_{S}. If the null hypothesis is rejected using a certain significance level α_{s}, i.e. $p_{\mathrm{val}} < \alpha_{\mathrm{s}}$, it will be concluded, that the higher model order is statistically not justifiable.

The test procedure, denoted as *Likelihood-Ratio t-Test with Bootstrapping*, is summarised in Algorithm 5.

6.3.2 Wald Amplitude Test with Monte Carlo Simulation

It has been shown, that the null distribution of the Wald test statistic does not follow the theoretical one. An explanation has been given by the presence of non-linear parameters, respectively by the non-linear model of the propagation. Under the null hypothesis, the respective parameter estimates do not follow the theoretically assumed normal distribution. As a consequence, the calculated test statistic does not follow the assumed distribution too.

An approximation of the null distribution can be drawn from stochastic simulation as e.g. Monte Carlo simulation, giving an empirical null distribution. The basic idea can be summarised as follows: consider a single path under the null hypothesis of the Wald amplitude test. Assign a distribution $\mathcal{D}$ to the parameter estimates of this path under the null hypothesis: $\hat{\boldsymbol{\varpi}}_{\mathcal{H}_0} \sim \mathcal{D}$. Draw B random samples from this distribution, resulting in a set of parameter estimates under the null hypothesis: $\left\{ \boldsymbol{\varpi}_1^\star, \ldots, \boldsymbol{\varpi}_B^\star \right\}$. For each set entry $\boldsymbol{\varpi}_b^\star$ calculate the test statistic $t_{\mathrm{W}_b}^\star$ of the Wald amplitude test, see equation (6.14). An approximation of the null
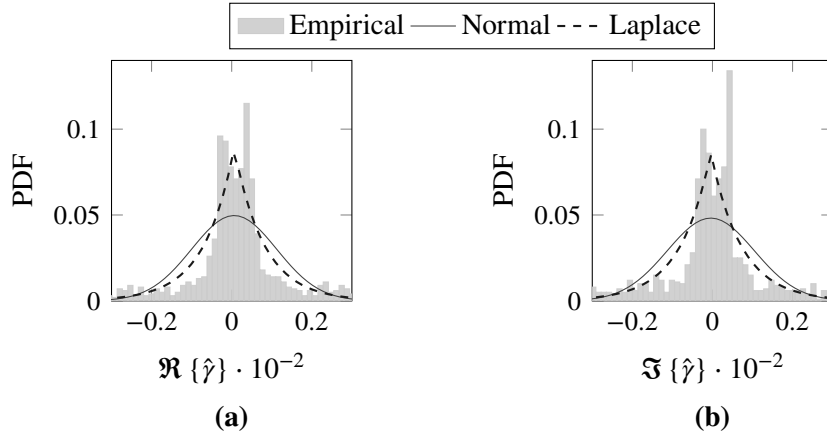
Figure 6.5: *Empirical PDFs of the (a) real and (b) imaginary part of the path weight estimates under the null hypothesis, and the PDF of the fitted normal and Laplace distribution.*

distribution is now given by the empirical distribution of all $t^\star_{W_1}, \dots, t^\star_{W_B}$ test statistics.

The remaining question is, how the distribution $\mathcal{D}$ of the parameter estimates under the null hypothesis has to be selected. Figure 6.4 shows the distribution of the parameter estimates for the valid path, i.e. the parameter distribution under the alternative hypothesis; and the distribution of the parameter estimates for the invalid path, i.e. the parameter distribution under the null hypothesis. The estimated non-linear parameters can be assumed as being independent uniformly distributed under the null hypothesis. The distribution of the linear parameters, i.e. distribution of the real and imaginary part of the path weight, is subject to clarification. Figure 6.5 shows the empirical distribution of the path weight estimates under the null hypothesis, and the PDFs of a fitted normal and Laplace distribution. The PDF $\mathcal{L}p$ of a Laplace distributed variable x is

$$\mathcal{L}p(x \mid \mu, \sigma) = \frac{1}{2\sigma} \cdot \exp\left\{ -\frac{|x - \mu|}{\sigma} \right\} \quad , \tag{6.22}$$

with location parameter μ and scale parameter $\sigma > 0$. Obviously, the Laplace distribution approximates the empirical distribution more properly, such that the path weight estimates will be assumed as being Laplace distributed under the null hypothesis.

The estimates of the real and imaginary part of the path weight may expose correlations due to the data model, such that a joint PDF has to be stated. The general PDF of a multivariate and symmetric Laplace distributed variable $\mathbf{x} \in \mathbb{R}^d$ is (for $|\boldsymbol{\Sigma}| = 1$) [ETT06]

$$\mathcal{L}p(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{2}{(2\pi)^{d/2}} \frac{K_{d/2-1}\left(\sqrt{2 (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})} \right)}{\left(\sqrt{\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})} \right)^{d/2-1}} \quad , \tag{6.23}$$

with K_m the modified Bessel function of the second kind and order m [AS72]. Under the null hypothesis of the Wald amplitude test, the path weight is assumed as vanishing: $\mathrm{E}\{\Re\{\hat{\gamma}\}\} = 0$ and $\mathrm{E}\{\Im\{\hat{\gamma}\}\} = 0$. Accordingly, the scale parameter is set to $\boldsymbol{\mu} = \mathbf{0}$ under the null hypothesis.

Algorithm 6 Monte Carlo Simulation-Based Wald Amplitude Test for *p*-th Path

---

Require: $\hat{\gamma}_p$, $\hat{\mathbf{C}}_\gamma$, B
Require: $\mathcal{D}_p$
   Compute: $t_{\mathrm{W}} = \left[\Re\left\{ \hat{\gamma}_p \right\}, \Im\left\{ \hat{\gamma}_p \right\} \right] \cdot \hat{\mathbf{C}}_{\gamma_p}^{-1} \cdot \left[\Re\left\{ \hat{\gamma}_p \right\}, \Im\left\{ \hat{\gamma}_p \right\} \right]^{\mathrm{T}}$
   for $b = 1, \ldots, B$ **do**
      Draw: $\boldsymbol{\varpi}_b^\star = \left[\varphi_b^\star, \tau_b^\star, \nu_b^\star, \Re\left\{ \gamma_b^\star \right\}, \Im\left\{ \gamma_b^\star \right\} \right]^{\mathrm{T}} \sim \mathcal{D}_p$
      Compute: $\hat{\mathbf{C}}_\gamma^\star$
      Compute: $t_{\mathrm{W}_b}^\star = \left[\Re\left\{ \gamma_b^\star \right\}, \Im\left\{ \gamma_b^\star \right\} \right] \cdot \hat{\mathbf{C}}_\gamma^{\star^{-1}} \cdot \left[\Re\left\{ \gamma_b^\star \right\}, \Im\left\{ \gamma_b^\star \right\} \right]^{\mathrm{T}}$
   end for
   Ranking: $t_{\mathrm{W}_{(1)}}^\star < t_{\mathrm{W}_{(2)}}^\star < \ldots < t_{\mathrm{W}_{(B)}}^\star$
   Get L: $\cdots t_{\mathrm{W}_{(L-1)}}^\star < t_{\mathrm{W}} < t_{\mathrm{W}_{(L)}}^\star \cdots$
   Compute: $p_{\mathrm{val}} = 1 - L/B$

The symmetric Laplace distribution for the bivariate case ($d = 2$) and zero location parameter is

$$\mathcal{L}p\left(\mathbf{x} | \mathbf{0}, \boldsymbol{\Sigma} \right) = \frac{1}{\pi \sqrt{|\boldsymbol{\Sigma}|}} \cdot K_0 \left(\sqrt{2 \mathbf{x}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{x}} \right) \quad . \tag{6.24}$$

The covariance matrix $\boldsymbol{\Sigma}$ is chosen to be the observed Cramér-Rao matrix (CRM) of the respectively estimated path weight: $\boldsymbol{\Sigma} = \hat{\mathbf{C}}_\gamma$.

The Monte Carlo simulation-based Wald amplitude test for the *p*-th path is summarised in Algorithm 6. Note, that this algorithm has to be run for each estimated path.

6.3.3 Experiments

Monte Carlo simulations as described in Section 6.2.3 are carried out to experimentally investigate the performance of the proposed model order estimation schemes. Hence, a simulation with a single path is carried out, but two paths are forced to be estimated. Furthermore, methods to control the FWER and the FDR are considered, too. The level of significance has been set to $\alpha_{\mathrm{s}} = 0.05$. The percentage of the positive decisions for the likelihood-ratio test and the Wald amplitude test are shown in Fig. 6.6 and Fig. 6.7, respectively, for both estimated paths. A positive decision means, that the path has been justified as valid by the test. Hence, Fig. 6.6(b) and Fig. 6.7(b) are the percentages of the false positive decisions.

The likelihood-ratio t-test shows a strong test performance for SNR above $-20\,\mathrm{dB}$ as no false negative decision is drawn for the valid path. This SNR threshold can be explained by the fact, that the estimator starts to work properly, i.e. detects the valid path, for this SNRs. Furthermore, the test conducts no false positive decision for the invalid path and the methods to control the FWER or the FDR have therefore no influence. For the Wald amplitude test with Monte Carlo simulation the SNR threshold of $-20\,\mathrm{dB}$ also applies, as no false negative decision is drawn for the valid path. However, a high percentage of false positive decisions for the invalid path is present, which is most properly reduced by the Bonferroni and the Šidak method.
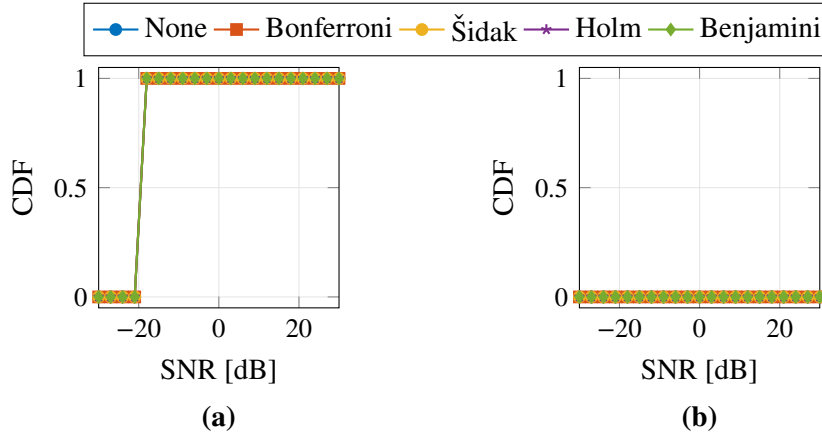

Figure 6.6: *Percentage of positive decisions of the likelihood-ratio t-test with bootstrapping for (a) the valid path (true positive decision) and (b) the invalid path (false positive decision).*
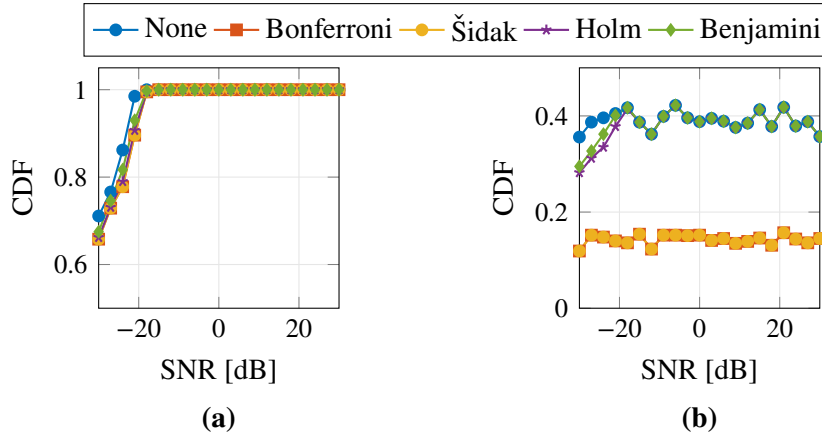

Figure 6.7: *Percentage of positive decisions of the Wald amplitude test with Monte Carlo simulation for (a) the valid path (true positive decision) and (b) the invalid path (false positive decision).*

6.4 Outline of the Algorithm for Model Order Estimation

Figure 6.8 shows the state diagram of the joint model order and parameter estimation approach. If no previous estimates are available, the initial model order is set to $P = 0$. If previous parameter estimates are available, they are validated by the Wald amplitude test (Algorithm 6). The number of valid paths P_{valid} is set as the initial model order, and the corresponding parameter estimates $\hat{\boldsymbol{\varpi}}_{P_{\text{valid}}}$ are passed to the stepwise regression search. In the stepwise regression search the model order is gradually increased until abortion. The parameter estimator from Chapter 5 is run to estimate the parameters for the considered model order. The gradual increase of the model order is exploited to save computation effort. The parameter estimates for model order $P - 1$ can be reused as initial parameter estimates for model order P. Hence, only a single path has to be detected by the initial path search scheme in Section 5.3.1. However, the joint optimisation of all estimates is still required. In conclusion, the estimator has not to be

completely re-run for each new model order. As abortion criterion for the stepwise regression search, the estimated models of order P and $P-1$ are compared using the likelihood-ratio test (Algorithm 5). If the test fails, the stepwise regression continues. If the test succeeds, the stepwise regression stops and the estimated model order is $\hat{P} = P - 1$. Also, the corresponding vector of parameter estimates $\hat{\boldsymbol{\omega}}_{\hat{P}}$ is given.

6.5 Summary

Stepwise regression search in conjunction with parametric hypothesis testing has been proposed to determine the order of the statistical model, i.e. the number of propagation paths/point targets to be considered in the estimation process. According to the step up search direction and the applied principle of Occam's razor, the model order is successively increased until a termination criterion is satisfied. As termination criterion parametric hypothesis testing is proposed, where models of adjacent order are compared. A null hypothesis proposes that the model of higher order is statistically not justifiable. If this null hypothesis cannot be rejected, the lower model order has to be preferred. The likelihood-ratio test and the Wald amplitude test have been proposed as parametric hypothesis tests. The likelihood-ratio test verifies the statistical significance of the improvement of the log-likelihood by taking into account an additional path, i.e. a higher model order. The Wald amplitude test verifies each path separately, by judging the statistical significance of the estimated path weight compared to zero. Both test have been compared in simulations. It turned out, that both test perform equally under asymptotic conditions. However, the likelihood-ratio test turned out to perform better for low SNR conditions. Hence, the likelihood-ratio test is proposed as termination criterion of the stepwise regression search. Additionally, the Bonferroni correction is used to control the FWER.

If paths from a previous estimation step e.g. due to tracking [27] are available, the model order selection problem slightly changes. The provided paths have to be evaluated, whether they are still valid or not, because paths can vanish. This evaluation of each path can be done by the Wald amplitude test. The Wald test is used, because the likelihood-ratio test requires the fitting of the restricted model. Hence, the model without the path under verification has to be estimated, which would increase the numerical effort. Paths, which are decided to be invalid by the Wald amplitude test, are dropped. The resulting model order is the initial model order for the subsequent stepwise regression search.
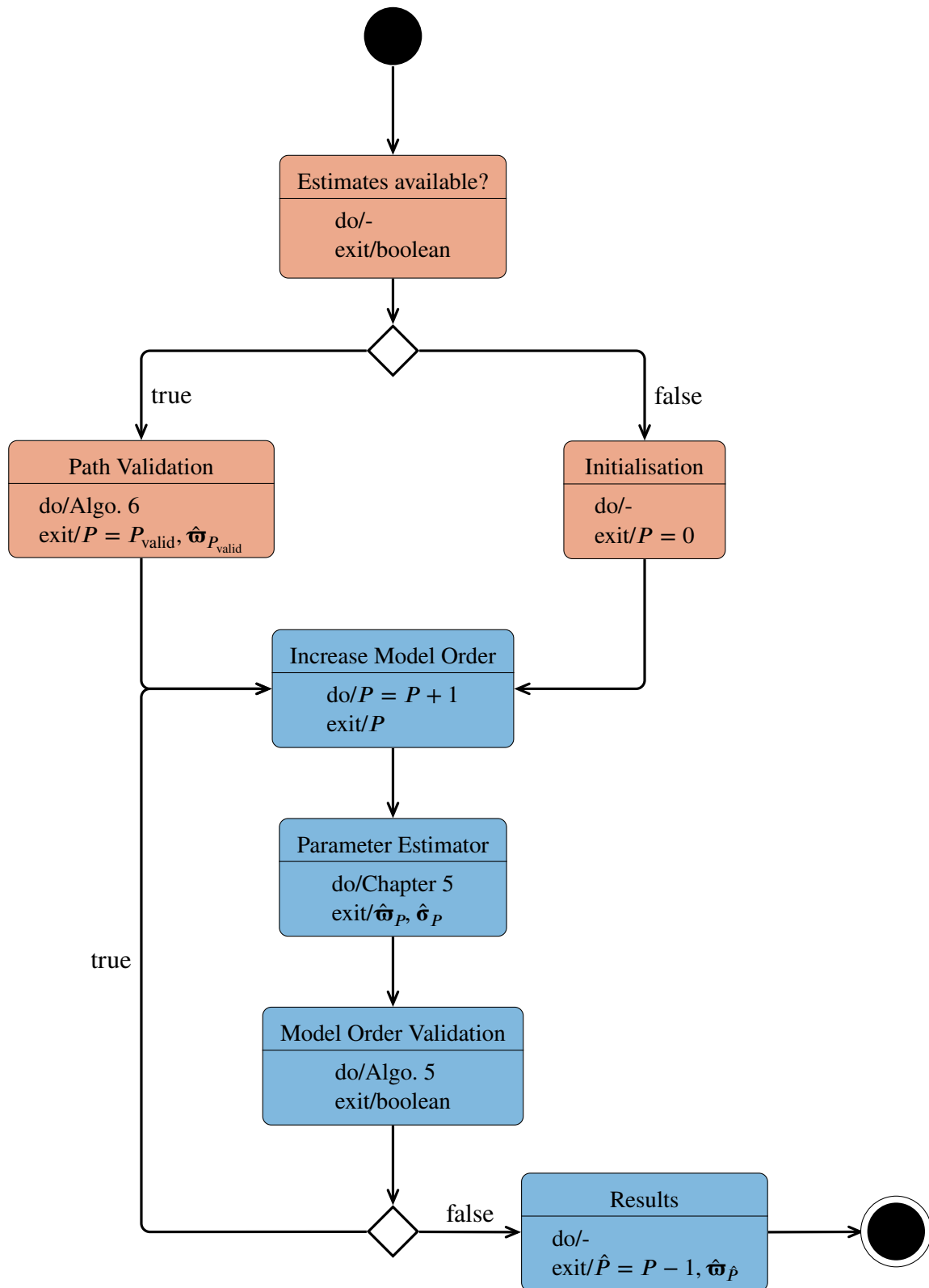
Figure 6.8: *State diagram of the approach to determine the model order $\hat{P}$ and jointly estimate the respective model parameters $\hat{\boldsymbol{\omega}}_{\hat{P}}$. The states of the initialisation procedure are coloured in red, and the states of the stepwise regression search are coloured in blue.*

CHAPTER 7

EXPERIMENTAL VERIFICATION

THE parameter estimator from Chapter 5 in conjunction with the model order estimator from Chapter 6 will be investigated by simulations, and will be applied to measurements as well. First, Monte Carlo simulations are carried out in Section 7.1, in order to experimentally investigate the estimator performance in conjunction with the optimisation scheme. The resulting root-mean-square error (RMSE) will be compared to theoretical foundations of ML estimators in Section 5.2. Next, the influence of model misspecification on the parameter estimation accuracy will be investigated in Section 7.2. Measurements with stationary and moving objects like corner reflectors, metal cylinders and a metal cube are considered in Section 7.3. The estimator is used to detect and localise point scatterers emerging from these objects. Last, in Section 7.4, the BinoMimo radar is introduced and the estimator is applied to measurements with this radar system. Limits of the applicability of the parameter estimator to observations by the BinoMimo radar are discussed, and alternative ways of processing the observations are outlined.

7.1 Estimator Performance and Root-Mean-Square Error

Monte Carlo simulations are carried out to experimentally investigate the performance of the parameter estimator. There, radar observations are generated from known model parameters, and the estimator is applied to infer the parameters from the generated observations. The 8×8 monostatic FMCW MIMO radar presented in [2] will be considered for the simulations, and calibration data of the radar are used to calibrate the device model. The settings of the radar are summarised in Tab. 7.1.

The simulations are carried out for various SNRs, whereas the SNR is defined as

$$SNR = \frac{\|\mathbf{x}\|^2}{NMK \cdot \eta} \quad . \tag{7.1}$$

Vector $\mathbf{x}$ is the noise-free observation, generated according to model (5.22); and $N = 8$, $M = 64$ and $K = 200$. In total $I = 1000$ Monte Carlo trials are ran for each SNR value.

The RMSE of the estimated parameters will be used to evaluate the performance of the parameter estimator. The RMSE of the parameter θ is calculated from the estimates of the

Table 7.1: *Settings of the monostatic FMCW MIMO radar for the Monte Carlo simulations*

Parameter	Setting	Parameter	Setting
MIMO channels	8×8	No. ramp periods N	8
f_c	$155\,\mathrm{GHz}$	W	$10\,\mathrm{GHz}$
T_M	$100\,\mathrm{\mu s}$	T_P	$150\,\mathrm{\mu s}$
f_S	$2\,\mathrm{MHz}$		

I Monte Carlo runs per SNR value.

$$\rho = \sqrt{\frac{1}{I} \sum_{i=1}^{I} \left(\theta_i - \hat{\theta}_i \right)^2}$$

θ_i denotes the known parameter in the i-th Monte Carlo run and $\hat{\theta}_i$ is its respective estimate by the parameter estimator. As known from Section 5.2 the RMSE should asymptotically, hence for a high SNR, attain the Cramér-Rao lower bound (CRLB).

7.1.1 Root-Mean-Square Error of Path Parameter Estimates

A single point target (a single propagation path, respectively) will be considered. The target parameters, i.e. the azimuth angle φ, the target range from the radar R, and the radial velocity v, are generated as uniformly distributed, random numbers: $\varphi \sim \mathcal{U}(-35°, 35°)$, $R \sim \mathcal{U}(0.1\,\mathrm{m}, 1.3\,\mathrm{m})$ and $v \sim \mathcal{U}\left(0\,\mathrm{m\,s^{-1}}, 0.5\,\mathrm{m\,s^{-1}}\right)$. The path weight is set to $\gamma = 1$. A new set of path parameters is randomly generated for each Monte Carlo run. Thermal noise will be considered as the only noise source, and leakage and phase noise will be neglected for simplicity.

The RMSE of the estimated path parameters and the corresponding CRLB are shown in Fig. 7.1. Obviously, the RMSE attains the CRLB for SNRs above $-30\,\mathrm{dB}$, indicating that the proposed estimator is an asymptotically efficient estimator.

Furthermore, one can observe that the RMSE curves for the non-linear parameters tend to saturate for low SNRs. For very low SNRs the RMSE becomes lower than the CRLB. However, in Section 5.2.3 it has been pointed out, that the CRLB denotes the lowest achievable bound for the RMSE. Hence, the empirical RMSE from the Monte Carlo simulation seems to be opposing with the theory. However, this is a known general issue in non-linear parameter estimation problems [TB07]. It has to be kept in mind, that the CRLB has been derived for an *unbiased* ML estimator. However, for very low SNRs the ML estimator cannot be assumed as unbiased, such that the CRLB is not a valid lower bound for the estimation variance in that region [TB07]. The ML estimator becomes biased for low SNRs as the location of the global maximum of the log-likelihood function becomes more and more influenced by the noise. Hence, the global maximum is not properly detectable. In the very low SNR case the location of the maximum becomes totally arbitrary, such that no useful information on the parameters are available any longer. The saturation of the RMSE of the non-linear parameters for low SNRs is due to
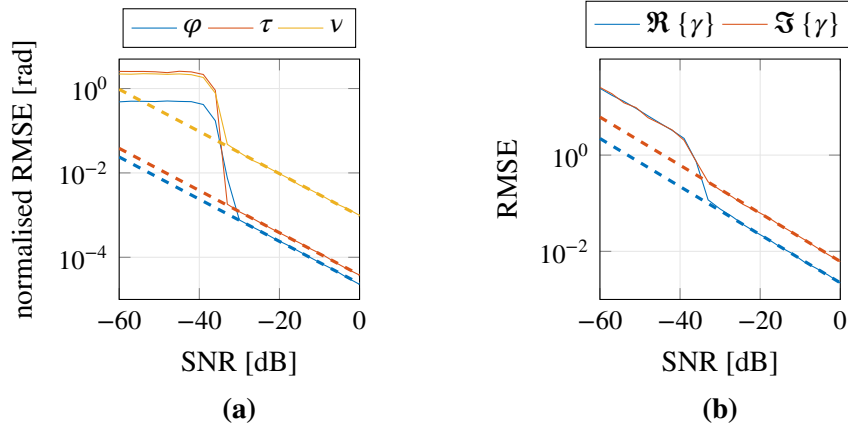
Figure 7.1: *Root-mean-square error (solid lines) of the estimated (a) non-linear parameters and (b) linear parameters of the path model. The CRLB for each parameter is indicated by dashed lines. The RMSE for the non-linear parameters has been normalised to 2π.*

the parameter bounds. The parameter bounds restrict the arbitrariness of the location of the maximum of the log-likelihoods, such that the estimation variance is upper bounded. As the linear parameters are unbounded, their RMSE does not converge for low SNRs. In conclusion the CRLB is only a valid under asymptotic, i.e. high SNR, conditions. Some other bounds are known from literature, which circumvent the problem of the CRLB for low SNRs as e.g. the Weiss-Weinstein bound, Ziv-Zakai bound or the Bhattacharyya bound [TB07].

7.1.2 Root-Mean-Square Error of Noise Parameter Estimates

Next, the RMSE of the estimates of the noise model parameters $\boldsymbol{\sigma}$ will be investigated by Monte Carlo simulations as well. The same settings of the radar as shown in Tab. 7.1 are used for the simulations. A single point target is considered, which is at azimuth angle $\varphi = 0°$ and at range $R = 1.5\,\text{m}$. The target moves with radial velocity of $v = 0.05\,\text{m s}^{-1}$ apart from the radar. The path weight is set to $\gamma = 1$. Various SNRs will be considered, such that the power η of the thermal noise varies. As a MIMO radar is considered, one set of noise parameter estimates $\hat{\boldsymbol{\sigma}}_m$ is available for each MIMO channel m. For convenience, an average RMSE over all MIMO channels is calculated for each noise parameter.

$$\rho = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\frac{1}{I}\sum_{i=1}^{I}\left(\sigma_{m,i} - \hat{\sigma}_{m,i}\right)^2}$$

First, thermal noise and leakage noise are considered as noise sources only. The parameters of the leakage noise model are set to $\varsigma = -60\,\text{dB Hz}$, $\varepsilon = 150$ and $\rho = -0.45$ for all MIMO channels. The resulting RMSE curves are shown in Fig. 7.2(a) as solid lines, and the respective CRLB are indicated by dashed lines. The estimator attains the CRLBs for SNRs above $-20\,\text{dB}$. For lower SNRs the leakage noise process cannot be observed, as the process is hidden by the thermal noise process. Therefore, parameters cannot be estimated and the
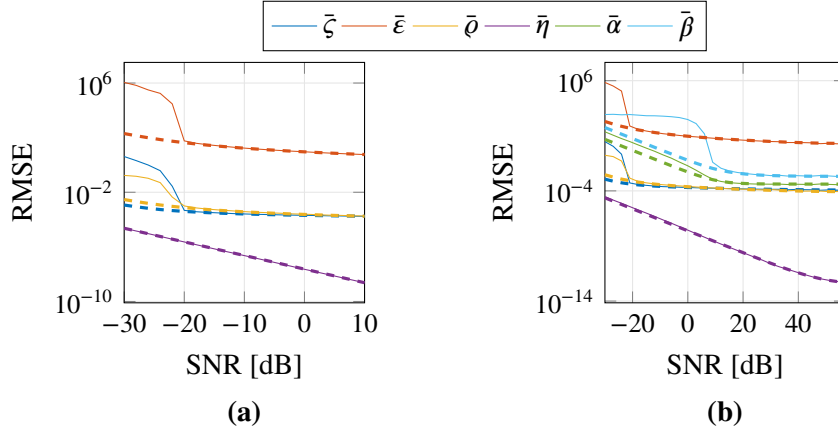
Figure 7.2: *Root-mean-square error (solid lines) of the parameters estimates of the covariance model, averaged over all MIMO channels. The CRLB for each parameter is indicated by dashed lines. The RMSE curves are drawn from Monte Carlo simulations (a) with and (b) without consideration of phase noise in the data generation and parameter estimation.*

estimation performance for the leakage model parameters breaks down. The estimator always attains the CRLB for the parameter η of the thermal noise model, as the thermal noise process is always properly observable.

Second, thermal noise, phase noise and leakage noise are considered as noise sources. The parameters of the leakage noise model are set to $\varsigma = -30\,\text{dB Hz}$, $\varepsilon = 150$ and $\varrho = -0.45$; and the parameters of the phase noise model are $\alpha = -27\,\text{dB}$ and $\beta = 0.01$. The resulting RMSE curves are shown in Fig. 7.2(b) as solid lines, and the respective CRLB are indicated by dashed lines. The RMSE of the parameters of the leakage noise attains the CRLB for SNRs above $-20\,\text{dB}$. For lower SNRs the leakage noise process is hidden by the thermal noise process and hence the parameters are not estimable. The same can be stated for the parameters of the phase noise model. If the SNR is too low, the phase nose process is hidden by the thermal noise process, such that the process is effectively not observed. As a consequence, the model parameter cannot be estimated.

7.1.3 Summary

Summarised, the Monte Carlo simulations indicate, that the estimator is capable to infer the model parameters from synthetically generated observations and asymptotically (for large SNRs) attains the Cramér-Rao lower bound. Hence, the estimator meets the ML property of asymptotic efficiency. Furthermore, the Monte Carlo simulations indicate, that the RMSE curves for the path parameters attain the CRLB for a SNR of approx. $-30\,\text{dB}$. A much higher SNR is required for the parameters of the leakage and phase noise model, so that their RMSE curves attain the CRLB. In other words, the noise processes have to be properly observed and should not be hidden by the thermal noise.

Table 7.2: *Simulation settings of the monostatic FMCW MIMO radar*

Parameter	Setting	Parameter	Setting
MIMO channels	8×8	T_S	500 ns
f_c	155 GHz	W	10 GHz
T_M	100 µs	T_P	150 µs

7.2 Estimator Performance Under Model Misspecification

The influence of model misspecification on the parameter estimation performance will be investigated in the following. Two cases of model misspecification will be considered: angle-delay coupling due to the narrowband modelling of the array response (see Section 4.4.3), and angle-Doppler coupling (see Section 4.5) and its ignorance and compensation by interpolation [10].

Monte Carlo simulations with $I = 1000$ runs will be carried out for the investigations. As figure of merit the RMSE (7.2a) will be calculated and compared to the CRLB. If the estimator works properly and the data model is sufficiently accurate the estimator is bias free and the RMSE should asymptotically attain the CRLB according to theory (see Section 5.2). Calibration data of the 8×8 monostatic FMCW TDM-MIMO radar presented in [2] will be used for the Monte Carlo simulations. The simulation settings of the radar are summarised in Tab. 7.1. Thermal noise with varying power is employed as only noise source. A single target will be considered, whose parameters (azimuth, range and velocity) are randomly generated from a uniform distribution: azimuth angle $\varphi \sim \mathcal{U}(-35°, 35°)$, target range $R \sim \mathcal{U}(0.1\,\text{m}, 1.3\,\text{m})$ and target velocity $v \sim \mathcal{U}(0\,\text{m s}^{-1}, 0.5\,\text{m s}^{-1})$. The path weight is set to $\gamma = 1$.

7.2.1 Angle-Delay Coupling

It has been pointed out in Section 4.4, that the array aperture size of the considered radar in conjunction with the signal bandwidth violates the narrowband assumption. Hence, a narrowband approximation of the array response is not conceivable, as the angle-delay coupling due to the large array aperture and the huge signal bandwidth is neglected there. However, if the narrowband approximation (due to e.g. its computational lower complexity) is nevertheless considered in the parameter estimator, the question arises, how this will effect the estimation performance. Monte Carlo simulations will be carried out to investigate this. The parameter estimator will be used to estimate the parameter values, whereas the wideband and the narrowband array model will be considered. The wideband array model of the radar will be used to generate the observations.

The resulting RMSE curves are shown in Fig. 7.3 together with the CRLB. The RMSE curves for the wideband model attain their respective CRLB. The RMSE curves for the narrowband model fail to do so and only the RMSE curve for the velocity attains their CRLB. An explanation for the failure is given by the model mismatch, because the observations are generated using the wideband array model, whereas the estimator considered the narrowband array model. As a result the estimator becomes biased and hence does not attain the CRLB (see Section 5.2). If

the narrowband approximation would be sufficiently valid, i.e. the array aperture size and/or signal bandwidth is small, the bias would be negligible. However, because the narrowband assumption is severely violated a strong bias is present.

The estimation performance for the azimuth angle φ and the delay τ are jointly degraded, if the narrowband array model is considered. The joint degradation arises, as both parameters influence the delay domain of the observations, i.e., both parameters introduce a delay in the observations. Hence, the parameters are coupled, which is account for by the data model, if the wideband array model is employed. Note, that the parameters are not coupled in physical reality. Under narrowband approximation the coupling is ignored in the data model, such that the coupling will be shifted to the estimates. Thus, if the estimation performance of one of the parameters degrades (e.g. the estimation performance of the azimuth due to the narrowband approximation), the estimator attempts to compensate this degradation by tuning the other parameter, resulting in a joined degradation of the estimation performance. A similar behaviour can be assigned to the path weight estimation. The estimator attempts to compensate the model mismatch by tuning the path weight estimates, to the cost of becoming biased and hence failing to attend the CRLB. As the Doppler (velocity) is not coupled to any other parameter by e.g. influencing the same dimension of the observations, the RMSE of the Doppler attains the CRLB also in the case of the narrowband array model.

The Monte Carlo simulations indicate, that the considered radar necessitates the usage of a wideband array model to properly describe the observations. Consideration of a narrowband array model in the estimator results in a strong model mismatch, such that the estimator performance severely degrades to a biased estimator. Also, narrowband approximation causes a coupling of the angle and delay estimates. However, the narrowband model can still be applied, if the bandwidth is reduced by e.g. pre processing of the measurements. Also, reducing the bandwidth enables the usage of a simpler model of the radar response, see Section 4.2.1. Therefore, simpler device models are applicable after bandwidth limitation, such that the numerical effort for the parameter estimator reduces. Hence, the computation time is much lower. Drawback is the loss in resolution, which can be compensated up to some extend by the high-resolution capability of the parameter estimation.

7.2.2 Angle-Doppler Coupling

The observations in each Monte Carlo ran are generated by considering the angle-Doppler coupling, and employing the wideband array model. Three different estimators are run to infer the model parameters from the generated observations. They differently cope with the angle-Doppler coupling: the proposed model-based consideration, compensation by interpolation to ensure time alignment [Som+02] and ignoring of the angle-Doppler coupling. The RMSE curves of the estimated path parameters for all three estimators are shown in Fig. 7.4 as solid linens, together with their respective CRLB (dashed lines). It can be observed, that the estimator does not attain the CRLB for the azimuth φ, if angle-Doppler coupling is ignored (see Fig. 7.4(a)). It can be concluded, that the estimator is biased due to the model mismatch. However, the estimator attains the CRLB for the velocity estimation, as the model mismatch seems to affect the angle estimation only. Furthermore, the delay estimation performance is affected too, which can be explained by the coupling of delay and azimuth due to the considered wideband array model in the observation generation. However, if the angle-Doppler coupling
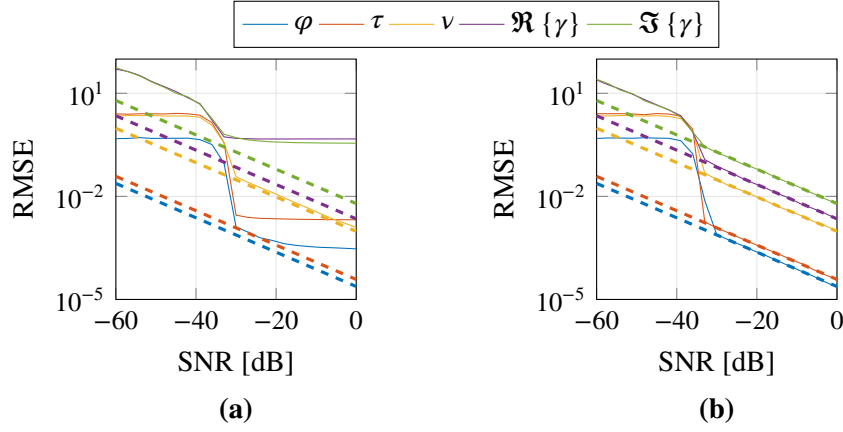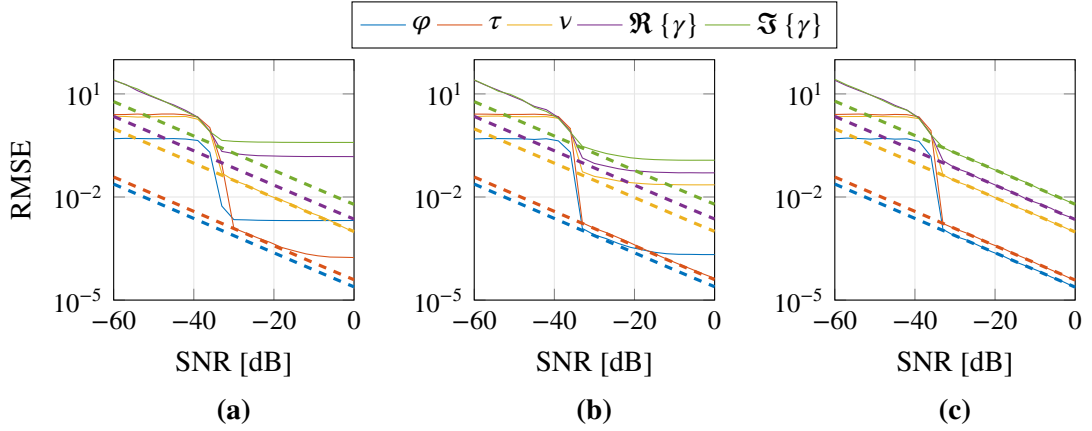
Figure 7.3: *RMSE curves (solid lines) of estimated path parameters drawn from Monte Carlo simulations. The dashed lines denote the CRLB, which indicates the lowest achievable RMSE of the parameter estimates (see Section 5.2). The RMSEs have been derived for the usage of (a) the narrowband array model and (b) the wideband array model in the parameter estimator, whereas the wideband array model has been considered for the observation generation in the Monte Carlo simulations. Therefore, a model mismatch is present for the narrowband estimation case, resulting in a biased estimator.*

Figure 7.4: *RMSE curves (solid lines) of estimated path parameters drawn from Monte Carlo simulations. The dashed lines denote the CRLB (see Section 5.2). The RMSE curves have been derived from parameter estimators, which (a) ignore, (b) compensate and (c) exploit the angle-Doppler coupling in the parameter estimator.*

is compensated by time-alignment, the delay estimation attempts the CRLB, as the accuracy of the azimuth estimation is improved, see Fig. 7.4(b). The RMSE of the azimuth estimation is improved by a factor of approx. 10, whereas the velocity estimation (Doppler) is degraded. If the correct model is used, the RMSE curves of all parameters attain the CRLB, because no model mismatch is present and the estimator is therefore not biased.

7.2.3 Summary

Based on the simulations it can be concluded, that model mismatches or model errors cause the estimator to become biased. This bias occurs, as the estimator attempts to tackle the model errors by further tuning of the parameters, or the global optimum of the objective function is shifted due to the model errors. Furthermore, joint degradation of the estimates of multiple parameters can occur, even if the model error is only present w.r.t. a single observation domain, e.g. the aperture domain (domain of the MIMO channels, respectively) for a wrongly chosen array model. The joint degradation is due to the coupling of the respective parameters, as they jointly influence the same domain of observation, e.g. azimuth angle and delay both influence the delay domain as explained by the wideband array model.

Quite often model errors can be neglected or are neglected in practise, as the SNR is low enough, such that the noise hides the model errors. Hence, degradation of the estimation performance due to model errors is not observed. However, model errors are not noise like. They are reproducible, if the observation setup remains fix, which is contrarily to noise. Therefore, model errors will cause a bias in the parameter estimates, if they become severe. Nevertheless, selection of a more proper model is not only suitable to avoid biased estimates, but also improves the variance of the estimates. Hence, a proper or more complicated model is always adorable if the estimation results shall be improved, typically to the cost of a higher computational complexity.

7.3 Estimation Results from Measurements

In order to demonstrate the capabilities of the developed parameter estimator, measurements will be considered now. The monostatic MIMO FMCW radar presented in [2] and shown in Fig. 7.5 is used for the measurements. The basic settings of the radar are summarised in Tab. 7.3. Various objects as e.g. corner reflectors, metal cylinders or a metal cuboid will be used as radar targets. These targets are either stationary or moving by slowly rotating them on a turntable. In the following, the localisation accuracy of the radar by investigating the angle and delay resolution capability will be proven. From the estimated delay $\hat{\tau}$ and the estimated azimuth $\hat{\varphi}$ the point target location $\hat{x}, \hat{y}$ in the azimuth plane can be estimated by

$$\hat{x} = \frac{\hat{\tau} \cdot c_0}{2} \cos{(\hat{\varphi})} \tag{7.2a}$$

$$\hat{y} = \frac{\hat{\tau} \cdot c_0}{2} \sin{(\hat{\varphi})} \quad . \tag{7.2b}$$

In addition, the influence of the phase noise model on the attenuation of model errors and the resulting detection of weak paths is investigated.

7.3.1 Stationary Scenario

Angle and Delay Resolution Capability

First, the angle and delay resolution capability will be verified by measurements with a single corner reflector (RCS: $-3\,\mathrm{dBsm}$) as target. The reflector is placed at a distance of approx.
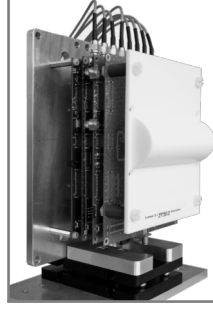
Figure 7.5: *Monostatic MIMO FMCW radar used for the measurements (Picture by André Dürr, Institute of Microwave Engineering, Ulm University.)*

Table 7.3: *Radar settings for the measurements*

Parameter	Setting	Parameter	Setting
f_c	154 GHz	W	10 GHz
T_M	100 µs	T_R	25 µs
T_W	25 µs	T_P	150 µs
T_S	40 ns		

3.58 m apart from the radar in broadside direction (azimuth of 0°). In order to realise different azimuth angles, the radar is rotated on a positioning device in the angular range of $-40°$ to $40°$ in 5° steps. The measurements are conducted in an anechoic chamber in order to have a well known environment, i.e. free of parasitic reflections. The localised reflector positions for each rotation position are shown in Fig. 7.7(a). Because the radar is rotated around its centre point, the corner reflector occurs as moving on a circle around the radar. Hence, the ground truth of the reflector positions occurs as a circular shape. It can be noted, that the parameter estimator perfectly estimates the target position.

Second, a measurement setup with two targets (RCS: 7 dBsm and 15.8 dBsm, distance: 3.59 m and 4.54 m, azimuth $-10°$ and 5°) will be considered. Again, the radar is rotated from $-40°$ to $40°$ in order to realise different focusing angles between targets and radar. The resulting ground truth and estimates are shown in Fig. 7.7(b). It can be seen, that the target locations are properly estimated, except for the radar rotation angles $-40°$ and $40°$. Note, that the targets have an initial angle w.r.t. the radar of $-10°$ and 5°. Hence, one of the targets is seen under a more extreme angle, if the radar is rotated to $\pm40°$. The array antennas have a quite low gain in this angular region, such that a low SNR is present. As a consequent, the target cannot be resolved properly.

Phase Noise and Model Error

In the following, the influence of the phase noise model on the remaining range profile after coherently subtracting all detected paths will be investigated. It will be shown, that the consideration of phase noise in the covariance model better whitens the remaining range profile, such that contributions of detected paths are sufficiently removed and weak paths become detectable.
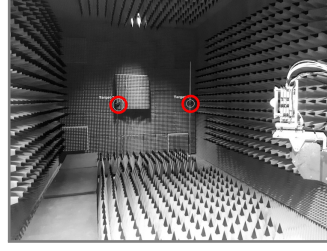
Figure 7.6: *Measurement setup with two corner reflectors as radar targets, and rotation of the radar to realise different focusing angles. (Picture by André Dürr, Institute of Microwave Engineering, Ulm University.)*
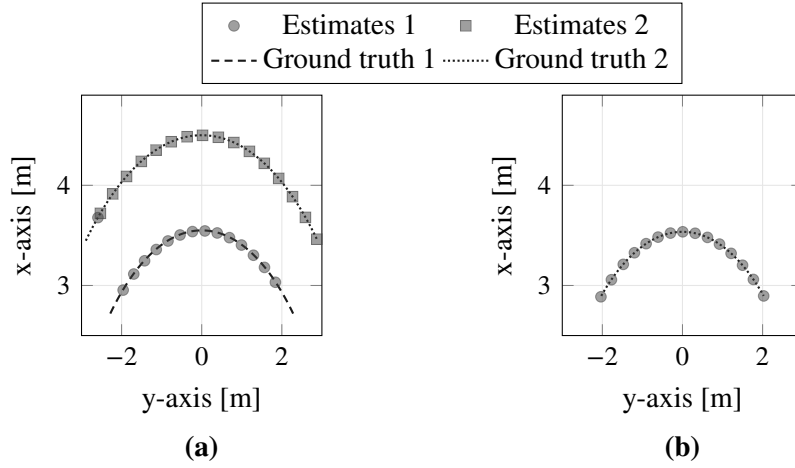

Figure 7.7: *Estimated positions of (a) two and (b) one corner reflector(s) and a stepwise rotated radar system.*

Corner Reflector   A single corner reflector with a RCS of 11.8 dBsm is used as target. The average range profile (averaged over the MIMO channels) for the case of phase noise consideration and phase noise neglection will be investigated. The measured, estimated and remaining range profile is shown in Fig. 7.8, with the corner reflector placed at location 0.8 m × 2.6 m. Note, that the number of paths to be estimated has been set to 1. The remaining range profile in Fig. 7.8(a), which is calculated by coherent subtraction of the parametrised data model using the parameter estimates, exhibit a strong remaining peak. This peak can be either due to phase noise or model errors, which cause the coherent subtraction to fail, such that a contribution from the path remains in the signal. The remaining contribution in Fig. 7.8(a) has an SNR of approx. 20 dB. By taken phase noise into account the range spectra becomes more white noise like, see Fig. 7.8(b). The remaining contributions are reduced to an SNR of approx. 3 dB.

Four Metal Cylinders   As radar target four metal cylinders (cylinder diameter: 16 mm) are used, which are mounted in a rectangular arrangement (dimension 25 cm × 19 cm) on a wooden plate, see [2]. Figure 7.9(a) shows a photography of the target setup, placed in an anechoic chamber. Eight sequentially captured MIMO snapshots are used for Doppler estimation. The estimated range, Doppler and azimuth angle are shown in Tab. 7.4. Measured, estimated and
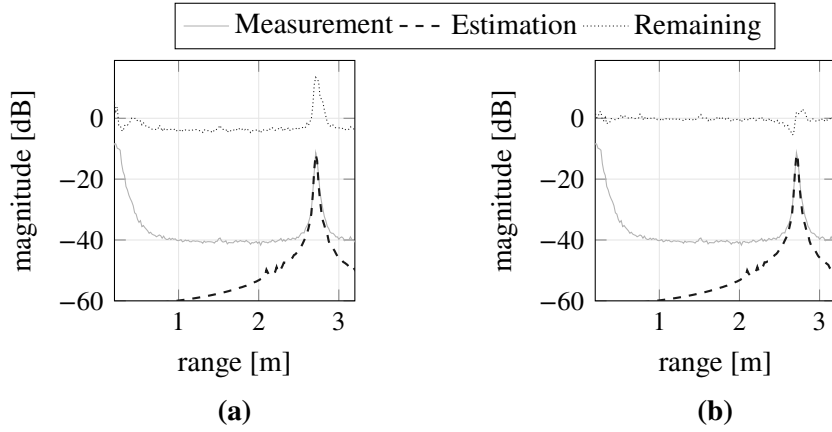
Figure 7.8: *Average range profile for a measurement with a corner reflector at 0.8 m × 2.6 m (distance: 2.72 m), and (a) neglection and (b) consideration of phase noise in the covariance model of the parameter estimator.*

remaining range profile and time domain signals are shown in Fig. 7.10, for the consideration and neglection of phase noise in the covariance model. Note, that the number of paths to be estimated has been set to 5. This is necessary in order to prevent the estimator from estimating too many paths, which happens if phase noise is neglected. It is obvious, that the consideration of phase noise significantly attenuates the model errors. Furthermore, an additional weak path (indicated by an arrow) is estimated at range of approx. 3 m, if phase noise is taken into account. The reason is, that, if phase noise is neglected, the remaining contributions are larger in amplitude than this path. Hence, the estimator attempts to resolve the model errors instead of detecting the really present path. However, if phase noise is taken into account, the model errors are sufficiently attenuated and the estimator can detect the weak path.

Summary Summarised, neglection of phase noise in the covariance model causes remaining signal contributions, which may dominate weaker paths in amplitude. As a consequence, weaker paths are not detectable. Furthermore, the estimator attempts to approximate the remaining signal contributions as propagation paths, which cause false detections and/or and overestimation of the number of propagation paths in the measurements. Not only phase noise but also modelling errors are covered by the phase noise model. Model errors can cause phase mismatches between measurement and data model, such that signal parts remain after the coherent subtraction. By considering phase noise in the covariance model, these remaining signal parts are attenuated. The attenuation of the model errors will improve the model order estimation accuracy, as remaining model errors can cause the estimation of additional, incorrect paths.

7.3.2 Non-stationary Scenario

Four Metal Cylinders

The plate with the mounted metal cylinders is rotated by a positioning device with an angular velocity of $50°\,\text{s}^{-1}$. Eight sequentially captured MIMO snapshots are used for Doppler
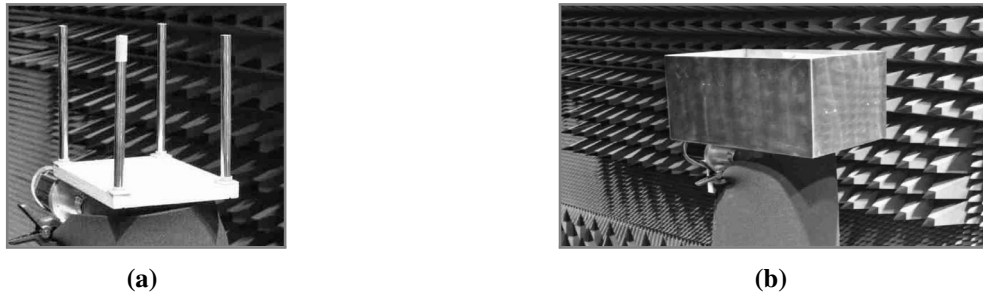
(a)

(b)

Figure 7.9: *Deployed radar targets, (a) four metal cylinders mounted on a wooden plate and (b) metal cuboid. The pictures are taken from [2].*

Table 7.4: *Estimated range, Doppler and azimuth angle of point targets, from measurement with four metal cylinders being the target.*

Azimuth [°]	Range [m]	Doppler [Hz]
1.31	1.43	0.25
−4.71	1.53	0.16
−0.21	1.74	$2.38 \cdot 10^{-2}$
5.51	1.65	$4.87 \cdot 10^{-2}$

Table 7.5: *Estimated range, Doppler and azimuth angle of point targets, from measurement with four rotating metal cylinders being the target.*

Azimuth [°]	Range [m]	Doppler [Hz]
3.74	1.73	−88.18
5.61	1.56	−132.89
−4.49	1.68	113.32
−3.8	1.5	95.43

estimation. Accordingly, the overall observation time is $8 \cdot 8 \cdot 150\,\mu s = 9.6\,ms$ and the total rotation angle of the plate is 0.48°. Due to the low total rotation angle, the angles of the point targets emerging from scattering at the object are static and the delays approximately vary in a linear manner, which can be described by the Doppler. Otherwise, if the plate is rotated faster or the total observation time is much longer, the assumption of fix target angles and a linearly varying delay are not ensured. In that case a dynamic model of the parameter variation and tracking of the parameters has to be employed [27]. The number of paths to be estimated is limited to 10, and the significance level is set to $\alpha_s = 0.05$. The estimated values for the range, Doppler and azimuth angle are shown in Tab. 7.5. Figure 7.11 shows the estimated target positions. The rectangular arrangement of the four metal cylinders is clearly visible.

The estimated range (delay, respectively) and azimuth angle are easy to validate, because they coincide with the initial target locations (see equation (7.2)). However, validating the estimated Doppler is more difficult, as the Doppler of a scattering point of a rotating object has to be taken into account [CS98, Wal80]. Because the metal cylinders (scattering points) are arranged in a rectangular shape and a rectangle is cyclic (all corners lie on a circle), the consideration of a
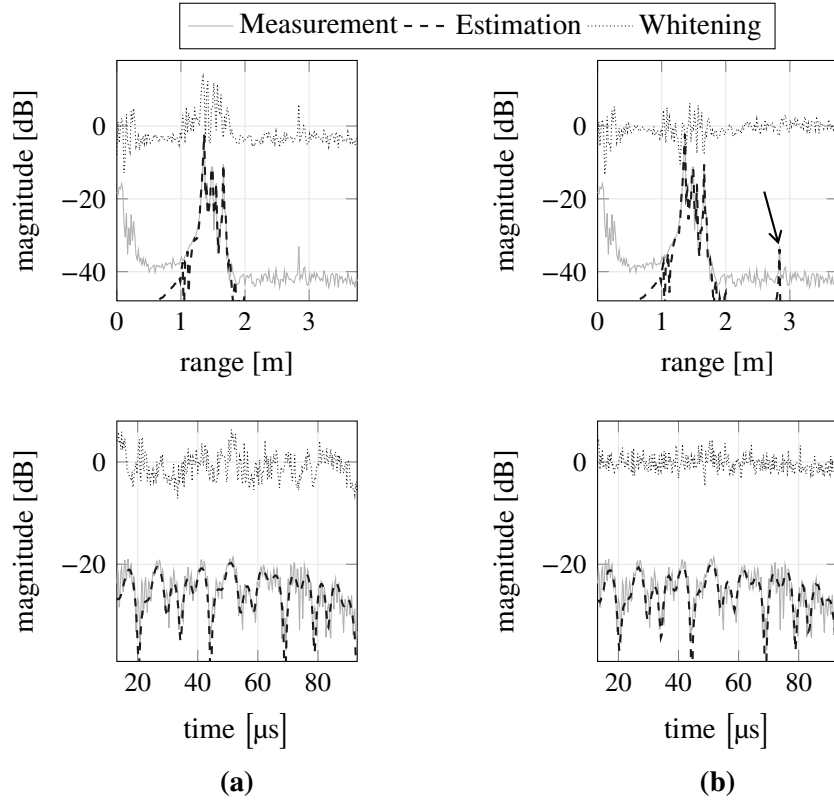
Figure 7.10: *Average range profiles (top) and time domain signal (bottom) of a measurement with four standing metal cylinders as radar target, and (a) neglection and (b) consideration of phase noise in the covariance model.*
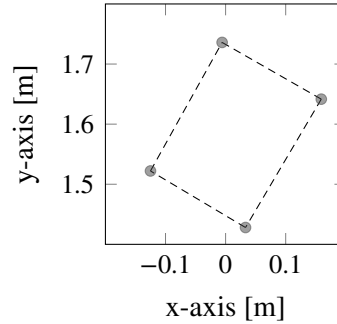

Figure 7.11: *Estimated point target locations from measurement with four rectangularly arranged metal cylinders in Fig. 7.9(a) as target.*

circle as moving path of the scatterers is sufficient. Be the rotation centre at location x_0, y_0 apart from the centre of a monostatic radar, see Fig. 7.12. The position of a scattering point of an rotating object at time t is

$$x(t) = r \cdot \cos\left(\omega t + \phi_0\right) + x_0 \tag{7.3a}$$

$$y(t) = r \cdot \cos\left(\omega t + \phi_0\right) + y_0 \tag{7.3b}$$

with r the distance of the scattering point to the centre of the rotating object, ω the angular
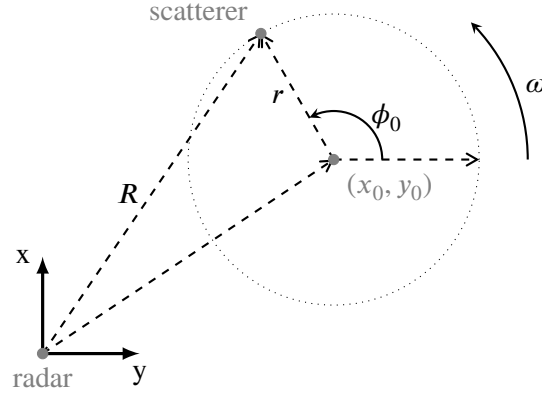
Figure 7.12: *Geometry of a rotating scatterer in the $x - y$ plane, which rotates with angular velocity ω around the centre point at (x_0, y_0).*

velocity of the rotating object and ϕ_0 the initial rotation angle. The distance of the scatter point to the radar at time t is

$$R(t) = \sqrt{x(t)^2 + y(t)^2} \quad . \tag{7.4}$$

From Section 3.2.1 it is known, that the Doppler shift v is proportional to the partial derivative of the path length. As a monostatic radar is considered, the path length is given by twice the distance.

$$v = 2 \cdot \frac{f_c}{c_0} \cdot \frac{d}{dt} R(t) \quad . \tag{7.5}$$

Calculating the derivative of the distance and plugging in to the Doppler formula yields

$$v = 2 \cdot \frac{f_c}{c_0} \cdot \frac{r \cdot \omega \cdot \left[y_0 \cos\left(\omega t + \phi_0\right) - x_0 \sin\left(\omega t + \phi_0\right) \right]}{\sqrt{\left[x_0 + r \cdot \cos\left(\omega t + \phi_0\right) \right]^2 + \left[y_0 + r \cdot \sin\left(\omega t + \phi_0\right) \right]^2}} \quad . \tag{7.6}$$

Considering the wooden plate with the four metal cylinders (the rotating object is the plate with the mounted cylinders, and each cylinder causes a scattering point), the parameters are: $x_0 = 0\,\text{m}$, $y_0 = 1.56\,\text{m}$, $r = 0.157\,\text{m}$ and $\omega = -0.87\,\text{rad s}^{-1}$ (negative sign due to clockwise rotation in the measurement). The initial rotation angle ϕ_0 for each scattering point is calculated by

$$\phi_0 = \text{atan} \left(\frac{\hat{y} - y_0}{\hat{x} - x_0} \right) \quad , \tag{7.7}$$

with $\hat{x}$ and $\hat{y}$ the estimated location of the scattering point using equation (7.2). The Doppler frequencies for $t = 9.6\,\text{ms}$ are $-83.97\,\text{Hz}$, $-137.02\,\text{Hz}$, $107.0\,\text{Hz}$ and $112.97\,\text{Hz}$ for the respective targets, which coincide quite well with the estimated values in Tab. 7.5.

Next, the impact of the consideration of phase noise in the covariance model will be briefly discussed. Figure 7.13 shows the range-Doppler spectra of the remaining and whitened signal for the consideration and neglection of phase noise in the covariance model. The advantage of
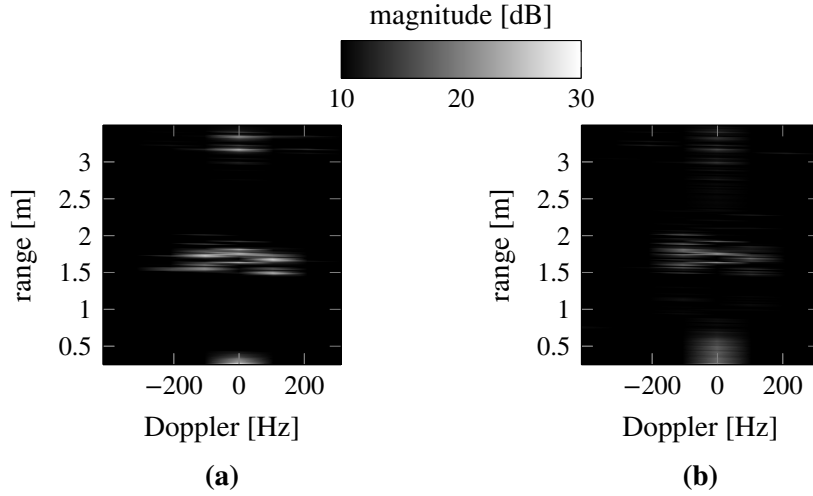
Figure 7.13: *Average range-Doppler spectra for a measurement with four rotating metal cylinders as target, and (a) neglection and (b) consideration of phase noise in the covariance model of the parameter estimator.*

phase noise consideration becomes visible, because the range-Doppler spectra in Fig. 7.13(b) does not exhibit strong remaining contributions as the spectra in Fig. 7.13(a). Hence, model errors, which are due to the phase noise, are attenuated. Because the model errors are attenuated two paths at range of 3.2 m and 3.3 m, and 0 Hz Doppler could be detected. If phase noise is neglected these paths stay in the remaining signal (see Fig. 7.13(a)), because the estimator attempts to resolve the model errors rather than detecting these paths.

Metal Cuboid

Measurements with a rotating metal cuboid, which is shown in Fig. 7.9(b), are considered next. The cuboid is of dimension 45 cm × 30 cm × 30 cm and is rotated with an angular velocity of $50\,°\,s^{-1}$ around his centre point. Sixteen MIMO snapshots are employed for Doppler estimation. The number of paths to be estimated is limited to 10, and the significance level is set to $\alpha_s = 0.05$. In Fig. 7.14 and Fig. 7.15 results for two different rotations of the cuboid are shown.

Figure 7.14(b) and Fig. 7.15(b) show the localisation results for the detected point targets. The geometry of the cuboid in the x-y-plane can be identified. In Fig. 7.14 the distances of the localised points targets coincide with the dimension of the cuboid. Hence, it can be concluded, that the respective propagation paths emerge from scattering at the edges of the cuboid. In Fig. 7.15 the shorter distance between two detected point targets is 20 cm, and does not coincide with the shortest cuboid dimension of 30 cm. An explanation is, that the localised point target on the right side does not emerge from scattering at the edge of the cuboid and is maybe due to a reflection from the metallic surface.

The measured (top) and from the estimation results reconstructed (bottom) average range-Doppler spectra are shown in Fig. 7.14(a) and Fig. 7.15(a). The range-Doppler spectra are averaged over the MIMO channels. Obviously, the range-Doppler spectra is properly reconstructed from the estimated parameters.
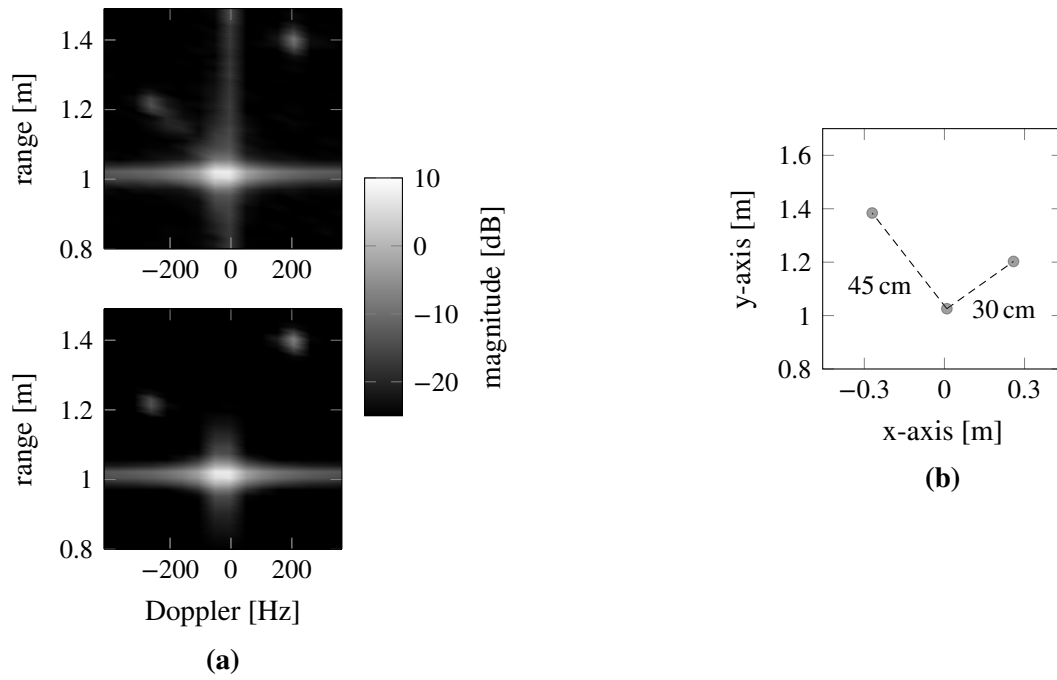
Figure 7.14: *Estimation results for the metal cuboid in Fig. 7.9(b) as target, (a) the measured (top) and reconstructed (bottom) average range-Doppler spectrum and (b) the localised scattering points.*

Figure 7.15: *Estimation results for the metal cuboid in Fig. 7.9(b) as target, (a) the measured (top) and reconstructed (bottom) average range-Doppler spectrum and (b) the localised scattering points.*
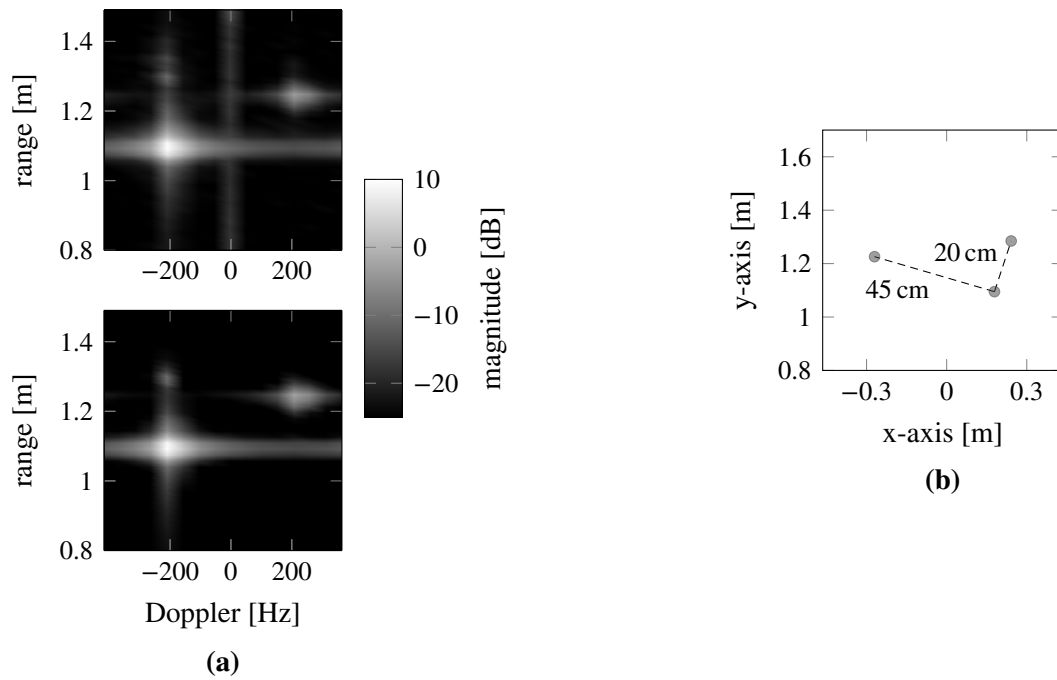
7.3.3 Summary

The applicability of the proposed parameter estimator to estimate parameters from measurements with a monostatic radar has been shown. Measurement scenarios with static and rotating targets, as well as point targets (corner reflector) and extended targets (metal cylinders and metal cuboid) have been used therefore. The estimator was capable to detect each scattering contribution from the targets. The estimated angle and delay values have been successfully used to localise the corresponding scattering sources, which fits quite well to the ground truth. Furthermore, the advantage of considering phase noise in the model of the covariance matrix has been demonstrated. It has been shown, that the consideration of phase noise attenuates remaining signal contributions due to phase noise in the hardware and model errors. As a result, the dynamic of weak paths w.r.t. remaining signals is improved, such that these paths become detectable.

7.4 Binocular MIMO Radar – An Outlook

7.4.1 Overview of the Radar System

In the BinoMimo project, funded by the Deutsche Forschungsgemeinschaft under grant no. TH 494/27-1, an advanced MIMO radar system is developed. The radar system is shown in Fig. 7.16. It comprises of two cooperative and spatially separated 8×8 MIMO radars, termed as sub-radars. The term binocular radar steams from the fact, that the sub-radars will focus an extended target under different angles due to their spatial displacement. Hence, they are "viewing" different scattering points, such that the radar exhibits the perception of depth w.r.t. the target as present in stereoscopic (binocular) vision.

Each sub-radar is a monostatic MIMO radar (co-located Tx and Rx array), having 8 Tx and 8 Rx channels. The sub-radars are mounted on a common platform and are driven by the same signal in order to ensure a coherent operation of the whole radar. Besides of the compact construction, the mounting on a common platform also reduces the hardware effort for a coherent operation. The sub-radars operate in a cooperative fashion as their transmit channels are orthogonal to each other due to a TDM accessing scheme. Furthermore, if the Tx channel of one sub-radar is radiating, the Rx channels of the other sub-radar are receiving. Hence, the whole radar is a 16×16 MIMO system (256 MIMO channels). As the two sub-radars are spatially separated they basically have different focusing angles to the target. Therefore, a bistatic radar architecture is present, if the transmit and receive array are from different sub-radars. Summarised, the novel radar system provides radar measurements with both co-located and distributed arrays, and advantages of both types of radar architectures can be exploited. Last, as both sub-radars operate coherently and if the target is far apart from the radar, such that both sub-radars have approx. the same focusing angle, the whole radar system becomes a monostatic radar. As a result, a giant virtual aperture is present, offering an angular resolution according to the Rayleigh criterion of better than $1°$ in broadside direction [Sch19].

7.4.2 Measurements with the Binocular MIMO Radar

Measurements with the binocular radar system and parameter estimation results will be considered next. The settings of the radar system are summarised in Tab. 7.6. As explained, the

(a)

(b)

Figure 7.16: *A* 16×16 *binocular MIMO radar system (BinoMimo radar) consisting of two cooperative and spatially separated* 8×8 *MIMO radars (sub-radar 1 and sub-radar 2). Because the sub-radars are cooperative, the binocular MIMO radar is capable to capture two monostatic radar observations (dashed lines) and a bistatic observation (dotted line), see (a). Picture (b) shows the setup up binocular radar system (picture taken from [Sch19]). The array centres of the sub-radars are separated by approx.* $L = 14$ cm.

Table 7.6: *Radar settings for the measurements*

Parameter	Setting	Parameter	Setting
f_c	152.5 GHz	W	$12 \cdot 844$ MHz
T_M	100 µs	T_R	25 µs
T_W	25 µs	T_P	150 µs
T_S	50 ns		

binocular radar system is capable to gather two mono- and two bistatic radar measurements. As the statistical model of the proposed parameter estimator (see Section 5.1) assumes measurements with a monostatic radar, the measurements have to be treated accordingly. Hence, the bistatic radar observations will not be considered subsequently and the corresponding measurements are excluded beforehand from the radar data cube (reduced from 256 to 128 MIMO channels). Only the monostatic measurements from each sub-radar will be considered. These measurements can be *jointly* or *separately* processed by the parameter estimator, depending on the assumption of co-located or distributed sub-radars. In the joint processing case it is assumed, that both sub-radars are focusing the target under the same angle. Hence, co-located sub-radars are assumed and the observations of both sub-radars are jointly processed by the estimator. In the separate processing case it is assumed, that both sub-radars are focusing the target under different angles. Hence, distributed sub-radars are assumed and the parameter estimator is applied to each sub-radar measurement separately.

Stationary Corner Reflector

The influence of the spatial separation of the sub-radars in relation to the distance of the target from the radar system will be investigated subsequently. The result of the beamforming in azimuth domain (see Section 5.3) will be used for the investigations. The beamforming result is calculated for the separate processing of the sub-radar measurements (i.e. assumption of distributed sub-radars), and the joint processing of the sub-radar measurements (i.e. assumption of co-located sub-radars). Measurements with a single target placed at 3 m, 4 m and 5 m apart

from the radar, and at angles −10°, 0° and 10° are considered. The calculated beamforming results are depicted in Fig. 7.17. First, it can be observed, that the width of the main lobe is smaller for the joint processing case compared to the separate processing case, indicating the higher resolution due to the larger virtual aperture. Second, it can be observed, that the maxima of the beamforming result occur at a biased position, if separate processing of the sub-radar measurements is considered. For the target being close to the radar the bias is severe, e.g. 0.5° for a target distance of 3 m. The reason is, that for a close target the sub-radars focus the target under different angles, hence, the sub-radars occur as spatially distributed w.r.t. the target. In other words, a spherical wave front occurs across the virtual aperture. As a result, different target direction are estimated. If the target is further apart, the bias vanishes as the sub-radars occur as co-located w.r.t. the target. Hence, a plane wave front is present across the virtual aperture. However, plane wave fronts can be assumed as always present across the virtual apertures of the sub-radars for the considered target distances. It can be concluded, that the consideration of the sub-radars as co-located or distributed depends on the relation of the sub-radar separation L and the target range R. If $L/R \ll 1$ holds, the sub-radars can be considered as co-located.

Based on this knowledge the beamforming result for the joint processing of the sub-radar measurements can be explained. First, the maximum is always at the position of the true target angle, but the beamforming results show strong side lobes. The position of the highest side lobe approximately coincides with the position of the beamforming maxima for separate processing, especially for the target being close. Because the sub-radars have different focusing angles for a close target, they cannot be jointly processed by assuming a common angle. Hence, a model error is present, leading to ambiguities as e.g. increased side lobes. These side lobes are not harmful in the single target case, as the maximum stays at the true target angle. However, in the multi target case the side lobes are harmful, as they may cause wrong detection by the beamforming-based initialisation scheme of the optimisation procedure (see Section 5.3). The maxima of the beamforming result stays at the position of the true target angle, because the beamformer attempts to compensate the model errors in a least-squares sense, resulting in a maxima at the position of an "average angle". Summarised, joint processing of the monostatic sub-radar measurements cannot be taken into account for close targets. As the target distance is unknown in advance, only separate processing will be considered in the following.

Stationary Metal Cylinders

Next, the localisation performance will be investigated. Therefore, the monostatic sub-radar measurements are separately processed by the parameter estimator. In total 8 MIMO snapshots are considered in the parameter estimator. The significance level for the model order estimation has been set to $\alpha_s = 0.05$. The locations of the detected point targets are calculated from the estimated azimuth and delay values using equation (7.2). The four metal cylinders depicted in Fig. 7.9(a) are used as radar target, which are stationary here. Three different rotation angles of the mounting platform are considered: 0°, 20° and 45°. The localisation results are shown in Fig. 7.18. It can be observed, that the localisation results from the processed monostatic measurements coincide, and the rectangular arrangement of the four metal cylinders can be identified. However, the localisation results do not totally agree with each other, as the metal cylinders have a diameter of 16 mm and each sub-radar may focus a different part of the cylinder, resulting in slightly different location results. Hence, by employing the measurements of both sub-radars the metal cylinders are recognised as slightly extended targets, rather than as point
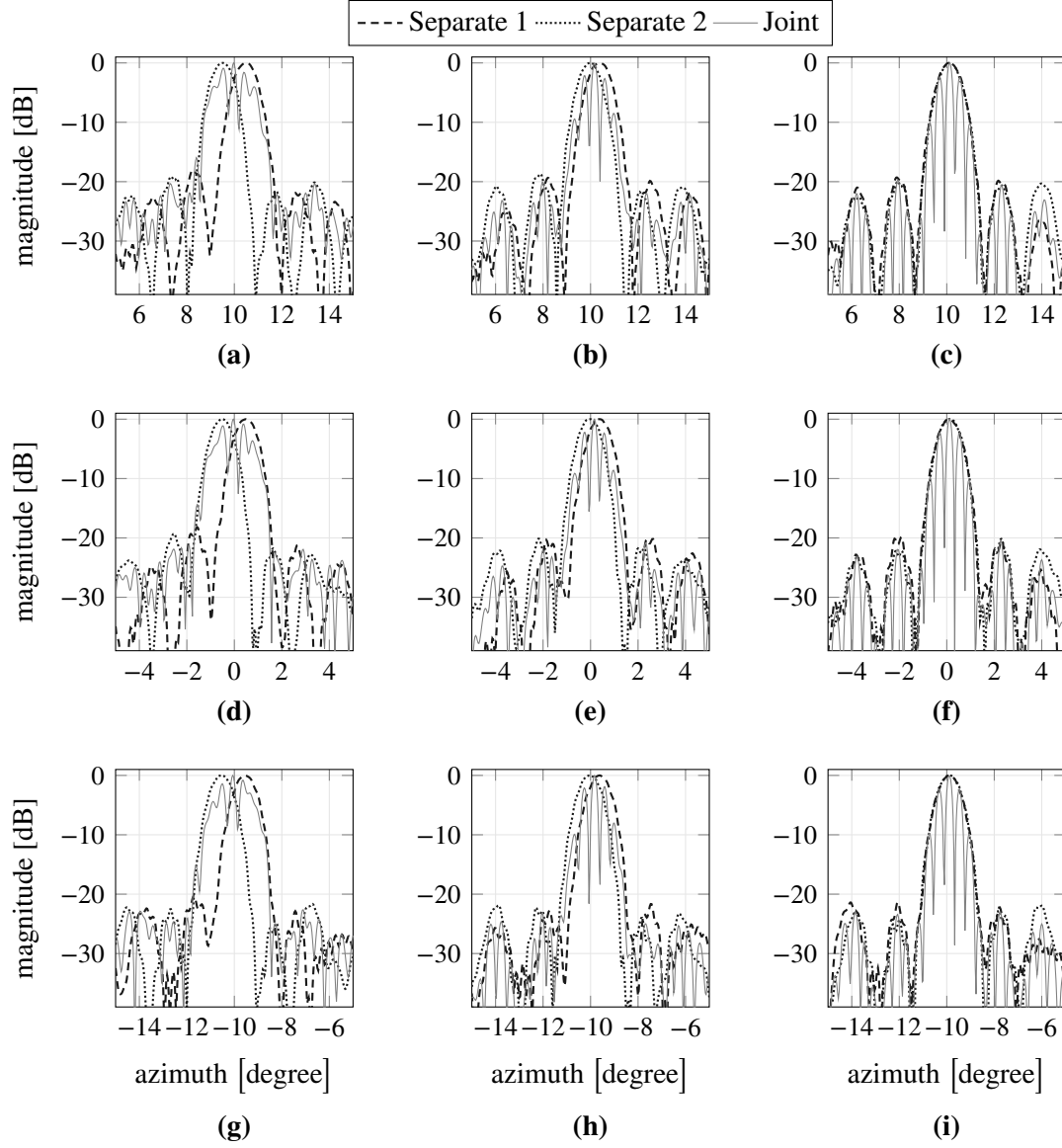
Figure 7.17: *Beamforming results in the azimuth domain using measurements with a corner reflector at distances 3 m (sub-figures (a), (d), (g)), 4 m (sub-figures (b), (e), (h)) and 5 m (sub-figures (c), (f), (i)); and angular positions −10°, 0° and 10°. The beamforming results are calculated for the separate processing (densely dashed and densely dotted line) and the joint processing (solid line) of the sub-radar measurements.*
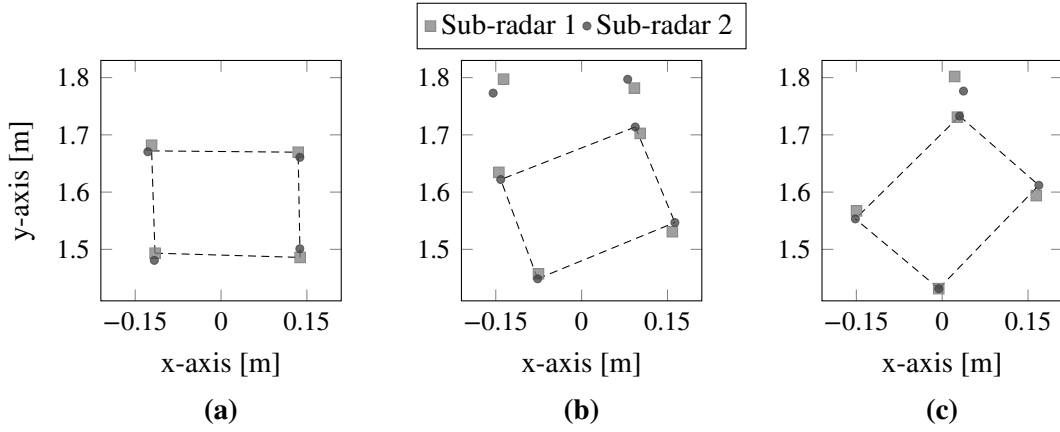
Figure 7.18: *Localisation results from separately processed sub-radar measurements for the target in Fig. 7.9(a). The rotation angles of the mounting platform are (a) 0°, (b) 20° and (c) 45°. The ground truth according to the target geometry is indicated by dashed lines.*

targets.

Rotating Metal Cube

The open metal cube (size: $20\,\text{cm} \times 20\,\text{cm} \times 20\,\text{cm}$) depicted in Fig. 7.19(a) will be considered next. The cube has three metal walls, two outside and one inside, such that 6 scattering edges are basically present in the x-y plane. The inside metal wall can be viewed, as the cube is open on two sides, which are turned to the radar. The cube is counter clock-wisely rotated with $50\,°\,\text{s}^{-1}$. Eight MIMO snapshots are employed for Doppler estimation, and the significance level is set to $\alpha_\text{s} = 0.05$. The localised scattering points by individually processing the monostatic sub-radar measurements are shown in Fig. 7.19(b). It is interestingly to note, that the 6 scattering edges of the cube are detectable only by considering the measurements of both sub-radars. This observation verifies the binocular idea, where more details of an extended target become detectable by using measurements of spatially distributed MIMO radars.

Rotating Wooden Box

Last, a wooden box of size $44\,\text{cm} \times 32\,\text{cm} \times 26\,\text{cm}$ (width $\times$ height $\times$ depth) is considered as target, see Fig. 7.20a. The box is counter-clock wise rotated with $35\,°\,\text{s}^{-1}$. Eight MIMO snapshots are used for Doppler estimation, and the significance level is set to $\alpha_\text{s} = 0.05$. Again, only the monostatic sub-radar measurements are considered, and are separately processed by the parameter estimator. The localised point scatterers are shown in Fig. 7.20b. Both sub-radars view the same scattering points of the wooden box, such that one sub-radar would be sufficient in that case.

7.4.3 Summary

A novel radar setup, the BinoMimo radar, has been briefly introduced. The radar consists of two spatially separated, and coherently and cooperatively operating MIMO sub-radars.
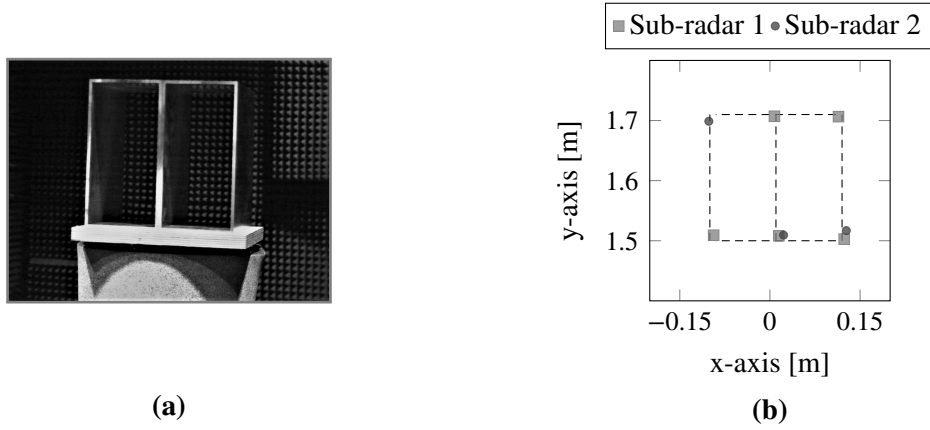
(a) (b)

Figure 7.19: *(a) picture of the open metal cube and (b) located point targets from the processed sub-radar measurements. The ground truth according to the target geometry is indicated by dashed lines.*
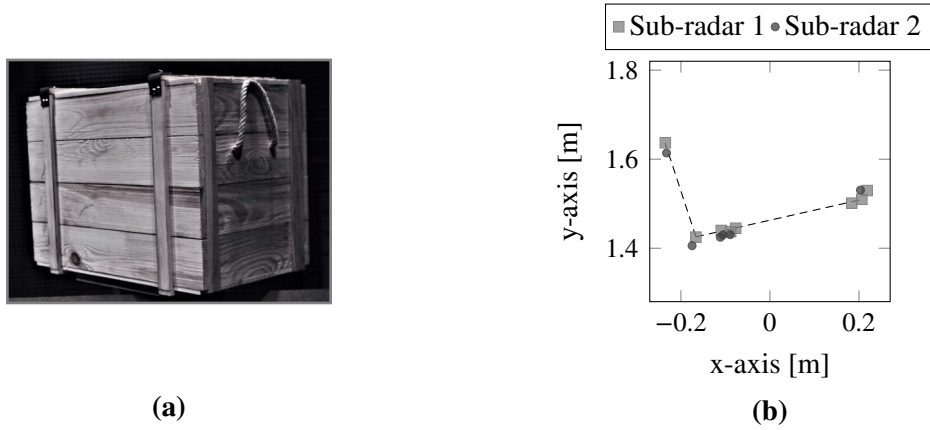

(a) (b)

Figure 7.20: *(a) picture of a wooden box and (b) located point scatterers from the processed sub-radar measurements. The ground truth according to the target geometry is indicated by dashed lines.*

The coherent operation allows joint processing of the measurements of both sub-radars, resulting in a virtual radar featuring a huge aperture. It has been shown, that the joint processing necessitates a target, which distance apart from the radar is much larger than the spatial separation of the sub-radars. Otherwise, spherical wave fronts occur across the large virtual aperture, resulting in erroneous direction estimates.

The developed parameter estimator has been applied to measurements taken by the BinoMimo radar. Because the target distance is unknown in advance, only separate processing of the sub-radar measurements has been considered to prevent the problem of spherical wave fronts. It could by shown, that more targets can be identified by evaluating the measurements of both sub-radars. Hence, the radar offers a more robust target detection and better recognition of target structures.

CHAPTER 8

SUMMARY AND OUTLOOK

PROCESSING of gathered radar observations has been treated as a model-based system identification problem in this thesis. The system under identification is the radar channel (or equivalently the wireless propagation channel) featuring radar targets, which is sensed by the radar system. This, however, is a quite uncommon treatment of radar signal processing. However, it provides a more general view on radar signal processing, which is not limited to the consideration of a certain radar hardware architecture, waveform of the radar signal or application. The considered system identification is model-based. A parametric model of the system under identification, termed as physical model, has been stated and the respective model parameters are subject to estimation. Accordingly, the system identification step becomes the parameter estimation step. The model parameters are related to the radar targets. By identifying the radar channel the targets are inherently detected and the estimated parameters can be exploited to localise or classify the respectively detected targets.

The sensing process causes the parameters of the model to be mapped to the observations. This mapping can be treated as a forward problem. The estimation of the model parameters from the observations is the reverse of this process. Hence, the parameter estimation is an inverse problem. The necessary basics of inverse problems have been revised, indicating, that a model of the observation process is required to solve the inverse problem. This model describes the embedding of the physical model and its respective parameters in the observations. As the embedding depends on the observer itself, this model cannot be the raw physical model. Therefore, a system model is required, which consists of the physical model and a model of the observer, termed as device model. The device model describes the observer, e.g. it is a model of the device response, and can depend on additional unknown parameters, which are subject to estimation too. These parameters are typically not of immediate interest and are therefore termed as nuisance parameters. Due to noise processes of the observer and uncertainties of the considered system model disturbances are present. These disturbances are treated as random, because of the lack of information and reasons of complexity. A statistical model, which subsists of the system model and a model of the noise processes, is required to account for these disturbances. As a consequence, the inverse problem casts into a stochastic inverse problem. The stochastic inverse problem is solved by Bayes' theorem, where a distribution of the model parameters after the observation process is assigned. Point estimators are used to deduce the intended parameters from their assigned distribution. Two well known point estimators, namely the MAP and MMSE point estimator, have been discussed and the MAP has been chosen due to its simpler calculation.

A physical model describing the propagation of EMWs in the radar channel has been derived from the principle of ray optics. Geometrical parameters as e.g. angles of incidence and departure or a propagation delay are assigned to each propagation path, which emerges from scattering at a point target. Each propagation path is modelled as a weighted multidimensional Dirac function, with a Dirac function for each geometrical dimension and dedicated parameter. The final physical model is given by the superposition of the propagation paths. Due to the Dirac function high-resolution, i.e., a resolution higher than the bandwidth or aperture determined resolution, is achievable to resolve each path. Subject to clarification is the modelling of the weight, which describes the time dispersion and variation of the wave due to interactions with scatterers. The weight has to be considered as a function, e.g. as an impulse response, under broadband considerations. A model of this function is required to keep the inverse problem unambiguous. However, to the best of the authors knowledge, such a model is not available. Hence, the narrowband assumption has been stated, where the weight becomes a scalar or, equivalently, a constant over the whole frequency band. This is a crucial simplification, because the considered radar system is a broadband system, which makes the narrowband assumption hard to be fulfilled.

A broadband, monostatic mmWave FMCW MIMO radar has been exemplary considered as observer, and a respective device model has been developed. The basic signal theory of FMCW radars has been reviewed. A linear response model has been proposed to describe the influence of the radar hardware on the observations. It has been turned out, that a linear device model is not sufficient due to hardware impairments, which cause strong signal distortions. Therefore, a novel non-linear device model has been proposed, which can be incorporated in the parameter estimation framework. It has been shown, that the signal distortions due to the hardware impairments are mitigated and the target detection accuracy is improved. Estimation of angles of incidence and departure of a propagation path is accomplished by using antenna arrays at Tx and Rx side of the observer. The response of the array in terms of the angles has to be known in order to estimate the respective angles. Commonly, a narrowband model is employed, where the array response is an angle and polarisation dependent scalar. However, because the considered radar is a broadband system (large signal bandwidth and large array aperture) angle-delay coupling occurs. Hence, the narrowband model is not sufficient. A broadband model has been introduced, which has been proven to outperform the narrowband model. Furthermore, the influence of TDM to access all MIMO channels has been discussed. The successive antenna switching results in a angle-Doppler coupling, because the Doppler phase evolves due to continuously moving targets. An extension of the Doppler model to cope with the angle-Doppler coupling has been proposed, which can be easily incorporated in the parameter estimator. Monostatic radars, i.e. radars with co-located Tx and Rx, suffer from leakage of the transmit signal into the receiver. Leakage can be due to the limited isolation between the Tx and Rx or due to hardware obstacles in the close vicinity of the radar (e.g. radome, car bumper, lens). Leakage can cause saturations of the receiver (especially in impulse radars) and short-range distortions of the receive signal, such that close and weak targets cannot be detected. A novel approach to mitigate the short-range distortions has been proposed. Background subtraction is used to mitigate the leakage. Remaining signal portions are treated as coloured, noise process and are termed as leakage noise. These signal portions are "suppressed" by whitening with a model of their covariance. As a result, short-range targets become detectable. Note, that saturation of the receiver cannot be avoided and corresponding signal distortions cannot be repaired by this approach! Radar observations are confounded by random noise. As explained, this turns

the inverse problem into a stochastic inverse problem. Here, three noise sources have been considered in the parameter estimation framework: thermal noise, leakage noise and phase noise of the radar hardware. Taking into account leakage noise and phase noise in the estimation process is a novelty in radar signal processing. The consideration of phase noise is additionally beneficial to treat uncertainties of the used system model, because model uncertainties cause similar errors as phase noise. Hence, by coping with phase noise model errors are additionally coped with. Statistical models, i.e. the distributions, of these noise sources have been proposed and the respective model parameters are nuisance parameters in the estimator.

Using the system model (physical model + device model) and the statistical model of all noise sources (thermal, leakage and phase noise), the statistical model of the radar observations, the likelihood, has been formulated. Based on the likelihood and prior informations regarding the model parameters, a MAP estimator has been derived. MAP estimation results in a high-dimensional objective function, which has to be optimised w.r.t. the model parameters. Because the objective is high-dimensional and also non-convex, its optimisation is computationally expensive. Exhaustive search has to be avoided and a computationally more attractive optimisation method has been presented. There, a coarse initial search for the optimum is followed by a gradient-based optimisation algorithm for refinement. Nevertheless, the computational effort is still high, which prevents the application of the estimator for online or real-time estimation. Instead, the parameter estimator can be considered as a benchmark regarding achievable resolution, mitigation of hardware impairments, coping with noise processes or issues related to the measurement principle.

Finally, an algorithm to estimate the number of targets, or respectively the number of propagation paths, has been presented. The stepwise regression search, where the number of considered targets is incremented until an abortion criterion is met, has been proposed. As abortion criterion parametric hypothesis tests, namely the Wald amplitude and the likelihood-ratio test, haven been taken into account. It has been shown, that the theoretic distributions of these tests do not apply, because the system model is partially non-linear in its parameters. Improvements, which use stochastic simulation techniques like bootstrap sampling and Monte Carlo simulations, have been used in order to get a more proper estimate of the distributions. Also, methods to cope with the problem of multi-hypothesis testing, which occur due to the stepwise regression search, have been discussed and the Bonferroni method has been chosen as best suited method.

The performance of the proposed estimator has been investigated by various simulations and practical experiments. The simulations indicate, that the estimator compensates the angle-Doppler and angle-delay coupling, which occur due to the TDM and the utilisation of a broadband radar. Therefore, the developed estimator outperforms standard estimators. Experiments have been conducted in an anechoic and various objects (corner reflector, metal cube, metallic cylinders) have been deployed as radar targets. The estimated ranges and angles have been used to localise the respective point targets. The localisation results show, that the estimator properly detects scatterers at the deployed targets. Extended targets become recognisable and the estimates can be applied to identify the geometric structure of extended targets. Last, the estimator has been applied to observations with a novel radar system, a binocular MIMO radar (the BinoMimo radar). This radar composes of two spatially separated, cooperative and coherent operating sub-radars, which are monostatic MIMO radars each. Hence, mono- as well as bistatic observations are gathered, and a suitable estimator is required in order to fully exploit the accompanying benefits of this radar, e.g. improved target detection performance or clutter

suppression. Limits of the applicability of the proposed estimator for the BinoMimo radar have been discussed. It has been shown, that the target range from the radar must be sufficiently large in order to jointly process the observations of both sub-radars (same angle of the target to each sub-radar). Otherwise, the sub-radar observations have to be processed individually, because the angle of the target to each sub-radar is different. Last, the benefit of the BinoMimo radar in terms of target recognition has been proven. It has been shown, that more point scatterers at an extended object are detectable, which improves the identifiability of the respective geometric structure.

A basic drawback of the developed estimator (or the considered radar itself) is the limitation to targets in the azimuth plane. ULAs are used at the co-located Tx and Rx. Hence, only the azimuth angle can be estimated, because the resulting virtual array is an ULA too. Furthermore, azimuth estimates become erroneous for targets out of the azimuth plane. As a result, such targets will be wrongly localised. A solution to this issue is the usage of an URA, which is capable to estimate azimuth and elevation. Hence, targets in the 3D space can be localised. A URA can be realised by a monostatic MIMO radar, where the ULAs at Tx and Rx are flipped by 90° to each other. As a result, a virtual URA is formed.

The implemented parameter estimator allows the estimation of two angles, and is flexible regarding the array model. Hence, the processing of bistatic radar observations with an ULA at Tx and Rx, i.e. two azimuth angles have to be estimated, is possible. Also, observations with a monostatic radar featuring a virtual URA as mentioned above can be processed. Note, that only the case of a virtual ULA, i.e. one azimuth angle has to be estimated, has been covered in this thesis. However, the required change of the statistical model, precisely the array model of the system model, is straightforward. The ability to process mono- and bistatic radar observations is a necessity to process all observations of the BinoMimo radar, because this radar captures both types of radar observations.

Last, the estimator does not assign or account for any relations of the detected point targets. Relations occur, when the point targets are due to scattering from a common target like an extended target. The developed estimator provides only estimates of point targets. Hence, no target identification or classification is conducted by the estimator. The developed estimator can be extended to do so, which requires a target model to assign geometric relations to the point targets. The identified geometric relations can be exploited to finally identify the object type.

BIBLIOGRAPHY

List of References

[3GP]   3GPP - 3GPP2 Spatial Channel Model Ad-hoc Group, 3GPP TR 25.996. *Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations*. Tech. rep. v6.1.0 (2003-09.

[AS72]   M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Vol. 9. Applied mathematics series. Dover Publications, 1972.

[ASS11]   T. Adali, P. J. Schreier and L. L. Scharf. 'Complex-Valued Signal Processing: The Proper Way to Deal With Impropriety'. In: *IEEE Transactions on Signal Processing* 59.11 (Nov. 2011), pp. 5101–5125.

[AW03]   R. J. Allard and D. H. Werner. 'The model-based parameter estimation of antenna radiation patterns using windowed interpolation and spherical harmonics'. In: *IEEE Transactions on Antennas and Propagation* 51.8 (Aug. 2003), pp. 1891–1906.

[And03]   S. J. Anderson. 'Inverse Problems in HF Radar'. In: *Coupling of Fluids, Structures and Waves in Aeronautics*. Ed. by Noel G. Barton and Jacques Periaux. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 143–157.

[Arn86]   J. M. Arnold. 'Geometrical theories of wave propagation: a contemporary review'. In: *IEE Proceedings* 133.2 (Apr. 1986), pp. 165–184.

[ABT13]   Richard C. Aster, Brian Borchers and Clifford H. Thurber. *Parameter Estimation and Inverse Problems*. second. Boston: Academic Press, 2013.

[ADA14]   Stefanie R. Austin, Isaac Dialsingh and Naomi S. Altman. 'Multiple Hypothesis Testing: A Review'. In: *Journal of the Indian Society of Agricultural Statistics* 68 (Jan. 2014), pp. 303–314.

[Avi+12]   E. Avignon-Meseldzija, Weixian Liu, Hongchuan Feng, S. Azarian, M. Lesturgie and Yilong Lu. 'Compensation of Analog Imperfections In a Ka-band FMCW SAR'. In: *9th European Conference on Synthetic Aperture Radar (EUSAR)*. Apr. 2012, pp. 60–63.

[Ayh+16]   S. Ayhan, S. Scherr, A. Bhutani, B. Fischbach, M. Pauli and T. Zwick. 'Impact of Frequency Ramp Nonlinearity, Phase Noise, and SNR on FMCW Radar Accuracy'. In: *IEEE Transactions on Microwave Theory and Techniques* 64.10 (Oct. 2016), pp. 3290–3301.

[Bal18]   B. Balaji. 'Quantum Radar: Snake Oil or Good Idea?' In: *International Carnahan Conference on Security Technology (ICCST)*. Oct. 2018, pp. 1–7.

[Bal97]   Constantine A. Balanis. *Antenna Theory: Analysis and Design*. Second. John Wiley Sons, 1997.

[Bea+90]   P. D. L. Beasley, A. G. Stove, B. J. Reits and B. As. 'Solving the problems of a single antenna frequency modulated CW radar'. In: *IEEE International Radar Conference*. May 1990, pp. 391–395.

[BRW17]   J. Bechter, F. Roos and C. Waldschmidt. 'Compensation of Motion-Induced Phase Errors in TDM MIMO Radars'. In: *IEEE Microwave and Wireless Components Letters* 27.12 (Dec. 2017), pp. 1164–1166.

[BRH14]   F. Belfiori, W. v. Rossum and P. Hoogeboom. 'Coherent MUSIC technique for range/angle information retrieval: application to a frequency-modulated continuous wave MIMO radar'. In: *IET Radar, Sonar Navigation* 8.2 (Feb. 2014), pp. 75–83.

[Ber06]   James Berger. 'The case for objective Bayesian analysis'. In: *Bayesian Anal.* 1.3 (Sept. 2006), pp. 385–402.

[BS94]   J.M. Bernardo and A.F.M. Smith. *Bayesian Theory*. Wiley Series in Probability & Statistics. Wiley, 1994.

[BE19]   A. Bhandari and Y. C. Eldar. 'Sampling and Super Resolution of Sparse Signals Beyond the Fourier Domain'. In: *IEEE Transactions on Signal Processing* 67.6 (2019), pp. 1508–1521.

[BF03]   D. W. Bliss and K. W. Forsythe. 'Multiple-input multiple-output (MIMO) radar and imaging: degrees of freedom and resolution'. In: *The Thrity-Seventh Asilomar Conference on Signals, Systems Computers*. Vol. 1. Nov. 2003, 54–59 Vol.1.

[BTB19]   M. Bonnin, F. L. Traversa and F. Bonani. 'Colored Noise in Oscillators. Phase-Amplitude Analysis and a Method to Avoid the Itô-Stratonovich Dilemma'. In: *IEEE Transactions on Circuits and Systems I: Regular Papers* 66.10 (Oct. 2019), pp. 3917–3927.

[Box76]   George E. P. Box. 'Science and Statistics'. In: *Journal of the American Statistical Association* 71.356 (1976), pp. 791–799.

[BV04]   Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. New York: Cambridge University Press, 2004.

[Bro+09]   Rasmus Bro, Richard A. Harshman, Nicholas D. Sidiropoulos and Margaret E. Lundy. 'Modeling multi-way data with linearly dependent loadings'. In: *Journal of Chemometrics* 23.7-8 (Jan. 2009), pp. 324–340.

[Bro03]   E. Brookner. 'Phased arrays around the world - progress and future trends'. In: *IEEE International Symposium on Phased Array Systems and Technology*. Oct. 2003, pp. 1–8.

[BB93]   M. C. Budge and M. P. Burt. 'Range correlation effects in radars'. In: *The Record of the 1993 IEEE National Radar Conference*. Apr. 1993, pp. 212–216.

[BA02]   K.P. Burnham and D.R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd ed. Springer Verlag, 2002.

[Bus82]   A. Buse. 'The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note'. In: *The American Statistician* 36.3 (1982), pp. 153–157.

[CW08]   E. J. Candes and M. B. Wakin. 'An Introduction To Compressive Sampling'. In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 21–30.

[CF13]   Emmanuel J. Candès and Carlos Fernandez-Granda. 'Super-Resolution from Noisy Data'. In: *Journal of Fourier Analysis and Applications* 19.6 (Dec. 2013), pp. 1229–1254.

[CF14]   Emmanuel J. Candès and Carlos Fernandez-Granda. 'Towards a Mathematical Theory of Super-resolution'. In: *Communications on Pure and Applied Mathematics* 67.6 (2014), pp. 906–956.

[Can05]   James V. Candy. *Model-Based Signal Processing*. Wiley-IEEE Press, Oct. 2005, p. 704.

[CG92]   C. K. Chen and W. A. Gardner. 'Signal-selective time-difference of arrival estimation for passive location of man-made signal sources in highly corruptive environments. II. Algorithms and performance'. In: *IEEE Transactions on Signal Processing* 40.5 (May 1992), pp. 1185–1197.

[CS98]   V. C. Chen and Shie Qian. 'Joint Time-Frequency Transform for Radar Range-Doppler imaging'. In: *IEEE Transactions on Aerospace and Electronic Systems* 34.2 (Apr. 1998), pp. 486–499.

[Chu+07]   P. J. Chung, J. F. Böhme, C. F. Mecklenbräuker and A. O. Hero. 'Detection of the Number of Signals Using the Benjamini-Hochberg Procedure'. In: *IEEE Trans. Signal Process.* 55.6 (June 2007), pp. 2497–2508.

[Coh+18]   D. Cohen, D. Cohen, Y. C. Eldar and A. M. Haimovich. 'SUMMeR: Sub-Nyquist MIMO Radar'. In: *IEEE Transactions on Signal Processing* 66.16 (2018), pp. 4315–4330.

[CL94]   Thomas Coleman and Yuying Li. 'On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds'. In: *Mathematical Programming* 67.2 (Oct. 1994), pp. 189–224.

[CL96]   Thomas F. Coleman and Yuying Li. 'An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds'. In: *SIAM Journal on Optimization* 6.2 (1996), pp. 418–445.

[Com14]   IEEE Antennas Committee. 'IEEE Standard for Definitions of Terms for Antennas'. In: *IEEE Std 145-2013 (Revision of IEEE Std 145-1993)* (Mar. 2014), pp. 1–50.

[CB93]   C.E. Cook and M. Bernfeld. *Radar Signals: An Introduction to Theory and Application*. Artech House radar library. Artech House, 1993.

[Cos+09]   J.P.C.L. da Costa, A. Thakre, F. Roemer and M. Haardt. 'Comparison of Model Order Selection Techniques for High-Resolution Parameter Estimation Algorithms'. In: *54th Internationales Wissenschaftliches Kolloquium*. Sept. 2009.

[CRK10]   M. Costa, A. Richter and V. Koivunen. 'Unified Array Manifold Decomposition Based on Spherical Harmonics and 2-D Fourier Basis'. In: *IEEE Transactions on Signal Processing* 58.9 (Sept. 2010), pp. 4634–4645.

[Cou+12]   N.R. Council, D.E.P. Sciences, B.M.S.T. Applications and V.U.Q. Committee on Mathematical Foundations of Verification. *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. National Academies Press, 2012.

[CSW02]   R. J. M. Cramer, R. A. Scholtz and M. Z. Win. 'Evaluation of an ultra-wide-band propagation channel'. In: *IEEE Transactions on Antennas and Propagation* 50.5 (May 2002), pp. 561–570.

[CC96]   Francisco Cribari-Neto and Gauss M. Cordeiro. 'On bartlett and bartlett-type corrections'. In: *Econometric Reviews* 15.4 (1996), pp. 339–367.

[Czi07]   Nicolai Czink. 'The Random-Cluster Model – A Stochastic MIMO Channel Model for Broadband Wireless Communication Systems of the 3rd Generation and Beyond'. PhD thesis. Technischen Universität Wien, 2007.

[DS17]   Masoumeh Dashti and Andrew M. Stuart. 'The Bayesian Approach to Inverse Problems'. In: *Handbook of Uncertainty Quantification*. Ed. by Roger Ghanem, David Higdon and Houman Owhadi. Cham: Springer International Publishing, 2017, pp. 311–428.

[DDV00]   L. De Lathauwer, B. De Moor and J. Vandewalle. 'A Multilinear Singular Value Decomposition'. In: *SIAM Journal on Matrix Analysis and Applications* 21.4 (2000), pp. 1253–1278.

[DEH19]   A. De Maio, Y. Eldar and A. Haimovich. *Compressed Sensing in Radar Signal Processing*. Cambridge University Press, 2019.

[Del+06]   G. Del Galdo, J. Lotze, M. Landmann and M. Haardt. 'Modelling and manipulation of Polarimetric antenna beam patterns via spherical harmonics'. In: *Signal Processing Conference, 2006 14th European*. Sept. 2006, pp. 1–5.

[DD94]   M.A. Doron and E. Doron. 'Wavefield modeling and array processing, Part I-Spatial Sampling'. In: *Signal Processing, IEEE Transactions on* 42.10 (Oct. 1994), pp. 2549–2580.

[DR10]   W. Dullaert and H. Rogier. 'Novel Compact Model for the Radiation Pattern of UWB Antennas Using Vector Spherical and Slepian Decomposition'. In: *Antennas and Propagation, IEEE Transactions on* 58.2 (Feb. 2010), pp. 287–299.

[DG00]   Ivo Düntsch and Günther Gediga. *Sets, Relations, Functions*. Methodos Publ., 2000.

[DBG08]   C. Duofang, C. Baixiao and Q. Guodong. 'Angle estimation using ESPRIT in MIMO radar'. In: *Electronics Letters* 44.12 (June 2008), pp. 770–771.

[Dür+20]   A. Dürr, R. Kramer, D. Schwarz, M. Geiger and C. Waldschmidt. 'Calibration-Based Phase Coherence of Incoherent and Quasi-Coherent 160-GHz MIMO Radars'. In: *IEEE Transactions on Microwave Theory and Techniques* 68.7 (2020), pp. 2768–2778.

[Dür+19]   A. Dürr, D. Schwarz, F. Roos, P. Hügler, S. Bucher, P. Grüner and C. Waldschmidt. 'On the Calibration of mm-Wave MIMO Radars Using Sparse Antenna Arrays for DoA Estimation'. In: *2019 16th European Radar Conference (EuRAD)*. 2019, pp. 349–352.

[Dür+18]   A. Dürr, B. Schweizer, J. Bechter and C. Waldschmidt. 'Phase Noise Mitigation for Multistatic FMCW Radar Sensor Networks Using Carrier Transmission'. In: *IEEE Microwave and Wireless Components Letters* 28.12 (Dec. 2018), pp. 1143–1145.

[DSW19]   A. Dürr, B. Schweizer and C. Waldschmidt. 'Leakage Phase Noise Mitigation for Monostatic FMCW Radar Sensors Using Carrier Transmission'. In: *2019 IEEE MTT-S International Microwave Symposium (IMS)*. 2019.

[Efr79]   B. Efron. 'Bootstrap Methods: Another Look at the Jackknife'. In: *The Annals of Statistics* 7.1 (1979), pp. 1–26.

[ETT06]   T. Eltoft, Taesu Kim and Te-Won Lee. 'On the multivariate Laplace distribution'. In: *IEEE Signal Processing Letters* 13.5 (May 2006), pp. 300–303.

[Eng84]   Robert F. Engle. 'Chapter 13 Wald, likelihood ratio, and Lagrange multiplier tests in econometrics'. In: *Handbook of Econometrics*. Vol. 2. Handbook of Econometrics. Elsevier, 1984, pp. 775–826.

[Eri01]   J. Eriksson. 'Multiple Signal Detection Under a Specified False Alarm Constraint'. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Vol. 5. May 2001, pp. 2873–2876.

[Fer16]   Carlos Fernandez-Granda. 'Super-resolution of point sources via convex programming'. In: *Information and Inference: A Journal of the IMA* 5.3 (2016), pp. 251–303.

[FH94]   J. A. Fessler and A. O. Hero. 'Space-Alternating Generalized Expectation-Maximization Algorithm'. In: *IEEE Transactions on Signal Processing* 42.10 (Oct. 1994), pp. 2664–2677.

[Fet+05]   G. Fettweis, M. Lohning, D. Petrovic, M. Windisch, P. Zillmann and W. Rave. 'Dirty RF: a new paradigm'. In: *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*. Vol. 4. Sept. 2005, 2347–2355 Vol. 4.

[Fet+07]   Gerhard Fettweis, Michael Löhning, Denis Petrovic, Marcus Windisch, Peter Zillmann and Wolfgang Rave. 'Dirty RF: A New Paradigm'. In: *International Journal of Wireless Information Networks* 14.2 (June 2007), pp. 133–148.

[FJ15]   J. Fink and F. K. Jondral. 'A Numerical Comparison of Chirp Sequence versus OFDM Radar Waveforms'. In: *IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*. Sept. 2015, pp. 1–2.

[Fis+04]   E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini and R. Valenzuela. 'MIMO Radar: An Idea Whose Time Has Come'. In: *IEEE Radar Conference*. Apr. 2004, pp. 71–78.

[Fis+04]   E. Fishler, A. Haimovich, R. Blum, R. Cimini, D. Chizhik and R. Valenzuela. 'Performance of MIMO Radar Systems: Advantages of Angular Diversity'. In: *38th Asilomar Conference on Signals, Systems and Computers*. Vol. 1. Nov. 2004, 305–309 Vol.1.

[Fis+06]   E. Fishler, A. Haimovich, R. S. Blum, L. J. Cimini, D. Chizhik and R. A. Valenzuela. 'Spatial Diversity in Radars - Models and Detection Performance'. In: *IEEE Transactions on Signal Processing* 54.3 (Mar. 2006), pp. 823–838.

[FJS02]   B.H. Fleury, P. Jourdan and A. Stucki. 'High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm'. In: *International Zurich Seminar on Broadband Communications, Access, Transmission, Networking*. 2002, pp. 31–39.

[Fle+99]   B.H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus and K. Ingeman Pedersen. 'Channel Parameter Estimation in Mobile Radio Environments using the SAGE Algorithm'. In: *IEEE Journal on Selected Areas in Communication* 17.3 (Mar. 1999), pp. 434–450.

[Fri09]   B. Friedlander. 'On the Relationship Between MIMO and SIMO Radars'. In: *IEEE Transactions on Signal Processing* 57.1 (Jan. 2009), pp. 394–398.

[Fue09]   Edgar Fuentes. *Example: Oblique Incidence*. Sept. 2009. URL: https://texample.net/tikz/examples/oblique-incidence/ (visited on 23/10/2020).

[Gal77]   Andrew Ronald Gallant. 'Testing a Nonlinear Regression Specification: A Nonregular Case'. In: *Journal of the American Statistical Association* 72.359 (Sept. 1977), pp. 523–530.

[GL06]   Dani Gamerman and Hedibert F. Lopes. *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference*. second. Boca Raton, Fla: CRC Press, 2006.

[Gil+05]   Kevin Gilholm, Simon Godsill, Simon Maskell and David Salmond. *Poisson models for extended target and group tracking*. 2005.

[Gol06]   Michael Goldstein. 'Subjective Bayesian Analysis: Principles and Practice'. In: *Bayesian Anal.* 1.3 (Sept. 2006), pp. 403–420.

[GP73]   G. Golub and V. Pereyra. 'The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate'. In: *SIAM Journal on Numerical Analysis* 10.2 (1973), pp. 413–432.

[GP03]   Gene Golub and Victor Pereyra. 'Separable Nonlinear Least Squares: the Variable Projection Method and its Applications'. In: *Inverse Problems* 19.2 (Feb. 2003), pp. 1–26.

[GV13]   Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. 4th ed. Baltimore, MD, USA: Johns Hopkins University Press, 2013.

[Goo19]   Joanna Goodrich. *How the Telemobiloskop Paved the Way for Modern Radar Systems*. 2019. URL: https://spectrum.ieee.org/the-institute/ieee-history/how-the-telemobiloskop-paved-the-way-for-modern-radar-systems (visited on 24/04/2020).

[GW10]   H. Griffiths and N. Willis. 'Klein Heidelberg—The First Modern Bistatic Radar System'. In: *IEEE Transactions on Aerospace and Electronic Systems* 46.4 (2010), pp. 1571–1588.

[GB17]   H.D. Griffiths and C.J. Baker. *An Introduction to Passive Radar*. Artech House radar library. Artech House Publishers, 2017.

[Gri+14]   M. Grimm, M. Allén, J. Marttila, M. Valkama and R. Thomä. 'Joint Mitigation of Nonlinear RF and Baseband Distortions in Wideband Direct-Conversion Receivers'. In: *IEEE Transactions on Microwave Theory and Techniques* 62.1 (Jan. 2014), pp. 166–182.

[Gri14]   Michael Grimm. 'Dirty RF Signal Processing for Mitigation of Receiver Front-end Non-linearity'. PhD thesis. Technische Universität Ilmenau, July 2014.

[Gro17]   T. Grosch. 'Correcting Nonlinear Modulation Error in Linear FMCW Radar Systems'. In: *IEEE Radar Conference (RadarConf)*. May 2017, pp. 1577–1581.

[Gua10]   Massimo Guarnieri. 'The Early History of Radar'. In: *Industrial Electronics Magazine, IEEE* 4 (Oct. 2010), pp. 36–42.

[GH08]   F. Gustafsson and G. Hendeby. 'On nonlinear transformations of stochastic variables and its application to nonlinear filtering'. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Mar. 2008, pp. 3617–3620.

[HTR03]   Martin Haardt, Reiner Thomä and Andreas Richter. 'Multidimensional High-Resolution Parameter Estimation with Applications to Channel Sounding'. In: *High-Resolution and Robust Signal Processing*. Ed. by Yingbo Hua, Alex Gershman and Qi Cheng. CRC Press, 2003. Chap. 5, pp. 253–332.

[Hab99]   Robert Haber. *Nonlinear system identification : input-output modeling approach*. Dordrecht Boston: Kluwer Academic Publishers, 1999.

[Had02]   Jacques Hadamard. 'Sur les problèmes aux dérivés partielles et leur signification physique'. In: *Princeton University Bulletin* 13 (1902), pp. 49–52.

[HBC08]   A. M. Haimovich, R. S. Blum and L. J. Cimini. 'MIMO Radar with Widely Separated Antennas'. In: *IEEE Signal Processing Magazine* 25.1 (2008), pp. 116–129.

[HTK06]   K. Haneda, J.-i. Takada and T. Kobayashi. 'A parametric UWB propagation channel estimation and its performance validation in an anechoic chamber'. In: *Microwave Theory and Techniques, IEEE Transactions on* 54.4 (June 2006), pp. 1802–1811.

[Han17]   Martin Hanke. *A Taste of Inverse Problems*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2017.

[HV10]   A. Hassanien and S. A. Vorobyov. 'Phased-MIMO Radar: A Tradeoff Between Phased-Array and MIMO Radars'. In: *IEEE Transactions on Signal Processing* 58.6 (June 2010), pp. 3137–3151.

[Hay94]   Simon Haykin. *Communication Systems*. third. New York: Wiley, 1994.

[HMS15]   Reinhard Heckel, Veniamin I. Morgenshtern and Mahdi Soltanolkotabi. *Super-Resolution Radar*. Aug. 2015. arXiv: 1411.6272 [cs.IT].

[Hei09]   Jeff Hein. *Example: Polarization state of light*. Aug. 2009. URL: https://texample.net/tikz/examples/polarization-state-of-light (visited on 23/10/2020).

[Hei09]   Jeff Hein. *Example: The 3dplot package*. Nov. 2009. URL: https://texample.net/tikz/examples/the-3dplot-package/ (visited on 23/10/2020).

[HK17]   F. Herzel and D. Kissinger. 'Phase Noise Analysis of a Homodyne Radar System Driven by a Phase-Locked Loop'. In: *IEEE International Symposium on Circuits and Systems (ISCAS)*. May 2017, pp. 1–4.

[Hit+18]   M. Hitzler, P. Grüner, L. Boehm, W. Mayer and C. Waldschmidt. 'On Monostatic and Bistatic System Concepts for mm-Wave Radar MMICs'. In: *IEEE Trans. Microw. Theory Techn.* 66.9 (Sept. 2018), pp. 4204–4215.

[Hit+17]   M. Hitzler, S. Saulig, L. Boehm, W. Mayer, W. Winkler, N. Uddin and C. Waldschmidt. 'Ultracompact 160-GHz FMCW Radar MMIC With Fully Integrated Offset Synthesizer'. In: *IEEE Transactions on Microwave Theory and Techniques* 65.5 (May 2017), pp. 1682–1691.

[Jac+15]   B. R. Jackson, S. Rajan, B. J. Liao and S. Wang. 'Direction of Arrival Estimation Using Directive Antennas in Uniform Circular Arrays'. In: *IEEE Transactions on Antennas and Propagation* 63.2 (Feb. 2015), pp. 736–747.

[JAA17]   S. Jardak, S. Ahmed and M. Alouini. 'Low Complexity Moving Target Parameter Estimation for MIMO Radar Using 2D-FFT'. In: *IEEE Transactions on Signal Processing* 65.18 (Sept. 2017), pp. 4745–4755.

[JC17]   Donald E. Jarvis and Corey D. Cooke. *Application of Statistical Linear Time-Varying System Theory to Modeling of High Grazing Angle Sea Clutter*. Tech. rep. AD1040982. Oct. 2017.

[Jay68]   E. T. Jaynes. 'Prior Probabilities'. In: *IEEE Transactions on Systems Science and Cybernetics* 4.3 (Sept. 1968), pp. 227–241.

[KS05]   Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*. Applied Mathematical Sciences. Springer New York, 2005.

[KW96]   Robert E. Kass and Larry Wasserman. 'The Selection of Prior Distributions by Formal Rules'. In: *J. Am. Stat. Assoc.* 91.435 (1996), pp. 1343–1370.

[Kay98]   S.M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall Signal Processing Series. PTR Prentice-Hall, Inc., 1998.

[Kel78]   Joseph B. Keller. 'Rays, waves and asymptotics'. In: *Bulletin of the American Mathematical Society* 84.5 (Sept. 1978), pp. 727–750.

[Kia+18]   A. Kiayani, M. Z. Waheed, L. Anttila, M. Abdelaziz et al. 'Adaptive Nonlinear RF Cancellation for Improved Isolation in Simultaneous Transmit–Receive Systems'. In: *IEEE Trans. Microw. Theory Techn.* 66.5 (2018), pp. 2299–2312.

[Kil15]   Per-Simon Kildal. *Foundations of Antenna Engineering: A Unified Approach for Line-of-Sight and Multipath*. Norwood: Artech House, 2015.

[KD09]   Armen Der Kiureghian and Ove Ditlevsen. 'Aleatory or epistemic? Does it matter?' In: *Structural Safety* 31.2 (2009). Risk Acceptance and Risk Communication, pp. 105–112.

[Kli62]   Morris Kline. *Electromagnetic Theory and Geometrical Optics*. Research Report. Corant Institute of Mathematical Science, New York University, 1962.

[Koc08]   J. W. Koch. 'Bayesian approach to extended object and cluster tracking using random matrices'. In: *IEEE Transactions on Aerospace and Electronic Systems* 44.3 (July 2008), pp. 1042–1059.

[Koi04]   P. Koivisto. 'Reduction of errors in antenna radiation patterns using optimally truncated spherical wave expansion'. In: *Progress In Electromagnetics Research* 48 (2004), pp. 313–333.

[KB09]   T. Kolda and B. Bader. 'Tensor Decompositions and Applications'. In: *SIAM Review* 51.3 (2009), pp. 455–500.

[Koz11]   Slawomir Koziel. *Computational Optimization, Methods and Algorithms*. Berlin Heidelberg: Springer Science & Business Media, 2011.

[Kru77]   Joseph B. Kruskal. 'Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics'. In: *Linear Algebra and its Applications* 18.2 (1977), pp. 95–138.

[Kyö+08]   Pekka Kyösti et al. 'WINNER II channel models'. In: *IST-4-027756 WINNER II D1.1.2 V1.2* (Feb. 2008).

[LD04]   M. Landmann and G. Del Galdo. 'Efficient antenna description for MIMO channel modelling and estimation'. In: *7th European Conference on Wireless Technology*. Oct. 2004, pp. 217–220.

[LKT12]   M. Landmann, M. Käske and R.S. Thomä. 'Impact of Incomplete and Inaccurate Data Models on High Resolution Parameter Estimation in Multidimensional Channel Sounding'. In: *IEEE Transactions on Antennas and Propagation* 60.2 (Feb. 2012), pp. 557–573.

[Lan08]   Markus Landmann. 'Limitations of experimental channel characterisation'. PhD thesis. Technische Universität Ilmenau, Nov. 2008.

[LF12]   Artur J. Lemonte and Silvia L.P." Ferrari. 'A note on the local power of the LR, Wald, score and gradient tests'. In: *Electron. J. Stat.* 6 (2012), pp. 421–434.

[LS07]   J. Li and P. Stoica. 'MIMO Radar with Colocated Antennas'. In: *IEEE Signal Processing Magazine* 24.5 (Sept. 2007), pp. 106–114.

[LS96]   Jian Li and Petre Stoica. 'Efficient Mixed-Spectrum Estimation with Applications to Target Feature Extraction'. In: *IEEE Transactions on Signal Processing* 44.2 (1996), pp. 281–295.

[LS08]   Jian Li and Petre Stoica. *MIMO Radar Signal Processing*. New York: John Wiley & Sons, 2008.

[LZG19]   Y. Liao, R. Zhao and L. Gao. 'Joint DOD and DOA Estimation in Bistatic MIMO Radar With Distributed Nested Arrays'. In: *IEEE Access* 7 (2019), pp. 50954–50961.

[LMY04]   K. Lin, R. H. Messerian and Yuanxun Wang. 'A Digital Leakage Cancellation Scheme for Monostatic FMCW'. In: *2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No.04CH37535)*. Vol. 2. June 2004, pp. 747–7502.

[LL98]   Zheng-She Liu and Jian Li. 'Implementation of the RELAX algorithm'. In: *IEEE Transactions on Aerospace and Electronic Systems* 34.2 (1998), pp. 657–664.

[Loè77]   M. Loève. *Probability Theory I*. 4th ed. Comprehensive Manuals of Surgical Specialties. Springer, 1977.

[Lue+20]   David Luengo, Luca Martino, Mónica Bugallo, Víctor Elvira and Simo Särkkä. 'A Survey of Monte Carlo Methods for Parameter Estimation'. In: *EURASIP Journal on Advances in Signal Processing* 2020.1 (May 2020).

[MSE08]   W.Q. Malik, C.J. Stevens and D.J. Edwards. 'Ultrawideband Antenna Distortion Compensation'. In: *IEEE Transactions on Antennas and Propagation* 56.7 (July 2008), pp. 1900–1907.

[Mau05]   Jürgen Maurer. 'Strahlenoptisches Kanalmodell für die Fahrzeug-Fahrzeug-Funk-kommunikation'. PhD thesis. Karlsruher Institut für Technologie, July 2005.

[Mel04]   W. L. Melvin. 'A STAP overview'. In: *IEEE Aerospace and Electronic Systems Magazine* 19.1 (2004), pp. 19–35.

[Mel+15]   A. Melzer, A. Onic, F. Starzer and M. Huemer. 'Short-Range Leakage Cancelation in FMCW Radar Transceivers Using an Artificial On-Chip Target'. In: *IEEE Journal of Selected Topics in Signal Processing* 9.8 (Dec. 2015), pp. 1650–1660.

[Mel+17]   A. Melzer, F. Starzer, H. Jäger and M. Huemer. 'Real-Time Mitigation of Short-Range Leakage in Automotive FMCW Radar Transceivers'. In: *IEEE Trans. Circuits Syst., II, Exp. Briefs* 64.7 (July 2017), pp. 847–851.

[MHL07]   A. Meta, P. Hoogeboom and L. P. Ligthart. 'Signal Processing for FMCW SAR'. In: *IEEE Transactions on Geoscience and Remote Sensing* 45.11 (Nov. 2007), pp. 3519–3532.

[ME11]   M. Mishali and Y. C. Eldar. 'Sub-Nyquist Sampling'. In: *IEEE Signal Processing Magazine* 28.6 (2011), pp. 98–124.

[Mol12]   Andreas Molisch. *Wireless Communications*. second. Wiley-IEEE Press, 2012.

[MS00]   Todd Moon and Winn Sterling. *Mathematical Methods and Algorithms for Signal Processing*. London: Prentice Hall, 2000.

[Mor94]   A. M. Morgan. 'Ultra-wideband impulse scattering measurements'. In: *IEEE Transactions on Antennas and Propagation* 42.6 (June 1994), pp. 840–846.

[MS12]   Jennifer L. Mueller and Samuli Siltanen. *Linear and Nonlinear Inverse Problems with Practical Applications*. USA: Society for Industrial and Applied Mathematics (SIAM), 2012.

[Nai09]   Prabhakar S. Naidu. *Sensor Array Signal Processing, Second Edition*. 2nd ed. CRC Press, June 2009.

[Nel01]   Oliver Nelles. *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer Berlin Heidelberg, 2001.

[Neu18]   Izaak Neutelings. *Electromagnetic Wave*. June 2018. URL: https://wiki.physik.uzh.ch/cms/latex:tikz:electromagnetic_wave (visited on 23/10/2020).

[NS10]   D. Nion and N. D. Sidiropoulos. 'Tensor Algebra and Multidimensional Harmonic Retrieval in Signal Processing for MIMO Radar'. In: *IEEE Transactions on Signal Processing* 58.11 (Nov. 2010), pp. 5693–5705.

[NW06]   Jorge Nocedal and Stephen Wright. *Numerical Optimization*. Berlin Heidelberg: Springer Science & Business Media, 2006.

[Orf16]   Sophocles J. Orfanidis. *Electromagnetic Waves and Antennas*. 2016.

[Owh+13]   H. Owhadi, C. Scovel, T. J. Sullivan, M. McKerns and M. Ortiz. 'Optimal Uncertainty Quantification'. In: *SIAM Review* 55.2 (2013), pp. 271–345.

[Pan17]   IEEE Radar Systems Panel. 'IEEE Standard for Radar Definitions'. In: *IEEE Std 686-2017 (Revision of IEEE Std 686-2008)* (Sept. 2017), pp. 1–54.

[Par+19] J. Park, S. Park, D. Kim and S. Park. 'Leakage Mitigation in Heterodyne FMCW Radar for Small Drone Detection With Stationary Point Concentration Technique'. In: *IEEE Transactions on Microwave Theory Techniques* 67.3 (2019), pp. 1221–1232.

[Par17] Thomas Parker. 'Finite-sample distributions of the Wald, likelihood ratio, and Lagrange multiplier test statistics in the classical linear model'. In: *Communications in Statistics - Theory and Methods* 46.11 (2017), pp. 5195–5202.

[Pay07] Payton Z. Peebles, Jr. *Radar Principles*. John Wiley & Sons, 2007.

[PMB04] Marius Pesavento, Christoph F. Mecklenbräuker and Johann F. Böhme. 'Multi-dimensional Rank Reduction Estimator for Parametric MIMO Channel Models'. In: *EURASIP Journal on Advances in Signal Processing* 2004.9 (Aug. 2004), p. 839148.

[PP12] K. B. Petersen and M. S. Pedersen. *The Matrix Cookbook*. 20121115th ed. Technical University of Denmark, Nov. 2012.

[PS01] Rik Pintelon and Johan Schoukens. *System Identification: A Frequency Domain Approach*. IEEE Press, 2001.

[Pip93] S. O. Piper. 'Receiver Frequency Resolution for Range Resolution in Homodyne FMCW Radar'. In: *Conference Proceedings National Telesystems Conference 1993*. June 1993, pp. 169–173.

[Pip95] S. O. Piper. 'Homodyne FMCW radar range resolution effects with sinusoidal non-linearities in the frequency sweep'. In: *Proceedings International Radar Conference*. May 1995, pp. 563–567.

[Pol98] D. N. Politis. 'Computer-Intensive Methods in Statistical Analysis'. In: *IEEE Signal Processing Magazine* 15.1 (Jan. 1998), pp. 39–55.

[Qui+06] Angela Quinlan, Jean-Pierre Barbot, Pascal Larzabal and Martin Haardt. 'Model Order Selection for Short Data: An Exponential Fitting Test (EFT)'. In: *EURASIP Journal on Advances in Signal Processing* 2007.1 (Oct. 2006), p. 071953.

[Ric13] Mark A. Richards. *Fundamentals of Radar Signal Processing*. Second. Madison: McGraw Hill Professional, 2013.

[RLT03] A. Richter, M. Landmann and R.S. Thomä. 'Maximum likelihood channel parameter estimation from multidimensional channel sounding measurements'. In: *57th IEEE Semiannual Vehicular Technology Conference (VTC-Spring)*. Vol. 2. Apr. 2003, 1056–1060 vol.2.

[RT05] A. Richter and R.S. Thoma. 'Joint Maximum Likelihood Estimation of Specular Paths and Distributed Diffuse Scattering'. In: *61st IEEE Vehicular Technology Conference (VTC Spring)*. Vol. 1. May 2005, 11–15 Vol. 1.

[Ric05] Andreas Richter. 'Estimation of Radio Channel Parameters'. PhD thesis. Technische Universität Ilmenau, Nov. 2005.

[RLT02] Andreas Richter, Markus Landmann and Reiner S. Thomä. 'A Gradient Based Method for Maximum Likelihood Channel Parameter Estimation from Multidimensional Channel Sounding Measurements'. In: *Multidimensional Channel Sounding Measurements," XXVIIth URSI General Assembly, Maastricht, NL*. 2002.

[Rob06]   C. Roblin. 'Ultra compressed parametric modelling of UWB antenna measurements'. In: *Antennas and Propagation, 2006. EuCAP 2006. First European Conference on*. Nov. 2006, pp. 1–8.

[SM11]   J. Salmi and A. F. Molisch. 'Propagation Parameter Estimation, Modeling and Measurements for Ultrawideband MIMO Radar'. In: *IEEE Transactions on Antennas and Propagation* 59.11 (Nov. 2011), pp. 4257–4267.

[SRK09]   J. Salmi, A. Richter and V. Koivunen. 'Detection and Tracking of MIMO Propagation Path Parameters Using State-Space Approach'. In: *IEEE Transactions on Signal Processing* 57.4 (2009), pp. 1538–1550.

[Sam+12]   P. F. Sammartino, D. Tarchi, J. Fortuny-Guasch, F. Oliveri and R. Giuliani. 'Phase compensation and processing in multiple-input-multiple-output radars'. In: *IET Radar, Sonar Navigation* 6 (Apr. 2012), pp. 222–232.

[SMS02]   S. Sando, A. Mitra and P. Stoica. 'On the Cramer-Rao bound for model-based spectral analysis'. In: *IEEE Signal Processing Letters* 9.2 (Feb. 2002), pp. 68–71.

[SD91]   L.L. Scharf and C. Demeure. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley series in electrical and computer engineering. Addison-Wesley Publishing Company, 1991.

[Sch80]   Martin Schetzen. *The Volterra and Wiener Theories of Nonlinear Systems*. New York: Wiley, 1980.

[SSR06]   B. Schiek, H.J. Siweris and I. Rolfes. *Noise in High-Frequency Circuits and Oscillators*. John Wiley & Sons, 2006.

[SKK12]   A. Schmitz, T. Karolski and L. Kobbelt. 'Using spherical harmonics for modeling antenna patterns'. In: *Radio and Wireless Symposium (RWS), 2012 IEEE*. Jan. 2012, pp. 155–158.

[Sch19]   Benedikt Schneele. 'Winkelschätzung für einen Verbund breitbandiger MIMO-Radare bei 160 GHz'. Master thesis. Ulm university, Nov. 2019.

[SS10]   P.J. Schreier and L.L. Scharf. *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.

[Sch09]   Dominique Schreurs. *RF Power Amplifier Behavioral Modeling*. The Cambridge RF and Microwave Engineering Series. Cambridge University Press, 2009.

[SJ17]   K. T. Selvan and R. Janaswamy. 'Fraunhofer and Fresnel Distances : Unified derivation for aperture antennas.' In: *IEEE Antennas and Propagation Magazine* 59.4 (2017), pp. 12–15.

[SW12]   Jin Keun Seo and Eung Je Woo. *Nonlinear Inverse Problems in Imaging*. John Wiley & Sons, Ltd, 2012.

[Sha17]   A. Shahmoradi. 'Multilevel Bayesian Parameter Estimation in the Presence of Model Inadequacy and Data Uncertainty'. In: *ArXiv e-prints* (Nov. 2017). arXiv: 1711.10599 [physics.data-an].

[Sil20]   SVG Silh. *Free SVG Image & Icon*. Nov. 2020. URL: https://svgsilh.com/ (visited on 18/11/2020).

[Sko61]   Merrill I. Skolnik. 'An Analysis of Bistatic Radar'. In: *IRE Transactions on Aerospace and Navigational Electronics* ANE-8.1 (Mar. 1961), pp. 19–27.

[Sko90]   Merrill I. Skolnik. *Radar Handbook*. 2nd ed. New York: McGraw Hill, 1990.

[SS89]   T. Söderström and P. Stoica. *System Identification*. Prentice Hall International Series In Systems And Control Engineering. Prentice Hall, 1989.

[Som+02]   G. Sommerkorn, A. Richter, R. Thomä and W. Wirnitzer. 'Antenna Multiplexing & Time Alignment for MIMO channel Sounding'. In: *XXXVII-th URSI General Assembly*. Aug. 2002.

[SW05]   Werner Sörgel and Werner Wiesbeck. 'Influence of the Antennas on the Ultrawideband Transmission'. In: *EURASIP J. Appl. Signal Process.* 2005 (Jan. 2005), pp. 296–305.

[Sor93]   J.M. Soriano. 'Global minimum point of a convex function'. In: *Applied Mathematics and Computation* 55.2 (1993), pp. 213–218.

[Sta15]   Stackexchange. *Drawing wave propagation using Tikz*. Aug. 2015. URL: https://tex.stackexchange.com/questions/262015/drawing-wave-propagation-using-tikz (visited on 23/10/2020).

[Sta15]   Stackexchange. *Drawing wave propagation using Tikz*. Aug. 2015. URL: https://tex.stackexchange.com/questions/262015/drawing-wave-propagation-using-tikz (visited on 23/10/2020).

[Sta19]   Stackexchange. *Stacking 3D Cubes with Spacing*. May 2019. URL: https://tex.stackexchange.com/questions/489712/stacking-3d-cubes-with-spacing (visited on 23/10/2020).

[Ste09]   Alwin Stegeman. 'On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode'. In: *Linear Algebra and its Applications* 431.1 (2009), pp. 211–227.

[SMB01]   M. Steinbauer, A.F. Molisch and E. Bonek. 'The double-directional radio channel'. In: *IEEE Antennas and Propagation Magazine* 43.4 (Aug. 2001), pp. 51–63.

[SN89]   P. Stoica and A. Nehorai. 'MUSIC, Maximum Likelihood, and Cramer-Rao Bbound'. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.5 (May 1989), pp. 720–741.

[SS04]   P. Stoica and Y. Selen. 'Model-order selection: a review of information criterion rules'. In: *IEEE Signal Processing Magazine* 21.4 (July 2004), pp. 36–47.

[Sto92]   A. G. Stove. 'Linear FMCW radar techniques'. In: *IEE Proceedings F - Radar and Signal Processing* 139.5 (Oct. 1992), pp. 343–350.

[Suh+17]   J. Suh, L. Minz, D. Jung, H. Kang, J. Ham and S. Park. 'Drone-Based External Calibration of a Fully Synchronized Ku-Band Heterodyne FMCW Radar'. In: *IEEE Transactions on Instrumentation and Measurement* 66.8 (2017), pp. 2189–2197.

[Sul15]   Tim Sullivan. *Introduction to Uncertainty Quantification*. Vol. 63. Springer International Publishing, 2015.

[SBL14]   H. Sun, F. Brigui and M. Lesturgie. 'Analysis and Comparison of MIMO Radar Waveforms'. In: *2014 International Radar Conference*. Oct. 2014, pp. 1–6.

[SC10]   A. Swain and D. Clark. 'Extended Object Filtering using Spatial Independent Cluster Processes'. In: *2010 13th International Conference on Information Fusion*. July 2010, pp. 1–8.

[SS98]   A. L. Swindlehurst and P. Stoica. 'Maximum Likelihood Methods in Radar Array Signal Processing'. In: *Proceedings of the IEEE* 86.2 (Feb. 1998), pp. 421–441.

[Tar04]   Albert Tarantola. *Inverse Problem Theory*. Society for Industrial and Applied Mathematics (SIAM), 2004.

[Tar05]   Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics (SIAM), 2005.

[Ter02]   Georg R. Terrell. 'The Gradient Statistic'. In: *Computing Science and Statistics* 34 (2002), pp. 206–215.

[Tho+01]   R. S. Thomä, D. Hampicke, A. Richter, G. Sommerkorn and U. Trautwein. 'MIMO vector channel sounder measurement for smart antenna system evaluation'. In: *European Transactions on Telecommunications* 12.5 (2001), pp. 427–438.

[Tho+19]   R. S. Thomä et al. 'Cooperative Passive Coherent Location: A Promising 5G Service to Support Road Safety'. In: *IEEE Communications Magazine* 57.9 (2019), pp. 86–92.

[TLS+04]   R.S. Thomä, M. Landmann, G. Sommerkorn et al. 'Multidimensional High-Resolution Channel Sounding in Mobile Radio'. In: *21st IEEE Instrumentation and Measurement Technology Conference (IMTC)*. Vol. 1. May 2004, 257–262 Vol.1.

[Tre01]   Harry L. Van Trees. *Detection, Estimation, and Modulation Theory, Radar-Sonar Signal Processing and Gaussian Signals in Noise -*. New York: John Wiley & Sons, 2001.

[TB07]   Harry L. Van Trees and Kristine L. Bell. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley-IEEE Press, 2007.

[TS88]   J. Tsao and B. D. Steinberg. 'Reduction of Sidelobe and Speckle Artifacts in Microwave Imaging: the CLEAN Technique'. In: *IEEE Transactions on Antennas and Propagation* 36.4 (1988), pp. 543–556.

[TF09]   T. Engin Tuncer and Benjamin Friedlander. *Classical and Modern Direction-of-Arrival Estimation*. Amsterdam Boston: Academic Press, 2009.

[UP09]   Mohammad Habib Ullah and AU Priantoro. 'A review on multiplexing schemes for MIMO channel sounding'. In: *International Journal of Computer Science and Network Security* (Jan. 2009).

[Van04]   Harry L. Van Trees. *Optimum Array Processing - Detection, Estimation, and Modulation Theory*. New York: John Wiley & Sons, 2004.

[Vas+20]   C. Vasanelli, F. Roos, A. Dürr, J. Schlichenmaier, P. Hügler, B. Meinecke, M. Steiner and C. Waldschmidt. 'Calibration and Direction-of-Arrival Estimation of Millimeter-Wave Radars: A Practical Introduction'. In: *IEEE Antennas and Propagation Magazine* (2020).

[Vuo89]   Quang H. Vuong. 'Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses'. In: *Econometrica* 57.2 (Mar. 1989), pp. 307–333.

[Wal+05]   L. J. Waldorp, H. M. Huizenga, A. Nehorai, R. P. P. P. Grasman and P. C. M. Molenaar. 'Model Selection in Spatio-Temporal Electromagnetic Source Analysis'. In: *IEEE Transactions on Biomedical Engineering* 52.3 (2005), pp. 414–420.

[WHG+02]   Lourens J. Waldorp, Hilde M. Huizenga, Raoul P. P. P. Grasman et al. 'Model Selection in Electromagnetic Source Analysis With an Application to VEFs'. In: *EEE Trans. Biomed. Eng.* 49 (2002), pp. 1121–1129.

[WHM21]   Christian Waldschmidt, Juergen Hasch and Wolfgang Menzel. 'Automotive Radar — From First Efforts to Future Systems'. In: *IEEE Journal of Microwaves* 1.1 (2021), pp. 135–148.

[Wal80]   J. L. Walker. 'Range-Doppler Imaging of Rotating Objects'. In: *IEEE Transactions on Aerospace and Electronic Systems* AES-16.1 (Jan. 1980), pp. 23–52.

[WK85]   M. Wax and T. Kailath. 'Detection of signals by information theoretic criteria'. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33.2 (Apr. 1985), pp. 387–392.

[Wei09]   Thomas Weise. *Global Optimization Algorithms – Theory and Application*. Germany: self-published, 2009.

[Wei11]   Thomas Weise. *Global Optimization Algorithms - Theory and Application*. 3rd ed. 2011.

[Whi+91]   M. W. Whitt, F. T. Ulaby, P. Polatin and V. V. Liepa. 'A general polarimetric radar calibration technique'. In: *IEEE Transactions on Antennas and Propagation* 39.1 (1991), pp. 62–67.

[WAS09]   W. Wiesbeck, G. Adamiuk and C. Sturm. 'Basic Properties and Design Principles of UWB Antennas'. In: *Proceedings of the IEEE* 97.2 (Feb. 2009), pp. 372–385.

[WS14]   W. Wiesbeck and L. Sit. 'Radar 2020: The future of radar systems'. In: *International Radar Conference*. 2014, pp. 1–6.

[Wil05]   N.J. Willis. *Bistatic Radar*. Electromagnetics and Radar. SciTech Publishing, 2005.

[Won+90]   K. M. Wong, Q. T. Zhang, J. P. Reilly and P. C. Yip. 'On Information Theoretic Criteria for Determining the Number of Signals in High Resolution Array Processing'. In: *IEEE Trans. Acoust., Speech, Signal Process.* 38.11 (Nov. 1990), pp. 1959–1971.

[XLS08]   L. Xu, J. Li and P. Stoica. 'Target Detection and Parameter Estimation for MIMO Radar Systems'. In: *IEEE Transactions on Aerospace and Electronic Systems* 44.3 (2008), pp. 927–939.

[Yan+13]   H. Yan, G. Shen, R. Zetik, O. Hirsch and R. S. Thomä. 'Ultra-Wideband MIMO Ambiguity Function and Its Factorability'. In: *IEEE Transactions on Geoscience and Remote Sensing* 51.1 (Jan. 2013), pp. 504–519.

[Yan17]   Xin-She Yang. *Nature-Inspired Algorithms and Applied Optimization*. Berlin, Heidelberg: Springer, 2017.

[Yoa10]   Benjamini Yoav. 'Simultaneous and selective inference: Current successes and future challenges'. In: *Biom. J.* 52.6 (2010), pp. 708–721.

[YPP10]   Y. Yu, A. P. Petropulu and H. V. Poor. 'MIMO Radar Using Compressive Sampling'. In: *IEEE Journal of Selected Topics in Signal Processing* 4.1 (2010), pp. 146–163.

[YI15]   Z. Yun and M. F. Iskander. 'Ray Tracing for Radio Propagation Modeling: Principles and Applications'. In: *IEEE Access* 3 (July 2015), pp. 1089–1100.

[Zac81]   Shelemyahu Zacks. *Parametric Statistical Inference: Basic Theory and Modern Approaches*. International series in nonlinear mathematics. Pergamon Press, 1981.

[Zat97]   M. Zatman. 'How narrow is narrowband? [adaptive array signal processing]'. In: *Thirty-First Asilomar Conference on Signals, Systems and Computers*. Vol. 2. Nov. 1997, 1341–1345 vol.2.

[Zha+10]   X. Zhang, L. Xu, L. Xu and D. Xu. 'Direction of Departure (DOD) and Direction of Arrival (DOA) Estimation in MIMO Radar with Reduced-Dimension MUSIC'. In: *IEEE Communications Letters* 14.12 (Dec. 2010), pp. 1161–1163.

[ZW88]   I. Ziskind and M. Wax. 'Maximum Likelihood Localization of Multiple Sources by Alternating Projection'. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.10 (Oct. 1988), pp. 1553–1560.

[ZB98]   A. M. Zoubir and B. Boashash. 'The Bootstrap and its Application in Signal Processing'. In: *IEEE Signal Processing Magazine* 15.1 (Jan. 1998), pp. 56–76.

List of Own Contributions

Journal Articles

[1]   D. Dupleich, N. Iqbal, C. Schneider, S. Häfner et al. 'Influence of system aspects on fading at mm-waves'. In: *IET Microwaves, Antennas Propagation* 12.4 (2018), pp. 516–524.

[2]   A. Dürr, D. Schwarz, S. Häfner et al. 'High-Resolution 160 GHz Imaging MIMO Radar using MMICs with On-Chip Frequency Synthesizers'. In: *IEEE Transactions on Microwave Theory and Techniques* (2019).

[3]   F. Fuschini, M. Häfner S. and Zoli et al. 'Item Level Characterization of mm-Wave Indoor Propagation'. In: *EURASIP Journal on Wireless Communications and Networking* 2016.1 (Jan. 2016), p. 4.

[4]   F. Fuschini, S. Häfner, M. Zoli et al. 'Analysis of In-Room mm-Wave Propagation: Directional Channel Measurements and Ray Tracing Simulations'. In: *Journal of Infrared, Millimeter, and Terahertz Waves* 38.6 (June 2017), pp. 727–744.

[5]   S. Häfner, A. Dürr, C. Waldschmidt and R. Thomä. 'Mitigation of RF Impairments of a 160-GHz MMIC FMCW Radar Using Model-Based Estimation'. In: *IEEE Transactions on Microwave Theory and Techniques* (2019), pp. 1–9.

[6]   S. Häfner, A. Dürr, C. Waldschmidt and R. Thomä. 'Mitigation of Leakage in FMCW Radars by Background Subtraction and Whitening'. In: *IEEE Microwave and Wireless Components Letters* 30.11 (Nov. 2020), pp. 1105–1107.

[7]   S. Häfner, Martin Käske and R. Thomä. 'On Calibration and Direction Finding with Uniform Circular Arrays'. In: *International Journal on Antennas and Propagation* (2019), p. 12.

[8] S. Häfner and R. Thomä. 'High Resolution Estimation of AoA, AoD and TdoA from MIMO Channel Sounding Measurements with Virtual Antenna Arrays: Maximum-Likelihood vs. Unitary Tensor-ESPRIT'. In: *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems* 7.2 (2018), pp. 27–34.

[9] S. Häfner and R. Thomä. 'Identification of Linear Time Invariant Systems Using FMCW Signals and Stretch Processing Receivers'. In: *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems* 8.2 (2019), pp. 35–39.

[10] S. Häfner and R. Thomä. 'Compensation of Motion-Induced Phase Errors and Enhancement of Doppler Unambiguity in TDM-MIMO Systems by Model-Based Estimation'. In: *IEEE Sensors Letters* 4.10 (Oct. 2020), pp. 1–4.

[11] N. Iqbal, J. Luo, R. Müller, G. Steinböck, C. Schneider, D. Dupleich, S. Häfner and R. S. Thomä. 'Multipath Cluster Fading Statistics and Modeling in Millimeter Wave Radio Channels'. In: *IEEE Transactions on Antennas and Propagation* (2019), pp. 1–1.

[12] N. Iqbal, J. Luo, C. Schneider, D. A. Dupleich, R. Müller, S. Häfner and R. S. Thomä. 'Investigating Validity of Wide-Sense Stationary Assumption in Millimeter Wave Radio Channels'. In: *IEEE Access* 7 (2019), pp. 180073–180082.

[13] C. Ling, X. Yin, R. Müller, S. Häfner et al. 'Double-Directional Dual-Polari-metric Cluster-Based Characterization of 70-77 GHz Indoor Channels'. In: *IEEE Transactions on Antennas and Propagation* 66.2 (Feb. 2018), pp. 857–870.

Conference Proceedings

[14] D. Dupleich, S. Häfner, R. Müller et al. 'Real-field performance of multiple-beam beam-former with polarization compensation'. In: *10th European Conference on Antennas and Propagation (EuCAP)*. Apr. 2016, pp. 1–5.

[15] D. Dupleich, S. Häfner, R. Müller et al. 'Real-field performance of hybrid MISO time reversal multi-beam beam-former at mm-Waves'. In: *11th European Conference on Antennas and Propagation (EUCAP)*. Mar. 2017, pp. 448–452.

[16] D. Dupleich, S. Häfner, C. Schneider et al. 'Double-directional and dual-polari-metric indoor measurements at 70 GHz'. In: *IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. Aug. 2015, pp. 2234–2238.

[17] D. Dupleich, N. Iqbal, C. Schneider, S. Häfner et al. 'Investigations on fading scaling with bandwidth and directivity at 60 GHz'. In: *11th European Conference on Antennas and Propagation (EUCAP)*. Mar. 2017, pp. 3375–3379.

[18] D. Dupleich, J. Luo, S. Häfner et al. 'A Hybrid Polarimetric Wide-Band Beam-former Architecture for 5G mm-Wave Communications'. In: *20th International ITG Workshop on Smart Antennas (WSA)*. Mar. 2016, pp. 1–8.

[19] D. Dupleich, R. Müller, Han Niu, S. Häfner, C. Schneider, J. Luo, G. D. Galdo and R. Thomä. 'Polarization in Spatial Channel Models at mm-Waves: a Correlation Based Approach'. In: *13th European Conference on Antennas and Propagation (EuCAP)*. Apr. 2019.

[20] S. Häfner, D. A. Dupleich, R. Müller et al. 'Characterisation of Channel Measurements at 70GHz in Indoor Femtocells'. In: *81st IEEE Vehicular Technology Conference (VTC Spring)*. May 2015, pp. 1–5.

[21] S. Häfner, A. Dürr, R. Thomä, C. Waldschmidt and G. D. Galdo. 'High-Resolution Parameter Estimation for Chirp-Sequence Radar Considering Hardware Impairments'. In: *11th German Microwave Conference (GeMiC)*. Mar. 2018, pp. 355–358.

[22] S. Häfner, M. Käske and R. Thomä. 'Polarisation-Angle-Delay Estimation for Blind Localisation Approaches under Multipath Propagation'. In: *9th European Conference on Antennas and Propagation (EuCAP)*. May 2015, pp. 1–6.

[23] S. Häfner, M. Käske, R. Thomä et al. 'Selection of Antenna Array Configuration for Polarimetric Direction Finding in Correlated Signal Environments'. In: *19th International ITG Workshop on Smart Antennas (WSA)*. Mar. 2015, pp. 1–8.

[24] S. Häfner, R. Müller and R. Thomä. 'Full 3D Antenna Pattern Interpolation Using Fourier Transform Based Wavefield Modelling'. In: *20th International ITG Workshop on Smart Antennas (WSA)*. Mar. 2016, pp. 1–8.

[25] S. Häfner, M. Röding, G. Sommerkorn et al. 'Contribution to Drone Detection by Exploiting Parameter Estimation for a Prototype mm-Wave Radar System'. In: *22nd International ITG Workshop on Smart Antennas (WSA)*. Mar. 2018, pp. 1–8.

[26] S. Häfner and R. Thomä. 'Application of Tensor Decomposition Methods to Antenna Array Calibration Measurements for Denoising and Narrowband Modelling'. In: *13th European Conference on Antennas and Propagation (EuCAP)*. Apr. 2019, pp. 1–5.

[27] S. Häfner and R. Thomä. 'Extended Kalman Filtering and Maximum-Likelihood Estimation for Point Target Localisation'. In: *13th German Microwave Conference (GeMiC)*. Mar. 2020.

[28] S. Häfner and R. Thomä. 'MAP and MMSE Based Parameter Estimation from Noisy Radar Observations'. In: *13th German Microwave Conference (GeMiC)*. Mar. 2020.

[29] N. Iqbal, D. Dupleich, C. Schneider, J. Luo, R. Müller, S. Häfner, G. D. Galdo and R. S. Thoma. 'Modeling of Intra-Cluster Multipaths for 60 GHz Fading Channels'. In: *12th European Conference on Antennas and Propagation (EuCAP 2018)*. Apr. 2018, pp. 1–5.

[30] N. Iqbal, J. Luo, D. Dupleich, S. Häfner, R. Müller, C. Schneider and R. S. Thomä. 'Second-order Statistical Characterization of the 60 GHz Cluster Fading Channels'. In: *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. Sept. 2018, pp. 241–245.

[31] N. Iqbal, J. Luo, C. Schneider, D. Dupleich, S. Häfner, R. Müller and R. Thomä. 'Frequency and Bandwidth Dependence of Millimeter Wave Ultra-Wide-Band Channels'. In: *11th European Conference on Antennas and Propagation (EUCAP)*. Mar. 2017, pp. 141–145.

[32] N. Iqbal, J. Luo, C. Schneider, D. Dupleich, R. Müller, S. Häfner and R. Thomä. 'Stochastic/deterministic Behavior of Cross Polarization Discrimination in mmWave Channels'. In: *IEEE International Conference on Communications (ICC)*. May 2017, pp. 1–5.

[33] N. Iqbal, J. Luo, Y. Xin, R. Mueller, S. Haefner and R. Thomae. 'Measurements Based Interference Analysis at Millimeter Wave Frequencies in an Indoor Scenario'. In: *2017 IEEE Globecom Workshops*. Dec. 2017, pp. 1–5.

[34] N. Iqbal, C. Schneider, J. Luo, D. Dupleich, R. Müller, S. Häfner and R. Thomä. 'On the Stochastic and Deterministic Behavior of mmWave Channels'. In: *11th European Conference on Antennas and Propagation (EUCAP)*. Mar. 2017, pp. 1813–1817.

[35] R. Müller, S. Häfner, D. Dupleich et al. 'Ultra-Wideband Channel Sounder for Measurements at 70 GHz'. In: *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*. May 2015, pp. 1–5.

[36] R. Müller, S. Häfner, D. Dupleich et al. 'Simultaneous multi-band channel sounding at mm-Wave frequencies'. In: *10th European Conference on Antennas and Propagation (EuCAP)*. Apr. 2016, pp. 1–5.

[37] R. Müller, M. Käske, P. Rauschenbach, G. Sommerkorn, C. Schneider, F. Wollenschläger, S. Häfner and R. Thomä. 'Design of a circular antenna array for MIMO channel sounding application at 2.53 GHz'. In: *8th European Conference on Antennas and Propagation (EuCAP)*. Apr. 2014, pp. 239–243.

[38] M. Röding, G. Sommerkorn, S. Häfner et al. 'Arc42 - A double-arch positioner for bistatic RCS measurements with four degrees of freedom'. In: *European Radar Conference (EURAD)*. Oct. 2017, pp. 425–428.

[39] M. Röding, G. Sommerkorn, S. Häfner et al. 'Fully Polarimetric Wideband RCS Measurements for Small Drones'. In: *11th European Conference on Antennas and Propagation (EUCAP)*. Mar. 2017, pp. 3926–3930.

[40] C. Schneider, M. Ibraheam, S. Häfner et al. 'On the reliability of multipath cluster estimation in realistic channel data sets'. In: *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*. Apr. 2014, pp. 449–453.

[41] G. Sommerkorn, M. Käske, C. Schneider, S. Häfner and R. Thomä. 'Full 3D MIMO channel sounding and characterization in an urban macro cell'. In: *XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*. Aug. 2014, pp. 1–4.

Misc

[42] S. Häfner and R. Thomä. 'Estimation of Radio Channel Parameters in Case of an Unknown Transmitter'. In: *arXiv* (2015). arXiv: 1512.03591 [cs.IT].

[43] G. Sommerkorn, S. Häfner, M. Röding and R. Thomä. 'Synthesis of Realistic Bistatic Range Profiles'. In: *COST Action CA15104 (IRACON), TD(17)03049*. Feb. 2017.

APPENDIX A

MATHEMATICAL NOTATION, FUNCTIONS, OPERATIONS AND RELATIONS

A.1 Notation

A.1.1 Variables

a, A, σ Scalars, defined as italic Greek or Latin letters

$\mathbf{a}, \boldsymbol{\sigma}$ Column vectors, defined as bold face lower-case Greek or Latin letters

$\mathbf{A}, \boldsymbol{\Sigma}$ Matrices, defined as bold face capital Greek or Latin letters

$\mathcal{A}, \mathcal{B}$ Tensors, defined as bold face calligraphic Latin letters

A, B   Sets, defined as serif Latin upper-case letters

A.1.2 Sets

$A = \{a_1, a_2, \ldots\}$ Set A containing the elements $a_1, a_2, \ldots$

$A \subseteq B$ Set A subset of set B

$A \subset B$ Set A proper subset of set B

A.1.3 Complex Numbers

$z = z_\mathrm{R} + J z_\mathrm{I} \in \mathbb{C}$ Cartesian form of complex number z

$z = |z| \cdot \exp\{J \arg\{z\}\} \in \mathbb{C}$ Polar form of complex number z

$z_\mathrm{R} = \Re\{z\} \in \mathbb{R}$ Real part of complex number z

$z_\mathrm{I} = \Im\{z\} \in \mathbb{R}$ Imaginary part of complex number z

$\arg\{z\} \in \mathbb{R}$ Phase angle of complex number z

$|z| \in \mathbb{R}_+$ Absolute value of complex number z

A.2 Functions

$\delta(t)$ Dirac delta function or Dirac impulse [Hay94]

$H(x)$ Heaviside step function [AS72]

$P_n(x)$ Legendre polynomial of n-th degree [AS72]

$\text{rect}\left(\frac{t}{T}\right)$ Rectangular function [Hay94]

$\text{sinc}(x)$ Sinc function [Hay94]

$j_n(x)$ Spherical Bessel function of first kind and n-th order [AS72]

A.2.1 Probability Distributions

$X \sim \mathcal{N}$ Normal distributed random variable X

$X \sim \mathcal{CN}$ Proper complex Normal distributed random variable X

$X \sim \mathcal{U}$ Uniform distributed random variable X

A.3 Operations and Relations

A.3.1 Matrix Relations

$\mathbf{A} \succeq \mathbf{0}$ Positive semi-definite matrix $\mathbf{A}$ (all Eigenvalues larger than or equal to zero)

$\mathbf{A} \succ \mathbf{0}$ Positive definite matrix $\mathbf{A}$ (all Eigenvalues larger than zero)

A.3.2 Scalar Operations

$x(t){*}y(t)$ Convolution of function $x(t)$ and $y(t)$ [Hay94]

$x(t) \;\circ\!\!\!\-\!\!\!\bullet\; X(f)$ Fourier transform of $x(t)$ [Hay94]

$X(f) \;\bullet\!\!\!\-\!\!\!\circ\; x(t)$ Inverse Fourier transform of $X(f)$ [Hay94]

$E\{X\}$ Expected value of random process X [Hay94]

A.3.3 Matrix and Vector Operations

$[\mathbf{a}]_{(i)} \in \mathbb{C}$ Selection of i-th entry of vector $\mathbf{a} \in \mathbb{C}^M$

$[\mathbf{A}]_{(i)} \in \mathbb{C}^M$ Selection of i-th column of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$

$[\mathbf{A}]_{(i,j)} \in \mathbb{C}$ Selection of i-th column and j-th row entry of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$

$||\mathbf{a}|| \in \mathbb{R}_+$ Euclidian norm of vector $\mathbf{a} \in \mathbb{C}^M$

$\|\mathbf{A}\|_F \in \mathbb{R}_+$	Frobenius norm of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$
$\|\mathbf{A}\| \in \mathbb{C}$	Determinant of matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$
trace$\{\mathbf{A}\} \in \mathbb{C}$	Trace of matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$

$$\text{trace}\,\{\mathbf{A}\} = \textstyle\sum_{m=1}^{M} [\mathbf{A}]_{(m,m)}$$

$\mathbf{a}^{\text{T}} \cdot \mathbf{b} \in \mathbb{C}$	Scalar product of vector $\mathbf{a} \in \mathbb{C}^M$ and $\mathbf{b} \in \mathbb{C}^M$
$\mathbf{a} \times \mathbf{b} \in \mathbb{C}^M$	Vector product of vector $\mathbf{a} \in \mathbb{C}^M$ and $\mathbf{b} \in \mathbb{C}^M$
vec $\{\mathbf{A}\} \in \mathbb{C}^{MN}$	Stacks columns of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ into a vector

$$\text{vec}\,\{\mathbf{A}\} = \left[[\mathbf{A}]_{(1)}^{\text{T}}, \dots, [\mathbf{A}]_{(N)}^{\text{T}} \right]^{\text{T}}$$

$\mathbf{A}^{\text{T}} \in \mathbb{C}^{N \times M}$	Transpose of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$
$\mathbf{A}^{\dagger} \in \mathbb{C}^{M \times N}$	Conjugate of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$
$\mathbf{A}^{\text{H}} \in \mathbb{C}^{N \times M}$	Hermitian (conjugate transpose) of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$
$\mathbf{A}^{-1} \in \mathbb{C}^{M \times M}$	Inverse of matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$
$\mathbf{A}^{+} \in \mathbb{C}^{N \times M}$	Moore–Penrose pseudo inverse of matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$
diag$\{\mathbf{A}\} \in \mathbb{C}^M$	Stacks main diagonal elements of matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$ into a vector

$$\text{diag}\{\mathbf{A}\} = \left[[\mathbf{A}]_{(1,1)}, [\mathbf{A}]_{(2,2)}, \dots, [\mathbf{A}]_{(M,M)} \right]^{\text{T}}$$

$\mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{NP \times MR}$ Kronecker product of matrix $\mathbf{A} \in \mathbb{C}^{N \times M}$ and $\mathbf{B} \in \mathbb{C}^{P \times R}$

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} [\mathbf{A}]_{(1,1)} \cdot \mathbf{B} & \dots & [\mathbf{A}]_{(1,M)} \cdot \mathbf{B} \\ \vdots & \ddots & \vdots \\ [\mathbf{A}]_{(N,1)} \cdot \mathbf{B} & \dots & [\mathbf{A}]_{(N,M)} \cdot \mathbf{B} \end{bmatrix}$$

$\mathbf{A} \Diamond \mathbf{B} \in \mathbb{C}^{NP \times M}$ Khatri–Rao product of matrix $\mathbf{A} \in \mathbb{C}^{N \times M}$ and $\mathbf{B} \in \mathbb{C}^{P \times M}$

$$\mathbf{A} \Diamond \mathbf{B} = \begin{bmatrix} [\mathbf{A}]_{(1)} \otimes [\mathbf{B}]_{(1)}, \dots, [\mathbf{A}]_{(M)} \otimes [\mathbf{B}]_{(M)} \end{bmatrix}$$

$\mathbf{A} \odot \mathbf{B} \in \mathbb{C}^{N \times M}$ Hadamard (element-wise) product of matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{N \times M}$

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} [\mathbf{A}]_{(1,1)} \cdot [\mathbf{B}]_{(1,1)} & \dots & [\mathbf{A}]_{(1,M)} \cdot [\mathbf{B}]_{(1,M)} \\ \vdots & \ddots & \vdots \\ [\mathbf{A}]_{(N,1)} \cdot [\mathbf{B}]_{(N,1)} & \dots & [\mathbf{A}]_{(N,M)} \cdot [\mathbf{B}]_{(N,M)} \end{bmatrix}$$

$\mathbf{A} \oslash \mathbf{B} \in \mathbb{C}^{N \times M}$ Hadamard (element-wise) division of matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{N \times M}$

$$\mathbf{A} \oslash \mathbf{B} = \begin{bmatrix} [\mathbf{A}]_{(1,1)} / [\mathbf{B}]_{(1,1)} & \dots & [\mathbf{A}]_{(1,M)} / [\mathbf{B}]_{(1,M)} \\ \vdots & \ddots & \vdots \\ [\mathbf{A}]_{(N,1)} / [\mathbf{B}]_{(N,1)} & \dots & [\mathbf{A}]_{(N,M)} / [\mathbf{B}]_{(N,M)} \end{bmatrix}$$

$\mathrm{T}\{\mathbf{a}\} \in \mathbb{C}^{M \times M}$ Constructor for hermitian Toeplitz matrix from vector $\mathbf{a} \in \mathbb{C}^{M}, [\mathbf{a}]_{(1)} \in \mathbb{R}$

$$\mathrm{T}\{\mathbf{a}\} = \begin{bmatrix} [\mathbf{a}]_{(1)} & [\mathbf{a}]_{(2)}^{\dagger} & \dots & [\mathbf{a}]_{(M-1)}^{\dagger} & [\mathbf{a}]_{(M)}^{\dagger} \\ [\mathbf{a}]_{(2)} & \ddots & \ddots & [\mathbf{a}]_{(M-2)}^{\dagger} & [\mathbf{a}]_{(M-1)}^{\dagger} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ [\mathbf{a}]_{(M)} & [\mathbf{a}]_{(M-1)} & \dots & [\mathbf{a}]_{(2)} & [\mathbf{a}]_{(1)} \end{bmatrix}$$

$\mathrm{diag}\{\mathbf{a}\} \in \mathbb{C}^{M \times M}$ Constructs a diagonal matrix from vector $\mathbf{a} \in \mathbb{C}^{M}$

$$\mathrm{diag}\{\mathbf{a}\} = \begin{bmatrix} [\mathbf{a}]_{(1)} & \dots & \dots & 0 \\ 0 & [\mathbf{a}]_{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & [\mathbf{a}]_{(M)} \end{bmatrix}$$

APPENDIX B

MATHEMATICAL DEFINITIONS

B.1 Mappings and Functions

B.1.1 Definitions

Definition 4. A *mapping* $\mathcal{M}$ is defined as a rule which associates elements of the set A with elements of the set B [MS00].

$$\mathcal{M} : A \mapsto B \tag{B.1}$$

Definition 5. A *function* f is a rule, which associates each point in the function domain $a \in A$ to one point in the codomain $b \in B$ [MS00].

$$f : A \mapsto B \tag{B.2}$$

The short hand notation is $b = f(a)$.

A function is *injective* or an one-to-one function, if $\forall a_1, a_2 \in A$ the equality $f(a_1) = f(a_2)$ implies $a_1 = a_2$ holds [DG00].

Definition 6. The *inverse function* f^{-1} of function f is defined as [DG00, MS00]

$$f^{-1} : B \mapsto A \quad . \tag{B.3}$$

The inverse is given, if $f^{-1}(f(a)) = a, \forall a \in A$ holds.

Definition 7. A *multivariate function* is a function which depends on multiple variables $a_n \in A_n$. Therefore, the function domain is a set of N-tuples $A^N \subseteq A_1 \times \dots A_N$.

$$f : A^N \mapsto B \tag{B.4}$$

The short hand notation is $b = f(a_1, \dots, a_N) = f(\mathbf{a})$, with point $\mathbf{a}$ in the vector space A^N.

Definition 8. If the codomain is also a vector space the function becomes a *vector-valued function*, which maps points $\mathbf{a} \in A^N$ to points $\mathbf{b} \in B^M$.

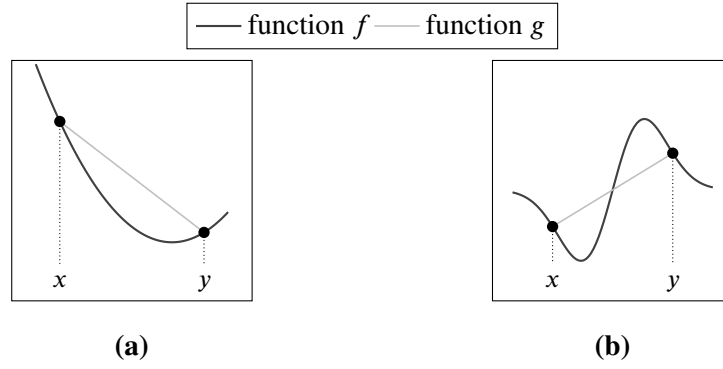$$\mathbf{f} : A^N \mapsto B^M \tag{B.5}$$

Figure B.1: *Example of a 1-dimensional function f being (a) convex and (b) non-convex, and linear function g.*

The short hand notation is $\mathbf{b} = \mathbf{f}(\mathbf{a})$.

B.1.2 Convex Function

Definition 9. A function $f : A \mapsto \mathbb{R}$ is said to be convex, iff its domain A is a convex set and for all $\mathbf{x}, \mathbf{y} \in A$ and $\forall \lambda \in [0, 1]$

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}) \tag{B.6}$$

holds [BV04]. Accordingly, a 1-dimensional function f with domain A can be said to be convex, if each function value $f(z)$ on the interval $z \in [x, y]$ is always lower or equal than the function value $g(z)$ of a linear function g connecting the interval edges $x, y \in A$. Figure B.1 shows an example of a 1-dimensional convex and non-convex function.

B.2 Description of Random Processes

B.2.1 Wide-Sense Stationary Processes

A random process $z(t)$ is wide-sense stationary, if its first-order moment m_z and its second-order moment ψ_{zz} do not vary over time [Hay94].

$$\mathrm{E}\{z(t)\} = m_z \tag{B.7a}$$

$$\mathrm{E}\{[z(t_1) - m] \cdot [z(t_2) - m]\} = \psi_{zz}(t_1, t_2) = \psi_{zz}(t_1 - t_2) \tag{B.7b}$$

If $z(t)$ is complex-valued, than $z(t)$ is wide-sense stationary, if its real and imaginary part are jointly wide-sense stationary. Wide-sense stationary in time domain implies, that distinct frequency bins of $Z(f) \bullet\!\!-\!\!\circ z(t)$ are uncorrelated.

$$\mathrm{E}\{Z(f_1) \cdot Z(f_2)\} = 0 \quad , \quad \forall f_1 \neq f_2 \tag{B.8}$$

Hence, the power spectral density $\Psi_{ZZ}(f_1, f_2)$ is given by

$$\Psi_{ZZ}(f_1, f_2) = \Psi_{ZZ}(f_1) \cdot \delta(f_2 - f_1) \quad . \tag{B.9}$$

B.2.2 Probability Theory and Probability Density Function

A Brief Introduction to Probability Theory

Random quantities are defined on a common *probability space*, which is a special *measure space* given by the triple $(\Omega, \mathcal{F}, P)$ [BS94, Loè77, Sul15]. The tuple $(\Omega, \mathcal{F})$ denotes a *measurable space*, defined by the non-empty set Ω and $\mathcal{F}$ a σ-algebra on Ω. Assuming $\mathcal{F}$ as a Borel σ-algebra $\mathcal{B}$ leads to the measurable space $(\Omega, \mathcal{B}(\Omega))$, where e.g. $\Omega \subseteq \mathbb{R}$.

In stochastic, the measurable space is denoted as event space. The set Ω is called the sample space, which denotes the set of all elementary results of e.g. an empirical observation. The set of events $\mathcal{F} \subseteq \Omega$ denotes subsets of the sample space, comprising all conceivable combinations or permutations of elementary events of Ω. In order to get a proper measure space a measure P on the measurable space $(\Omega, \mathcal{F})$ is introduced. Because the measure space should be a probability space, the measure is a probability measure. Introducing the function $P(B)$ as a probability measure on $\mathcal{F}$, which fulfils $P(\Omega) = 1$. This function assigns each event $B \subseteq \mathcal{F}$ a likelihood of happening [BS94, Loè77]: $P : \mathcal{F} \mapsto [0, 1]$.

Example B.1. For illustration purpose of the above definition, consider tossing a six-sided die once. Accordingly, the sample space is $\Omega := \{1, 2, 3, 4, 5, 6\}$. For the set of events all subsets, which are of interest, can be considered, e.g. $\mathcal{F} := \{\Omega, 1, 3, 5\}$. According to the probability measure, a probability can be assigned to each subset in $\mathcal{F}$. In case of an ideal and faire die, these probabilities may be e.g. $P(1) = \frac{1}{6}$, or $P(1, 3, 5) = \frac{1}{2}$.

Probability Density Function

Stochastic processes are defined as a collection of random quantities or variables. Consider a random quantity X. This quantity is defined as a function on the probability space [BS94, Loè77], assigning a variable $X(\omega) \subseteq \mathbb{R}$ of the observation space to each elementary outcome $\omega \in \Omega$ of the sample space [ABT13, BS94, Zac81].

$$X : \Omega \mapsto \mathbb{R} \tag{B.10}$$

The *expectation value* E of the random variable X is defined as [Zac81]

$$\mathrm{E}\{X\} = \int_{\Omega} X(\omega)\, dP(d\omega) \tag{B.11}$$

with $dP(d\omega)$ the probability of event ω happening. Assuming $dP(d\omega)$ as continuous, the probability can be described in terms of a non-negative density function $p(\omega)$.

$$p(\omega) = \frac{dP(d\omega)}{d\omega} \tag{B.12}$$

Function $p(\omega)$ is denoted as probability density function (PDF).

Definition 10. A PDF $p(\omega)$ is said to be *proper*, if $p(\omega) \geq 0, \forall \omega \in \Omega$ and $\int_{\omega \in \Omega} p(\omega) \, d\omega = 1$ are satisfied. Otherwise, the PDF is said to be *improper*.

The probability measure, which assigns a probability of the happening of event B, is given by integral of the PDF over the event's space.

$$P(B) = \int_B p(\omega) \, d\omega \tag{B.13}$$

Hence, a PDF is a convenient way to model random processes by assigning a probability of happening to their quantities.

Probability Density of Complex Random Signals

Defining a stationary, complex random multivariate vector $\mathbf{y} = \mathbf{y}_R + J\mathbf{y}_I \in \mathbb{C}^N$. Vector $\mathbf{y}_R \in \mathbb{R}^N$ and $\mathbf{y}_I \in \mathbb{R}^N$ denote the random real and imaginary part, respectively. The respective PDF is given by the joint probability density function of the real and imaginary part.

$$p(\mathbf{y}) = p\left(\mathbf{y}_R + J\mathbf{y}_I\right) \triangleq p\left(\mathbf{y}_R, \mathbf{y}_I\right) \tag{B.14}$$

Next, the complex augmented vector $\bar{\mathbf{y}}$ is introduced, comprising the complex random variable and its complex conjugate.

$$\bar{\mathbf{y}} = \left[\mathbf{y}^T, \mathbf{y}^H\right]^T \tag{B.15}$$

The mean vector and the augmented mean vector are given by their respective expectation.

$$\boldsymbol{\mu} = \mathrm{E}\{\mathbf{y}\} \tag{B.16a}$$

$$\bar{\boldsymbol{\mu}} = \mathrm{E}\{\bar{\mathbf{y}}\} \tag{B.16b}$$

The second-order statistic is given by the augmented covariance matrix [ASS11]

$$\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}} = \mathrm{E}\left\{(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}) \cdot (\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}})^H\right\} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{yy}} & \tilde{\boldsymbol{\Sigma}}_{\mathbf{yy}} \\ \tilde{\boldsymbol{\Sigma}}_{\mathbf{yy}}^\dagger & \boldsymbol{\Sigma}_{\mathbf{yy}}^\dagger \end{bmatrix} \quad , \tag{B.17}$$

with the regular covariance matrix

$$\boldsymbol{\Sigma}_{\mathbf{yy}} = \mathrm{E}\left\{(\mathbf{y} - \boldsymbol{\mu}) \cdot (\mathbf{y} - \boldsymbol{\mu})^H\right\} = \boldsymbol{\Sigma}_{\mathbf{yy}}^H \tag{B.18}$$

and the complementary covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_{\mathbf{yy}} = \mathrm{E}\left\{(\mathbf{y} - \boldsymbol{\mu}) \cdot (\mathbf{y} - \boldsymbol{\mu})^T\right\} = \tilde{\boldsymbol{\Sigma}}_{\mathbf{yy}}^T \quad . \tag{B.19}$$

If the complementary covariance matrix is zero ($\tilde{\boldsymbol{\Sigma}}_{\mathbf{yy}} = \mathbf{0}$), the random variable is uncorrelated with its complex conjugate. Such a complex random variable is stated as *proper*, otherwise as *improper* [SS10]. Propriety implies that the covariance and cross-covariance matrices of the

real and imaginary part fulfil

$$\Sigma_{y_R y_R} = \Sigma_{y_I y_I} \tag{B.20a}$$

$$\Sigma_{y_R y_I} = -\Sigma_{y_I y_R}^T \quad . \tag{B.20b}$$

Another property, which is commonly assigned to complex random variables, is *circularity*. A complex random variable is circular, if its PDF is rotationally invariant in the complex plane. Define the rotated version of **y** in the complex plane.

$$\acute{\mathbf{y}} = \exp\{J\varphi\} \cdot \mathbf{y} \tag{B.21}$$

The random vector is circular, if $\mathbf{y} \equiv \acute{\mathbf{y}}$ [ASS11]. Basically, propriety requires that the second-order statistics (second moment) is rotationally invariant, whereas circularity requires that all moments are rotationally invariant.

Special Case: Normal Distribution Consider the complex random vector as normal distributed. Hence, its real and imaginary parts are jointly normal distributed. The respective PDF is [SS10]

$$\mathbf{y} \sim p(\mathbf{y}) = \mathcal{N}\left(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}\right) = \frac{1}{\pi^N \sqrt{\left|\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}\right|}} \cdot \exp\left\{-\frac{1}{2}(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}})^H \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}^{-1}(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}})\right\} \quad . \tag{B.22}$$

Furthermore, assume the complex random vector **y** as proper. The PDF simplifies, because the covariance matrix $\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$ becomes of block diagonal structure. The determinant reduces to [PP12]

$$\left|\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}\right| = \left|\begin{bmatrix} \boldsymbol{\Sigma}_{yy} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy}^\dagger \end{bmatrix}\right| = \left|\boldsymbol{\Sigma}_{yy}\right| \cdot \left|\boldsymbol{\Sigma}_{yy}^\dagger\right| = \left|\boldsymbol{\Sigma}_{yy}\right|^2 \quad . \tag{B.23}$$

Also, the argument of the exponential simplifies due to the block-diagonal structure of $\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$.

$$\frac{1}{2}(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}})^H \cdot \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}^{-1} \cdot (\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}) = (\mathbf{y} - \boldsymbol{\mu})^H \cdot \bar{\boldsymbol{\Sigma}}_{yy}^{-1} \cdot (\mathbf{y} - \boldsymbol{\mu}) \tag{B.24}$$

Plugging in yields the PDF of a proper complex normal distributed vector [SS10].

$$\mathbf{y} \sim p(\mathbf{y}) = \mathcal{CN}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{yy}\right) = \frac{1}{\pi^N \left|\boldsymbol{\Sigma}_{yy}\right|} \cdot \exp\left\{-(\mathbf{y} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu})\right\} \tag{B.25}$$

$\mathcal{CN}$ denotes the proper complex normal distribution, with the first and second argument denoting the mean and variance, respectively. Quite often, measurement noise is assumed as complex normal distributed, whereas circularity is implicitly assumed. Circularity of a normal distributed random vector implies propriety and also zero-mean [SS10].

APPENDIX C

Derivations for FMCW Radars

C.1 Degradation of the Resolution

The delayed, periodically radiated transmit signal at the stretch processing receiver is

$$s_P(t) * \delta(t - \tau) = \sum_{n=-\infty}^{\infty} s\left(t - n \cdot T_P\right) * \delta(t - \tau) = \sum_{n=-\infty}^{\infty} s\left(t - n \cdot T_P - \tau\right) \quad . \qquad \text{(C.1)}$$

The output signal $x(t)$ of the stretch processor is given by the mixing of the transmit and echo signal

$$x(t) = s_P(t)^{\dagger} \cdot \left[s_P(t) * \delta(t - \tau)\right] = \sum_{n_1=-\infty}^{\infty} s\left(t - n_1 \cdot T_P\right)^{\dagger} \cdot \sum_{n_2=-\infty}^{\infty} s\left(t - n_2 \cdot T_P - \tau\right)$$

$$= \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} s\left(t - n_1 \cdot T_P\right)^{\dagger} \cdot s\left(t - n_2 \cdot T_P - \tau\right) \quad . \qquad \text{(C.2)}$$

The respective complex baseband signal is

$$x(t) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \text{rect}\left(\frac{t - \frac{T_M}{2} - n_1 \cdot T_P}{T_M}\right) \cdot \text{rect}\left(\frac{t - \frac{T_M}{2} - \tau - n_2 \cdot T_P}{T_M}\right)$$

$$\cdot \exp\left\{-j2\pi\left(f_0 \cdot (t - n_1 \cdot T_P) + \frac{W}{2T_M} \cdot (t - n_1 \cdot T_P)^2\right)\right\}$$

$$\cdot \exp\left\{j2\pi\left(f_0 \cdot (t - n_2 \cdot T_P - \tau) + \frac{W}{2T_M} \cdot (t - n_2 \cdot T_P - \tau)^2\right)\right\} \quad . \qquad \text{(C.3)}$$

To simplify the above equation, the following lemma is necessary:

Lemma 1. *Multiplication of two rectangular functions with the same width T, but different*

shifts $a, b \in \mathbb{R}$ is

$$\operatorname{rect}\left(\frac{t-a}{T}\right) \cdot \operatorname{rect}\left(\frac{t-b}{T}\right) = \operatorname{rect}\left(\frac{t - \frac{1}{2} \cdot (a+b)}{T - |a-b|}\right) \quad .$$

Accordingly, the output signal of the stretch processor simplifies to

$$
\begin{aligned}
x(t) = {} & \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \operatorname{rect}\left(\frac{t - \frac{1}{2} \cdot [\tau + T_P\left(n_2 - n_1\right) + T_M]}{T_M - \left|T_P\left(n_2 - n_1\right) + \tau\right|}\right) \\
& \cdot \exp\left\{ J 2\pi \left(f_0 n_1 T_P + \frac{W}{T_M} T_P n_1 t - \frac{W}{2 T_M} n_1^2 T_P^2 \right) \right\} \\
& \cdot \exp\left\{ J 2\pi \left(-n_2 f_0 T_P - f_0 \tau + \frac{W}{2 T_M} n_2^2 T_P^2 + \frac{W}{2 T_M} \tau^2 - \frac{W}{T_M} T_P n_2 t - \frac{W}{T_M} t\tau + \frac{W}{T_M} T_P n_2 \tau \right) \right\} \\
= {} & \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \operatorname{rect}\left(\frac{t - \frac{1}{2} \cdot [\tau + T_P(n_1 + n_2) + T_M]}{T_M - \left|T_P\left(n_2 - n_1\right) + \tau\right|}\right) \\
& \cdot \exp\left\{ J 2\pi \left(\left(n_2 - n_1\right) f_0 T_P + \left(n_2^2 - n_1^2\right) \frac{W}{2 T_M} T_P^2 - f_0 \cdot \tau + \frac{W}{2 T_M} \tau^2 + n_2 \frac{W}{T_M} T_P \tau \right) \right\} \\
& \cdot \exp\left\{ J 2\pi \left(\left(n_2 - n_1\right) \frac{W}{T_M} T_P t - \frac{W}{T_M} t\tau \right) \right\} \quad .
\end{aligned}
\tag{C.4}
$$

The argument of the first exponential is constant w.r.t. the time and therefore does not result in a beat frequency. Hence, this term can be neglected. The argument of the second exponential results in the beat frequency, which, according to equation (4.4), is

$$
f_b = \frac{d}{dt}\left(\left(n_1 - n_2\right) \cdot W \cdot \frac{T_P}{T_M} \cdot t - \frac{W}{T_M} \cdot t \cdot \tau \right) = W \cdot \left(\left(n_1 - n_2\right) \cdot \frac{T_P}{T_M} - \frac{\tau}{T_M} \right) \quad . \tag{C.5}
$$

In order derive the resulting beat frequency, the relation between the variables n_1 and n_2 has to be examined. Assuming a non-ambiguous delay, hence $0 \leq \tau < T_P$ holds, the beat frequency must fulfil $0 \leq |f_b| < W T_P / T_M$. Plugging into the above formula for the beat frequency yields

$$
0 \leq \left| \left(n_1 - n_2\right) \cdot \frac{T_P}{T_M} - \frac{\tau}{T_M} \right| < \frac{T_P}{T_M} \quad . \tag{C.6}
$$

Because the quotient $\frac{\tau}{T_P}$ has to fulfil $0 \leq \frac{\tau}{T_P} < 1$ and $\left(n_1 - n_2\right) \in \mathbb{Z}$ holds, the following criteria for the variable n_1 and n_2 is given

$$
0 \leq n_1 - n_2 \leq 1 \quad . \tag{C.7}
$$

Accordingly, the indices n_1 and n_2 can have the following relations: $n_1 = n_2$ and $n_1 = n_2 + 1$. Hence, two terms remain for the double summation. Furthermore, it will be assumed, that the period equals the modulation time ($T_M \approx T_P$) and the signal recovery time and wait time is

zero: $T_R \approx 0$ and $T_W \approx 0$. The output signal of the stretch processor becomes

$$
\begin{aligned}
x(t) \approx &\sum_{n=-\infty}^{\infty} \text{rect}\left(\frac{t - \frac{1}{2} \cdot [\tau + T_M \cdot (2n+1)]}{T_M - \tau}\right) \cdot \exp\left\{j2\pi\left(-\frac{W}{T_M}t\tau - f_0\tau + \frac{W}{2T_M}\tau^2 + Wn\tau\right)\right\} \\
&+ \sum_{n=-\infty}^{\infty} \text{rect}\left(\frac{t - \frac{1}{2} \cdot [\tau + T_M \cdot (2n+2)]}{\tau}\right) \\
&\cdot \exp\left\{j2\pi\left(Wt - \frac{W}{T_M}t\tau + f_0(T_M - \tau) + \frac{W}{2T_M}\tau^2 + Wn\tau - (2n+1)\frac{WT_M}{2}\right)\right\} \\
= &\sum_{n=-\infty}^{\infty} \text{rect}\left(\frac{t - \frac{\tau}{2} - nT_M}{\tau}\right) \\
&\cdot \exp\left\{j2\pi\left(Wt - \frac{W}{T_M}t\tau + f_0(T_M - \tau) + \frac{W}{2T_M}\tau^2 + Wn\tau - (2n+1)\frac{WT_M}{2}\right)\right\} \\
&+ \sum_{n=-\infty}^{\infty} \text{rect}\left(\frac{t - \frac{T_M+\tau}{2} - nT_M}{T_M - \tau}\right) \cdot \exp\left\{j2\pi\left(-\frac{W}{T_M}t\tau - f_0\tau + \frac{W}{2T_M}\tau^2 + Wn\tau\right)\right\} \quad .
\end{aligned}
$$

(C.8)

Note, that the first exponential in equation (C.4) has been neglected in the above equation, because the term is constant w.r.t. the time. The beat frequencies for both summands in equation (C.8) can be calculated according to equation (4.4). The resulting beat frequencies are

$$
f_{b_1} = \frac{W}{T_M} \cdot \tau
$$

(C.9a)

$$
f_{b_2} = W - \frac{W}{T_M} \cdot \tau \quad ,
$$

(C.9b)

which are present at disjunct time instances in each period (see Fig. 4.3). Beat frequency f_{b_1} results from the frequency shift between the received and the actually transmitted chirp, and lasts $T_M - \tau$ per period. Beat frequency f_{b_2} results from the frequency shift between the received and the next transmitted chirp, and lasts τ per period. Commonly, the second beat frequency is not captured by the ADC, because $f_S \ll W$ and $\tau \ll T_M$. Hence, the effectively processed beat frequency is f_{b_1}. Accordingly, the received baseband signal for a single period is

$$
x(t) \approx \text{rect}\left(\frac{t - \frac{T_M+\tau}{2}}{T_M - \tau}\right) \cdot \exp\left\{j2\pi\left(\frac{W}{2T_M}\tau^2 - \frac{W}{T_M}t\tau - f_0\tau\right)\right\} \quad .
$$

(C.10)

C.2 Stretch Processing of Multiple Received FMCW Signals

The signal at the receiver is given by the convolution of the transmit signal model (4.32) and the propagation model (4.35).

$$
\begin{aligned}
s(t)*h(t) &= \sum_p \gamma_p \cdot \delta\left(t - \tau_p\right) * \sum_n \sum_m c_{n,m} \left[\cos(\phi_{n,m}^+(t)) + \cos(\phi_{n,m}^-(t))\right] \\
&= \sum_n \sum_m \sum_p \gamma_p \cdot c_{n,m} \cdot \left[\cos\left(\phi_{n,m}^+\left(t - \tau_p\right)\right) + \cos\left(\phi_{n,m}^-\left(t - \tau_p\right)\right)\right]
\end{aligned}
\tag{C.11}
$$

The corresponding phase terms are

$$
\phi_{n,m}^+(t - \tau_p) = \phi_{n,m}^+(t) - \theta_{n,m}\left(t, \tau_p\right) - 2\pi\tau_p n f_{\text{LO}}
\tag{C.12a}
$$

$$
\phi_{n,m}^-(t - \tau_p) = \phi_{n,m}^-(t) + \theta_{n,m}\left(t, \tau_p\right) - 2\pi\tau_p n f_{\text{LO}}
\tag{C.12b}
$$

$$
\theta_{n,m}(t, \tau_p) = 2\pi\tau_p \left(m f_{\text{RO}} + m\frac{W_{\text{RO}}}{T_{\text{M}}}t - m\frac{W_{\text{RO}}}{2T_{\text{M}}}\tau_p\right) \quad .
\tag{C.12c}
$$

Plugging equation (C.11) into equation (4.34) yields the baseband signal.

$$
\begin{aligned}
x(t) &= \sum_p \gamma_p \cdot \int_t^{t+T} \sum_{n_1} \sum_{m_1} \sum_{n_2} \sum_{m_2} \cdot c_{n_2,m_2} \cdot \left[\cos\left(\phi_{n_2,m_2}^+\left(t' - \tau_p\right)\right) + \cos\left(\phi_{n_2,m_2}^-\left(t' - \tau_p\right)\right)\right] \\
&\quad \cdot c_{n_1,m_1} \cdot \left[\cos\left(\phi_{n_1,m_1}^+(t')\right) + \cos\left(\phi_{n_1,m_1}^-(t')\right)\right] \, dt'
\end{aligned}
\tag{C.13}
$$

Using trigonometric identity [1] the phase terms of the resulting cosines are the sum/difference of $\phi_{n_1,m_1}^\pm(t)$ and $\phi_{n_2,m_2}^\pm(t)$. The constant terms $\theta_{n,m}(t, \tau_p)$ and $2\pi\tau_p n f_{\text{LO}}$ will be ignored for now. The resulting phase terms are listed in equation (C.15). Keeping in mind that $n_1, n_2, m_1, m_2 \geq 1$ holds, only phase terms with the difference of the sum indices ($n_1 - n_2$ and $m_2 - m_1$) will remain after integration (low-pass filtering). Other phase terms, which result in a frequency above the cut-off frequency of the low-pass filter, will vanish or at least become negligible. Finally, the baseband signal is

$$
\begin{aligned}
x(t) &\approx \sum_p \frac{\gamma_p}{2} \cdot \sum_{n_1} \sum_{m_1} \sum_{n_2} \sum_{m_2} c_{n_1,m_1} \cdot c_{n_2,m_2} \cdot \cos\left(\phi_{n_1,m_1}^+(t) - \phi_{n_2,m_2}^+(t - \tau_p)\right) \\
&\quad + \sum_p \frac{\gamma_p}{2} \cdot \sum_{n_1} \sum_{m_1} \sum_{n_2} \sum_{m_2} c_{n_1,m_1} \cdot c_{n_2,m_2} \cdot \cos\left(\phi_{n_1,m_1}^-(t) - \phi_{n_2,m_2}^-(t - \tau_p)\right) \quad .
\end{aligned}
\tag{C.14}
$$

---

[1] $\cos x \cos y = 0.5 \cdot [\cos(x - y) + \cos(x + y)]$

$$\phi^+_{n_1,m_1}(t) + \phi^+_{n_2,m_2}(t) = 2\pi t\left(n_1 + n_2\right) f_{\text{LO}} + \pi\left(m_1 + m_2\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t\left(m_1 + m_2\right) f_{\text{RO}}$$
$$\text{(C.15a)}$$

$$\phi^-_{n_1,m_1}(t) + \phi^-_{n_2,m_2}(t) = 2\pi t\left(n_1 + n_2\right) f_{\text{LO}} - \pi\left(m_1 + m_2\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 - 2\pi t\left(m_1 + m_2\right) f_{\text{RO}}$$
$$\text{(C.15b)}$$

$$\phi^+_{n_1,m_1}(t) - \phi^+_{n_2,m_2}(t) = 2\pi t\left(n_1 - n_2\right) f_{\text{LO}} + \pi\left(m_1 - m_2\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t\left(m_1 - m_2\right) f_{\text{RO}}$$
$$\text{(C.15c)}$$

$$\phi^-_{n_1,m_1}(t) - \phi^-_{n_2,m_2}(t) = 2\pi t\left(n_2 - n_1\right) f_{\text{LO}} + \pi\left(m_2 - m_1\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t\left(m_2 - m_1\right) f_{\text{RO}}$$
$$\text{(C.15d)}$$

$$\phi^+_{n_1,m_1}(t) + \phi^-_{n_2,m_2}(t) = 2\pi t\left(n_1 + n_2\right) f_{\text{LO}} + \pi\left(m_1 - m_2\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t\left(m_1 - m_2\right) f_{\text{RO}}$$
$$\text{(C.15e)}$$

$$\phi^-_{n_1,m_1}(t) + \phi^+_{n_2,m_2}(t) = 2\pi t\left(n_1 + n_2\right) f_{\text{LO}} + \pi\left(m_2 - m_1\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t\left(m_2 - m_1\right) f_{\text{RO}}$$
$$\text{(C.15f)}$$

$$\phi^-_{n_1,m_1}(t) - \phi^+_{n_2,m_2}(t) = 2\pi t\left(n_1 - n_2\right) f_{\text{LO}} - \pi\left(m_1 + m_2\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 - 2\pi t\left(m_1 + m_2\right) f_{\text{RO}}$$
$$\text{(C.15g)}$$

$$\phi^+_{n_1,m_1}(t) - \phi^-_{n_2,m_2}(t) = 2\pi t\left(n_1 - n_2\right) f_{\text{LO}} + \pi\left(m_1 + m_2\right) \frac{W_{\text{RO}}}{T_{\text{M}}} t^2 + 2\pi t\left(m_1 + m_2\right) f_{\text{RO}}$$
$$\text{(C.15h)}$$

C.3 FMCW Signal and Doppler Shift

Consider a time-variant delay, which varies linearly.

$$\tau(t) = \tau_0 + \frac{v_0}{c_0} t \tag{C.16}$$

Plugging into equation (4.12) yields the following expression for the phase of the baseband signal.

$$\begin{aligned}
\frac{1}{2\pi}\phi(t) &= \frac{W}{2T_{\text{M}}}\tau(t)^2 - \frac{W}{T_{\text{M}}}\tau(t)t - f_0\tau(t) \\
&= \frac{W}{2T_{\text{M}}}\left[\tau_0^2 + 2\frac{v_0}{c_0}\tau_0 t + \left(\frac{v_0}{c_0}t\right)^2\right] - \frac{W}{T_{\text{M}}}\tau_0 t - \frac{W}{T_{\text{M}}}\frac{v_0}{c_0}t^2 - f_0\tau_0 - \frac{v_0}{c_0}f_0 t \tag{C.17}
\end{aligned}$$

If the velocity of the scatterer is small, i.e. $|v_0| \ll c_0$ holds, the following approximations can be drawn.

$$\frac{v_0}{c_0}\frac{W}{T_\mathrm{M}}\tau_0 t - \frac{W}{T_\mathrm{M}}\tau_0 t = \frac{W}{T_\mathrm{M}}\tau_0 t \left(\frac{v_0}{c_0} - 1\right) \quad \approx -\frac{W}{T_\mathrm{M}}\tau_0 t \qquad \text{(C.18a)}$$

$$\frac{W}{2T_\mathrm{M}}\left(\frac{v_0}{c_0}\right)^2 t^2 - \frac{W}{T_\mathrm{M}}\frac{v_0}{c_0}t^2 = \frac{W}{T_\mathrm{M}}\frac{v_0}{c_0}t^2\left(\frac{v_0}{2c_0} - 1\right) \approx -\frac{W}{T_\mathrm{M}}\frac{v_0}{c_0}t^2 \qquad \text{(C.18b)}$$

The residual video phase term $W\tau_0^2/2T_\mathrm{M}$ will be neglected too, because it is commonly very small [FJ15]. Accordingly, the phase term (C.17) reduces to

$$\frac{1}{2\pi}\phi(t) \approx -\frac{W}{T_\mathrm{M}}\tau_0 t - \frac{W}{T_\mathrm{M}}\frac{v_0}{c_0}t^2 - f_0\tau_0 - \frac{v_0}{c_0}f_0 t \quad . \qquad \text{(C.19)}$$

Commonly, it is assumed, that the Doppler shift w.r.t. the frequency f_0 dominates (narrowband approximation of the Doppler shift, see Section 3.2.1).

$$\frac{W}{T_\mathrm{M}}\frac{v_0}{c_0}t^2 \ll \frac{v_0}{c_0}f_0 t \qquad \text{(C.20)}$$

Approximation (C.20) necessitates that

$$\frac{W}{f_0} \ll \frac{T_\mathrm{M}}{t} \qquad \text{(C.21)}$$

holds. Hence, a short observation time has to be assumed. If so, the phase term (C.17) finally reduces to [FJ15]

$$\phi(t) \approx -2\pi\left(f_0\frac{v_0}{c_0}t + \frac{W}{T_\mathrm{M}}\tau_0 t + f_0\tau_0\right) \quad . \qquad \text{(C.22)}$$