TECHNISCHE UNIVERSITÄT DRESDEN

Experimental Investigations of Millimeter Wave Beamforming

Tobias Kadur

der Fakultät Elektrotechnik und Informationstechnik
der Technischen Universität Dresden

zur Erlangung des akademischen Grades eines

Doktoringenieurs
(Dr.-Ing.)

genehmigte Dissertation

Vorsitzender:	Prof. Dr.-Ing. Hubert Lakner
Gutachter:	Prof. Dr.-Ing. Dr. h.c. Gerhard Fettweis
	Prof. Dr.-Ing. Hans Dieter Schotten
Tag der Einreichung:	06.05.2019
Tag der Verteidigung:	03.12.2019

Acknowledgment

First and foremost, I want to thank the almighty God for guiding me and giving me strength and blessings every single day. Without His blessings, this work would have not been created.

I want to express my deep gratitude to my supervisor Prof. Gerhard Fettweis for giving me the chance of performing this work. Already in undergraduate courses, he was raising my interest in communications, which peaked in my dissertation. He provided me with guidance throughout the course of the thesis. Without his interesting ideas and extraordinary, continuous support, this work would not have been possible. I would also like to thank Prof. Hans Dieter Schotten for reviewing this work as my second assessor.

Special thanks goes to Dr. Wolfgang Rave who gave important research impulses. I thank him for sharing his in-depth knowledge and reviewing papers and this dissertation.

I want to thank my colleagues Amanda, Song and Mostafa for interesting discussions and reviewing this work. Additionally, thankfulness goes to Marcus Wagner for reviewing this thesis.

Finally, I want to thank my family, especially my wife Tine and our daughter Joela for supporting me and keeping my back free in stressful situations during writing this thesis.

Dresden, April 2019 Tobias Kadur

Abstract

The millimeter wave (mmW) band, commonly referred to as the frequency band between 30 GHz and 300 GHz, is seen as a possible candidate to increase achievable rates for mobile applications due to the existence of free spectrum. However, the high path loss necessitates the use of highly directional antennas. Furthermore, impairments and power constraints make it difficult to provide full digital beamforming systems.

In this thesis, we approach this problem by proposing effective beam alignment and beam tracking algorithms for low-complex analog beamforming (ABF) systems, showing their applicability by experimental demonstration.

After taking a closer look at particular features of the mmW channel properties and introducing the beamforming as a spatial filter, we begin our investigations with the application of detection theory for the non-convex beam alignment problem. Based on an M-ary hypothesis test, we derive algorithms for defining the length of the training signal efficiently. Using the concept of black-box optimization algorithms, which allow optimization of non-convex algorithms, we propose a beam alignment algorithm for codebook-based ABF based systems, which is shown to reduce the training overhead significantly. As a low-complex alternative, we propose a two-staged gradient-based beam alignment algorithm that uses convex optimization strategies after finding a subregion of the beam alignment function in which the function can be regarded convex. This algorithm is implemented in a real-time prototype system and shows its superiority over the exhaustive search approach in simulations and experiments.

Finally, we propose a beam tracking algorithm for supporting mobility. Experiments and comparisons with a ray-tracing channel model show that it can be used efficiently in line of sight (LoS) and non line of sight (NLoS) scenarios for walking-speed movements.

Publications of the Author

Journal Publications & Book Chapters

[1] Hsiao-Lan Chiang, Wolfgang Rave, **Tobias Kadur** and Gerhard Fettweis. Hybrid Beamforming Based on Implicit Channel State Information for Millimeter Wave Links. In *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 326-339, May 2018.

Conference Publications

[1] **Tobias Kadur**, Wolfgang Rave and Gerhard Fettweis. Application of Multiple Hypothesis Testing for Beam Selection. In *Proceedings of IEEE International Conference on Communications (ICC)*, Shanghai, 2019, accepted for publication.

[2] **Tobias Kadur**, Hsiao-Lan Chiang and Gerhard Fettweis. Experimental Validation of Robust Beam Tracking in a NLoS Indoor Environment. In *Proceedings of IEEE International Conference on Telecommunications (ICT)*, Saint Malo, 2018.

[3] **Tobias Kadur**, Wolfgang Rave, Hsiao-Lan Chiang and Gerhard Fettweis. Experimental validation of a robust beam alignment algorithm in an indoor environment. In *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*, Barcelona, 2018.

[4] **Tobias Kadur**, Hsiao-Lan Chiang and Gerhard Fettweis. Efficient Gradient-Based Beam Alignment for Low-Cost mmW Devices. In *Proceedings of the Smartcom (Smartcom'17)*, Rome, 2017.

[5] **Tobias Kadur**, Hsiao-Lan Chiang and Gerhard Fettweis. Effective Beam Alignment Algorithm for Low Cost Millimeter Wave Communication. In *Proceedings of IEEE Vehicular Technology Conference (VTC-Fall)*, Montreal, QC, 2016.

[6] **Tobias Kadur**, Ivan Gaspar, Nicola Michailow and Gerhard Fettweis. Non-Data Aided Frequency Synchronization Exploiting ICI in Non-Orthogonal Systems. In *Proceedings of IEEE Vehicular Technology Conference (VTC-Fall)*, Vancouver, BC, Canada 2014.

[7] Hsiao-Lan Chiang, **Tobias Kadur** and Gerhard Fettweis. Frequency-Selective Hybrid Beamforming Based on Implicit CSI for Millimeter Wave Systems. In *Proceedings of IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, 2018.

[8] Hsiao-Lan Chiang, Wolfgang Rave, **Tobias Kadur** and Gerhard Fettweis. A low-complexity beamforming method by orthogonal codebooks for millimeterwave links. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017.

[9] Hsiao-Lan Chiang, **Tobias Kadur** and Gerhard Fettweis. Analyses of orthogonal and non-orthogonal steering vectors at millimeter wave systems. In *Proceedings of IEEE International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Coimbra, 2016.

[10] Hsiao-Lan Chiang, Wolfgang Rave, **Tobias Kadur** and Gerhard Fettweis. Full rank spatial channel estimation at millimeter wave systems. In *Proceedings of International Symposium on Wireless Communication Systems (ISWCS)*, Poznan, 2016.

[11] Hsiao-Lan Chiang, Wolfgang Rave, **Tobias Kadur** and Gerhard Fettweis. Low-complexity spatial channel estimation and hybrid beamforming for millimeter wave links. In *Proceedings of IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Valencia, 2016.

[12] Yaning Zou, Per Zetterberg, Ulf Gustavsson, Tommy Svensson, Ali Zaidi, **Tobias Kadur**, Wolfgang Rave and Gerhard Fettweis. Impact of Major RF Impairments on mm-Wave Communications Using OFDM Waveforms. In *Proceedings of IEEE Globecom Workshops (GC Workshops)*, Washington, DC, 2016.

Contents

List of Abbreviations

3GPP	Third Generation Partnership Project
ABF	analog beamforming
ADC	analog to digital converter
AGC	automatic gain control
AoA	angle of arrival
AoD	angle of departure
AP	access point
BER	bit error rate
CDF	cumulative density function
CFO	carrier frequency offset
CP	cyclic prefix
CS	compressed sensing
CSI	channel state information
D-MPS	mode pursuing sampling method for a discrete variable space
DAC	digital to analog converter
DAB	digital audio broadcasting
DBF	digital beamforming
DECT	digital enhanced cordless telecommunications
DFT	discrete Fourier transform
DL	downlink
DTFT	discrete time Fourier transform
DVB-T	terrestrial digital video broadcasting

EIRP	equivalent isotropic radiation power
ESPRIT	estimation of signal parameters via rotational invariance techniques
ES	exhaustive search
FDD	frequency division duplex
FFT	fast Fourier transform
FPGA	field programmable gate array
GLRT	generalized likelihood ratio test
GPS	global positioning system
GSCM	geometry-based stochastic channel model
HBF	hybrid beamforming
HPBW	half power beamwidth
i.i.d.	independently and identically distributed
IC	integrated circuit
ITU	International Telecommunication Union
LoS	line of sight
LTE	long term evolution
LVDS	low voltage differential signaling
MAP	maximum a posteriori
MAC	media access control layer
MCS	modulation coding schemes
MIMO	multiple input multiple output
MISO	multiple input single output
MiWaveS	Beyond 2020 Heterogeneous Wireless Network with Millimeter Wave Small Cell Access and Backhauling
ML	maximum likelihood
mmW	millimeter wave
MMSE	minimum mean square error

MPC	multi path component
MUSIC	multiple signal classification
MVUE	maximum variance unbiased estimator
NCP-SC	null cyclic prefix single carrier
NGMN	Next Generation Mobile Network Alliance
NI	National Instruments
NLOS	non line of sight
NLoS	non line of sight
NR	new radio
OFDM	orthogonal frequency-division multiplexing
OLF	overlap factor
OMP	orthogonal matching pursuit
PAPR	peak-to-average power ratio
PDF	probability density function
PHY	physical layer
pn	pseudo noise
PXI	PCI extensions for instrumentation
QoS	quality of service
QAM	quadrature amplitude modulation
Quadriga	QUAsi Deterministic RadIo channel GenerAtor
RF	radio frequency
RF	radio frequency
SC	single carrier
SC	single carrier
SC-FDE	single carrier with frequency domain equalization
SINR	signal to noise ratio
SISO	single input single output

SIMO	single input multiple output
SM	spatial multiplexing
SMA	SubMiniature version A
SNR	signal to noise ratio
SV	Saleh-Valenzuela
SVD	singular value decomposition
TDM	time division multiplex
TDD	time division duplex
TTI	transmission time interval
TUD	Technische Universität Dresden
UD	user device
UE	user equipment
UL	uplink
ULA	uniform linear array
VI	virtual instrument
WLAN	wireless local area network
WPAN	wireless personal area network

Mathematical Notation

Operators and Functions

$\mathrm{Tr}\left(\mathbf{A}\right)$	trace of matrix $\mathbf{A}$
$\mathrm{rank}\left(\mathbf{A}\right)$	rank of matrix $\mathbf{A}$
$\mathbf{A}^{\mathrm{T}}$	transpose of matrix $\mathbf{A}$
$\mathbf{A}^{H}$	complex conjugate of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of matrix $\mathbf{A}$
$\mathbb{E}\left\{\cdot\right\}$	expectation value
$\|\cdot\|_2$	Euclidian norm
$\|\cdot\|_{\mathrm{F}}$	Frobenious norm
$P\left(A\right)$	probability of event A
$\Phi\left(\cdot\right)$	cumulative probability density
$p\left(\cdot\right)$	probability density
$\mathrm{Q}\left(\cdot\right)$	Q function
$\mathrm{Q}^{-1}\left(\cdot\right)$	inverse Q function
$\lfloor\cdot\rfloor$	floor function
$a \mod b$	modulo operation, equals $a - \left\lfloor \frac{a}{d} \right\rfloor d$

Mathematical Symbols

P_{tx}	transmit power
λ	wavelength
ϕ_{A}	angle of arrival
ϕ_{D}	angle of departure
ρ	(power) pathloss
τ_m	excess delay
A_{antenna}	aperture area
a_m	complex amplitude of multi path component (MPC)
c	speed of light
D	propagation distance
d_{con}	size of the continuous array
d_{s}	antenna element spacing
f	frequency
G_{rx}	receive antenna (power) gain
G_{tx}	transmit antenna (power)gain
L	no. of rays per cluster
M	no. of clusters
$N_{\mathrm{rf,tx}}$	number of RF chains for the transmitter
$N_{\mathrm{rf}}^{\mathrm{tx}}$	transmitter RF chains
N_{rx}	number of receiver antenna elements
N_{tx}	number of antenna elements of the transmitter
N_{tx}	number of transmitter antenna elements
N_f	transmitter codebook size
N_w	receiver codebook size
P_{rx}	receive power
P_{tx}	transmit power
$s(\mathbf{x}, t)$	electromagnetic wave
$W(\phi)$	array factor
y	received signal
z	received signal after correlation
$\mathbf{a}_{\mathrm{rx}}$	receive array propagation vector
$\mathbf{a}_{\mathrm{tx}}$	transmit array propagation vector
$\mathbf{f}_{n_f}$	transmit precoding vector
$\mathbf{k}(\phi)$	wave number vector
$\mathbf{r}$	received signal
$\mathbf{r}_{\mathrm{tx}}$	position of the transmitter
$\mathbf{s}$	transmitted signal
$\mathbf{w}_{n_f}$	receive precoding vector
$\mathbf{x}$	position vector
$\mathbf{A}$	array propagation matrix
$\mathbf{D}$	propagation path gain matrix

$\mathbf{F}_{\mathrm{bb}}$	baseband precoding matrix
$\mathbf{F}_{\mathrm{rf}}$	RF precoding matrix
$\mathbf{H}$	channel matrix (coupling coefficients between antenna elements)
$\mathbf{W}_{\mathrm{bb}}$	baseband combining matrix
$\mathbf{W}_{\mathrm{rf}}$	RF combining matrix
$\tilde{\mathbf{A}}_{\mathrm{rx}}$	codebook of array propagation vectors of the virtual channel, receiver side
$\tilde{\mathbf{A}}_{\mathrm{tx}}$	codebook of array propagation vectors of the virtual channel, transmitter side
$\tilde{\mathbf{F}}_{\mathrm{rf}}$	analog beam-codebook of the transmitter
$\tilde{\mathbf{W}}_{\mathrm{rf}}$	analog beam-codebook of the receiver
$\tilde{\mathbf{W}}_{\mathrm{rf}}$	analog beam-codebook of the receiver

List of Figures

List of Tables

Chapter 1

Introduction

When Jagadish Chandra Bose conducted one of the first communications experiments using electromagnetic waves at 60 GHz in 1895 [SSS98], he might have been surprised knowing that the considerable commercial use of electromagnetic waves of the same frequency did not occur until one century later. In fact, he already developed a working high-frequency communications system including a spark transmitter, a coherer, a dielectric lens, and horn antennas. However, radio communications since then instead focused on lower frequencies in the range between 500 MHz and 6 GHz [RHDM14, p.5]. The introduction of the 2nd generation of cellular communications (2G) has influenced human interaction by providing reliable wireless speech transmission and reception over long distances using small devices. Subsequently, the 3rd (3G) and 4th (4G) generation of cellular wireless communications concentrated on providing reliable mobile data transmission. The advancement of communications and increasing acceptance of this technology led to a dramatic increment of devices and users. It also caused the coexistence of several wireless technologies, among them cellular networks, wireless local area network (WLAN), Bluetooth and digital enhanced cordless telecommunications (DECT), which have to share the spectrum with broadcasting services like digital audio broadcasting (DAB) and terrestrial digital video broadcasting (DVB-T), to name just a few of the wireless standards.

The increasing demand for data throughput due to improved hand-held devices was traditionally met by improving spectral efficiency using advanced signal processing schemes to achieve high order modulation transmission, low out-of-band radiation, and multiple input multiple output (MIMO) techniques. Using the formula of the achievable rate

$$R = B \cdot \sum_{i=1}^{N_{\mathrm{st}}} \log_2 \left(1 + \mathrm{SNR}_i\right), \tag{1.1}$$

we see that the rate is linearly dependent on the bandwidth B and on the number of spatially separated, parallel transmitted streams N_{st}. The signal to noise ratio (SNR), in contrast, has only a logarithmic influence on the rate. While the bandwidth is limited due to the increased number of coexisting communications standards and devices, the number of parallel transmitted streams can be exploited using MIMO technologies. However, this number is upper bounded by the rank of the individual channel matrix. Also the achieved SNR is limited by the regulations on the transmit power and by the mobile power sources in hand-held devices if using omni-directional antennas. Thus, nowadays it seems that there is little scope for improvement using these methods, therefore exploiting less congested spectrum may be necessary to fulfill future throughput demands.

Due to the availability of large portions of unused frequencies, the use of the millimeter wave (mmW) band, defined as the spectrum between 6 GHz and 300 GHz [HGPR+16] is regarded as one of the solutions to the shortage of bandwidth. Especially the band between 28 GHz - 73 GHz is considered to be a suitable candidate for cellular communications. For this reason, the 5th generation (5G) of cellular communications [3GP17] reconsiders mmW for cellular communications in the current standardization process, while the standards IEEE 802.15.3c [IEEb] and IEEE 802.11ad [IEEa] already considered these frequencies for the use in wireless personal area network (WPAN) and WLAN systems prior to this development. However, the known laws of propagation of electromagnetic waves have different manifestations at high frequencies, e.g., diffraction plays a subordinate role. Most importantly, the free-space path loss is much higher in the range of 60 GHz compared to sub-6 GHz frequencies. Thus, high directive antennas have to be used, which will be demonstrated by simple link-budget calculation here.

We compare the link budget and receive SNRs for an long term evolution (LTE) signal at 2 GHz and a mmW signal at 60 GHz center frequency at a transmission distance of $D = 50\,\mathrm{m}$ (pico cell) assuming no to use a directive antenna. For simplicity, we just consider the free space path loss here and leave the investigation of influences like atmospheric effects to the later parts of the thesis. Given a maximum allowed transmit power $P_{\mathrm{tx}} = 23\,\mathrm{dBm}$ for an LTE pico cell [v1417b], which we also use for the mmW link here for fair comparison and a bandwidth of $B_{\mathrm{LTE}} = 20\,\mathrm{MHz}$ for the LTE link and $B_{\mathrm{mmW}} = 1\,\mathrm{GHz}$ for the mmW link, the received SNRs are

$$\mathrm{SNR}_{\mathrm{LTE}}[\mathrm{dB}] = P_{\mathrm{tx}}[\mathrm{dBm}] - \rho_{\mathrm{LTE}}[\mathrm{dB}] - P_{\mathrm{n,LTE}}[\mathrm{dBm}] + G_{\mathrm{a}}[\mathrm{dBi}] \qquad = 71.6\,\mathrm{dB}, \tag{1.2}$$

$$\mathrm{SNR}_{\mathrm{mmW}}[\mathrm{dB}] = P_{\mathrm{tx}}[\mathrm{dBm}] - \rho_{\mathrm{mmW}}[\mathrm{dB}] - P_{\mathrm{n,mmW}}[\mathrm{dBm}] + G_{\mathrm{a}}[\mathrm{dBi}] \qquad = 5\,\mathrm{dB}, \tag{1.3}$$

where the path-losses are $\rho_{\mathrm{LTE}} = 52.4\,\mathrm{dB}$ and $\rho_{\mathrm{mmW}} = 102\,\mathrm{dB}$, respectively for the LTE and mmW transmission according to the free space path-loss formula [Fri46] and where the combined gain of the transmit and receive antenna $G_{\mathrm{a}}[\mathrm{dBi}] = G_{\mathrm{tx}}[\mathrm{dBi}] + G_{\mathrm{rx}}[\mathrm{dBi}]$ is zero for the considered case of omni-directional transmission. The thermal noise power is given with $P_{\mathrm{n}} = 10\log_{10}\left(B\,10^{\left(\frac{-174}{10}\right)}\right)$ dBm as a function of the bandwidth B. Thus, the noise power of the LTE and mmW transmission is $P_{\mathrm{n,LTE}} = -101\,\mathrm{dBm}$ and $P_{\mathrm{n,mmW}} = -84\,\mathrm{dBm}$, respectively. It is self-evident that the large difference in the link budget plays a major role for achieving high throughput. For practical system, high order modulation is used to achieve a good spectral efficiency. Many wireless standards use quadrature amplitude modulation (QAM) for transmission. In order to transmit uncoded 16 QAM at a bit error rate (BER) $\mathrm{BER} = 10^{-5}$, the receiver needs to achieve at least a SNR of 19.4 dB [Kam08, section 11.4.6, p.388]. While this is not an issue for the LTE link shown in (1.2), 15 dB path loss has to be compensated for the mmW link in order to reach the targeted SNR in our example.

This can be achieved by using directional antennas at the transmitter and the receiver. The gain of the antenna in the lossless case can be given according to [Bal05, p. 51] as

$$G_{\mathrm{tx/rx}} = \frac{4\pi}{\Omega_{\mathrm{A_{tx/rx}}}}, \tag{1.4}$$

where $\Omega_{\mathrm{A_{tx/rx}}}$ is the beam-solid angle. For uniform linear array (ULA) arrays located along the x-axis, the radiation intensity is equally distributed over the elevation domain, while the radiation is concentrated only at specific azimuthal ranges. The needed antenna gain can be achieved by the directional antenna of the base-station solely or be shared by the directional antennas of the base-station and the user device. We follow the latter approach because it might be easier to realize the lower gains of both antennas in practical applications compared to the necessity of realizing one highly directional antenna in practice. Assuming idealized beam patterns for the ULA with radiation intensity solely nonzero for an azimuthal range of $\phi \in \{0, \phi_{\mathrm{beamwidth}}\}$, we can formulate the radiation intensity $F_{\mathrm{n}}(\theta, \phi)$ as

$$F_{\mathrm{n}}(\theta, \phi) = \begin{cases} 1, & \text{if } 0 \leq \phi \leq \phi_{\mathrm{beamwidth}} \\ 0, & \text{otherwise.} \end{cases} \tag{1.5}$$

Using this definition, we calculate the beamwidth from the beam-solid angle as [Bal05, p.38]

$$\Omega_{\mathrm{A_{tx/rx}}} = \int\limits_{0}^{\phi_{\mathrm{beamwidth}}} \int\limits_{0}^{\pi} F_{\mathrm{n}}(\theta, \phi) \sin(\theta) \, \mathrm{d}\theta \, \mathrm{d}\phi, \tag{1.6}$$

and thus get a simple relation for the beamwidth in dependency on the gain as

$$\phi_{\mathrm{beamwidth}} = \frac{\Omega_{\mathrm{A,tx/rx}}}{2}. \tag{1.7}$$

Assuming equal gains at transmitter and receiver $G_{\mathrm{tx}} = G_{\mathrm{rx}}$, we can transform (1.3) to

$$G_{\mathrm{tx/rx}}[\mathrm{dB}] = \frac{1}{2} \left(\mathrm{SNR}_{\mathrm{mmW,target}}[\mathrm{dB}] + \rho_{\mathrm{mmW}}[\mathrm{dB}] + P_{\mathrm{n,mmW}}[\mathrm{dBm}] - P_{\mathrm{tx}}[\mathrm{dBm}] \right). \tag{1.8}$$

Using (1.4) in (1.8), we can find

$$\phi_{\mathrm{beamwidth_{tx,rx}}} = 2\pi \sqrt{\frac{P_{\mathrm{tx}}}{\mathrm{SNR}_{\mathrm{mmW}} \cdot \rho_{\mathrm{mmW}} \cdot P_{\mathrm{n,mmW}}}}. \tag{1.9}$$

Using a target SNR of 19.4 dB, we can see that the ideal beamwidths with no side lobes and enabling 16 QAM mmW transmission needs to be

$$\phi_{\mathrm{beamwidth_{tx}}} = \phi_{\mathrm{beamwidth_{rx}}} = 68.6^\circ. \tag{1.10}$$

Note that this is a quite optimistic calculation. Assuming a more realistic scenario with $P_{\mathrm{tx}} = 10\,\mathrm{dBm}$ [RMG11] and an additional common noise figure (NF) of 10 dB [RHDM14, section 5.11.1, p.324], we get $\phi_{\mathrm{beamwidth_{tx}}} = \phi_{\mathrm{beamwidth_{rx}}} = 4.9^\circ$.

As directional transmission becomes essential at mmW, new signal processing topics have to be considered. First, if the direction of transmission is limited to a small angular region, it has to be ensured that the transmitter and the receiver find a beam to point to each other in order to establish a connection, a problem commonly called *beam alignment*. Second, the support of mobility requires the ability to follow the beams according to the trace of the movement. The related problem is commonly called *beam tracking*.

1.1 Contribution of this Work

Although considerable work has been carried out to solve this problem for several system designs theoretically, practical evaluations are scarce. In this thesis, we investigate a simple, but realistic system architecture. We develop robust beam alignment and beam tracking algorithms from a theoretical standpoint that are suitable for the codebook-based analog beamforming (ABF) architecture. Subsequently, we use a demonstrator with the implemented algorithms to evaluate their performance and suitability for the targeted application experimentally. The rest of this thesis is organized as follows:

- ▷ In Chapter 2 we give an overview of the special characteristics of the mmW channel and evaluate possible channel models. Based on the necessity of using directional antennas, we provide an introduction to the concept of beamforming as a spatial filter. Furthermore we discuss ABF, hybrid beamforming (HBF), digital beamforming (DBF) systems and existing beam alignment and beam tracking algorithms.

- ▷ In Chapter 3, we develop an optimal definition for the training signal length for each beam pair for the beam alignment problem. First, we show that the beam alignment problem can be interpreted as a hypothesis test. This allows us to define the training length optimally under the assumption of given channel state information (CSI). Subsequently, we derive two heuristic algorithms to define the training length without prior CSI knowledge. The work had been published in [KRF19].

- ▷ In Chapter 4, we introduce the design of the mmW testbed, which we will use in subsequent chapters to demonstrate our algorithms experimentally. We show the design steps and the philosophy behind the real-time demonstrator, which consists of two nodes with ABF capabilities and is one of the first testbeds to test cellular joined mmW physical layer (PHY) and media access control layer (MAC) layer algorithms in real-time and closed-loop. The testbed has been designed and built as part of the EU-funded project MiWaveS. The author contributed with the implementation of the beam alignment and beam tracking algorithms in the demonstrator. Parts of the testbed description have been published in [MiW17].

- ▷ In Chapter 5, we derive two beam alignment algorithms for ABF systems. We theoretically prove the superiority over exhaustive search. Afterward, we describe the implementation of the algorithms into the testbed and prove its applicability experimentally. These results have been published in [KCF16, KCF17, KRCF18].

- ▷ In Chapter 6, we develop a beam tracking algorithm for the practical testbed and prove by simulations and experimentally that it is working using a mounted robot in a room. These results have been published in [KCF18].
- ▷ In Chapter 7 we summarize our work and show directions of future works.

Chapter 2

Fundamentals and State of the Art

In order to motivate the need for beamforming techniques, we introduce the particularities of the millimeter wave (mmW) band in this chapter. We furthermore show appropriate channel models, which bear account to their special features and can be used for theoretical investigations. Additionally, we show how the concept of beamforming and related techniques can be used to mitigate challenging channel properties. Finally, we give an overview of prior work related to this topic.

2.1 The Millimeter Wave Band

MmW wireless channels are distinguished from wireless channels using traditional frequency bands in various aspects, making it necessary to rethink signal processing techniques. The most important physical effects that play a major role in the different interaction of the environment with high frequency electromagnetic waves from 60 GHz to 80 GHz compared to frequencies below 6 Gz will be discussed in the following.

Path Loss. The free space path loss equation presented by Friis [Fri46] describes the propagation of electromagnetic waves with no obstructions for line of sight (LoS) communication depending on the wavelength λ for a lossless antenna

$$\frac{P_{\mathrm{rx}}}{P_{\mathrm{tx}}} = G_{\mathrm{tx}} G_{\mathrm{rx}} \left(\frac{\lambda}{4\pi D} \right)^2, \tag{2.1}$$

where P_{tx} and P_{rx} are the transmit and receive power, respectively, G_{tx} and G_{rx} are the linear power gains of the transmit and receive antennas compared to an isotropic antenna and D is the propagation distance. Using the well known equation $c = \lambda f$ relating the speed of light c with the wavelength λ and frequency f, it is obvious that the free path loss is inversely proportional to the frequency squared

$$\frac{P_{\mathrm{rx}}}{P_{\mathrm{tx}}} \propto \frac{1}{f^2}, \tag{2.2}$$

assuming that the gain of the antennas remain constant (e.g. isotropic antennas $G_{\mathrm{tx}} = G_{\mathrm{rx}} = 1$). However, this reflects only a part of the truth, since the gain of a perfectly matched antenna of the effective area A_{antenna} increases with increasing frequency, as shown in [RHDM14, section 3.2, p.102], [Fri46],

$$G_{\max} = A_{\text{antenna}} \frac{4\pi}{\lambda^2}. \tag{2.3}$$

Thus it becomes clear that the path loss is related proportional to the square of the frequency for constant aperture areas $A_{\text{antenna,tx}}$ and $A_{\text{antenna,rx}}$ of the transmitter and receiver, respectively,

$$\frac{P_{\text{rx}}}{P_{\text{tx}}} = \frac{f^2 A_{\text{antenna,tx}} A_{\text{antenna,rx}}}{c^2 D^2}. \tag{2.4}$$

Using (2.2), it can be concluded that the free space path loss of the mmW band is some order of magnitude larger than in traditional frequency bands when omni-directional antennas are considered. Hence directional antennas have to be used to compensate the path loss. When considering the same physical aperture area, the path loss deficiency can be even over-compensated in the mmW band as the antennas of the same physical size have a naturally larger gain at higher frequencies than at lower frequencies [HGPR+16], as shown in (2.4). Nevertheless, it has to be considered that the antennas still suffer from losses, for example insertion loss and impedance mismatch. Experimental works suggest that the free space path loss does not hold true exactly in reality [RSM+13]. To give an enhanced description for the path loss, several authors proposed a generalized log-distance path loss model for omni-directional antennas to provide better a fit to path loss models [HGPR+16, TNMR14, GTC+14a]. Unless otherwise stated, throughout this thesis the omni-directional path-loss model derived in [TNMR14],

$$\rho[\text{dB}] = 20 \log_{10} \frac{4\pi}{\lambda} + 10 n_{\text{pl}} \log_{10} \frac{d}{d_0}, \tag{2.5}$$

will be used, where $\rho[\text{dB}] = 10 \log_{10} \frac{P_{\text{rx}}}{P_{\text{tx}}}$ is the path loss and the path loss exponents are given for LoS scenarios $n_{\text{pl}} = 2.1$ and $n_{\text{pl}} = 3.3$ for non line of sight (NLoS) scenarios and $d_0 = 1\text{m}$. Omni-directional antennas are assumed for this model, the antenna gain of using an antenna array is inherently considered using amultiple input multiple output (MIMO) channel model like that one introduced in the next section.

Atmospheric and Weather Effects. Measurements show that the molecular absorption of electromagnetic waves by the atmosphere generally increases with frequency [RMG11]. Additionally, there are several peaks of attenuation caused by increased absorption due to resonance frequencies of molecules of the atmosphere at some frequencies in the mmW band, for example an increase up to 15.5 dB between $57 - 64$ GHz mainly due to the resonance frequency of the oxygen molecule [RHDM14, section 3.2.2, p.107]. Additionally, the weather has a significant effect on the attenuation of electromagnetic waves at the mmW band as the physical dimensions of the raindrops and snowflakes are in the order of the wavelength. According to [RHDM14, section 3.2.2, p.107], the worst case attenuation is around 15 dB/km for mmW band communications. Throughout this thesis, we ignore the effects of atmospheric and weather-related attenuation for simplicity,

Diffraction. While diffraction is a phenomenon that helps traditional cellular systems to cope with NLoS scenarios, the impact of diffraction decreases with decreasing wavelength [ZMS+13], [RHDM14, section 3.2.4, p.111]. Thus, shadowing is a more severe phenomenon, and propagation of electrical waves can be modeled more precisely by optical propagation theory. The effect of the frequency dependence of diffraction can be illustrated by the equation of the Fresnel zones [RHDM14, section 3.2.7, p.121],

$$r_n = \sqrt{\frac{n\lambda D_1 D_2}{D_1 + D_2}} = \sqrt{\frac{ncD_1 D_2}{f(D_1 + D_2)}}, \tag{2.6}$$

where r_n is the radius of the n^{th} Fresnel zone, λ is the wavelength of the plane wave, and D_1 and D_2 are distances between transmitter and obstacle and obstacle and receiver, respectively. The first Fresnel zone is a circle around the LoS path representing the zones where still energy can be transmitted even if the direct LoS path is blocked and no NLoS path is available due to the path-length increase. The radius of the Fresnel zone scales with the square root of the wavelength and is hence inverse proportional with frequency. Thus, small objects which did not harm the LoS transmission at traditional frequency bands can block the first Fresnel zone at mmW frequencies completely. This can be illustrated by a simple example. Consider a LoS radio connection of 100 m and an obstacle blocking the LoS path directly in the middle, i.e. in 50 m distance from both transmitter and receiver. According to [PHAH97], 60% of the first Fresnel zone has to be unobstructed in order to receive some meaningful signal. Using (2.6), an obstacle of $0.24\,\mathrm{m}^2$ is able to obstruct the a 60 GHz connection, while an obstacle of $7\,\mathrm{m}^2$ is needed to cut off a 2 GHz radio transmission.

Reflection and Scattering. As diffraction is weaker for mmW frequencies, the role of reflection and scattering becomes dominant in the high frequency wireless channel. Measurements [ZMS+13, AWW+13] show that, despite the large roughness compared to the mmW wavelength, most outdoor and indoor materials are considerably reflective at these frequencies. Consequentially, a lot of materials such as tinted glass and foliage have a large penetration loss of more than 40 dB [RHDM14, section 3.2.5, p.112] at mmW frequencies. The relative roughness of the materials make scattering a more powerful effect.

The dominance of reflection over diffraction in combination with the large propagation loss causes the mmW channel to be sparse [HGPR+16, RMRGPH15]. The effect of the sparsity of the channel in the angular and delay domain has three main reasons.

- It is a consequence of the high path-loss that the attenuation of the energy of the rays varies largely with different traveling distances. Thus multi path components (MPCs), which experienced more than one reflections between transmitter and receiver are almost undetectable due to enhanced attenuation caused by the large travel distance compared to LoS MPCs.

- The decreased diffraction phenomenon compared to lower frequencies decreases the number of MPCs received over NLoS.

- The usage of directional antennas at the transmitter and receiver stimulates the channel only in a narrow angular range. Thus, no energy is transmitted in the direction of a large portion of possible reflectors.

The sparse nature of the channel is experimentally validated in several measurement campaigns [HGPR+16, ALS+14].

2.2 Channel Model

In order to be able to develop algorithms for wireless communications, a channel model has to be used which reflects reality in the crucial aspects of wireless transmission. Optimally, the model is as complex as necessary and as simple as possible to be able to exploit distinct characteristics of the channels algorithmically. We present common approaches of channel modeling for mmW MIMO channels based on [Gus14]. First, we highlight the difference between a propagation channel model and a radio channel model. The propagation channel model describes the effect of the environment on a transmitted electromagnetic wave of a particular frequency for a given transmitter and receiver position, without considering the impact of the antenna. In contrast, the radio channel model describes the combined effect of antennas and environment to the transmitted signal [Gus14, chapter 3, p. 13]. In our investigation, we first take a look into methods of modeling the propagation channel, before deriving a suitable radio channel model for our purpose.

Deterministic Ray Tracing Channel Model. Deterministic channel modeling approaches derive the main channel characteristics for deterministic side-specific scenarios. One approach is to model the electromagnetic propagation by using the Maxwell equations analytically. In contrast, ray-tracing is a powerful approach for channel modeling using some common simplifications of plane wave-based propagation. This approach is particularly suitable for high-frequency ranges where the electromagnetic propagation becomes quasi-optical and can be quite accurately modeled by the reflections and scattering of the rays. The main problem of this approach is the site-specificity, which requires to model each channel environment separately. However, the approach is advantageous for the interpretation of channel measurements and the confirmation of dominant phenomena of side-specific channels in the mmW band, which can be further exploited in algorithm design.

Stochastic Correlation Based Channel Model. A more abstract and generalizing way of modeling the channel is to reproduce the statistical behavior of the channel. Stochastic correlation based channel models aim to model the correlation between all transmitter and receiver antenna elements of a MIMO channel. In the general case, this approach is often too complicated because the correlation matrix is subject to a large number of parameters [Gus14]. For this reason, it is widely accepted to apply simplifications to the model. A common abstraction is to assume all entries of the channel matrix to be circularly complex independently and identically distributed (i.i.d.) Gaussian random variables [Fos98, Tel99]. However, this model is only precise for environments with large receive and transmit antenna distances and rich scattering [RS11], which is a problem for mmW outdoor communications due to the relative roughness of the surrounding materials [RHDM14, section 3.2.6, p.119 and section 3.2.7, p.120].

Stochastic Geometry Based Channel Model. An intermediate stage between side-specific deterministic and generalized stochastic correlation based channel modeling is found in stochastic geometry based channel models [Gus14]. These models use stochastically placed ensembles of point scatterers to describe the channel so that it is possible to design the parameters of the model to represent physical propagation mechanisms. Compared to deterministic models, the advantage of these models is larger flexibility combined with the ability to model essential channel properties. In this way, it is possible to model the sparse characteristics and dependency of the angular characteristic of the channel on reflectors by using stochastically modeled environmental scatterers. As a further benefit, the directional transmission and reception is an inherent part of the modeling, and thus the influence of the directional antenna can easily be considered.

Most models for the mmW band are based on the Saleh-Valenzuela (SV) model, which was originally designed to describe indoor single input single output (SISO) channels in the time-delay domain for traditional frequency bands [SV87]. It models the arrival time of the paths in clusters, where the power of the clusters decay exponentially with delay [Mol05, section 7.3.3, p.130]. This model was adapted to the angular domain by [WJ02, SJJS00]. In [SMB01] the channel is described as a sum of several MPCs, each assigned to a delay, angle of departure (AoD), angle of arrival (AoA) and a complex amplitude. Most common channel models in mmW research use an extended version of the SV model. Usually the MPCs are grouped in clusters, where the respective delay AoD, AoA and gain being subject to large-scale propagation effects like path loss and reflection [RHDM14, section 3.7.1.1, p. 152], often based on measurements or ray tracing simulations. The clusters consist of several rays, which vary around the cluster-specific parameters by different stochastically modeled distributions and are called small-scale parameters. These intra-clusters aim to model small-scale propagation effects, most notably the important scattering phenomenon. The extended SV model is often categorized as geometry based stochastic model [AH16, Pay17], while the work [ABB+07] classifies this model as a non-geometric statistical model. The reason is that in the classical extended SV model, the AoD and AoA are modeled directly by some distributions, without considering scatterers explicitly. However, this model can easily be transformed into a full geometric based statistical model as stated below.

While the channel model for the early IEEE802.15.3c mmW standard [IEEb] models only AoD, i.e. it is a single input multiple output (SIMO) channel model, the IEEE802.11ad [IEEa] standard models an indoor MIMO channel at 60 GHz. The large scale parameters are calculated mainly using ray tracing, while the small scale parameters and the amplitude are modeled by empirical distributions [Gus14]. Thus, the parameter sets for the channel model are side specific for room types like conference room, cubicle office environment and living room.

The QUAsi Deterministic RadIo channel GenerAtor (Quadriga) [JRB+17] and 3GPP model [3GP17, v1417a] are both cluster based channel models, inspired by the WINNER model [KMH+07, KMH+]. The WINNER model uses stochastically distributed scatterers to model the large-scale propagation effects of the channel and is capable of considering mobility by modeling Doppler spread and time evaluation of the spatial model. The small-scale parameters are modeled statistically, while the distributions are parametrized motivated by measurements. The Quadriga model additionally offers a quasi-deterministic change of the channel due to the time-varying environments or mobility of the user. While these models are quite accurate and also ensure general applicability of the model through the random deployment of point scatterers, the complexity of designing the model is quite high.

In the following, we use a variant of the SV model, which is extended to the AoA and AoD domain. The propagation channel in angular-delay domain (without the effect of the antenna [Gus14]) is described as a superposition of M_c clusters each consisting of M_p rays, and is a function of the AoD ϕ_D, AoA ϕ_A, excess delay τ and the positions of the transmitter $\mathbf{x}_\mathrm{tx} = [x_\mathrm{tx}, y_\mathrm{tx}, z_\mathrm{tx}]^\mathrm{T}$ and receiver $\mathbf{x}_\mathrm{rx} = [x_\mathrm{rx}, y_\mathrm{rx}, z_\mathrm{rx}]^\mathrm{T}$ [WJ02]

$$
\begin{aligned}
h_\mathrm{p}(\mathbf{x}_\mathrm{tx}, \mathbf{x}_\mathrm{rx}, \phi_\mathrm{D}, \phi_\mathrm{A}, \tau) = \frac{1}{\sqrt{M_\mathrm{c} M_\mathrm{p}}} \sum_{m_\mathrm{c}=1}^{M_\mathrm{c}} a_{m_\mathrm{c}} \sum_{m_\mathrm{p}=1}^{M_\mathrm{p}} a_{m_\mathrm{c},m_\mathrm{p}} \delta(\tau - \tau_{m_\mathrm{c}} - \tau_{m_\mathrm{c},m_\mathrm{p}}) \\
\times \delta(\phi_\mathrm{D} - \phi_{\mathrm{D},m_\mathrm{c}} - \phi_{\mathrm{D},m_\mathrm{c},m_\mathrm{p}}) \delta(\phi_\mathrm{A} - \phi_{\mathrm{A},m_\mathrm{c}} - \phi_{\mathrm{A},m_\mathrm{c},m_\mathrm{p}})
\end{aligned}
\tag{2.7}
$$

where a_{m_c}, τ_{m_c}, $\phi_{\mathrm{D},m_\mathrm{c}}$, $\phi_{\mathrm{A},m_\mathrm{c}}$ are the large scale complex amplitude (that includes the effect of the path loss), delay, AoD and AoA of the m_c^th cluster, respectively and $a_{m_\mathrm{c},m_\mathrm{p}}$, $\tau_{m_\mathrm{c},m_\mathrm{p}}$, $\phi_{\mathrm{D},m_\mathrm{c},m_\mathrm{p}}$, $\phi_{\mathrm{A},m_\mathrm{c},m_\mathrm{p}}$ are the small scale deviation of the complex amplitude, delay, AoD and AoA of the m_p^th ray of the the m_c^th cluster.

Used Channel Model. The above-described propagation channel is frequency selective. Following the discussion above, we can apply some reasonable simplifications to this channel model:

- The severeness of the effect of angular spread is decreasing with an increasing beamwidth of the antenna array. In our investigations, we use antenna arrays with relatively large beamwidth (Sec. 4.2.2) compared to angular spread measurements conducted in [SWA+13] for example. For this reason, we use the simplification of one ray per cluster (MPC) throughout this work.

- As already mentioned, frequency selectivity is not assumed to be dominant in the mmW channel due to the limited effect of diffraction, the high path loss and the use of directional antennas. The sensitivity to frequency selectivity can be reduced by the use of current modulation and equalization techniques like orthogonal frequency-division multiplexing (OFDM) or single carrier with frequency domain equalization (SC-FDE). However, due to the sparse nature of the channel and limited delay spread [RHDM14, section 3.3.1, p. 129], time domain equalization techniques may also be used. The common approach of repeated application of a beamforming algorithm to different frequency-flat subcarriers does not generate additional insight compared to a beamforming algorithm designed for a frequency-flat channel. Thus, as we focus on algorithms exploiting the degree of freedom of spatial and angular propagation in this work, we use the simplification of a common excess delay for all MPCs and assume a single ray for every cluster.

The propagation channel model in (2.7) thus simplifies to

$$h_{\mathrm{p}}(\mathbf{x}_{\mathrm{tx}}^{i}, \mathbf{x}_{\mathrm{rx}}^{j}, \phi_{\mathrm{D}}, \phi_{\mathrm{A}}) = \frac{1}{\sqrt{\rho M_{\mathrm{p}}}} \sum_{m_{\mathrm{p}}=1}^{M_{\mathrm{p}}} a_{m_{\mathrm{p}}} \delta(\phi_{\mathrm{D}} - \phi_{\mathrm{D},m_{\mathrm{p}}}) \delta(\phi_{\mathrm{A}} - \phi_{\mathrm{A},m_{\mathrm{p}}}), \tag{2.8}$$

by additionally limiting it to a 2D channel model (only azimuthal angle considered), where we split the large scale and small scale effect of the complex amplitude into the large scale path loss ρ, defined in (2.5) and the complex small scale amplitude of the MPCs, defined as $a_{m_{\mathrm{p}}} \sim \mathcal{CN}(0,1)$ [AALH13]. Thus the expected value of the power of the channel over the sum of all MPCs and several realizations is

$$\mathbb{E}\left\{ \|h(\phi_{\mathrm{D}}, \phi_{\mathrm{A}})\|_2^2 \right\} = \frac{1}{\rho}, \tag{2.9}$$

where ρ is dependent on the distance and defined in (2.5). The corresponding radio channel, which includes the effect of the antenna arrays, can be modeled depending on the relative position of the transmitter and receiver antenna array elements $\mathbf{x}_{\mathrm{tx}}^{i}$ and $\mathbf{x}_{\mathrm{rx}}^{j}$ with respect to the array phase center (see [Gus14, PH13])

$$\begin{aligned} h_{i,j}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}}) &= \int_{\phi_{\mathrm{D}}} \int_{\phi_{\mathrm{A}}} h(\mathbf{x}_{\mathrm{tx}}^{i}, \mathbf{x}_{\mathrm{rx}}^{j}, \phi_{\mathrm{D}}, \phi_{\mathrm{A}}) f_{i,\mathrm{tx}}(\phi_{\mathrm{D}}) f_{j,\mathrm{rx}}(\phi_{\mathrm{A}}) \mathrm{e}^{j\mathbf{k}\mathbf{x}_{\mathrm{tx}}^{i}} \mathrm{e}^{j\mathbf{k}\mathbf{x}_{\mathrm{rx}}^{j}} d\phi_{\mathrm{D}} d\phi_{\mathrm{A}} \\ &= \int_{\phi_{\mathrm{D}}} \int_{\phi_{\mathrm{A}}} \frac{1}{\sqrt{\rho M_{\mathrm{p}}}} \sum_{m_{\mathrm{p}}=1}^{M_{\mathrm{p}}} a_{m_{\mathrm{p}}} \delta(\phi_{\mathrm{D}} - \phi_{\mathrm{D},m_{\mathrm{p}}}) \delta(\phi_{\mathrm{A}} - \phi_{\mathrm{A},m_{\mathrm{p}}}) f_{i,\mathrm{tx}}(\phi_{\mathrm{D}}) f_{j,\mathrm{rx}}(\phi_{\mathrm{A}}) \\ &\times \mathrm{e}^{j\mathbf{k}_{\mathrm{tx}}^{m_{\mathrm{p}}} \mathbf{x}_{\mathrm{tx}}^{i}} \mathrm{e}^{j\mathbf{k}_{\mathrm{rx}}^{m_{\mathrm{p}}} \mathbf{x}_{\mathrm{rx}}^{j}} d\phi_{\mathrm{D}} d\phi_{\mathrm{A}}, \end{aligned} \tag{2.10}$$

where $f_{i,\mathrm{tx}(\phi_{\mathrm{D}})}$ and $f_{j,\mathrm{rx}(\phi_{\mathrm{D}})}$ are the radiation patterns of the i^{th} transmitter antenna element and the j^{th} receiver antenna element, respectively. The wave number vector in the x-y plane is defined as

$$\mathbf{k}(\phi) = \frac{2\pi}{\lambda} [\sin(\phi), \cos(\phi)]^{\mathrm{T}}, \tag{2.11}$$

as we assume plane waves in this channel model. Using the simplification of plane waves and the narrow-band assumption (the bandwidth is significantly smaller than the carrier frequency $f_{\mathrm{BW}} << f_{\mathrm{c}}$), we can simplify (2.10) to [ABB+07]

$$h_{i,j}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}}) = \frac{1}{\sqrt{\rho M_{\mathrm{p}}}} \sum_{m_{\mathrm{p}}=1}^{M_{\mathrm{p}}} a_{m_{\mathrm{p}}} f_{i,\mathrm{tx}}(\phi_{\mathrm{D},m_{\mathrm{p}}}) f_{j,\mathrm{rx}}(\phi_{\mathrm{A},m_{\mathrm{p}}}) \mathrm{e}^{j\mathbf{k}_{\mathrm{tx}}^{m_{\mathrm{p}}}(\phi_{\mathrm{D}}) \mathbf{x}_{\mathrm{tx}}^{i}} \mathrm{e}^{j\mathbf{k}_{\mathrm{rx}}^{m_{\mathrm{p}}}(\phi_{\mathrm{A}}) \mathbf{x}_{\mathrm{rx}}^{j}}. \tag{2.12}$$

Defining the array propagation vector for a uniform linear array (ULA) with antenna elements spacing d_{s} as

$$\begin{aligned} \mathbf{a}_{\mathrm{rx}}(\phi_{\mathrm{A}}) &= \frac{1}{\sqrt{N_{\mathrm{rx}}}} \left[\mathrm{e}^{\mathbf{k}_{\mathrm{rx}}^{m_{\mathrm{p}}}(\phi_{\mathrm{A}}) \mathbf{x}_{\mathrm{rx}}^{i}}, \quad \ldots \quad, \mathrm{e}^{\mathbf{k}_{\mathrm{rx}}^{m_{\mathrm{p}}}(\phi_{\mathrm{A}}) \mathbf{x}_{\mathrm{rx}}^{j}} \right]^{T} & (2.13) \\ &= \frac{1}{\sqrt{N_{\mathrm{rx}}}} \left[1, e^{\frac{j2\pi}{\lambda} d_{\mathrm{s}} \sin(\phi_{\mathrm{A}})}, \ldots, e^{d_{\mathrm{s}} \frac{j2\pi}{\lambda} (N_{\mathrm{rx}}-1) d_{\mathrm{s}} \sin(\phi_{\mathrm{A}})} \right]^{T}, & (2.14) \end{aligned}$$

and a similar expression for $\mathbf{a}_{\mathrm{tx}}$, the MIMO matrix $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{rx}} \times N_{\mathrm{tx}}}$ consists of the complex coupling coefficients between all N_{tx} transmitter antennas and N_{rx} receiver antenna elements and thus can be given as

$$\mathbf{H} = \begin{pmatrix} h_{1,1} & ... & h_{1,N_{\mathrm{rx}}} \\ ... & h_{i,j} & ... \\ h_{N_{\mathrm{tx}},1} & ... & h_{N_{\mathrm{tx}},N_{\mathrm{rx}}} \end{pmatrix} \tag{2.15}$$

$$= \sqrt{\frac{N_{\mathrm{rx}} N_{\mathrm{tx}}}{\rho M_{\mathrm{p}}}} \sum_{m_{\mathrm{p}}=1}^{M_{\mathrm{p}}} a_{m_{\mathrm{p}}} \mathbf{a}_{\mathrm{rx}}(\phi_{\mathrm{A},m_{\mathrm{p}}}) \mathbf{a}_{\mathrm{tx}}^{H}(\phi_{\mathrm{D},m_{\mathrm{p}}}), \tag{2.16}$$

as the AoA and AoD can be assumed to be constant over all antenna elements of the array. An equivalent but simpler representation of the channel matrix can be given by a product of three matrices

$$\mathbf{H} = \sqrt{\frac{N_{\mathrm{tx}} N_{\mathrm{rx}}}{\rho M_{\mathrm{p}}}} \mathbf{A}_{\mathrm{rx}} \mathbf{D} \mathbf{A}_{\mathrm{tx}}^{H}, \tag{2.17}$$

where $\mathbf{A}_{\mathrm{tx}} = [\mathbf{a}_{\mathrm{tx}}(\phi_{\mathrm{D},1}), ..., \mathbf{a}_{\mathrm{tx}}(\phi_{\mathrm{D},M_{\mathrm{p}}})] \in \mathbb{C}^{N_{\mathrm{tx}} \times M_{\mathrm{p}}}$ and $\mathbf{A}_{\mathrm{rx}} = [\mathbf{a}_{\mathrm{rx}}(\phi_{\mathrm{A},1}), ..., \mathbf{a}_{\mathrm{rx}}(\phi_{\mathrm{A},M_{\mathrm{p}}})] \in \mathbb{C}^{N_{\mathrm{rx}} \times M_{\mathrm{p}}}$ are array propagation matrices containing all M_{p} array propagation vectors, and $\mathbf{D} = \mathrm{diag}(a_1, a_2, ..., a_{M_p})^T \in \mathbb{C}^{M_{\mathrm{p}} \times M_{\mathrm{p}}}$ is the propagation path gain matrix containing all complex amplitudes of the MPCs. In the later chapters of the thesis, we use the distance of adjacent antenna elements to be $d_{\mathrm{s}} = \frac{\lambda}{2}$ in order to avoid grating lobes. The angles of departure ϕ_D and arrival ϕ_A are assumed to be uniformly distributed in the range of $\phi_D, \phi_A \in (-\frac{\pi}{4}, \frac{\pi}{4})$.

Time Varying Channel. The model of the channel must include time variations of the AoAs, AoDs and path gains to model realistic cellular scenarios. Additionally, temporal and spatial fading has to be taken into account. The most significant effects to consider are the Doppler shift, resulting from the relative speed of the transmitter and receiver and the change of the channel due to mobility, the change of AoDs and AoAs of existing MPCs, the disappearance of existing MPCs due to blocking and the advent of new MPCs due to reflections. In traditional SISO radio channel models for omni-directional antennas, the change of the AoDs and AoAs has hardly any influence on the model, as the common simplification of isotropic rich scattering can be applied. To model the change of the AoDs and AoAs in a realistic way in situations where this assumption does not apply, some geometry-based stochastic channel models (GSCMs) such as [KMH+, v1417a] propose the fairly complex process of changing the arrangement of the scatterers. To make the modeling as simple as possible, but still describe reality with sufficient accuracy, we follow the approach of [HKG+14], which we explain below. We split the temporal variation of the channel in several subsequent channel versions, called channel blocks, which we assume to be constant for a certain channel block time. We define the channel at time instant (or channel block time) n as

$$\mathbf{H}_n = \sqrt{\frac{N_{\mathrm{tx}} N_{\mathrm{rx}}}{\rho M_{\mathrm{p}}}} \mathbf{A}_{\mathrm{rx},n} \mathbf{D}_n \mathbf{A}_{\mathrm{tx},n}^{H} \tag{2.18}$$

where $\mathbf{D} = \text{diag}(a_{n,1}, a_{n,2}, ..., a_{n,M_p})^T \in \mathbb{C}^{M_p \times M_p}$ is the propagation path gain matrix at channel block n. The temporary correlation of the channel is modeled based on [HKG+14]

$$\mathbf{H}_{n+1} = \sqrt{\frac{N_{\text{tx}} N_{\text{rx}}}{\rho M_{\text{p}}}} \mathbf{A}_{\text{rx},n+1} \mathbf{D}_{n+1} \mathbf{A}_{\text{tx},n+1}^H \tag{2.19}$$

with

$$\mathbf{D}_{n+1} = c \cdot \mathbf{D}_n + \sqrt{1 - c^2} \mathbf{B}_{n+1} \tag{2.20}$$

$$\mathbf{A}_{\text{tx},n+1} = \mathbf{A}_{\text{tx},n}(\boldsymbol{\phi}_{\text{tx}} + \Delta\boldsymbol{\phi}_{\text{tx}}) \tag{2.21}$$

$$\mathbf{A}_{\text{rx},n+1} = \mathbf{A}_{\text{rx},n}(\boldsymbol{\phi}_{\text{rx}} + \Delta\boldsymbol{\phi}_{\text{rx}}), \tag{2.22}$$

where $\mathbf{B}_{n+1}$ is an independent diagonal matrix with entries from $\mathcal{CN}(0, 1)$. The parameter c is defined as the time correlation between the gain of the same path at channel block n and channel block $(n + 1)$

$$c = \mathbb{E}\left\{a_{m_{\text{p}},n} a_{m_{\text{p}},n+1}^*\right\}. \tag{2.23}$$

The time correlation coefficient between two 2D channel coefficients of subsequent timing can be shown to follow the Jakes model under the assumption that the angle difference between the AoAs and the direction of the velocity is uniformly distributed between 0 and π [Mol05, section 5.6, p.84]. This is the case for each channel realization (i.e., for one defined velocity direction of the user device (UD)), if we assume to have many scatterers around the receiver and omni-directional transmit and receive antennas. Thus the AoA can be expected uniformly distributed in this case. As we discussed above, these are invalid assumptions for the sparse mmW channel. However, we have two reasons to use this simple relation also under these channel conditions for each MPC individually. First, we can take this relation as a worst-case approximation to compute the coherence time. Second, the uniform distribution of the difference of the AoA and direction of velocity can be justified over *several* realizations of the channel, i.e., an uniformly distributed velocity direction of the UD for each realization, which we assume at this point. Another reason for choosing this option is the unavailability of a statistical description of the correlation in the case of a deterministic velocity direction of the user and a sparse channel. Following this argumentation and [HKG+14], the coefficient follows the Bessel function of the first kind of order zero

$$c = J_0(2\pi f_{\text{D}} \Delta T), \tag{2.24}$$

where f_D is the maximum Doppler frequency and ΔT is the channel block length. The Doppler frequency given by $f_D = \frac{f_c v}{c}$, where f_c is the carrier frequency, v the speed of the UD and c the speed of light. The angle variations $\Delta\boldsymbol{\phi}_{\text{tx}}$ and $\Delta\boldsymbol{\phi}_{\text{rx}}$ are uniformly distributed in the range of $\mathcal{U}(-3°, 3°)$.

Ray Tracing Channel Model. The drawback of the described time-varying channel model is the relatively simple parametrization, which represents reality only to a certain extent. It is likely that the channel behaves friendlier, since the velocity direction of the

Table 2.1: Ray tracing parameters [LLH94, PYP08]

Parameter	**Value**
relative permittivity ϵ_{r} of granite	5.5+j 0.2
relative permittivity ϵ_{r} of glass	6.24+j 0.17
relative permittivity ϵ_{r} of plaster	3.08+j 0.055
roughness standard deviation σ_{h}	$0.6 \cdot 10^{-3}$m

receiver is doesn't usually change randomly for every channel block. To have a simulative reference for beam tracking measurements, we developed a ray tracking simulation. Therefore, we model the room which was used for experiments described in the later chapters. In our simulation, we focus on a one-bounce model, which means that only first order reflections are taken into account. That is a reasonable assumption given the high path-loss, the reflection loss and the tendency of the surface to be electrical rough and causing scattering at the high frequencies of the mmW band. Beneficial for the modeling is the fact that we can ignore statistical representations of the shadowing effect, solely using optical laws for the calculations. We use the circumstance that there is precisely one possible transmission path between transmitter and receiver using the reflection of one specific plane. Given the position of the transmitter and the receiver, thus the AoD and AoA can be determined analytically.

At the interface between the air and the material, the gain and the phase of the electromagnetic wave is changed. For perfectly smooth materials, this effect can be described by the Fresnel reflection coefficients [ZVM10,Smu94,LLH94], which depend on the relative permeability μ' and the relative complex dielectric constant $\epsilon_{\mathrm{r}} = \epsilon' + j\epsilon''$ of the materials. For simplicity, we can approximate the values of the air to be $\mu'_{\mathrm{air}} \approx 1$ and $\epsilon_{\mathrm{r,air}} \approx 1$. Consequently, the reflection coefficient of the electrical field for orthogonal and parallel polarization (with respect to the wave of incidence) is given by [Smu94]

$$R_{\mathrm{o}} = \rho_{\mathrm{s}} \frac{\cos\alpha - \sqrt{\mu_{\mathrm{pr}}\epsilon_{\mathrm{r}} - \sin^2\alpha}}{\cos\alpha + \sqrt{\mu_{\mathrm{pr}}\epsilon_{\mathrm{r}} - \sin^2\alpha}} \tag{2.25}$$

$$R_{\mathrm{p}} = \rho_{\mathrm{s}} \frac{\mu_{\mathrm{pr}}\epsilon_{\mathrm{r}}\cos\alpha - \sqrt{\mu_{\mathrm{pr}}\epsilon_{\mathrm{r}} - \sin^2\alpha}}{\mu_{\mathrm{pr}}\epsilon_{\mathrm{r}}\cos\alpha + \sqrt{\mu_{\mathrm{pr}}\epsilon_{\mathrm{r}} - \sin^2\alpha}}, \tag{2.26}$$

where μ' and ϵ_{r} are the relative permeability and complex relative permittivity of the wall material and ρ_{s} is the scattering loss factor. While we can assume $\mu' \approx 1$ for all dielectric materials, the values of the permittivity are material-specific and frequency dependent. In our simulation, we use ϵ_{r} values taken from measurements for 60 GHz for specific materials given in literature [LLH94], which are summarized in Tab. 2.1. For conductive materials like iron, the permittivity is given by $\epsilon_{\mathrm{r}} \approx 1$, while the permeability is usually large in the order of $\mu' \approx 10^5$ [Smu94], which are frequency independent values and result in $R_{\mathrm{o}} = -1$ and $R_{\mathrm{p}} = 1$. Due to the relative roughness of typical surfaces compared to the wavelength, a multiplicative scattering loss factor [Smu94, ZVM10]

$$\rho_{\mathrm{s}} = \exp\left\{-0.5\left(\frac{4\pi\sigma_{\mathrm{h}}\cos\alpha}{\lambda}\right)^2\right\} \tag{2.27}$$

has to be considered, where σ_{h} is the standard deviation of the roughness in meters. We assume a standard deviation of the roughness of 0.6 mm for all materials motivated by measurements in [LLH94] for granite at 62 GHz and remarks in [Smu94]. Additionally, we take into account the phase shift due the total traveling distance of the MPC d_{mpc}

$$p_{\mathrm{l}} = \exp\{-j\frac{2\pi}{\lambda}d_{\mathrm{mpc}}\} \tag{2.28}$$

and the Doppler shift due to the relative movement direction of the UD with respect to the incident ray at the receiver depending on the sample distance Δt

$$p_{\mathrm{D}} = \exp\{j\frac{2\pi}{\lambda}v_{\mathrm{rel}}\Delta t\}. \tag{2.29}$$

Thus, we calculate the AoD, AoA and complex amplitude of each ray and obtain a propagation channel model. We transform these parameters to a radio channel model similarly as done above, which we can use for our investigations by exchanging the complex amplitude of each ray (all reflected rays of every plane and the LoS ray) in (2.15) with

$$\alpha_{\mathrm{rel}} = 1/\sqrt{\rho} \cdot R \cdot p_{\mathrm{l}} \cdot p_{\mathrm{D}}. \tag{2.30}$$

Note that we ignore the different time of arrival in the radio channel model for the same reasons as mentioned above, as we investigate only analog beamforming techniques in the following chapters. We also include the effect of the path loss to the complex amplitude to take into account the different distance the MPC travel at our simulation.

2.3 Spatial Filtering – The Concept of Beamforming

Instead of explaining antenna array processing by the phase differences of antenna array elements induced by a plane wave from a certain direction, we introduce the concept of array processing as the process of filtering in the spatial domain or wavenumber domain, respectively. This representation has the benefit that the concept of the well-known Fourier transform can be used to relate space and wavenumber domain. Consequently, we can interpret the wavenumber as the spatial frequency. Thus it is possible to find a beneficial design of the array, acting as a spatial filter, which exploits the spatial characteristics of the channel. In this section, we mainly follow the argumentation of [JD92, p. 41ff].

An electromagnetic plane wave is a function of time and space and can be represented as

$$s(\mathbf{x}, t) = \exp\{j(\omega t - \mathbf{k}_0\mathbf{x})\}, \tag{2.31}$$

where $\mathbf{x} = [x, y, z]^{\mathrm{T}}$ is the position vector, $\omega_0 = 2\pi f_0$ is the angular frequency and $\mathbf{k} = [k_x, k_y, k_z]^T$ is the wavenumber vector, which becomes $\mathbf{k} = \frac{2\pi}{\lambda}[\sin(\phi_0)\cos(\theta_0), \sin(\phi_0)\sin(\theta_0), \sin(\theta_0)]^{\mathrm{T}}$ considering the assumed case of a plane wave. The azimuth angle is ϕ_0 and the elevation angle is θ_0. The electromagnetic wave can be also described in dependency of a spatial frequency variable using the Fourier transform, where the wavenumber vector $\mathbf{k}_0$ is the spatial frequency variable. Note

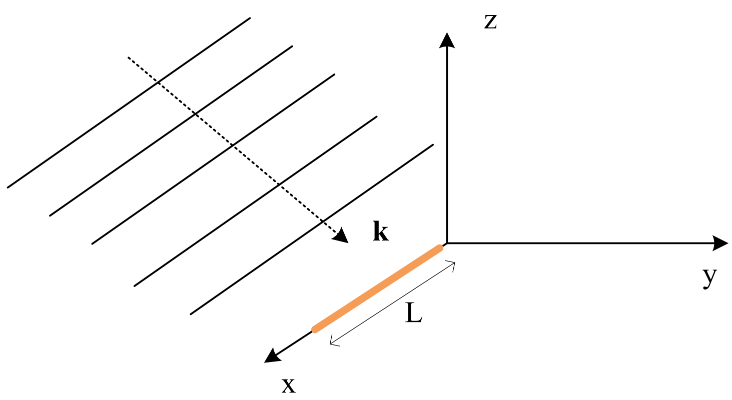


Figure 2.1: Plane wave of direction $\mathbf{k}$ hits a linear aperture located at the x-axis.

that the terms spatial frequency and wavenumber are often used synonymously in the literature. The relation between the description of the wave $s(\mathbf{x}, t)$ in spatial-temporal domain and wavenumber-frequency domain is given by the multidimensional Fourier transform over the space and time variables

$$S(\mathbf{k}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\mathbf{x}, t) \exp\{-j\omega t + j\mathbf{k}\mathbf{x}\} d\mathbf{x} dt \tag{2.32}$$

$$s(\mathbf{x}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\mathbf{k}, \omega) \exp\{j\omega t - j\mathbf{k}\mathbf{x}\} d\mathbf{k} d\omega. \tag{2.33}$$

We first assume that the wave is observed by a linear continuous aperture positioned at the x-axis and ranging from $x_{\min} = -\frac{d_{\text{con}}}{2}$ to $x_{\max} = \frac{d_{\text{con}}}{2}$ (Fig. 2.1). The output of the aperture can be given as

$$z(\mathbf{x}, t) = w(\mathbf{x}) s(\mathbf{x}, t), \tag{2.34}$$

where the aperture function $w(\mathbf{x})$ is defined as

$$w(x) = \begin{cases} 1, & \|x\|_2 \leq \frac{d_{\text{con}}}{2} \\ 0, & \text{otherwise.} \end{cases} \tag{2.35}$$

Converting the wave to the space-time domain $S(\mathbf{k}, \omega)$ and the aperture to the wavenumber domain $W(\mathbf{k})$, the receive signal can be represented as the convolution of the wave equation and aperture function in wavenumber-frequency domain

$$Z(\mathbf{k}, \omega) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} W(\mathbf{k} - \mathbf{l}) S(\mathbf{l}, \omega) d\mathbf{l}. \tag{2.36}$$

Every wave can be reproduced by a sum of plane waves [JD92, p. 40]. A plane wave can be represented in the wavenumber-frequency domain as

$$S(\mathbf{k}, \omega) = S(\omega) \delta(\mathbf{k} - \mathbf{k}_0), \tag{2.37}$$

leading to a output signal of

$$z(\mathbf{k},t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\mathbf{k}-\mathbf{l}) S(\mathbf{l},\omega) d\mathbf{l} \exp\{j\omega t\} d\omega \tag{2.38}$$

$$z(\mathbf{k},t) = s(t) W(\mathbf{k}-\mathbf{k}_0). \tag{2.39}$$

It becomes clear that the output of the antenna depends only on the aperture function $W(\mathbf{k})$ and the AoA of the wave in $\mathbf{k}_0$ for a given frequency.

Beamwidth. Solving the Fourier transform for the exemplary linear aperture in Fig. 2.1 gives

$$W(k_x) = \int_{-\frac{d_{\text{con}}}{2}}^{\frac{d_{\text{con}}}{2}} \exp\{jk_x x\} dx \tag{2.40}$$

$$W(k_x) = \frac{1}{jk_x} \left(\exp\{j\frac{d_{\text{con}} k_x}{2}\} - \exp\{j\frac{-d_{\text{con}} k_x}{2}\} \right) \tag{2.41}$$

$$W(k_x) = \frac{2}{k_x} \sin\left(\frac{d_{\text{con}} k_x}{2} \right) \tag{2.42}$$

$$W(k_x) = \frac{\sin(k_x {}^{d_{\text{con}}}/_2)}{{}^{k_x}/_2}, \tag{2.43}$$

which reveals that the aperture function is similar to a sinc function in wavenumber domain. The dependence of the aperture function on the direction of the plane wave can be clearly seen by exchanging the wavenumber component of the x-axis by its definition $k_x = \sin\phi$. This reduces the problem to two dimensions, inherently assuming $\theta = \pi$ for the elevation angle without loss of generality

$$W(\phi) = \frac{\sin(\sin(\phi) d_{\text{con}} \frac{\pi}{\lambda})}{\sin(\phi)\frac{\pi}{\lambda}}. \tag{2.44}$$

Thus it becomes obvious that the width of the mainlobe depends only on the size of the array d_{con}.

Sampling and Aliasing – Discrete Arrays. The relation of the aperture function of a discrete array can be derived from the aperture function for the continuous array by applying the spatial version of the discrete time Fourier transform (DTFT). Defining the distance of the antenna elements as before with d_s (i.e the spatial sampling period), the aperture function of an infinite large linear array with discretized locations of the antenna elements at the positions $x_n = nd_\text{s}$ is given by

$$W(\check{k}_x) = \sum_{n=-\infty}^{\infty} w_n \exp\{j\check{k}_x n\}. \tag{2.45}$$

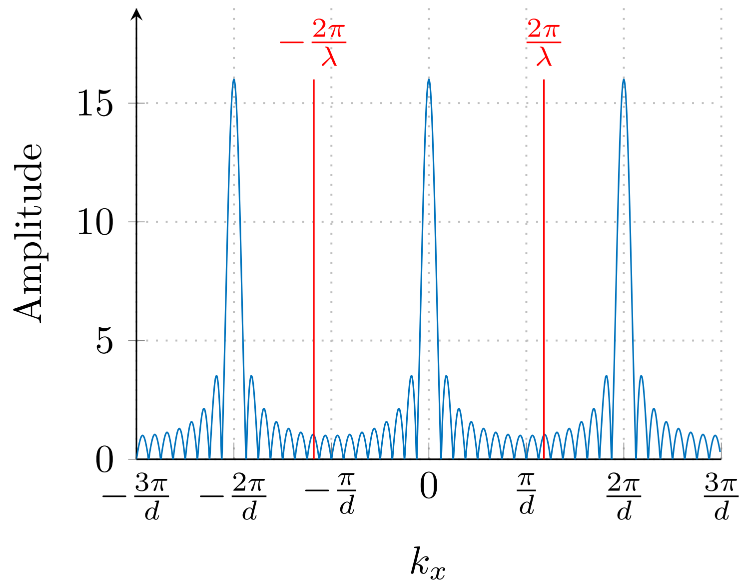


Figure 2.2: The effect of the antenna spacing to aliasing ($d = 0.59\lambda$).

Consequently, the aperture function for a linear array with finite number of N sensors can be given as

$$W(\check{k}_x) = \sum_{n=1}^{N} w_n \exp\{j\check{k}_x n\}. \tag{2.46}$$

Similar to the definition of the time-frequency DTFT, we normalize the digital spatial frequency by the sampling distance $\check{k} = kd_\text{s}$, which ranges from $-\pi$ to π. One result of the sampling is the existence of images in the spatial frequency domain with periodicity of $\check{k}_\text{p} = 2\pi$. Thus the output signal $z(t)$ is now defined by a circular convolution between the space sampled wave function $S_\text{s}(k_x, \omega)$ in wavenumber frequency domain and the aperture function

$$z(\mathbf{k}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\frac{\pi}{d_\text{s}}}^{\frac{\pi}{d_\text{s}}} W(k_x - l_x) S_\text{s}(l_x, \omega) dl_x \exp\{j\omega t\} d\omega. \tag{2.47}$$

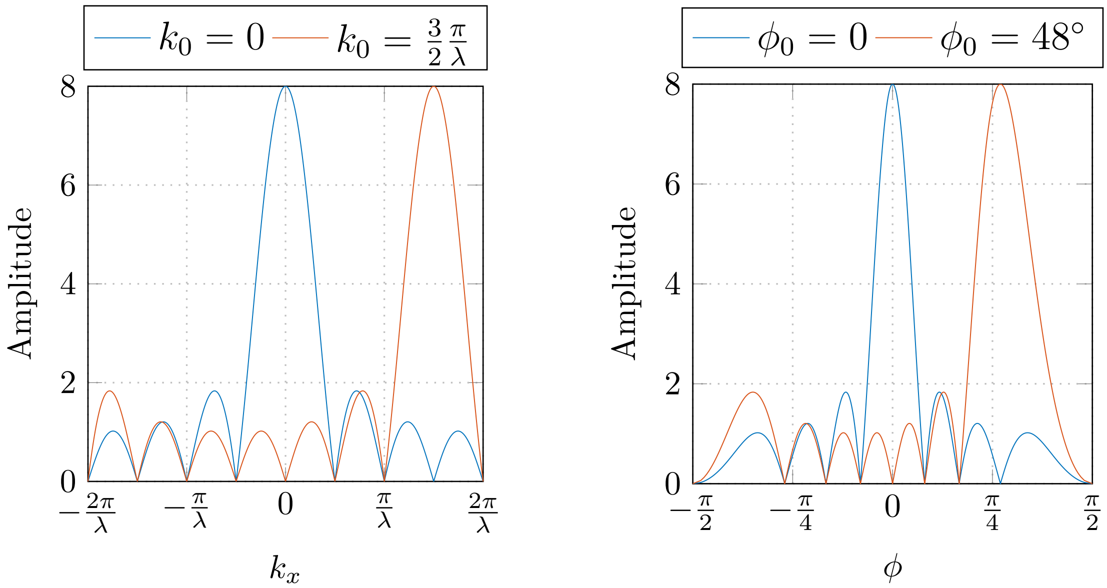
As can be seen in (2.47), the mainlobe of the aperture is steered towards the direction of the incident wave with

$$W(k - k_0) = \sum_{n=1}^{N} w_n \exp\{j(k_x - k_0) n d_\text{s}\} \tag{2.48}$$

which can be formulated as the Dirichlet kernel [JD92, p. 121] if $w_n = 1,\ \forall n$

$$W(k - k_0) = \frac{\sin\left(\frac{N d_\text{s}}{2}(k_x - k_0)\right)}{\sin\left(\frac{d_\text{s}}{2}(k_x - k_0)\right)} \tag{2.49}$$

omitting a common phase factor. The aperture function for the continuous linear array and thus every image of the spatially sampled ULA has infinite spatial bandwidth as can



(a) Beam-pattern in k domain. The beamwidth remains constant in this domain for all different wavenumber values k_0.

(b) Beam-pattern in ϕ domain. The beamwidth is varying in this domain for different wavenumber values ϕ_0.

Figure 2.3: Performance against deflection coefficient (differential SNR).

be seen in (2.43). However, most of the spatial bandwidth consists of side lobes. As we are mainly interested in the main lobe, the aperture function can be considered band-limited. If a main lobe of the neighboring image interferes with the original image at the visible range, the signal is received from two directions, which results in ambiguity. The images cause aliasing if the spatial sampling interval d_s is too large for the spatial bandwidth $(k_\mathrm{max} - k_\mathrm{min})$ of the wave. As the wavenumber vector is defined as $k_{x,0} = \frac{2\pi}{\lambda}\sin(\phi)$, the physical obtainable range of spectral frequency is $-\frac{2\pi}{\lambda_0} \leq k_{x,0} \leq \frac{2\pi}{\lambda_0}$. The periodicity of the images of the aperture function $W(k - k_0)$ is $k_\mathrm{p} = \frac{2\pi}{d_\mathrm{s}}$ as seen in (2.47). To avoid the interference of images, one image should cover the whole range of the physical obtainable range, therefore $\frac{4\pi}{\lambda_0} = \frac{2\pi}{d_\mathrm{s}}$, which results in $d_\mathrm{s} = \frac{\lambda_0}{2}$.

It is worth mentioning that the beamwidth is constant for every steering wavenumber k_0 in the spatial frequency domain but varies with the corresponding angle $\phi_0 = \sin^{-1}\left(\frac{k_{x,0}\lambda}{2\pi}\right)$, as can be observed in Figs. 2.3a and 2.3b. The aperture function for the discrete array is conventionally called array factor or *array pattern*.

Beamforming using a DFT codebook. The analogy between the space-wavenumber relation using antenna arrays and the well-known time-frequency relation means that we can apply common signal processing concepts of the time-frequency space to beamforming. Above, we derived that the antenna spacing has to fulfill $d_\mathrm{s} \leq \frac{\lambda}{2}$ to avoid images in the wavenumber domain in the visible range. Interpreting the visible range as our bandwidth, the spatial sampling rule $d_\mathrm{s} \leq \frac{\lambda}{2}$ can be interpreted as spatial Nyquist sampling. To avoid confusion, we have to take notice of the fact that the wavenumber $\mathbf{k}$ corresponds to the angular frequency in the time-frequency relation.

We explain this relation by using an example. Consider a uniformly distributed linear array at the x-axis. Thus we only have to consider one dimension of $\mathbf{k}$. We define the maximum (linear) spatial frequency of the incoming wave as $\tilde{\nu}_\mathrm{max} = \frac{k_\mathrm{max}}{2\pi} = \frac{1}{\lambda_0}$. According to the Shannon-Nyquist theorem, the sampling frequency of a signal should be twice as large as the maximal frequency of the signal $\tilde{\nu}_\mathrm{s} \geq 2\tilde{\nu}_\mathrm{max}$ and thus the antenna spacing should be $d_\mathrm{s} \leq \frac{2}{\lambda_\mathrm{max}}$. In case of a monochromatic wave, $\lambda_\mathrm{max} = \lambda_0$ and the antenna spacing equals the result we derived above.

Using Nyquist sampling, we know that the sampled signal can be perfectly reconstructed in the range of the signal [PM96, p. 14]. Therefore, it is clear theoretically that we can restore the incoming signal using the information of all the antennas of an antenna array fulfilling the Nyquist sampling theorem. The common problem is to estimate the direction of the incoming wave, which means we just need one parameter of the signal, not the full reconstruction.

As stated earlier, the spatial frequency is related to the AoA. In case of a ULA on the x-axis it applies $k_x = k_{x,0} \sin\phi$. A straight forward way to estimate this parameter would be to translate the information obtained by the antenna array to the spatial frequency domain, which can be easily done by multiplying a discrete Fourier transform (DFT) matrix to the vector of antenna signals. As the DFT matrix has full rank (and thus an inverse operation can be applied), no information is lost in this process.

We explain this using again a ULA with N antennas. As before, $s(\mathbf{x}, t)$ is the signal of the incoming wave in dependency of the space and time. Using the Nyquist spaced antennas of the ULA at $\mathbf{x} = [0, d_\mathrm{s}, ..., (N-1)d_\mathrm{s}]$, and fixing the time $t = t_0$, the antenna output signals can be stacked into the vector $\mathbf{s}$. Applying the N-point DFT, the signal becomes

$$\mathbf{s}_\mathrm{DFT} = \mathbf{W}_{\mathrm{DFT},N} \cdot \mathbf{s} \tag{2.50}$$

with the N-point DFT matrix using $W_N = \exp\left(\frac{-j2\pi}{N}\right)$

$$\mathbf{W}_{\mathrm{DFT},N} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ 1 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix}. \tag{2.51}$$

The benefit of this description is that the spatial frequency components of the signal can be directly obtained.

However, two (connected) phenomena complicate the full exploitation of this domain. Treating the incoming signal as periodic, the DFT outputs a discrete spectrum with equidistant N bins with index m and spatial frequency $\frac{k_{x,m}}{2\pi} = \tilde{\nu}_m = m\frac{1}{d_s N} = m\frac{2}{\lambda N}$ in the obtainable spatial frequency domain (the visible range). That prevents us from obtaining the exact spatial frequency (and thus angle of arrival) of an incoming monochromatic wave.

Second, the spatial sampling with a limited ULA acts as the application of a rectangular window with the antenna signals, resulting in a convolution of a Dirichlet kernel with the spatial frequency components. This effect is commonly called *Leakage Effect* and is well investigated in the time-frequency relation. As we stated before, all information necessary to reconstruct the signal perfectly is available after applying the DFT. The remaining question is how to extract it from the result.

From the theory of the DFT, we can obtain the shape of the Dirichlet kernel (using a time-shifted DFT definition of a summation over $-\frac{N}{2}, \cdots, \frac{N}{2}$) which causes the leakage as [Lyo04]

$$X(m) = \frac{\sin(\pi m)}{\sin(\frac{\pi m}{N})}, \tag{2.52}$$

which can be approximated using a sinc function

$$X(m) \approx N\frac{\sin(\pi m)}{\pi m} = N\mathrm{sinc}(m). \tag{2.53}$$

This Dirichlet kernel is convolved with the (continuous) spatial frequency signal of the incoming wave and sampled at the frequency bins m. In case we consider again a monochromatic incoming wave, whose spatial spectrum can be described as a shifted Dirac, this convolution gives a shifted Dirichlet kernel.

$$X(m, m_{\mathrm{shift}}) = \frac{\sin(\pi(m - m_{\mathrm{shift}}))}{\sin(\frac{\pi(m-m_{\mathrm{shift}})}{N})}, \tag{2.54}$$

with the index $m_{\mathrm{shift}} = \frac{k_x \lambda_0 N}{\pi} = 2N\sin(\phi)$ corresponding to the normalized spatial frequency of the signal, where the wavenumber of the signal can be obtained by $k_x = \frac{m_{\mathrm{shift}}\pi}{\lambda_0 N}$. Eq. (2.52) can thus be reformulated as

$$X(k_{x,m}, k_{x,0}) = \frac{\sin(\frac{\lambda_0 N}{4}(k_{x,m} - k_{x,0}))}{\sin(\frac{\lambda_0}{4}(k_{x,m} - k_{x,0}))}. \tag{2.55}$$

In order to obtain the correct wavenumber $k_{x,0}$ we can minimize the error between the result of the DFT $\mathbf{s}_{\mathrm{DFT}}$ and the sampled version of the Dirichlet kernel. As the first is complex while the latter is real, it is sufficient to minimize the error of the absolute value

$$\min_{k_{x,0}} \left(|\mathbf{s}_{\mathrm{DFT}}| - |X(k_{x,m}, k_{x,0})|\right)^2. \tag{2.56}$$

Due to the side-lobes this might be a non-convex problem but can be evaluated exhaustively with a given precision. An example of this procedure in the time-frequency domain is given in Fig. 2.4.

Therefore, we showed that it is possible to estimate the spatial frequency (and thus angle of arrival) using the antenna output signals which are processed by a DFT. As the DFT matrix consists of phase rotations only, it can be implemented using phase shifters in practice. The Dirichlet kernels correspond to an shifted versions of the aperture function (2.49), with angles corresponding to equidistant spatial frequencies. However, the application of a DFT matrix to the antenna output signals would correspond to N phase-shifts *per antenna element* simultaneously. Applying the phase shifts successively leads to a loss of the mutual phase relation of elements of $\mathbf{s}_{\mathrm{DFT}}$ in practice (due to synchronization). Nevertheless, as the described method is only dependent on the absolute value of $\mathbf{s}_{\mathrm{DFT}}$, it can be still applied in the case of successive training.

However, this method has still some drawbacks. First, we explained the estimation of the angle of arrival with only one incoming wave. Assuming multiple incoming waves, the expression becomes more complicated, especially, as the waves may have a mutual phase offset and the number of incoming waves has to be known for the optimization. Second, the phases and thus the beam-shapes have to be perfect for applying this method; otherwise, the result might not fit the Dirichlet kernel function or Sinc function. Third, it is evident that the smaller values of $\mathbf{s}_{\mathrm{DFT}}$ might be distorted a lot due to the noise in real-world applications. Finally, the angle of incidence of a wave might be estimated, but this information gives no benefit in case of a DFT codebook with limited steering angles.

Therefore, it is also justified to use oversampled codebooks with more entries than N entries. Although the same information can be obtained using a DFT codebook, it might be more robust to the noise and imperfections of the beam-shape and allows the exploitation of the angle of incidence by having more options of choosing a beam pointing closely in the direction. Note that there are more sophisticated algorithms, which have a superior performance of estimating the angle of arrival in a noisy environment by using the antenna signals directly (e.g., MUSIC, ESPRIT [JSDM91]). However, these algorithms need to transform each antenna element signal to the baseband and digitize it (digital beamforming).

Formula and Relation to the Sum and Delay Beamformer. In the following section, we show that the derived array factor is the array pattern of the well-known sum-and-delay beamformer for a monochromatic plane wave stimulus. Following the discussion above, we can also use (2.39) to describe the output signal of a discrete array, which is linearly located on the x-axis for the input of a plane wave. The output signal of the array is given by

$$z(t) = W(k_x - k_0)\exp\{j\omega t\}. \tag{2.57}$$

Equivalently, using (2.48), the output signal can be also given by

$$z(t) = \sum_{n=1}^{N} w_n \exp\{j(k_x - k_0)nd_{\mathrm{s}}\}\exp\{j\omega t\} \tag{2.58}$$

$$z(t) = \sum_{n=1}^{N} w_n \exp\{j\omega \underbrace{\frac{k_x}{\omega}nd_{\mathrm{s}}}_{-\Delta_n} - jk_0 nd_{\mathrm{s}}\}\exp\{j\omega t\}. \tag{2.59}$$

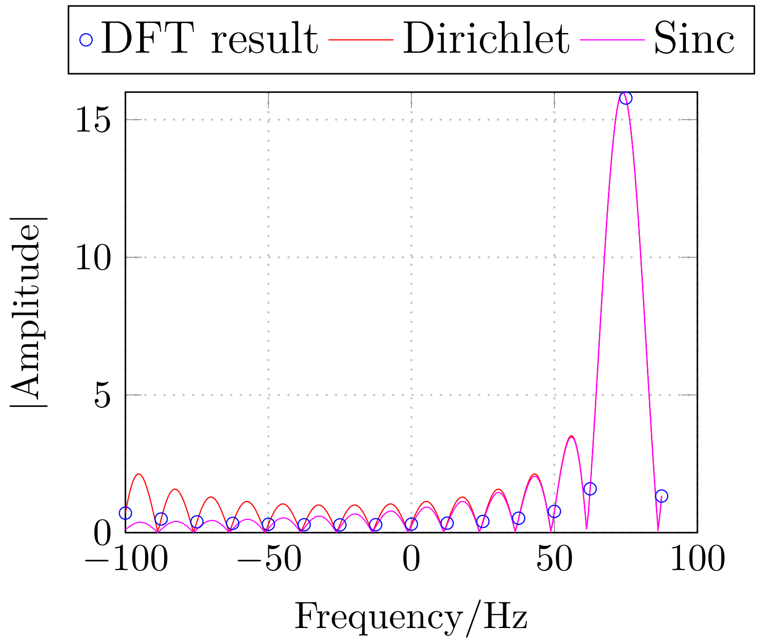


Figure 2.4: Example of interpolating a result of a DFT using a Dirichlet kernel in the time-frequency domain. Example use $N = 16$, $f_\mathrm{s} = 200$Hz. The unknown frequency of a complex sine wave with frequency $f = 73.8608$ Hz can be obtained by minimizing the error between the absolute of the DFT of the signal and the absolute Dirichlet kernel.

By interpreting

$$\frac{k_x}{\omega} n d_\mathrm{s} = \frac{\psi}{c} n d_\mathrm{s} = -\Delta_n \tag{2.60}$$

as a time delay and with using the definition of the wavenumber vector x coordinate $k_x = \frac{\psi\omega}{c}$, where c is the speed of light and ψ is the traveling direction of the plane wave, the output signal can be summarized as

$$z(t) = \sum_{n=1}^{N} w_n s(t - \Delta_n - \frac{k_0}{\omega} n d_\mathrm{s}), \tag{2.61}$$

which is the definition of the sum-and-delay beamformer, where the input signals are delayed for every antenna element.

Spatio-temporal Filter. Until now, we considered the process of array processing as a filtering in the wavenumber-frequency domain, which is a multiplication of the array factor and the wave function in the spatio-temporal domain and a convolution of the array factor and the wave function in the wavenumber domain. Another interpretation is to consider the beamforming process as a filter in the spatio-temporal domain (which might be obvious taking into account the structure of the sum and delay beamformer), resulting in a convolution in the spatio-temporal domain and a multiplication in the wavenumber-frequency domain. The corresponding formula for the continuous aperture would be then

$$z(\mathbf{k},t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\mathbf{k},\omega) F(\mathbf{k},\omega) d\mathbf{l} \exp\{j\omega t\} d\omega. \tag{2.62}$$

By identifying $H(\mathbf{k},\omega) = W(\mathbf{k}-\mathbf{l})$ we get

$$z(\mathbf{k},t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\mathbf{k}-\mathbf{l}) F(\mathbf{l},\omega) d\mathbf{l} \exp\{j\omega t\} d\omega \tag{2.63}$$

which is the same relation as (2.38). Thus these two views of the problem are equivalent.

The array factor of the linear antenna can be decomposed in vectors which will help us to separate the influence of the direction of the incoming plane wave and the steering of the mainlobe of the array. If we assume again that the array is located on the x-axis, we just have to consider the x-component of the wavenumber vector. The array factor is given in (2.48). If we assume the weights to be $w_n = 1$ for all n, we can write

$$W(k_x, k_0) = \sum_{n=1}^{N} \exp\{j(k_x - k_0) n d_\mathrm{s}\} \tag{2.64}$$

$$W(\phi, \phi_0) = \sum_{n=1}^{N} \exp\{j\frac{2\pi}{\lambda_0}(\sin\phi - \sin\phi_0) n d_\mathrm{s}\}. \tag{2.65}$$

The sum can also be presented using matrix algebra, which we do here for the receive array. If we use the array propagation vector definition in (2.13) for the case of the ULA located at the x-axis with N antennas, which describes the effect of the incoming wave direction on the array

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{N}} \left[\mathrm{e}^{j\frac{2\pi}{\lambda_0}\sin\phi}, \quad \dots \quad, \mathrm{e}^{jN\frac{2\pi}{\lambda_0}\sin\phi} \right]^T, \tag{2.66}$$

and additionally describe the effect of the steering of the mainlobe of the array by the steering vector

$$\mathbf{w}(\phi_0) = \frac{1}{\sqrt{N}} \left[\mathrm{e}^{j\frac{2\pi}{\lambda_0}\sin\phi_0}, \quad \dots \quad, \mathrm{e}^{jN\frac{2\pi}{\lambda_0}\sin\phi_0} \right]^T, \tag{2.67}$$

the array factor is given by

$$W(\phi, \phi_0) = \mathbf{w}^H \mathbf{a}. \tag{2.68}$$

The array factor for a transmit array can be decomposed in a similar way.

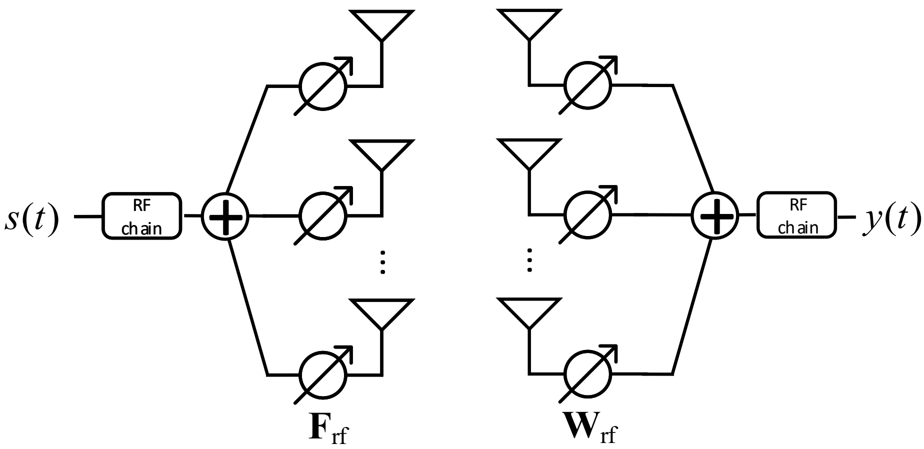
2.3.1 Beamforming Architectures

Mostly, discrete arrays are used in the context of wireless communications because the electrical steering of such arrays is relatively easy to realize. Especially at mmW frequencies, the primary motivation for using directional antennas is to increase the effective area of the antenna, resulting in an increase of gain. Additionally, the directionality also makes it possible to communicate in different directions at the same time and thus to use the spatial resources in parallel. In the next sections, we discuss several types of system designs for directional communications. The main difference in design lies in the number of radio frequency (RF) chains, associated with the antenna array that determines the possibility of parallel communications from a signal processing perspective.

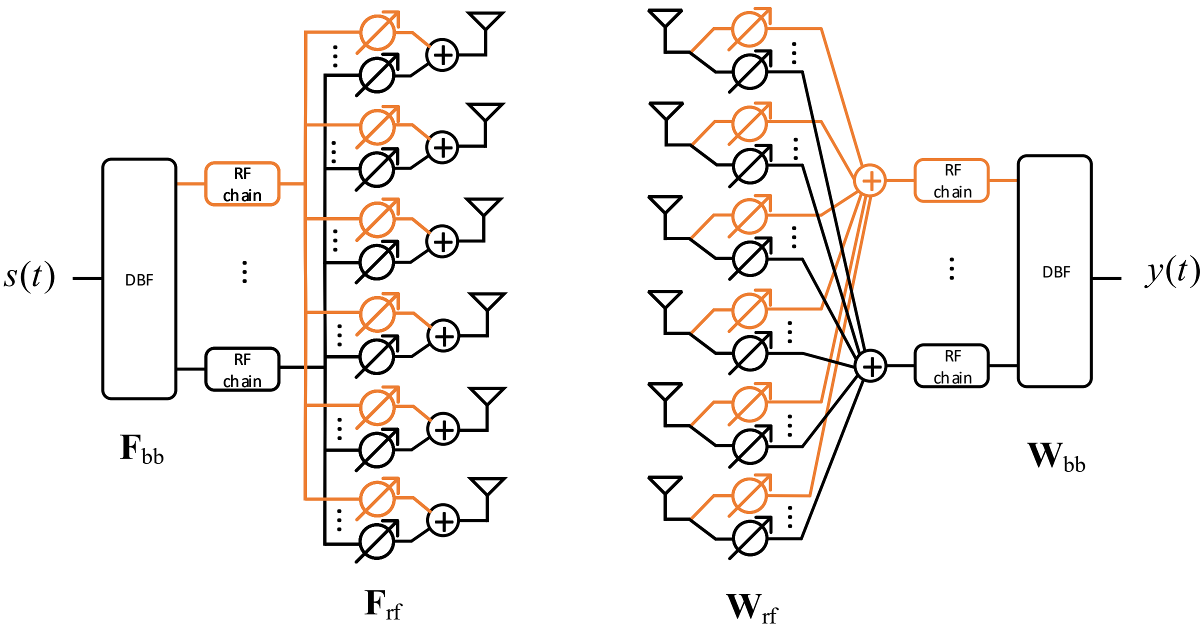
Analog Beamforming. The simplest system design for directional communications is analog beamforming (Fig. 2.5a) with one RF chain associated to the antenna and where the signal of each antenna element is summed up after the phase shift is applied. That corresponds to the simplified array factor at (2.65). As follows from the description of the array factor above, the direction of the main lobe can be adapted to the direction of an incoming wave optimally by applying a phase shift at each antenna element. Phase shifters can be produced relatively power efficient for traditional frequencies up to 6 GHz. Additionally, the analog beamforming system design has the advantage that not every signal has to be converted down and sampled by the analog to digital converter (ADC) in order to perform the steering, but can be processed in passband directly and just one signal has to be converted down to baseband after the individual antenna signals summed up. However, this scheme only allows the simultaneous transmission of one signal introduced by one electromagnetic wave. Also, it is not possible to use more advanced beam patterns, i.e., for canceling an interfering wave, which would require a non-constant amplitude factor w_n change in the array factor [VT02, p.517]. As high precision phase shifters for mmW frequencies are costly and power-inefficient [YKZ+16], several adaptations were made to further simplify the scheme. Several works suggest to reduce the resolution of the phase shifters down to 1-bit [MRRGP+16] to make the implementation more effective. Another common approach is to allow only certain directions of the main lobe, which can be realized by fixed-wired networks as the Butler matrix. The set of all valid array steering vectors for steering the individual antenna element phases are then called a codebook.

Hybrid Beamforming. A hybrid beamforming system (Fig. 2.5b) in contrast has more than one RF chain, but less RF chains than the number of antenna elements, i.e. $N_{\mathrm{rf,tx}} < N_{\mathrm{tx}}$ and $N_{\mathrm{rf,rx}} < N_{\mathrm{rx}}$. Two sub-categories are existing: either the system consists of several sub-arrays with analog beamforming, where every antenna element is connected to exactly one RF chain after the phase shifter, or the system is designed in such a way that every antenna element is connected to *all* RF chains [HIXR15]. While the latter version of the hybrid beamforming has the disadvantage of higher complexity, it offers the usage of all antenna elements for each beam, which significantly increases the gain and decreases the beamwidth of the resulting beam.

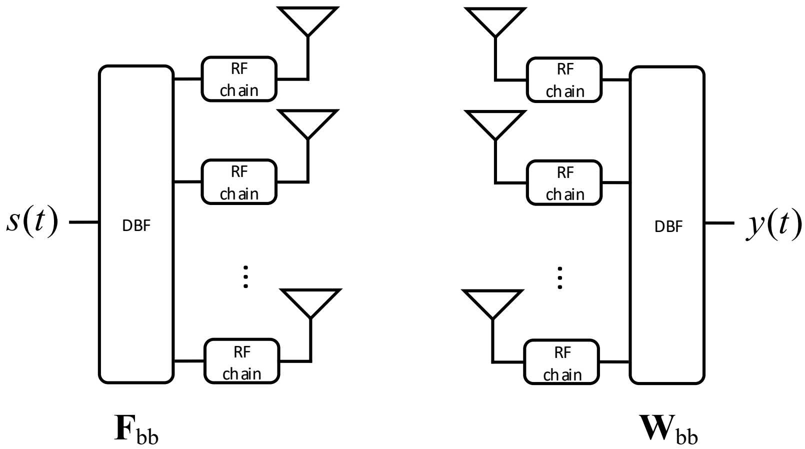
The main benefit of hybrid beamforming (HBF) is that this design offers more freedom for achieving beneficial beam patterns with non-constant weighting factors w_n in (2.61) and also generally allows the use of parallel spatial transmission with either different beams or using traditional MIMO techniques such as spatial multiplexing (SM). Considering the sparsity of MPC in the mmW channel, it is likely that the channel matrix $\mathbf{H}$ has only a few dominant eigenvalues.



(a) Analog beamforming.



(b) Hybrid beamforming.



(c) Digital beamforming.

Figure 2.5: Different beamforming systems.

Therefore only a limited number of parallel data streams can be transmitted, which also limits the need for RF chains. In fact, [SY15] showed that the hybrid beamforming transmitter and receiver achieve the same rate as the (optimal) digital beamforming transmitter and receiver constructed by using the common method of singular value decomposition (SVD) of the channel matrix, if the number of RF chains for the transmitter and the receiver is at least double the number of the data streams N_{st}. The number of data streams is usually determined by the number of separable MPC M_{p}, i.e. $2 \cdot M_{\mathrm{p}} < N_{\mathrm{rf,tx}}$ and $2 \cdot M_{\mathrm{p}} < N_{\mathrm{rf,rx}}$. Also for the case of $M_{\mathrm{p}} = N_{\mathrm{rf}}$, algorithms have been developed, which achieve near the full digital optimal performance, where the optimal preprocessing and combining, in terms of achievable rate, is defined by the left and right singular vectors of the SVD of the channel matrix.

Digital Beamforming. If every antenna element has its own RF chain, the system is called digital beamforming (DBF) (Fig. 2.5c). In such a system, which is used in most modern sub 6 GHz MIMO systems, every antenna element signal is converted down to the baseband and sampled. Thus the weighting factors w_n in (2.61) and the phase shifts can be applied directly at the digital domain, and no (analog) passband processing is necessary. The technique offers the largest flexibility, and conventional MIMO techniques can be used. However, down-converters for the mmW band and wideband ADCs are expensive, power consuming and subject to impairments. Additionally, the large number of required antenna elements at high frequencies compared to traditional cellular frequencies prohibits the use of this system at mmW bands.

2.4 System Model

In the rest of the thesis, we use a point-to-point hybrid beamforming system model with one transmitter and one receiver. Consider the transmitter and receiver to have each one ULA consisting of N_{tx} and N_{rx} antenna elements, respectively. We assume for the moment the transmitter to have $N_{\mathrm{rf}}^{\mathrm{tx}}$ and the receiver to have $N_{\mathrm{rf}}^{\mathrm{rx}}$ RF chains. N_{st} parallel data streams are transmitted. The precoding at the transmitter as well as the combining at the receiver consists of a digital part and an analog part. The received signal $\mathbf{y} \in \mathbb{C}^{N_{\mathrm{st}} \times 1}$ can be represented at sample n_{s}

$$\mathbf{y}[n_{\mathrm{s}}] = \mathbf{W}_{\mathrm{bb}}^{H}\mathbf{W}_{\mathrm{rf}}^{H}\mathbf{H}\mathbf{F}_{\mathrm{rf}}\mathbf{F}_{\mathrm{bb}}\mathbf{s}[n_{\mathrm{s}}] + \mathbf{W}_{\mathrm{bb}}^{H}\mathbf{W}_{\mathrm{rf}}^{H}\mathbf{n}[n_{\mathrm{s}}], \tag{2.69}$$

where $\mathbf{F}_{\mathrm{bb}} \in \mathbb{C}^{N_{\mathrm{rf}}^{\mathrm{tx}} \times N_{\mathrm{st}}}$ is the baseband precoding matrix, $\mathbf{F}_{\mathrm{rf}} \in \mathbb{C}^{N_{\mathrm{tx}} \times N_{\mathrm{rf}}^{\mathrm{tx}}}$ is the RF precoding matrix summarizing the influence of the phase shifters, $\mathbf{W}_{\mathrm{bb}} \in \mathbb{C}^{N_{\mathrm{rf}}^{\mathrm{rx}} \times N_{\mathrm{st}}}$ is the baseband combining matrix, $\mathbf{W}_{\mathrm{rf}} \in \mathbb{C}^{N_{\mathrm{rx}} \times N_{\mathrm{rf}}^{\mathrm{rx}}}$ is the RF combining matrix and $\mathbf{s} \in \mathbb{C}^{N_{\mathrm{st}} \times 1}$ is the transmitted signal vector with $\|\mathbf{s}\|_2 = P_{\mathrm{tx}}$, where P_{tx} is the transmit power and $\mathbf{n} \in \mathbb{C}^{N_{\mathrm{rx}} \times 1}$ is the thermal noise. In the rest of the thesis, we use a narrow-band channel model introduced in (2.15). Each RF precoder and combining matrix consists of $N_{\mathrm{rf}}^{\mathrm{tx}}$ and $N_{\mathrm{rf}}^{\mathrm{rx}}$ column vectors,respectively, which are equivalent to the steering vectors presented above

$$\mathbf{F}_{\mathrm{rf}} = \left[\mathbf{f}_0, \quad ..., \quad \mathbf{f}_{N_{\mathrm{rf}}^{\mathrm{tx}}-1}\right] \tag{2.70}$$

$$\mathbf{W}_{\mathrm{rf}} = \left[\mathbf{w}_0, \quad ..., \quad \mathbf{w}_{N_{\mathrm{rf}}^{\mathrm{rx}}-1}\right]. \tag{2.71}$$

The steering vectors are taken from a codebook of size K and L, which include all possible steering vectors

$$\tilde{\mathbf{F}}_{\mathrm{rf}} = \{\mathbf{f}_k \in \mathbb{C}^{N_{\mathrm{tx}}\times 1}, k = 0, ..., K-1, \|\mathbf{f}_k\|_2^2 = 1\} \tag{2.72}$$

$$\tilde{\mathbf{W}}_{\mathrm{rf}} = \{\mathbf{w}_l \in \mathbb{C}^{N_{\mathrm{rx}}\times 1}, l = 0, ..., L-1, \|\mathbf{w}_l\|_2^2 = 1\}. \tag{2.73}$$

In the case of analog beamforming the elements of the steering vectors $\mathbf{f}_k$ and $\mathbf{w}_l$ will contain only scaled phase-shifts. Equation (2.69) can be simplified to

$$y_{k,l}[n_{\mathrm{s}}] = \mathbf{w}_l^H \mathbf{H}\mathbf{f}_k s[n_{\mathrm{s}}] + \mathbf{w}_l^H \mathbf{n}[n_{\mathrm{s}}]. \tag{2.74}$$

We assume to use training signals of good correlation properties of length N_{s}, $\{s[n_1], ..., s[n_{N_{\mathrm{s}}}]\}$

$$\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = P_{\mathrm{tx}}\mathbf{I}_{N_{\mathrm{s}}}, \tag{2.75}$$

thus the correlated observation can be described as

$$\begin{aligned} z_{k,l}[n_{\mathrm{s}}] &= \frac{s^*[n_{\mathrm{s}}]}{|s[n_{\mathrm{s}}]|^2} y_{k,l}[n_{\mathrm{s}}] \\ &= \mathbf{w}_l^H \mathbf{H}\mathbf{f}_k + \underbrace{\mathbf{w}_l^H \mathbf{n}[n_{\mathrm{s}}]}_{\nu[n_{\mathrm{s}}]}, \end{aligned} \tag{2.76}$$

where the effective noise is still Gaussian distributed with $\nu[n_{\mathrm{s}}] \sim \mathcal{CN}(0, \sigma^2)$ due to the analog beamforming constraints of $\|\mathbf{f}(i)\|_2^2 = \frac{1}{N_{\mathrm{tx}}}, \forall i$ and $\|\mathbf{w}(i)\|_2^2 = \frac{1}{N_{\mathrm{rx}}}, \forall i$.

2.5 Theoretical Prior Work

The overall goal of wireless communications is to achieve a high data rate and resilient transmission over the air. Thus, the channel conditions have to be estimated first in order to pre-distort (at the transmitter) or equalize (at the receiver) the signal. In sub 6 GHz transmission systems, which use omni-directional signaling, the delay-frequency domain was the primary domain which had been taken into consideration for estimation and equalization. By using directional antennas in mmW communications, the spatial domain also has to be taken into account. This section focuses on algorithms to estimate the spatial domain of the channel and use the estimation for precoding and combining.

Beamforming, Beam steering, Beam alignment, and Beam tracking. In this thesis, we use the term beamforming as a general term for all MIMO techniques which rely on a physical interpretation of beams for precoding and combining. The term beam steering is used as umbrella term for beamforming techniques in the context of analog beamforming (ABF), consisting of beam alignment, which describes the process of initial alignment of the beams for communications of one transmitter and receiver and beam tracking, which defines the process of beam-adaptation in a dynamic channel environment due to movement or shadowing.

2.5.1 Channel Estimation

Before the channel can be equalized, relevant parameters of the channel have to be estimated. While the whole channel matrix is commonly estimated directly for sub 6 GHz MIMO applications, this is not feasible for large antenna arrays. Furthermore, the sparsity of the channel at the mmW band limits the number of parameters which have to be estimated in order to be able to describe the channel, as, e.g., the AoD, AoA and the strength of the MPC. In general, two estimation strategies are worth mentioning, *explicit* and *implicit* channel estimation. The first method aims to estimate the above-mentioned parameters for the propagation channel directly (explicit channel state information (CSI)). Since the channel can only be experienced in the context of directional antennas, this requires some effort to exclude the effect of the antennas in order to obtain the parameters of the propagation channel. In order to use the parameters for precoding, they have to be translated back to a (possibly different) radio channel (for different antennas) with the same propagation channel. While this approach is easy to debug and interpret for optimization algorithms, as direct channel parameters are obtained, the drawback is the high complexity of these conversions. Implicit Channel estimation measures parameters of the radio channel (implicit CSI, e.g., which beam of a codebook receives the highest energy), without translating it to the parameters of the propagation channel. The drawback is that the optimization for the precoding algorithm might not be as obvious as before. In both cases there might be some ambiguity due to the limitations of the array (e.g., grating lobes).

Early works for the use of beamformers focus on the estimation of the AoA for digital beamforming, which is also relevant for modern radar techniques. Besides using a simple ”sweep” of the beam in the angular domain, a more advanced version, the Capon beamformer [Cap69] was introduced, which use a Lagrangian function to optimize the resolution of the AoA estimation. The famous multiple signal classification (MUSIC) and estimation of signal parameters via rotational invariance techniques (ESPRIT) algorithms [JSDM91] use subspaces of the covariance matrix of the received signals to estimate the exact AoA of a propagation channel with improved precision. However, these algorithms need to process the baseband signals of all the antenna elements, which make them infeasible to be used in the context of mmW antennas.

Early algorithms for mmW communications focus on approaches, which are implementable with analog beamforming systems. As only the sum signal is received and thus the effect of the array cannot be canceled by postprocessing, these algorithms are not able to estimate the propagation channels, but only the parameters of the radio channel. In contrast, HBF systems allow a translation of parameters of the radio channel to parameters of the propagation channel. In order to avoid too much overhead, many channel estimation techniques try to exploit the sparsity in the spatial frequency and delay domain, borrowing ideas of compressed sensing (CS). The original idea of CS is to solve under-determined systems of linear equations using a minimum of observations. By approximating the channel consisting of array propagation vectors with continuous AoDs and AoAs by a channel matrix with quantized AoDs and AoAs (or spatial frequencies respectively) of the transmit and receive array propagation vectors, CS algorithms can be applied for channel estimation. This spatially quantized channel representation is also called virtual channel [HGPR+16]. Following the steps of [AALH14], we can define the received signal after correlation according to (2.74)

$$\mathbf{z} = \mathbf{w}_{\mathrm{bb}}^H \mathbf{w}_{\mathrm{rf}}^H \mathbf{H} \mathbf{f}_{\mathrm{rf}} \mathbf{f}_{\mathrm{bb}} + \underbrace{\mathbf{w}_{\mathrm{bb}}^H \mathbf{w}_{\mathrm{rf}}^H \mathbf{n}}_{\nu} . \tag{2.77}$$

By performing the training sequentially, using all possible analog precoders and combiners of the codebook, we can stack the results into a matrix

$$\mathbf{Z} = \tilde{\mathbf{W}}^H \mathbf{H} \tilde{\mathbf{F}} + \boldsymbol{\nu}, \tag{2.78}$$

where $\mathbf{Z} \in \mathbf{C}^{L \times K}$ contains all observations, $\tilde{\mathbf{F}} = \tilde{\mathbf{F}}_{\mathrm{rf}} \tilde{\mathbf{F}}_{\mathrm{bb}}$ and $\tilde{\mathbf{W}} = \tilde{\mathbf{W}}_{\mathrm{rf}} \tilde{\mathbf{W}}_{\mathrm{bb}}$ with $\tilde{\mathbf{F}}_{\mathrm{bb}} = [\mathbf{f}_{\mathrm{bb},1}, ..., \mathbf{f}_{\mathrm{bb},K}]$ and $\tilde{\mathbf{W}} = [\mathbf{w}_{\mathrm{bb},1}, ..., \mathbf{w}_{\mathrm{bb},L}]$ [HGPR+16], and $\boldsymbol{\nu} \in \mathbb{C}^{L \times K}$ is a Gaussian distributed noise matrix, with each element $[\boldsymbol{\nu}]_{l,k} \sim \mathcal{CN}(0, \sigma^2)$; $l = \{1, ..., L\}$, $k = \{1, ..., K\}$.

For simplicity, the channel matrix can be approximated by allowing only array-propagation vectors, which are discretized in the angular domain $\phi_{\mathrm{D}} \in \{-\frac{\pi}{2}, -\frac{\pi}{2} + \frac{\pi l}{V_{\mathrm{tx}}}, ..., \frac{\pi}{2}\}$ using V_{tx} steps for the transmitter side. In similar form this can be applied for the receiver side, using V_{rx} steps. The channel matrix can be written as a function of the array propagation codebook matrix for the transmitter side $\tilde{\mathbf{A}}_{\mathrm{tx}} = \left[\mathbf{a}_{\mathrm{tx}}(\phi_{\mathrm{A},l}), ..., \mathbf{a}_{\mathrm{tx}}(\phi_{\mathrm{A},V_{\mathrm{rx}}-1})\right]$ and the receiver side $\tilde{\mathbf{A}}_{\mathrm{rx}} = \left[\mathbf{a}_{\mathrm{rx}}(\phi_{\mathrm{A},l}), ..., \mathbf{a}_{\mathrm{rx}}(\phi_{\mathrm{A},N_{\mathrm{rx}}})\right]$

$$\mathbf{H} \approx \tilde{\mathbf{H}} = \tilde{\mathbf{A}}_{\mathrm{rx}} \tilde{\mathbf{D}} \tilde{\mathbf{A}}_{\mathrm{tx}}^H, \tag{2.79}$$

where $\tilde{\mathbf{D}} \in \mathbf{C}^{V_{\mathrm{rx}} \times V_{\mathrm{tx}}}$ is a diagonal matrix with only a few nonzero elements, representing the MPCs of the channel. Stacking all observations into a vector, (2.78) can be modified to

$$\mathbf{z} = \mathrm{vec}\{\tilde{\mathbf{W}}^H \tilde{\mathbf{H}} \tilde{\mathbf{F}}\} + \mathrm{vec}\{\boldsymbol{\nu}\}. \tag{2.80}$$

Using a property of the Kronecker product, the formula can be represented in the following form

$$\mathbf{z} = \left(\tilde{\mathbf{F}}^{\mathrm{T}} \otimes \tilde{\mathbf{W}}^H\right) \mathrm{vec}\{\tilde{\mathbf{H}}\} + \mathrm{vec}\{\boldsymbol{\nu}\} \tag{2.81}$$

$$\mathbf{z} = \underbrace{\left(\tilde{\mathbf{F}}^{\mathrm{T}} \otimes \tilde{\mathbf{W}}^H\right) \left(\tilde{\mathbf{A}}_{\mathrm{t}}^* \otimes \tilde{\mathbf{A}}_{\mathrm{r}}\right)}_{\boldsymbol{\Phi}} \underbrace{\mathrm{vec}\{\mathbf{D}\}}_{\mathbf{d}} + \mathrm{vec}\{\boldsymbol{\nu}\} \tag{2.82}$$

$$\mathbf{z} = \boldsymbol{\Phi} \mathbf{d} + \mathrm{vec}\{\boldsymbol{\nu}\}. \tag{2.83}$$

As $\mathbf{d}$ is sparse, (2.82) has the form of a standard CS problem, where $\left(\tilde{\mathbf{A}}_{\mathrm{t}}^* \otimes \tilde{\mathbf{A}}_{\mathrm{r}}\right)$ is the dictionary matrix and $\left(\tilde{\mathbf{F}}^{\mathrm{T}} \otimes \tilde{\mathbf{W}}^H\right)$ is the sensing matrix, which can be combined to the equivalent sensing matrix $\mathbf{\Phi}$. The sensing problem can be formulated in two different ways, depending on whether the number of MPCs is known or not. If the number is known the problem is to minimize the error between the observation and the weights of the paths

$$\arg\min \|\mathbf{z} - \Phi\mathbf{d}\|_2 \quad \text{subject to} \quad \min \|\mathbf{d}\|_0 = M_\mathrm{p}. \tag{2.84}$$

If the number is unknown, another approach is to try to estimate $\mathbf{d}$ such that the number of active paths is minimal (as we know the channel is sparse) by not exceeding a certain error

$$\min \|\mathbf{d}\|_0 \quad \text{subject to} \quad \|\mathbf{z} - \Phi\mathbf{d}\|_2 < \sigma. \tag{2.85}$$

The problem can be solved using standard CS algorithms. An adaptive CS technique was used in [AALH14] and also the greed algorithm orthogonal matching pursuit (OMP) can be used to get the solution [MRRA$^+$15, ALH15].

2.5.2 Precoding and Combining

The overall goal of spatial precoding and combining is to find the precoding and combing matrices that maximize the achievable rate under certain constraints. Without loss of generality, this problem can be stated in the following form, given perfect (explicit) CSI assuming Gaussian signaling

$$\begin{aligned}
\max_{\mathbf{W}_\mathrm{rf},\mathbf{W}_\mathrm{bb},\mathbf{F}_\mathrm{rf},\mathbf{F}_\mathrm{bb}} & \log_2 \det\left(\mathbf{I}_{N_\mathrm{st}} + \mathbf{R}_\mathrm{n}^{-1}\mathbf{W}_\mathrm{bb}^H\mathbf{W}_\mathrm{rf}^H\mathbf{H}\mathbf{F}_\mathrm{rf}\mathbf{F}_\mathrm{bb}\mathbf{R}_\mathrm{s}\mathbf{F}_\mathrm{bb}^H\mathbf{F}_\mathrm{rf}^H\mathbf{H}^H\mathbf{W}_\mathrm{rf}\mathbf{W}_\mathrm{bb}\right)\\
& \text{s.t.}\operatorname{Tr}\left(\mathbf{F}_\mathrm{rf}\mathbf{F}_\mathrm{bb}\mathbf{R}_\mathrm{s}\mathbf{F}_\mathrm{bb}^H\mathbf{F}_\mathrm{rf}^H\right) = P_\mathrm{tx}\\
& \mathbf{w}_{\mathrm{rf},i} \in \tilde{\mathbf{W}}_\mathrm{rf},\ \forall i \in \{0,...,N_\mathrm{rf}^\mathrm{rx}-1\}\\
& \mathbf{f}_{\mathrm{rf},j} \in \tilde{\mathbf{F}}_\mathrm{rf},\ \forall i \in \{0,...,N_\mathrm{rf}^\mathrm{rx}-1\}
\end{aligned} \tag{2.86}$$

where $\mathbf{R}_\mathrm{s} = P_\mathrm{tx}\mathbf{I}_{N_\mathrm{st}}$ is the covariance matrix of the signal and $\mathbf{R}_\mathrm{n} = \sigma^2\mathbf{W}_\mathrm{bb}^H\mathbf{W}_\mathrm{rf}^H\mathbf{W}_\mathrm{rf}\mathbf{W}_\mathrm{bb}$ the noise covariance matrix and $\mathbf{f}_{\mathrm{rf},i}$ the ith column of $\mathbf{F}_\mathrm{rf}$ and $\mathbf{w}_{\mathrm{rf},i}$ the ith column of $\mathbf{W}_\mathrm{rf}$. We assume a total power constraint similar to the one presented in [CRKF18]. This general form of the problem can be translated to a digital beamforming problem, by exchanging the RF precoding and combining matrices $\mathbf{F}_\mathrm{rf}$ and $\mathbf{W}_\mathrm{rf}$ with identity matrices and increasing the dimensions of the baseband matrices $\mathbf{F}_\mathrm{bb}$ and $\mathbf{W}_\mathrm{bb}$ to cover all antenna elements with $N_\mathrm{st} = 1$. In the case of an analog system, the baseband matrices are exchanged with identity matrices, while the precoding matrices become vectors. Thus in the case of analog beamforming, the problem

$$\begin{aligned}
\max_{\mathbf{w}_\mathrm{rf},\mathbf{f}_\mathrm{rf}} & \log_2\left(1 + P_\mathrm{tx}\mathbf{R}_\mathrm{n}^{-1}\mathbf{w}_\mathrm{rf}^H\mathbf{H}\mathbf{f}_\mathrm{rf}\mathbf{f}_\mathrm{rf}^H, \mathbf{H}^H\mathbf{w}_\mathrm{rf}\right)\\
& \text{s.t.}\|\mathbf{w}_\mathrm{rf}\|_\mathrm{F} = 1\\
& \|\mathbf{f}_\mathrm{rf}\|_\mathrm{F} = 1\\
& \mathbf{w}_\mathrm{rf} \in \tilde{\mathbf{W}}_\mathrm{rf}\\
& \mathbf{f}_\mathrm{rf} \in \tilde{\mathbf{F}}_\mathrm{rf}
\end{aligned} \tag{2.87}$$

where $\|\mathbf{w}_{\mathrm{rf}}\|_{\mathrm{F}} = 1$ and $\|\mathbf{f}_{\mathrm{rf}}\|_{\mathrm{F}} = 1$ is always fulfilled, can be translated into a similar problem of maximizing the receive power

$$\begin{aligned} &\max_{\mathbf{w}_{\mathrm{rf}},\mathbf{f}_{\mathrm{rf}}} \left(\left|\mathbf{w}_{\mathrm{rf}}^{H}\mathbf{H}\mathbf{f}_{\mathrm{rf}}\right|^{2}\right) \\ &\mathrm{s.t.}\mathbf{w}_{\mathrm{rf}} \in \tilde{\mathbf{W}}_{\mathrm{rf}} \\ &\mathbf{f}_{\mathrm{rf}} \in \tilde{\mathbf{F}}_{\mathrm{rf}} \end{aligned} . \tag{2.88}$$

Solving this optimization problem is challenging as it is not convex in general; thus [ARAS+14] proposed to optimize the transmitter and receiver side individually by minimizing the distance of the constrained precoder and the optimal digital precoder, which can be obtained by SVD of the channel matrix first. This problem can also be solved by the OMP algorithm. In the second step, the combiner is chosen to minimize the distance to the minimum mean square error (MMSE) receiver, constructed using the estimated precoder.

2.5.3 Joint Channel Estimation and Combining

Both, explicit channel estimation and precoding use parameters of the propagation channel but can only observe the radio channel including the effect of the antenna. Thus the question arises, whether the estimation of radio channel parameters can be used for the precoding directly. The benefit of using radio channel parameters directly to pass information between the channel estimation and precoding algorithm lies in a decreased complexity, while the result of the estimation is bound to the same radio channel. As analog beamforming systems have only access to the parameters of the radio channel as digital processing of parallel received signals is not possible, they already make use of the joined estimation and precoding inherently by choosing the beam with the highest receive power. Note that since the parameters of the radio channel (e.g., beam number with the largest receive power) correlate with parameters of the propagation channel (e.g.AoA), the correlation coefficient between the parameters increases with increasing antenna size as the beamwidth decreases. A simple algorithm is the exhaustive search (ES) algorithm, which scans thoroughly the (accessible) angular (or spatial frequency) space for strong MPC using all possible beam pair combination of the transmitter and receiver codebook and finally chooses the one with the largest receive power. This algorithm can be realized with the analog beamforming system and was therefore introduced for mmW systems. In the standards IEEE 802.15c [IEEb] and IEEE 802.11ad/ay [IEEa], a tree search algorithm is employed, which uses different beam widths successively in order to avoid a time-consuming search with fine angular resolved beams.

The philosophy of joined estimation and precoding without explicit CSI can also be applied to HBF systems. The work in [CRKF17] shows that channel estimation and precoding can be done jointly if the resolution of the array propagation matrices $\mathbf{A}_{\mathrm{tx}}$ and $\mathbf{A}_{\mathrm{rx}}$ are equal to the number of transmitter and receiver antennas respectively and if they are sampled in spatial frequency domain instead of the angular domain as before. More specifically,

$$\tilde{\mathbf{A}}_{\mathrm{rx}} = \left[\mathbf{a}_{\mathrm{tx}}(k_{\mathrm{A},0}), ..., \mathbf{a}_{\mathrm{tx}}(k_{\mathrm{A},l}), ..., \mathbf{a}_{\mathrm{tx}}(k_{\mathrm{A},N_{\mathrm{tx}}-1})\right], \tag{2.89}$$

where $k_{\mathrm{tx},l} = \frac{-2\pi}{\lambda} + \frac{4\pi l}{\lambda N_{\mathrm{tx}}-1}$ with $l = 0, .., N_{\mathrm{tx}} - 1$. $\tilde{\mathbf{A}}_{\mathrm{tx}}$ is analogously defined. These matrices have the form of DFT matrices and thus have full rank and are orthonormal. As the array response vectors corresponding to the nonzero MPC span the same subspace as the singular vectors of the channel matrix $\mathbf{H}$ (or equivalently, the left-singular vectors of $\mathbf{H}$ and the array propagation vectors of dictionary $\tilde{\mathbf{A}}_{\mathrm{rx}}$ are linear dependent, an analogous relation is valid for the right-singular vectors of the channel matrix and $\tilde{\mathbf{A}}_{\mathrm{tx}}$) [CKRF16], the array propagation vectors corresponding to the MPC, which are determined by the channel estimation, can be directly used for the precoding design [CKRF16]. If DFT matrices are also used as RF precoding and combining matrices $\tilde{\mathbf{F}}_{\mathrm{rf}}$ and $\tilde{\mathbf{W}}_{\mathrm{rf}}$, the problem reduces to a beam-selection problem of choosing the beams with the largest receive power [CRKF17]. However, in reality the AoDs and AoAs are continuous. The described technique can still be used in the continuous case, but the array propagation vectors corresponding to the largest receive power do not exactly correspond to the array propagation vector corresponding to the AoDs and AoAs, and thus the subspace of these array propagation vectors equals not exactly to the subspace of the channel matrix, which leads to an estimation error. Correspondingly, the precoding can also be erroneous. Thus taking only the receive power of the beams as a measure for choosing the beams in a HBF system will not always lead to the optimal solution. As the received power is nevertheless a strong indicator of the quality of a beam choice, this problem can be solved by trying several combinations of beams with relatively high receive power (the limitation to the dictionary of array propagation vectors $\tilde{\mathbf{A}}_{\mathrm{rx}}$ and $\tilde{\mathbf{A}}_{\mathrm{tx}}$ acts as a sampling in the spatial frequency domain, therefore the observations are subject to the leakage or smoothing effect, which results in the fact that more observations have nonzero receive power than MPC are existing), and applying the SVD of the effective channel, which is visible to the digital system and is the channel matrix multiplied with the RF precoder and combiner using candidate beams

$$\mathbf{H}_{\mathrm{eff}} = \mathbf{W}_{\mathrm{rf}}^{H}\mathbf{H}\mathbf{F}_{\mathrm{rf}}, \tag{2.90}$$

with $\mathbf{W}_{\mathrm{rf}} = \left[\mathbf{w}_{\mathrm{rf},1}, ..., \mathbf{w}_{\mathrm{rf},N_{\mathrm{s,rx}}}\right]$ and an analogues expression for $\mathbf{F}_{\mathrm{rf}}$. The mutual information can be maximized by choosing the beams $\mathbf{F}_{\mathrm{rf}}$ and $\mathbf{W}_{\mathrm{rf}}$ which cause the effective channel to have the largest singular values. This requires a significant amount of complex SVD calculations. To address this problem, [CRKF18] proposed to use the low-complex Frobenius norm or the determinant of the effective channel instead as a performance criterion in order to identify promising beam candidate combinations. The work also shows that the optimal baseband precoder $\mathbf{F}_{\mathrm{bb}}^{\mathrm{opt}}$ and combiner $\mathbf{W}_{\mathrm{bb}}^{\mathrm{opt}}$ can be calculated using the left and right-singular values of the effective channel matrix $\mathbf{H}_{\mathrm{eff}}$ and the optimal RF precoding matrix $\mathbf{F}_{\mathrm{rf}}^{\mathrm{opt}}$ and combining matrix $\mathbf{W}_{\mathrm{rf}}^{\mathrm{opt}}$ [CRKF18].

2.6 Summary

In this chapter, we laid the foundation of the discussion of later chapters, by showing the features of the mmW channel, which distinguishes itself from communication channels used in conventional sub 6 GHz frequency bands. We showed the necessity to develop new propagation and radio channel models, the latter also taking into account the directionality of the antenna arrays, and presented viable approaches for such models. Based on the necessity to work with high gain antennas, we presented fundamental antenna array theory with special regard to the interpretation of an antenna array as a spatial filter. We used this theory to compare and evaluate three different system designs for MIMO systems. Using this knowledge, we presented a system and signal model and discussed the main goals and challenges of wireless communications in the context of cellular systems. Finally, we gave an extensive review of known concepts for channel estimation and precoding for digital, hybrid and analog beamforming in literature.

Chapter 3

Hypothesis Test for Beam Selection

In order to design promising precode and combine vectors for the antenna arrays to increase the system throughput, the channel has to be estimated. While conventional MIMO systems can obtain the channel coupling factors for each antenna element pair individually, the aforementioned design restrictions force HBF and ABF systems to rather use predefined beam patterns, where the angular stimulation of the channel is characterized to get insight in the channel properties. To acquire sufficient information about the environment, all beam pairs are usually trained exhaustively.

In order to determine the beam pair achieving the best channel conditions, the beam pair combination with the largest receive power provides the optimal achievable rate in the case of an ABF system. The sparsity and low diffraction of the signal in the mmW band combined with the usage of large antenna arrays provides an accumulation of the received power in a small angular sector and allows to draw direct conclusions on the angles of the dominant paths also for HBF systems. Usually, the second and third strongest paths are also of interest. This information can be used to reconstruct the channel matrix [SY15, AALH14] or can be used directly for choosing the best precoder and combiner for communications [CRKF18, KCF16]. However, it is not initially clear, how long the training signal has to be in order to provide reliable information about the beam pair with the largest received power. A possible trade-off between the length of the training sequence and performance loss taking into consideration the second strongest beam pair was not investigated yet to the best of the authors' knowledge. Hypothesis testing in the context of array processing is presented in [JD92], but for a radar-like application assuming DBF systems. The work [BHR+15] uses hypothesis testing in a different context without searching for the minimum training length.

In this chapter, we approach this problem by investigating the behavior of a fixed length M-ary hypothesis test for channel estimation and precoding for a simple analog beamforming scenario and one MPC.

3.1 System Model

In order to motivate the special form of the M-ary fixed length test compared to the standard hypothesis test [Kay98], [MV12], we introduce the beam selection problem for

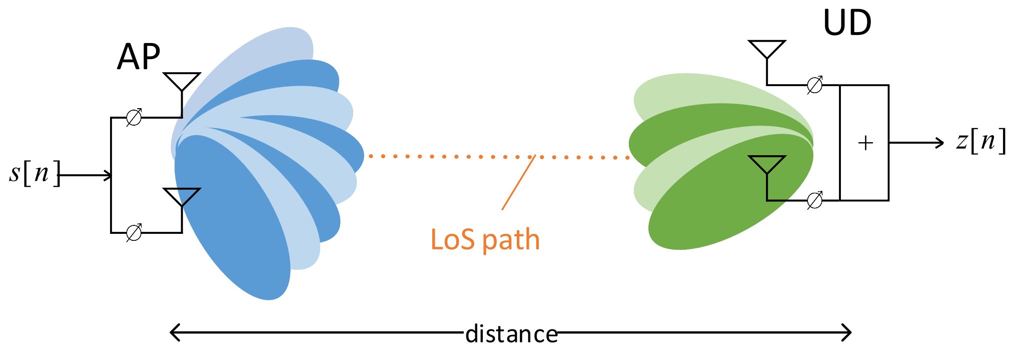


Figure 3.1: System model. Transmitter and receiver with a codebook of 8 and 4 beams, respectively that communicate over a single path, e.g. (Source: [KRF19] © 2019 IEEE).

a point-to-point mmW link. We limit ourselves to the case of an ABF transmitter and receiver using sets of beams in different directions and a channel with one dominant MPC. A sketch of the system is shown in Fig. 3.1. We use the received power as a selection criterion, where the best receive-transmit beam pair is directly used for data transmission as this allows us to use reasonable benchmarks for the performance.

Using (2.74) and the definitions of the RF codebooks $\tilde{\mathbf{F}}_{\mathrm{rf}}$ and $\tilde{\mathbf{W}}_{\mathrm{rf}}$, respectively, we reformulate the received signal as

$$y_m[n_{\mathrm{s}}] = \mathbf{w}_l^H \mathbf{H} \mathbf{f}_k s[n_{\mathrm{s}}] + \mathbf{w}_l^H \mathbf{n}_{\mathrm{c}}[n_{\mathrm{s}}], \tag{3.1}$$

where the beam pair is indexed using the combined index $m = (k-1)L+l$, $\mathbf{H} \in \mathbb{C}^{N_{rx} \times N_{tx}}$. In order to get a reasonable estimate of the strength of the effective channel coefficient using beam pair combination m, we calculate the sample-wise complex cross correlation $z_m^{\mathrm{c}}[n_{\mathrm{s}}]$ between the received signal y_m and the training signal sample $s[n_{\mathrm{s}}]$

$$\begin{aligned} z_m^{\mathrm{c}}[n_{\mathrm{s}}] &= \frac{s^*[n_{\mathrm{s}}]}{|s[n_{\mathrm{s}}]|^2} y_m[n_{\mathrm{s}}] \\ &= \underbrace{\mathbf{w}_l^H \mathbf{H} \mathbf{f}_k}_{A_m \mathrm{e}^{j\phi}} + \underbrace{\mathbf{w}_l^H \mathbf{n}_{\mathrm{c}}[n_{\mathrm{s}}]}_{\nu_{\mathrm{c}}[n_{\mathrm{s}}] \sim \mathcal{CN}(0, \sigma_{\mathrm{c}}^2)} \ . \end{aligned} \tag{3.2}$$

The result of the correlation consist of a noise free coefficient $A_m \mathrm{e}^{j\phi}$ and noise. ν_c is still additive white Gaussian noise with variance σ_{c}^2 due to the restriction $\|\mathbf{w}_l\|_2 = 1$. We call $A_m \mathrm{e}^{j\phi}$ the effective complex channel coefficient between the transmitter and receiver using beam pair $m = (k-1)L + l$. The correlation result $z_m^{\mathrm{c}}[n_{\mathrm{s}}]$ is thus a noisy observation of this coupling coefficient. It is reasonable to assume that the variances for all observations of different beam pair combinations m are the same, as the bandwidth, which dominates the thermal noise, stays constant over all beam pairs.

We further assume perfect synchronization in order to convert the received signal to the real domain, which makes the following calculations easier. We perform the conversion by estimating the phase information of the correlated observation. Thus, knowing the argument of the correlated training signal

$$\hat{\phi} = \angle\left(\sum_{n_\mathrm{s}=1}^{N_\mathrm{s}} z_m^{\mathrm{c,nf}}[n_\mathrm{s}]\right) \approx \angle \mathbf{w}_l^H \mathbf{H}\mathbf{f}_k, \tag{3.3}$$

applying the phase rotation and only taking the real part, we define the equivalent real-valued observation sample $z_m[n_\mathrm{s}]$

$$\begin{aligned} \Re(z_m^{\mathrm{c,nf}}[n_\mathrm{s}]) + \Re(\nu_\mathrm{c}[n_\mathrm{s}]) &\approx \Re(z_m^{\mathrm{c}}[n_\mathrm{s}]\mathrm{e}^{-j\hat{\phi}}) \\ z_m[n_\mathrm{s}] = A_m + \nu[n_\mathrm{s}], &\quad z_m[n_\mathrm{s}] \sim \mathcal{N}(A_m, \sigma^2). \end{aligned} \tag{3.4}$$

Throughout this chapter, we call A_m the (effective) coupling coefficient or amplitude of signal $z_m[n_\mathrm{s}]$.

The noise variance of the real signal is half of the variance of the complex signal $\sigma^2 = \frac{\sigma_\mathrm{c}^2}{2}$. In this way, we can deal with real values instead of complex or power values [KCF16], which allows us to use real Gaussian probability density functions (PDFs) for the calculation of the selection probability. Note that, due to the effect of side lobes, it is likely that other nonzero amplitude levels besides the strongest one are existing.

3.2 M-ary Hypothesis Test Problem Description

The goal of a beam searching procedure is to choose the coupling coefficient associated with that beam pair combination m which will lead to the largest achievable throughput. As we assume to have only one RF chain for the transmitter and receiver respectively, this corresponds to the search of the largest value of all M coupling coefficients for this system model.

We can formulate the beam selection problem as an M-ary hypothesis test by assigning one hypothesis to each of the events that coupling coefficient $1, ..., M$ is the largest one

$$\begin{aligned} &\mathcal{H}_1 : z_1[n_\mathrm{s}] = A_1 + \nu[n_\mathrm{s}],\ \&\&\ A_1 \equiv A_{\max} \\ &\quad ... \\ &\mathcal{H}_M : z_M[n_\mathrm{s}] = A_M + \nu[n_\mathrm{s}],\ \&\&\ A_M \equiv A_{\max}. \end{aligned} \tag{3.5}$$

Taking advantage of the length of the training signals, we define sufficient statistics for each hypothesis by taking the mean of the observations of the N_s samples

$$\begin{aligned} \bar{z}_m &= \frac{1}{N_\mathrm{s}} \sum_{n_\mathrm{s}=1}^{N_\mathrm{s}} z_m[n_\mathrm{s}] \\ &= A_m + \frac{1}{N_\mathrm{s}} \sum_{n_\mathrm{s}=1}^{N_\mathrm{s}} \nu[n_\mathrm{s}] \\ &= A_m + \bar{\nu}. \end{aligned} \tag{3.6}$$

The variance of the noise of the sufficient statistics $\bar{\nu} \sim \mathcal{N}(0, \frac{\sigma^2}{N_\mathrm{s}})$ is reduced by a factor $\frac{1}{N_\mathrm{s}}$ compared to a single sample of the observation. The reliability of detecting the estimate of the largest coupling coefficient can be determined by the variance of the noisy observations. Additionally, the variance is directly dependent on the training length. The problem we want to discuss in this chapter is to determine the *minimal* length of the training signal which is sufficient to achieve a specified selection probability or another performance measure of the selection process. Note that the definition of the hypothesis test (3.5) differs from the standard formulation of an M-ary hypothesis test with genie knowledge in [Kay98] and [MC78, section 5.2, p.110] . The difference comes from the fact that we have to compare M signals in parallel instead of categorizing one signal to M categories.

3.3 Selection Performance Criterion

The selection probability is defined as the probability that we correctly choose signal z_i, if $\mathcal{H}_i$ is true

$$\begin{aligned} P_\mathrm{s} &= \sum_{i=1}^{M} P\left(\hat{\mathcal{H}}_i|\mathcal{H}_i\right) P\left(\mathcal{H}_i\right) \\ &= \sum_{i=1}^{M} P\left(\bar{z}_i = \max_{k\in\{1\ldots M\}} \bar{z}_k\right) P\left(\mathcal{H}_i\right), \end{aligned} \tag{3.7}$$

which is equivalent to choosing the maximum estimate of the coupling coefficient when using the maximum likelihood (ML) rule. We assume equal prior probabilities $P\left(\mathcal{H}_i\right) = \frac{1}{M}$, which correspond to the common case of uniformly distributed AoAs. As the conditional probability is equal for every hypothesis i due to the symmetry of the hypothesis test, we can focus our study on the selection probability of a special case. Therefore, we can assume $A_1 > \max_{i\in\{1\ldots M\}} A_i$ without loss of generality which means that $\mathcal{H}_1$ is true. So we can rewrite (3.7) as

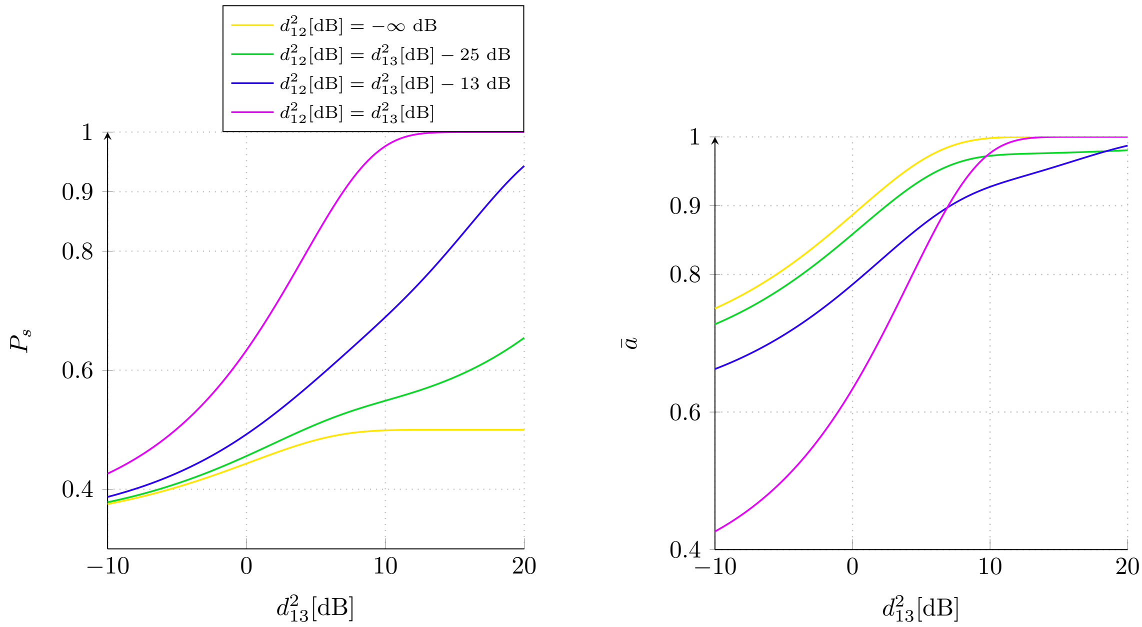
$$P_\mathrm{s} = P\left(\hat{\mathcal{H}}_1|\mathcal{H}_1\right). \tag{3.8}$$

In this chapter, we are interested in the mean normalized performance of the channel we can achieve by using the selection based on the training signal. The selection probability does only provide a measure of the success of a detection problem, ignoring the severeness of different error events. Therefore we introduce a more sophisticated performance criterion based on the average signal receive strength of the channel given the selection probabilities, the normalized received amplitude

$$\bar{a} = \frac{\bar{A}}{A_\mathrm{max}} = \frac{\sum_{\substack{i\in\{1,..,M\}\\ i\neq j}} P\left(\bar{z}_i = \max_{j\in\{1,..,M\}} \bar{z}_j\right) A_i}{A_\mathrm{max}}, \tag{3.9}$$

which is the ratio between the mean coupling coefficient achieved by the selection problem and the maximum amplitude.

We now derive the selection probability for the M-ary hypothesis test defined in (3.5). A simple mathematical expression using Q functions for the selection probability and error probability of the binary analysis is given in [Kay98]. If we increase the number of signals to $M > 2$, the relation becomes more complicated and is given by an integral over a multivariate Gaussian distribution. By exploiting the fact that the signals $\bar{z}_m$ are



(a) P_s against d_{13}^2[dB]. **(b)** $\bar{a}$ against d_{13}^2[dB].

Figure 3.2: Performance in terms of selection probability P_s of the best beam pair and normalized amplitude $\bar{a}$ against deflection coefficient (differential SNR) for $M = 3$ beam pairs (one observation). It can be seen that the overall performance $\bar{a}$ is good (the system receives a strong signal), even if P_s is low for some amplitude distributions (yellow and green line). This motivates the use of $\bar{a}$ as the performance indicator. (Source: [KRF19] © 2019 IEEE).

mutually independent, we can rewrite the multi-dimensional integral as a product of M one-dimensional integrals. Using notation of the Q-function, the selection probability of an M-ary hypothesis test defined in (3.5) under the assumption that $\mathcal{H}_1$ is true can be given by a one-dimensional integration of the product of Gaussian PDF and $(M-1)$ Q-functions in the form

$$\begin{aligned} P_\mathrm{s} &= P\left(\hat{\mathcal{H}}_1 | \mathcal{H}_1\right) \\ &= P\left(\bar{z}_1 = \max_{i \in \{1 \dots M\}} \bar{z}_i\right) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} e^{\frac{-(\bar{z}_1 - A_1)^2}{2\bar{\sigma}^2}} \cdot \prod_{i \in \{2 \dots M\}} \mathrm{Q}\left(\frac{A_i - \bar{z}_1}{\bar{\sigma}}\right) \mathrm{d}\bar{z}_1. \end{aligned} \tag{3.10}$$

The complete derivation of how to get this result is given in the Appendix A. Using variable substitution, (3.10) can be rewritten using the roots of the deflection coefficients $d_{1i} = \frac{\sqrt{N_\mathrm{s}}(A_1 - A_i)}{\sigma}$ and $x = \frac{\sqrt{N_\mathrm{s}}(\bar{z}_1 - A_1)}{\sigma}$

$$P_\mathrm{s} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x)^2}{2}} \cdot \prod_{i \in \{2 \dots M\}} \mathrm{Q}\left(-d_{1i} - x\right) \mathrm{d}x. \tag{3.11}$$

An expression of similar form to (3.10) is given in [Kay98] for an M-ary test of signals (signal correlator). While [Kay98] uses mutual uncorrelated signals for the derivation of the formula, we rely on the fact that the noise realizations of the individual signals z_i

are mutual uncorrelated. The same formula can be also used to calculate the probability of error that we detect the wrong amplitude using $x_k = \frac{\sqrt{N}(z_k - A_k)}{\sigma}$, i.e. to select another amplitude $A_k,\ k \neq 1$

$$\begin{aligned} P_{\mathrm{e,k}} &= P\left(\hat{\mathcal{H}}_k | \mathcal{H}_1\right) \\ &= P\left(\bar{z}_k = \max_{i\in\{1\ldots M\}} \bar{z}_i\right) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x_k)^2}{2}} \cdot \prod_{i\in\{1\ldots M\}\setminus k} \mathrm{Q}\left(-d_{ki} - x_k\right) \mathrm{d}x_k \ . \end{aligned} \tag{3.12}$$

Note that the deflection coefficient is a function of the length of the training signal and thus also the selection and error probabilities are functions of N_{s}. Therefore, the total probability of missed selection of the strongest signal is found as

$$P_{\mathrm{e}} = \sum_{k=2}^{M} P_{\mathrm{e,k}} \ . \tag{3.13}$$

Although the selection probability equation in (3.11) has a similar form as the error rate computation of differential quadrature phase shift keying in Nakagami-m fading [Bea06], a general solution to the integral for arbitrary numbers of M has not been found to the best of the author's knowledge.

In order to clarify the difference between the selection probability P_{s} and the mean normalized amplitude $\bar{a}$, we plotted these parameters for a ternary hypothesis test example ($M = 3$) for different values of the deflection coefficients in Figs. 3.2a and 3.2b, respectively. The deflection coefficient d_{1i}^2 is defined by the square of the difference between coupling coefficient one and coupling coefficient i related to the variance $\bar{\sigma}^2$, which is dependent on the length of the training signal. It can be understood as a measure of how good the value of two coupling coefficients can be distinguished from each other and be interpreted as differential signal to noise ratio (SNR). Because we consider the special case $A_1 = A_{\max}$, d_{1i} is always positive. The large deflection coefficient can thus be interpreted as a large separation of amplitudes A_1 and A_i, meaning that A_i is much smaller than A_1. Another interpretation of a large deflection coefficient is a small variance $\bar{\sigma}^2$, resulting from a long training sequence. Thus the two coefficients d_{12}^2 and d_{13}^2 define the relation between the amplitudes A_1, A_2 and A_3. In this example we set $A_3 = 0$ and $\sigma^2 = 1$. We plot the deflection coefficients at a logarithmic scale, as they can be interpreted as differential SNR values. Fig. 3.2a shows the selection probability P_{s} in dependency of the logarithmic deflection coefficient $(d_{13}^2)_{\mathrm{dB}}$, which can be interpreted as differential SNR $\frac{A_1^2 - A_3^2}{\bar{\sigma}^2}$ and defines how well the observations $\bar{z}_1$ and $\bar{z}_3$ can be separated. The value of the second amplitude is defined by the deflection coefficient d_{12}^2. We show the case of 4 different values of this coefficient

$$(d_{12}^2)_{\mathrm{dB}} = \begin{cases} -\infty\,\mathrm{dB}, & A_2 = A_1 \\ (d_{13}^2)_{\mathrm{dB}} - 25\,\mathrm{dB}, & A_2 = 10^{\frac{-25}{20}}(A_3 - A_1) + A_1 \\ (d_{13}^2)_{\mathrm{dB}} - 13\,\mathrm{dB}, & A_2 = 10^{\frac{-13}{20}}(A_3 - A_1) + A_1 \\ (d_{13}^2)_{\mathrm{dB}}, & A_2 = A_3 \ . \end{cases} \tag{3.14}$$

It is worth to note that the selection probability in (3.8) depends only on the deflection coefficient, i.e. the ratio between the squared amplitudes and is independent of the absolute values of the amplitudes. Fig. 3.2a shows that P_{s} rises with increasing deflection coefficients d_{12}^2 and d_{13}^2. The reason of this effect is that the observation $\bar{z}_1$ is better separable from other observations with increasing deflection coefficients. We can see that the cases of $d_{12}^2 = -\infty$ and $d_{12}^2 = d_{13}^2$ define the upper and lower bound of the selection probability performance. In contrast to P_{s}, the absolute values of the amplitudes have a large impact on the performance criterion $\bar{a}$ in (3.9). We can interpret this criterion as a weighted sum of selection probabilities. Thus not only the separability of the observation, but also the weighting of the selection event is crucial for the performance. Therefore the case of $d_{12}^2 = d_{13}^2$, i.e. $A_2 = A_3$, is transformed from the lower bound of P_{s} to the upper bound for $\bar{a}$. Similarly, the case $d_{12}^2 = -\infty$, i.e. $A_1 = A_2$, is transformed from the upper bound of P_{s} to the lower bound of $\bar{a}$ for small values of the deflection coefficient d_{13}^2. This means that we can achieve a better performance if we have some competing amplitudes $A_i > A_M$. However, the existence of the competing amplitudes cause a down shift of the increase of the performance with rising deflection coefficient for large values of d_{13}^2. The reason for this is that for amplitudes $A_1 > A_2 > A_3$ it is more probable to select signal $\bar{z}_1$ or $\bar{z}_2$ with expectation values A_1 and A_2 than in the case of $A_1 > A_2 = A_3$, but it is harder to select only signal $\bar{z}_1$. The intersection value d_{13}^2 of the cases $A_1 > A_2 > A_3$ and $A_1 > A_2 = A_3$ is dependent on the absolute values of the amplitudes and can only be found numerically. However, the case of a single nonzero amplitude $A_1 \neq 0,\ A_i = 0\ i = 2, ..., M$ is a worst case scenario for small deflection coefficients and is an appropriate first assumption for the problem at hand due to the sparsity of the mmW channel.

3.4 Composite Hypothesis Test Algorithms

In this section, we develop an algorithm to find the minimum necessary length of the training sequence N_{s} in order to achieve the desired performance of the mean amplitude $\bar{a}_{\mathrm{target}}$. If we assume to have genie knowledge of all the expectation values of the coupling coefficients A_i and the variance σ^2, but the match between the coupling coefficient values and the receive signals $\bar{z}_m$ is assumed to be unknown, we can directly use the definition of the criterion $\bar{a}$ (3.9) and the definition of the selection probability (3.10) and (3.12) to get a relation between $\bar{a}$ and N_{s}

$$\bar{a} = \frac{A_1}{A_1} P_{\mathrm{s}}(N_{\mathrm{s}}) + \frac{A_2}{A_1} P_{\mathrm{e},2}(N_{\mathrm{s}}) + ... + \frac{A_M}{A_1} P_{\mathrm{e},M}(N_{\mathrm{s}}). \tag{3.15}$$

Due to the unavailability of a closed-form solution of the integral of the selection probability P_{s}, this relation can only be solved numerically in order to determine the minimum value of N_{s}. Thus the minimum length for the training signal can be calculated using the selection probability of the fixed length test. However, the correct values A_i and σ^2 are not available at the receiver. Thereby the M-ary hypothesis test becomes a *composite* M-ary hypothesis test [1]. A common approach is to use ML estimates for the unknown parameters

[1] A test is called a composite test if the PDF is assumed to be known up to some parameters

input : $\bar{a}_{\text{target}}$, $\text{SNR}_{\text{assumed}}$
output: Decision to hypothesis $\hat{\imath}_{\max}$

stage 1: Estimation of Parameters;
$\tilde{A}_1 = \sqrt{\text{SNR}_{\text{assumed}}}$, $\tilde{\sigma} = 1$;
begin
  Determine the minimal $\tilde{N}_{\text{s}}$ for which applies
  $\bar{a}_{\text{target}} \overset{!}{\leq} =$
  $\int\limits_{-\infty}^{\infty} \frac{\sqrt{\tilde{N}_{\text{s}}}}{\sqrt{2\pi\tilde{\sigma}}^2} e^{\frac{-(\bar{z}_1 - \tilde{A}_1)^2 \tilde{N}_{\text{s}}}{2\tilde{\sigma}^2}} \cdot \text{Q}^{(M-1)} \left(\frac{-\sqrt{\tilde{N}_{\text{s}}} \bar{z}_1}{\tilde{\sigma}} \right) \mathrm{d}\bar{z}_1$;
  Transmit first training signal of length $\tilde{N}_{\text{s}}$;
  $\hat{A}_i = \frac{1}{\tilde{N}_{\text{s}}} \sum\limits_{1}^{\tilde{N}_{\text{s}}} z_i[\tilde{n}_s]$;
  $\hat{\sigma}^2 = \frac{1}{M} \sum\limits_{1}^{M} \left(\frac{1}{\tilde{N}_{\text{s}} - 1} \sum\limits_{1}^{\tilde{N}_{\text{s}}} (z_{\hat{\imath}_{\min}}[\tilde{n}_s] - \hat{A}_{\hat{\imath}_{\min}})^2 \right)$;
end
stage 2: M-ary Test;
begin
  Determine the number of samples N_{s}, which are necessary to achieve the performance based on the estimated parameters
  $\bar{a}_{\text{target}} \overset{!}{\leq} P_{\text{s}}(\hat{A}_i, N_{\text{s}}, \hat{\sigma}) + \sum_{i \in \{2 \ldots M_{\text{new}}\}} \frac{\hat{A}_i}{\hat{A}_{\max}} P_{\text{e},i}(\hat{A}_i, N_{\text{s}}, \hat{\sigma})$;
  Transmit second training signal of length;
  $N_{\text{s},2} = \min(N_{\text{s}} - \tilde{N}_{\text{s}}, 0)$;
  $\hat{\imath}_{\max} = \arg\max_{\forall i} \bar{z}_i$;
end

Algorithm 1: Composite M-ary Test with single extension

in the selection probability, which is called the generalized likelihood ratio test (GLRT) approach. Although the GLRT is well studied for the binary hypothesis test, the application of this approach to an M-ary test is not easy because some estimated parameters are usually nested [Kay98]. However, the problem at hand does not have this complication, as we have M independent signals with each one having an unknown parameter.

Using the GLRT composite test approach, we first have to provide reasonably good estimates $\hat{A}_i$ and $\hat{\sigma}^2$ in order to determine the required length of the training signal N_{s} from (3.15), given the performance criterion. However, it is not an easy task to determine the effect of the variance of the estimates to the detection probability and thus to the solution of (3.15). The issue is that N_{s} must be found from a version of (3.15) with estimated parameters $\hat{A}_i$ and $\hat{\sigma}^2$, while the actual values of the selection probability P_{s} and thus the normalized amplitude $\bar{a}$ are subject to the actual values.

The unavailability of a closed-form expression of P_{s} prevents an exact investigation of this problem. However, it is clear if we want to design the length of the training sequence in this way that we have a small variance of the estimates to provide a robust solution of (3.15), but not exceeding the length of the training sequence we need for the solving

(3.15) with the correct values. For this reason, we pursue a simple heuristic approach for balancing these two requirements. Therefore we split the algorithm in an estimation stage followed by a detection stage.

In the estimation stage, we initially use reasonable but optimistic apriori assumptions about the values of A_i and σ^2 which are typical for some environments in order to determine the length of the initial training signal. In order to avoid too long training, we determine the length of the training signal needed for achieving the desired performance given assumed channel conditions. For this reason, we assume to have only one nonzero coupling coefficient, which was identified as the worst case assumption for small deflection coefficients above. The relation of the squared amplitude and the noise variance is taken from a meaningful assumption of a typical SNR ($\mathrm{SNR}_{\mathrm{assumed}}$), which can be determined by the free space path loss equation. After we determined the needed initial training length $\tilde{N}_s$ based on these assumptions, we probe each beam-combination using a training signal of this length.

Subsequently, we use the observations $\bar{z}_i$ to get an estimate of amplitude and variance. The amplitude is estimated using the ML estimate

$$\hat{A}_i = \frac{1}{\tilde{N}_\mathrm{s}} \sum_{1}^{\tilde{N}_\mathrm{s}} z_i[\tilde{n}_s], \tag{3.16}$$

while the variance is estimated using the maximum variance unbiased estimator (MVUE)

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{1}^{M} \left(\frac{1}{\tilde{N}_\mathrm{s} - 1} \sum_{1}^{\tilde{N}_\mathrm{s}} (z_{\hat{i}_\mathrm{min}}[\tilde{n}_s] - \hat{A}_{\hat{i}_\mathrm{min}})^2 \right). \tag{3.17}$$

Then we plug the estimates in (3.15) to determine the needed length of the training signal N_s in order to achieve the targeted performance $\bar{a}_\mathrm{target}$. Following this, we continue training for the missing samples $N_{\mathrm{s},2} = N_\mathrm{s} - \tilde{N}_\mathrm{s}$, but only if the length of the training sequence based on the estimates N_s is larger than the length of the training sequence $\tilde{N}_s$ we already probed during the first stage. In the detection phase, the combined N_s samples are used to determine the beam pair with the largest amplitude.

3.4.1 Extended Composite Hypothesis Test

A drawback of the algorithm consists in the fact that the ability to estimate amplitudes in low SNR situations is limited by the length of the initial training length $\tilde{N}_\mathrm{s}$ of the first stage. In order to be able to meet the targeted performance $\bar{a}_\mathrm{target}$ also in those situations, we can use the additional training signal in the second stage of length $N_{\mathrm{s},2}$ to increase the quality of the estimates $\hat{A}_i$ and $\hat{\sigma}$. Using the improved estimates, we can calculate the expected performance again

$$\hat{\bar{a}} = \frac{\hat{A}_1}{\hat{A}_\mathrm{max}} P_\mathrm{s}(\hat{A}_i, N_\mathrm{s}, \hat{\sigma}) + \sum_{i \in \{2 \ldots M_\mathrm{new}\}} \frac{\hat{A}_i}{\hat{A}_\mathrm{max}} P_{\mathrm{e},i}(\hat{A}_i, N_\mathrm{s}, \hat{\sigma}) \tag{3.18}$$

and verify that the targeted performance $\bar{a}_\mathrm{target}$ can be met. If the new estimates show that this is not the case, the length of the training signal which will achieve the performance

input : $\bar{a}_{\text{target}}$, $\text{SNR}_{\text{assumed}}$
output: Decision to hypothesis $\hat{i}$

stage 1: Estimation of Parameters;
$\tilde{A}_1 = \sqrt{\text{SNR}_{\text{assumed}}}$, $\tilde{\sigma} = 1$;
begin
  Determine the minimal $\tilde{N}_{\text{s}}$ for which applies
  $\bar{a}_{\text{target}} \overset{!}{\leq} =$
  $\int\limits_{-\infty}^{\infty} \frac{\sqrt{\tilde{N}_{\text{s}}}}{\sqrt{2\pi\sigma}^2} e^{\frac{-(\bar{z}_1-\tilde{A}_1)^2 \tilde{N}_{\text{s}}}{2\tilde{\sigma}^2}} \cdot \mathrm{Q}^{(M-1)} \left(\frac{-\sqrt{\tilde{N}_{\text{s}}}\bar{z}_1}{\sigma} \right) \mathrm{d}\bar{z}_1$;
  Transmit first training signal of length $\tilde{N}_{\text{s}}$;
  $\hat{A}_i = \frac{1}{\tilde{N}_{\text{s}}} \sum\limits_{1}^{\tilde{N}_{\text{s}}} z_i[\tilde{n}_s]$;
  $\hat{\sigma}^2 = \frac{1}{M} \sum\limits_{1}^{M} \left(\frac{1}{\tilde{N}_{\text{s}}-1} \sum\limits_{1}^{\tilde{N}_{\text{s}}} (z_{\hat{i}_{\min}}[\tilde{n}_s] - \hat{A}_{\hat{i}_{\min}})^2 \right)$;
end
stage 2: M-ary Test;
while $\bar{a}_{\text{target}} \overset{!}{\leq} \frac{\hat{A}_1}{\hat{A}_{\max}} P_{\text{s}}(N_{\text{s}}) + \sum_{i \in \{2 \dots M_{\text{new}}\}} \frac{\hat{A}_i}{\hat{A}_{\max}} P_{\text{e},i}(N_{\text{s}})$ **do**
  $M_{\text{new}} | i \in \{1, 2, ..., M_{\text{new}}\} \; \forall i, d_{1i} < d_{\text{th}} \; d_{\text{th}} = \frac{A_1 - (A_1 - \bar{\sigma})}{\bar{\sigma}} = 2$;
  Determine the number of samples N_{s}, which are necessary to achieve the performance based on the estimated parameters
  $\bar{a}_{\text{target}} \overset{!}{\leq} P_{\text{s}}(\hat{A}_i, N_{\text{s}}, \hat{\sigma}) + \sum_{i \in \{2 \dots M_{\text{new}}\}} \frac{\hat{A}_i}{\hat{A}_{\max}} P_{\text{e},i}(\hat{A}_i, N_{\text{s}}, \hat{\sigma})$;
  Transmit second training signal of length;
  $N_{\text{s},2} = \min(N_{\text{s}} - \tilde{N}_{\text{s}}, 0)$;
  Update Estimates: $\hat{A}_i = \frac{1}{N_{\text{s}}} \sum\limits_{1}^{N_{\text{s}}} z_i[\tilde{n}_s]$;
  $\hat{\sigma}^2 = \frac{1}{M} \sum\limits_{1}^{M} \left(\frac{1}{\tilde{N}_{\text{s}}-1} \sum\limits_{1}^{\tilde{N}_{\text{s}}} (z_{\hat{i}_{\min}}[\tilde{n}_s] - \hat{A}_{\hat{i}_{\min}})^2 \right)$;
  check condition;
end

Algorithm 2: Extended composite M-ary Test

can be calculated as before in (3.18) and the second step will be repeated with the cost of one additional feedback from the receiver to the transmitter. As we can expect a large number of amplitudes to be close to zero due to the sparse characteristics of a mmW channel, we additionally decrease the number of candidates. Therefore we only consider the coupling coefficients in the next stage, which are in $2\bar{\sigma}$ range of the maximum estimate of the maximum coupling coefficient for the next iteration of the hypothesis test. Thus N_{s} is the same for every candidate beam pair, but the number of candidate beam pairs M is decreased in every iteration. That means that we just have to consider the signals which are not yet clearly distinguishable because of the noise process. The repetition of the stage continues until the expected performance $\hat{\bar{a}}$ equals the desired performance $\bar{a}_{\text{target}}$. The extension of the test is given in Algorithm 2.

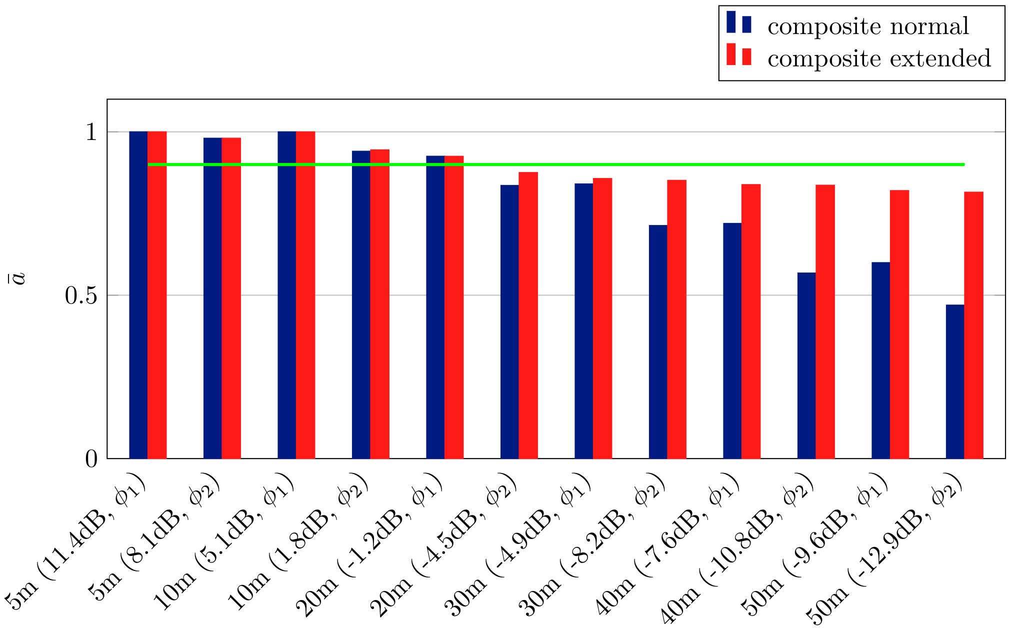


Figure 3.3: Example of M-ary hypothesis test for $M = 32$. For ϕ_1, AoD,AoA=0°, for ϕ_2, AoD=60° and AoA=90°. Performance in terms of mean normalized received signal strength $\bar{a}$ after training for both discussed algorithms. The extended test is almost reaching the targeted performance of $\bar{a}_{\text{target}} = 0.9$ for every SNR value. (Source: [KRF19] © 2019 IEEE).

3.5 Application

We illustrate the algorithm by tackling the initial problem of detecting the best beam pair combination of a transmitter and a receiver in a mmW scenario (Fig. 3.1). We consider a small-cell scenario where one access point (AP) has to serve one UD in the radius of 50 m. We assume the AP and UD have ULAs of $N_{\text{tx}} = 8$ and $N_{\text{rx}} = 4$ antenna elements, respectively, with a codebook of the same size, resulting in $M = 32$ coupling coefficients. In order to get a good initial assumption of the SNR, we calculate the maximum and minimum of the SNR which can occur, given an assumed transmit power of $P_t = 35$ dBm. We use the path-loss formula given in [3GP17] for $f_c = 60$ GHz. Supposing a bandwidth of 1 GHz, minimum and maximum SNR for the system for a distance range between 5m and 50m considering the antenna gain and possible misalignment loss of $P_{\text{mis}} = 6$ dB is given by $\text{SNR}_{\min} = -13$dB and $\text{SNR}_{\max} = 9$dB. We assume a $\text{SNR}_{\text{assumed}} = 0$ dB for the first stage and try to achieve a performance is $a_{\text{target}} = 0.9$, which results in an initial training length of $\tilde{N}_s = 13$.

We performed a simulation of the performance of the fixed length test for different distances between the transmitter and receiver. Also we investigated the case that both nodes where aligned in such a way that maximum of the beam patterns could be used (ϕ_1: AoD, AoA= 0°), and the case that the AoD and AoA do not allow to use the maximum of the best suitable beam shape (ϕ_2: AoD= 60°, AoA= 90°). The coupling coefficients were generated by using (3.1) with a simple LoS geometric channel.

In Fig. 3.3, 3.4 and 3.5 we show the outcome of the algorithm for different distances

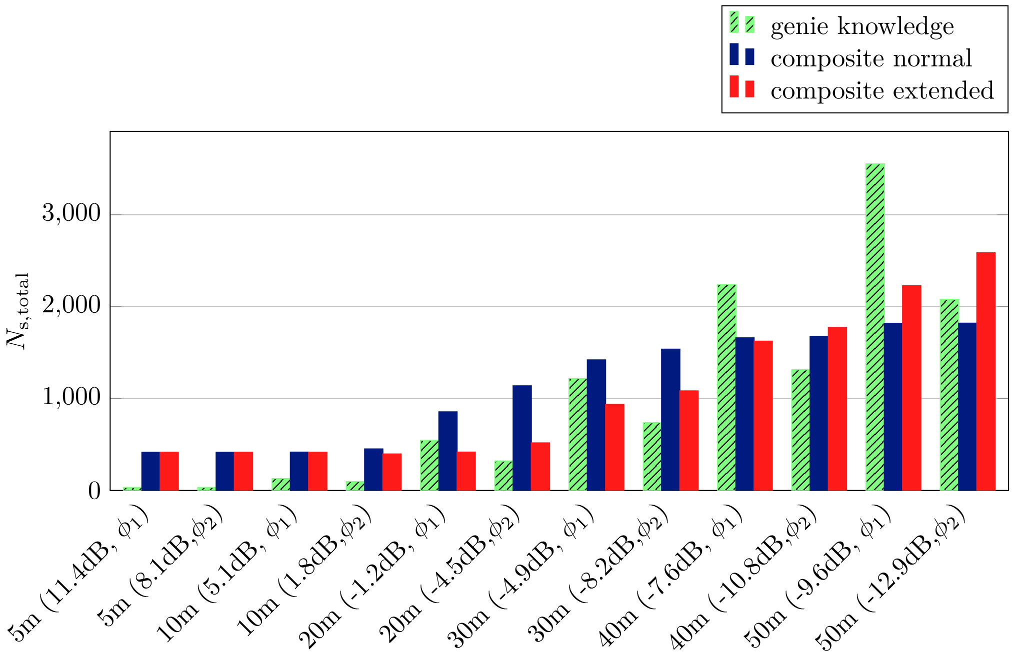


Figure 3.4: Example of M-ary hypothesis test for $M = 32$. For ϕ_1, AoD,AoA=0°, for ϕ_2, AoD=60° and AoA=90°. Total training length (total number of training samples transmitted) for different SNR values. The amount of samples needed for achieving the targeted performance with genie knowledge of the parameters A_i and σ is varying not only with the SNR but also with the mutual relation of the amplitudes A_i. (Source: [KRF19] © 2019 IEEE).

between transmitter and receiver and AoD and AoA combinations ϕ_1 and ϕ_2. The SNR corresponding to the largest coupling coefficient is also depicted on the abscissa. We performed Monte Carlo simulations using 1000 different realizations of the Gaussian noise process disturbing the training signal for each distance and AoD and AoA combination ϕ_1 and ϕ_2. In Fig. 3.3 we show that the targeted performance of $\alpha_{\mathrm{target}} = 0.9$ can be achieved for SNR values above the assumed SNR of 0 dB using the presented normal composite algorithm (blue) with two stages. Although the length of the training sequence is increasing with decreasing SNR value (Fig. 3.4) to a certain degree, the achieved normalized performance is successively decreasing towards lower SNR values. This is because the calculation of the total training length $N_{\mathrm{s,total}}$ is based on estimates of the parameters $\hat{A}_i$ and $\hat{\sigma}$ achieved using fixed initial training sequence length $\tilde{N}_{\mathrm{s}}$. In this way, the detection of amplitudes at low SNR is limited, due to a small training sequence length and thus limited performance of the estimation and detection. The proposed extended algorithm (red), however, essentially achieves the targeted performance by repeating the estimation procedure until estimates of sufficient quality are given. Still, it can happen that a coupling coefficient with the true maximum value is discarded from the set of candidates in the second stage of the algorithm, which explains the small deviation from the targeted performance. The cost of achieving the performance can be measured in terms of the total number of transmitted training samples, which is the sum of all samples send over all beam pair combinations, depicted in Fig. 3.4, and the number of feedbacks, which has

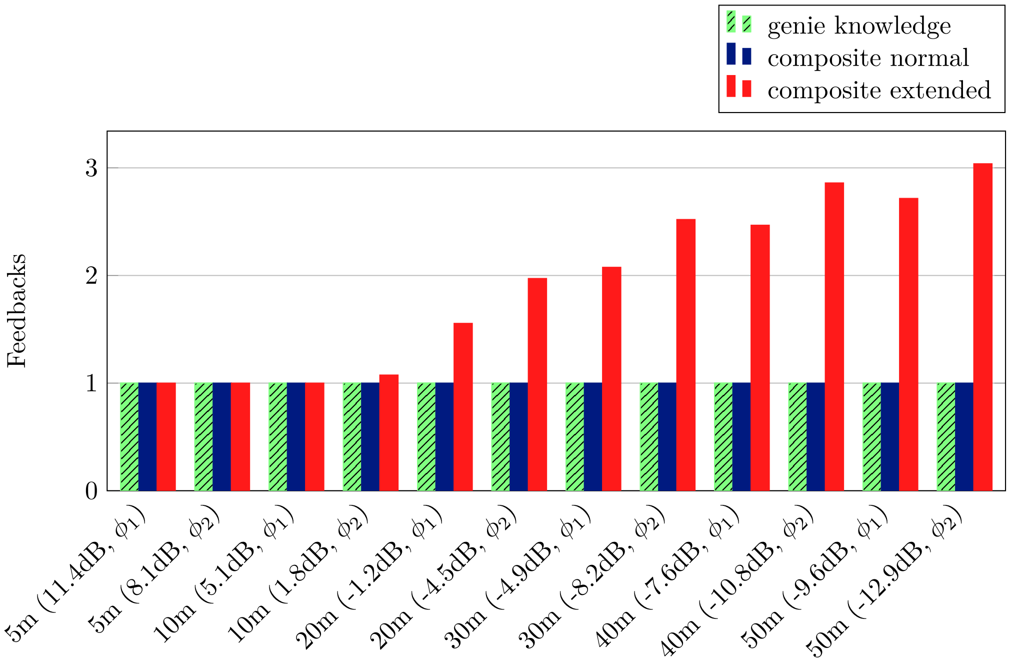


Figure 3.5: Number of required feedbacks for M-ary hypothesis test for $M = 32$. While the normal composite test uses always one feedback due to the two staged nature of the algorithm, the number of feedbacks for the extended algorithm is rising with decreasing SNR in order to increase the estimation quality of the parameters. (Source: [KRF19] © 2019 IEEE).

to be sent from the receiver back to the transmitter to resend a training sequence (Fig. 3.5).

We compare the total length of the training sequence of the composite test with the length of the sequence for a genie test assuming perfect knowledge of the amplitudes A_i and variance σ^2, which is designed to achieve the targeted performance. It can be seen that the length of the genie test is varying not only with the distance but also with the number of competing amplitudes. In general, less probing is needed in the case of angle combination ϕ_2, as this AoA and AoD combination results in larger competing coupling coefficients, which also contribute to the normalized amplitude in the error event. The needed length of the training sequence for the composite test is rising with decreasing SNR but stays constant at a certain small SNR level. That can be explained by the fact that the predefined estimation training signal length $\tilde{N}_s$ limits the estimation quality. In contrast to the genie test, the length of the composite test remains almost equal in the case of more competing amplitudes. The extended composite test needs a comparably large amount of training samples as needed by the standard composite test. However, Fig. 3.5 shows that up to 3 feedbacks are needed for the extended algorithm for the smallest SNR condition in contrast to one feedback for the standard algorithm. The choice of which composite algorithm to use is dependent on the application. While the proposed extended composite algorithm almost achieves the prespecified targeted performance, the additional feedbacks may cause latency issues in some applications as the signal has to be sent with high coding effort or a different, already established link.

3.6 Summary

In this chapter, we bring the beam selection problem to a mathematical context by proving that the beam selection problem can be formulated as an M-ary hypothesis test. Investigating the properties of the test, we gave some insight into the challenges of the beam selection problem. Assuming only knowledge of the expected SNR range at the receiver, we designed two composite tests using the GLRT approach and showed that the theory of the hypothesis test could be used to design the minimum needed length of the training signal for achieving the desired selection performance in a beam selection scenario.

Chapter 4

Millimeter Wave Testbed

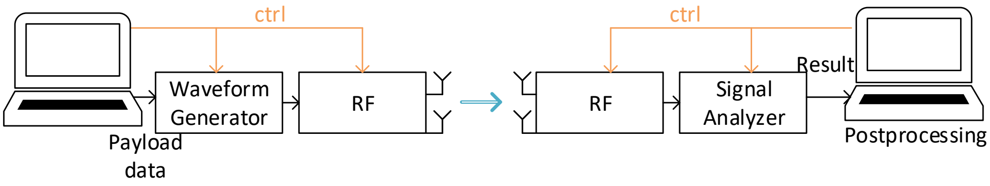
As presented in the last chapters, reasonable theoretical work has been done to enable mmW communications. This work has however been scarcely considered in practice for cellular systems. The standards IEEE 802.15c [IEEb] and IEEE 802.11.ad [IEEa] with improvement IEEE 802.11ay propose a wireless personal area network (WPAN) and wireless local area network (WLAN) standard, respectively. Nonetheless, not many devices are available yet. The WLAN approach also imposes different requirements in terms of reliability and mobility support than a cellular system. Most notably, a cellular system has to reliably provide a minimum quality of service (QoS), while WLAN systems are less strict in this domain. Currently, the use of mmW cellular communications is going to be standardized as 5G new radio (NR) by 3GPP [v1518b]. Nevertheless, the scarce public availability of performance results of mmW beamforming algorithms in a cellular context motivated the generation of a mmW testbed.

In this chapter, we will introduce the mmW real-time demonstrator platform, which was designed and created in the scope of the EU-funded project Beyond 2020 Heterogeneous Wireless Network with Millimeter Wave Small Cell Access and Backhauling (MiWaveS). The project contained 14 contributing partners, amongst them 2 (National Instruments (NI) and Technische Universität Dresden (TUD)) contributed to the platform, which will be presented here. The goal of this section is to give an understanding of the framework, used by the author for the implementation of the beam alignment and beam tracking algorithms, which will be introduced in detail in the later chapters. In order to provide transparency in the contribution of the author, we first give an introduction and background information on the structure of the demonstrator and summarize the achievements of the author in section 4.3. The different tasks for the parties were:

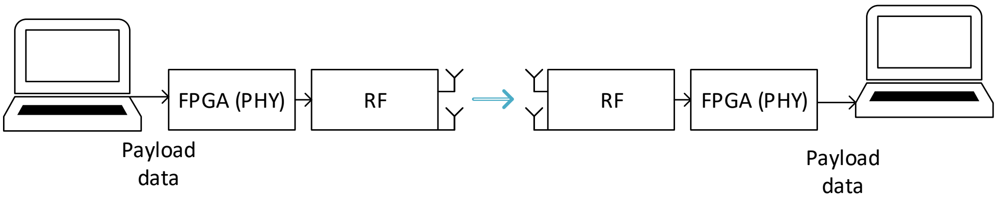
- NI: System and concept design of the demonstrator. physical layer (PHY) and media access control layer (MAC) design and implementation in FPGA and real-time computer.

- TUD (the work of the author): Implementation of beam alignment and beam tracking algorithms as part of the MAC real-time computer implementation.

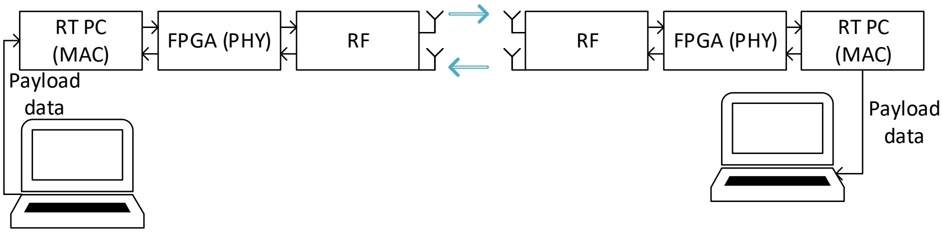
In order to understand the necessity of a testbed, it is important to distinguish different levels of demonstrations, which differ in terms of complexity and properties. We assume



(a) Single shot demonstrator.



(b) Single link real time demonstrator.



(c) Closed loop real time demonstrator.

Figure 4.1: Different types of demonstrators.

a single-link demonstrator with one transmitter and one receiver in the context of mmW communications. Corresponding visualizations are presented in Fig. 4.1.

Single Shot Demonstrator. A single-shot demonstrator allows to transmit and receive a predefined waveform over the air. Possible adjustments of the RF settings and the beams used for transmission and reception have to be defined beforehand and cannot be adjusted in real time. The preprocessing and all post-processing are computed offline and commonly performed using high precision floating point mathematical operations. The demonstrator is able to fundamentally test the ability of the communications of the channel, using the real RF hardware and antennas. Using high-quality analog equipment, also channel measurements are possible.

Single Link Real Time Demonstrator. If the creation of the transmit signal, the preprocessing and the postprocessing is done in real-time, the demonstrator is able to show continuous stream transmission, which is closer to real applications. As the post- and preprocessing is commonly done in fixed point arithmetics, it also shows that PHY and beamforming algorithms can be realized in practice with limited precision arithmetics.

That can be demonstrated, e.g., by a real-time video link transmission. However, the missing feedback sets limitations for the realization of precoding algorithms which need feedback and limits the scenario to be static.

Closed Loop Real Time Demonstrator. A demonstrator implementing a closed loop real-time transmission is most demanding, as not only the preprocessing and post-processing have to be realized in fixed-point arithmetic, but they also have to be fast enough to enable a closed loop processing. Additionally, some MAC processing is required to schedule resources between downlink (DL) and uplink (UL) phases, especially in time division duplex (TDD) systems. Thus the practical use of beamforming algorithms can be demonstrated including the effect of time-delayed channel feedback. Also, it is possible to demonstrate beam tracking algorithms in a non-static environment.

4.1 Prior Experimental Work

There are already experiments for mmW communications described in the literature. However, most of the results are obtained for a unidirectional non-real time prototype, targeting WLAN like applications, or are poorly described in terms of the underlying algorithms [IKS+15a, KLO+15, KLO+15]. Cellular systems have different requirements for QoS, the stability of the communication also in non line of sight (NLOS) and mobility scenarios compared to WLAN systems, which is rarely considered in practice yet.

Demonstrators for tracking experiments were used in the 28 GHz regime by [OOA+15, OIA+16], while demonstrators for the 70 GHz band for cellular type application were used in [IKS+15b, IKO+15, YIS+16, IYK+17] (closed loop single carrier (SC) communications, but one-sided search).

The unavailability of a real-time closed-loop demonstrator motivates the creation of such a system, where beam alignment and beam tracking algorithms can be implemented and tested under real-world constraints like RF impairments and the unavailability of instantaneous CSI. Due to the complexity of such a system and the unavailability of more advanced RF equipment, it was decided that the system should use an analog beamforming design.

4.2 General Structure of the Demonstrator

Each node of the demonstrator consists of two main parts: the baseband and the transceiver and antenna part. Additionally, an external PC is used to visualize the status of the system. However, the external PC does not influence communication. The structure of these parts can be summarized as:

- **Digital Baseband System**
  - Several field programmable gate arrays (FPGAs) process main parts of the received data of the ADC and prepare the modulation of the transmit signal

via the digital to analog converter (DAC). ADC and DAC are each directly connected to the FPGAs via a plug connection. All the modules are part of a chassis and can communicate via PCI extensions for instrumentation (PXI).

  - One real-time computer (NI) is used per node to perform the MAC layer functionality and some PHY layer control functionality. As it is also connected to the PXI bus, the communication between the real-time computer and the FPGAs can be done in real-time.

- **Transceiver and Antenna System**
  - A transceiver (up and down-converter), which converts the baseband signal to the passband for the transmission and vice versa for the reception path.
  - A steerable phased array consisting of physically separated 12 antenna elements for the transmission and 12 elements for the reception (see Fig. 4.2c).
  - An interface card, which converts the control commands of the baseband to a signal which is suited to drive the phased array and select the respective beam. It is also responsible for the transmission of the TDD switching signal.
  - A power supply for the interface card and the antenna.
- **External Computer System**
  - The external computer hosts the LabVIEW based user interface that monitors and adjusts MAC parameters, and is connected via Ethernet with the PXI chassis.

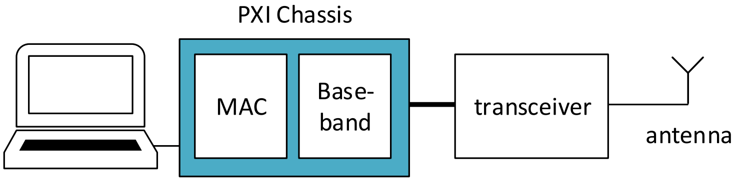
The main parts will explain in more detail in the following sections.

4.2.1 Digital Baseband System

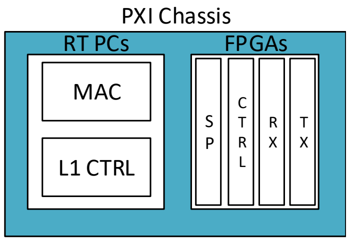
The baseband of the system was implemented using the programming language LabVIEW. In order to fulfill the real-time requirements, a large part of the PHY was implemented on FPGAs, while the MAC and a small portion of the PHY was implemented using a real-time computer. Both components, FPGAs, and real-time computer were combined in a chassis, which is equipped with a backplane enabling high-speed wired communications between the individual components.

NI PXI 1085 chassis. The chassis connects the real-time PC and FPGAs and clock modules (see Fig. 4.4). The chassis has 18 slots. The modules can communicate with a rate of 1.6 GB/s in both directions.

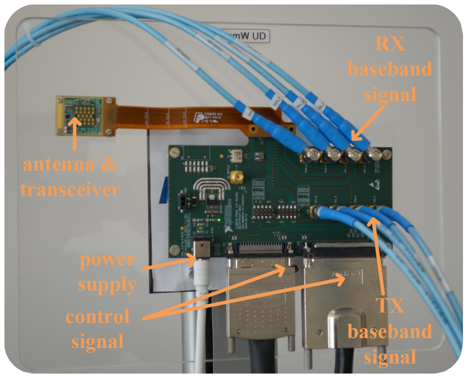
Signal Generator NI PXI 5652. The signal generator generates the sampling clocks for the ADC and DAC.



(a) Overview of the structure of one node.



(b) Baseband structure of the demonstrator.



(c) Transceiver and antenna of the demonstrator.

Figure 4.2: Structure of the demonstrator

ADC NI 5771. The ADC receives single-ended I/Q samples with a sampling rate of 1.5 GHz. The resolution is 8 bit. The ADC module is directly attached to a FPGA which performs the equalization and synchronization.

DAC AT 1212. The DAC generates samples at a rate of 1.24 GHz and provides a differential output. The DAC is directly attached to a FPGA, which is responsible for the modulation and pulse shaping.

FPGA NI 7975R. The NI 7975R card connects a Xilinx Kintex 7 FPGA to the PXI chassis. In this demonstrator, this device is used to perform heavy computational tasks like modulation, demodulation, coding, and decoding. In total, 7 of these (or FPGA cards with comparable performance) are used per node of the demonstrator.

I/O Module NI 6583. The module connects the baseband with the phased-array antenna control board using 35 single ended and 19 low voltage differential signaling

(LVDS) channels. The module can operate at a frequency up to 200 MHz, which makes it possible to transmit the information corresponding to the beam index close to real-time.

Meinberg GPS Clock. For each node of the demonstrator, we use a Meinberg global positioning system (GPS) clock which provides a high-precision 10 MHz reference signal to the PXI chassis and all modules. By using that clock, we mitigate the carrier frequency mismatch that results from deviations of the sampling clocks of the ADCs and DACs of different nodes. The remaining frequency offset can be covered by standard carrier frequency offset (CFO) correction algorithms for single carrier systems.

The demonstrator uses null cyclic prefix single carrier (NCP-SC) modulation, which has been regarded a promising candidate waveform for mmW cellular communications [GTC+14b, CGK+13] before the decision of the 3GPP consortium to use OFDM for NR [v1417a]. The advantage of single carrier modulation over OFDM in the context of mmW communications is especially the lower sensitivity to CFO and lower peak-to-average power ratio (PAPR), which enables the use of cheaper power amplifiers. Although just one power amplifier is used in the demonstrator, possibly every antenna element needs a power amplifier due to the high loss of the beamforming network or the phase shifters and the large power gains for larger transmission distances, which are needed for outdoor cellular ranges. One also has to keep in mind that, since much larger bandwidths than in traditional cellular systems may be used, the decrease of cost by using nonlinear amplifiers may be significant. Nonetheless, to the best of the author's knowledge, no specific cost analysis is available in the literature.

One advantage of using nulls in the guard interval instead of the cyclic prefix (CP) is the availability of short transmission breaks in each fast Fourier transform (FFT) block. These short periods could be used for switching the beams. However, as the implementation and hardware of the prototype system allow the steering in a timescale of several μ seconds, a time duration of 10 blocks is used to switch the beam, corresponding to 6.8μs. One single-carrier block consists of 512 samples, which is also the FFT size. A part of the block, namely 32 samples are taken as the guard interval. One hundred fifty blocks are grouped as a time division multiplex (TDM) slot, which has similar properties as the 3GPP long term evolution (LTE) transmission time interval (TTI). The beams can be changed in each TDM slot. 200 TDM slots yield a radio frame as largest entity which is processed by the MAC protocol.

Fig. 4.3 shows a simplified illustration of the data flow through the MAC and PHY protocol. At the AP the beam steering algorithm decides for which beams to probe next. The scheduler assigns the data slots and the modulation coding schemes (MCS) for the transmission in the uplink and the downlink. After Turbo encoding, modulation and pulse-shaping, the radio frame is transmitted. The UD receives the signal and processes it digitally at the FPGA, where also the channel synchronization and equalization takes place. After the Turbo decoder, the downlink channel is estimated for each beam-pair which were probed during the radio frame and stored in the beam steering database. These measurements will be used for the beam steering later on. After the control messages are

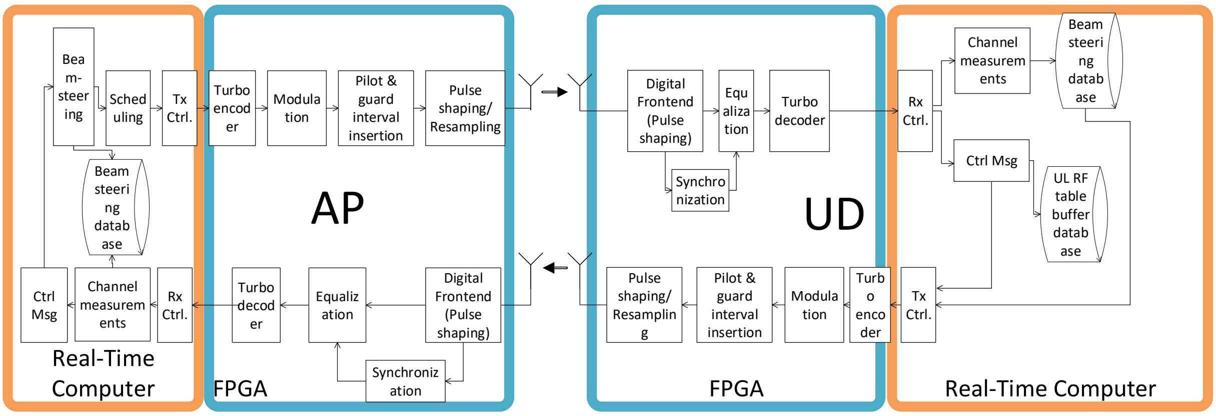


Figure 4.3: Simplified presentation of the PHY and MAC processing of the AP and UD.

decoded, the next uplink radio frame is generated, which undergoes the same procedure of turbo encoding and modulation at the PHY implementation of the FPGA of the UD. After reception of the uplink signal and processing of the PHY at the receiver of the AP, the channel of each pilot TDM slot (beam pair) will be measured and stored in the beam steering database. Note that the previously taken downlink measurement will be transmitted back to the AP by a robust control channel using one-sided beamforming and stored in the database. The beam steering algorithm can access this data in the next step. The round-trip delay amounts to three radio frames. The beam steering algorithm has to decide the beams for three radio frames in advance.

4.2.2 Transceiver and Antenna

In the demonstrator, an off-the-shelf integrated circuit (IC) of Sibeams (SIL6340) was used, which contains the up- and down-conversion, phase shifters, twelve separate antenna elements for the transmitter and receiver, respectively. The transceiver IC was integrated on a circuit board, which also contains the reference clock. The usage of independent integrated clocks for the AP and UD node placed special demands on the implemented synchronization algorithm. The transceiver board is connected with a flex cable to an interface board, which provides power supply and control and signal interface via standard SubMiniature version A (SMA) connectors. As the output of the ADC is single-ended and the input of the interface board needs a differential signal input, this has to be adjusted using Baluns, and the level is adjusted using low noise amplifiers. The twelve antenna elements are placed in two horizontal rows with five elements each and a third row with two elements. Therefore, five elements form the beam characteristics for the azimuthal beamwidth while the three rows define the beamwidth in the elevation. The transceiver is used in TDD mode, which means that either the transmitting or the receiving part is active. The switching between the transmit and receive chain takes about 2 μs. Each phase shifter has a resolution of 2 bits. To set up the phase shifters, after reception of the control signal, 100 ns are needed. Additionally to the usage of all antenna elements, also only one element can be used to provide a beam which covers about 60° of the azimuthal range. The transceiver is able to operate to at all four IEEE 802.11ad bands, while we

Table 4.1: System parameters

Parameters	**Values**
Sampling rate DAC	1250 MHz
Sampling rate ADC	1500 MHz
Bandwidth/ Symbol rate	750 MHz
Block length	512 symbols
Block duration	682.7 ns
Guard length per block	21 symbols
Guard duration per block	42.7 ns
Slot length	150 blocks
Slot duration	102.4 μs
Guard length per slot	10 blocks
Pilot length per slot	1 blocks
Data length per slot	138 blocks
Radio frame length	200 slots
Radio frame duration	20.48 ms
Duplex	TDD
Modulation Scheme	Single Carrier
Carrier Frequency	62.64 GHz
Azimuthal Antenna Elements (TX & RX)	5
Antenna Codebook Size	25
Antenna Scan Range	$\pm$ 60°
HPBW	36 °

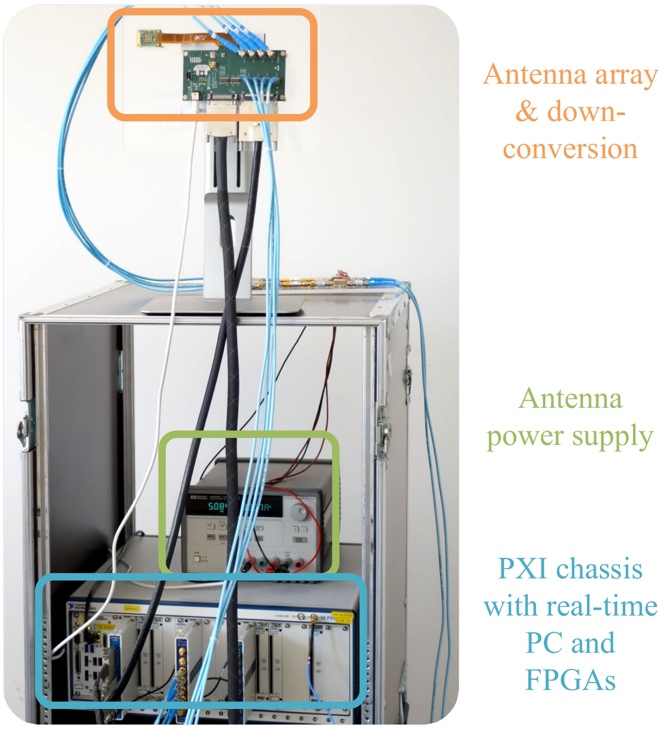


Figure 4.4: One node of the demonstrator. Eight cables re needed to connect the differential I and Q paths of the transmitter and receiver.

use the 62.64 GHz carrier frequency here.

4.3 Author's Contribution

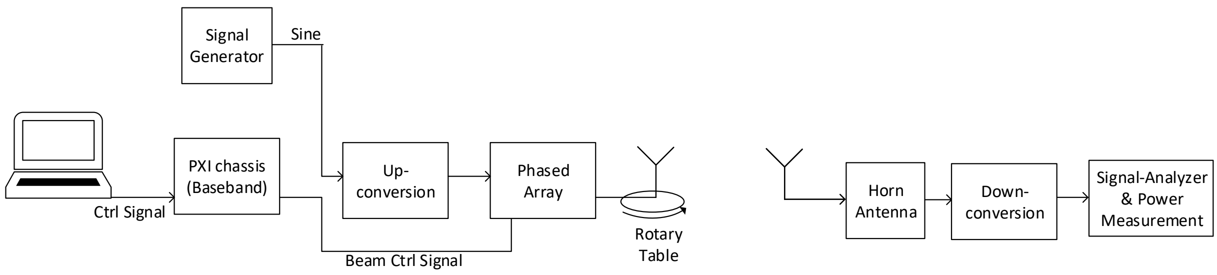
The author's contribution to the demonstrator was to implement the beam steering algorithms for static (beam alignment) and dynamic (beam tracking) scenarios.

4.3.1 Antenna Measurements

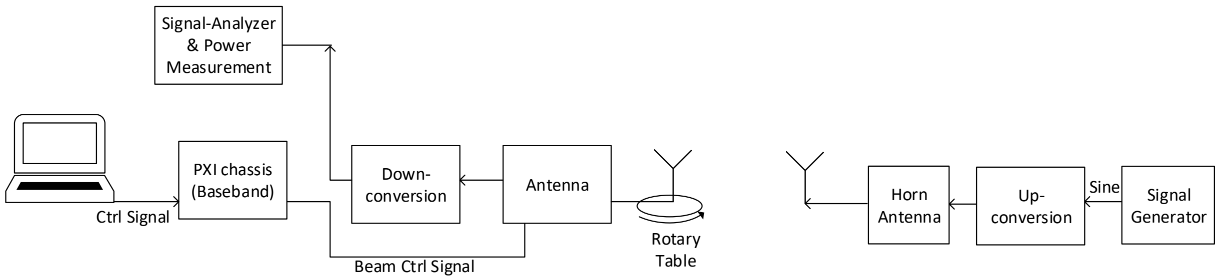
In order to design the beam algorithms and to define how robust they have to be in terms from imperfections with regard to the beam patterns, we performed a power measurement in an anechoic chamber. For this test, we used a predefined beam pattern codebook of 25 azimuthal beams, whose main lobes are roughly equidistantly distributed over the angular range of $-60°$ to $60°$ in $5°$ steps. Additionally, to the directional beams, the beam-shape of a single antenna element was measured. It served later on as fall-back candidate for control message transmission as a so-called "wide beam". For the measurements, we used a reference horn antenna of 25 dBi gain as transmit antenna (for Rx measurements) and receive antenna (for Tx measurements) with a small beamwidth. We measured the patterns using a single sine signal at 500 MHz offset to the center frequency of 62.64 GHz. The distance between the reference horn antenna and the phased array antenna was chosen in such a way that both antennas operated in the far field. With a maximal aperture size less than 3 cm, the far field distance amounts to $d_{\text{farfield,min}} = \frac{2D^2}{\lambda} \approx 0.4m$ [Ban99], so that the distance of 1m between phased array and horn antenna is sufficient for both antennas to receive the electromagnetic waves in the far-field.

The setup for measuring the transmit beam patterns is depicted in Fig. 4.5a. In order to have full control to the input signal, we generated the sine with a signal generator and inserted it directly to the input of the phased array transceiver board, while the antenna board was controlled by the baseband. The reference horn antenna received the transmitted signal and a frequency analyzer stored the received power for each frequency bin inside the targeted bandwidth with a resolution of 30 kHz. Subsequently, the frequency value point having the maximum power was chosen as the received value in order to compensate for the effect of the carrier frequency offset. This measurement was repeated by rotating the phased array in steps of $2°$ azimuthal angle from $-85°$ to $85°$ for one beam. This procedure was reiterated for each of the 25 transmit beams and the wide beam.

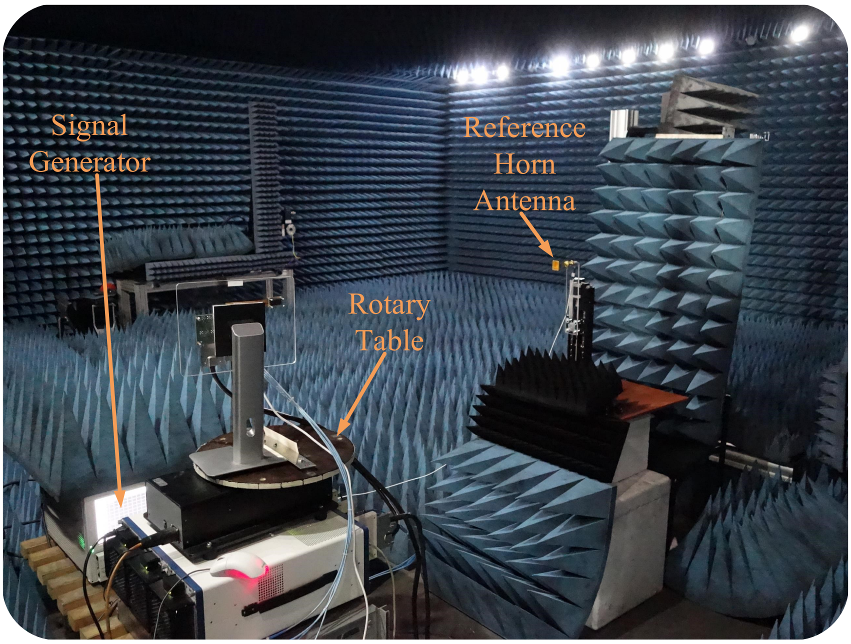
The setup for the receive beam measurements is sketched in Fig. 4.5b. The transmit signal was fed into the reference horn antenna, and the phased array received the signal, subsequently post-processed by the signal analyzer in a similar way as described before. Also, the antenna was rotated in the same way as before, and the measurement was repeated for all elements of the receive beam codebook. Fig 4.5c shows a picture of the setup in the anechoic chamber.



(a) Transmission beam characterization setup.



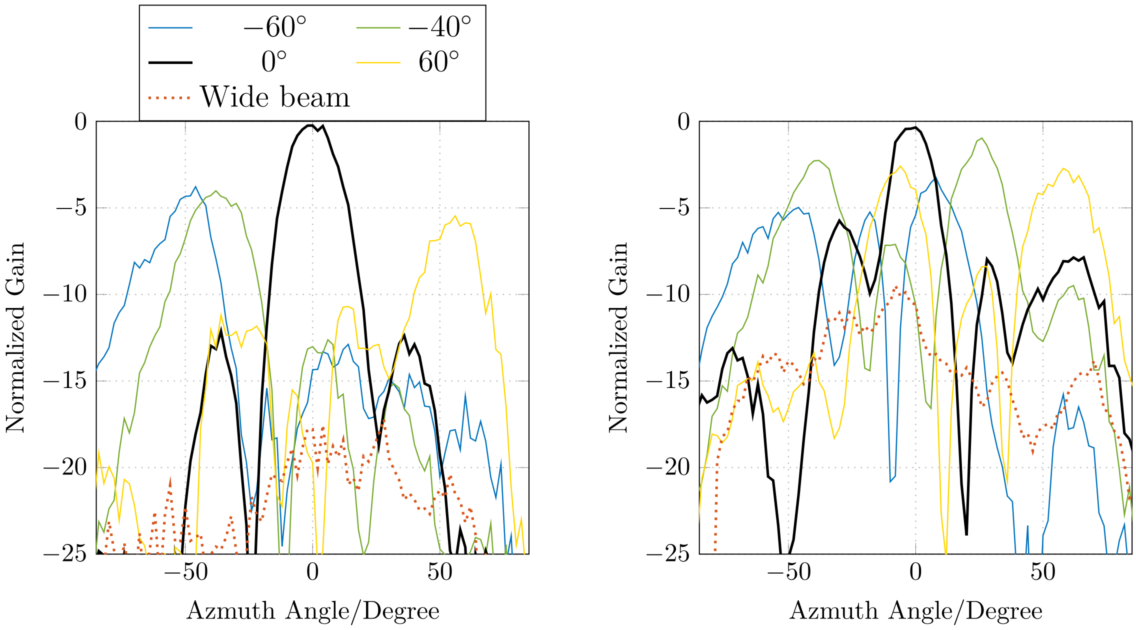
(b) Receiver beam characterization setup.



(c) Picture of antenna measurement setup in anechoic chamber.

Figure 4.5: Characterization of phased array.

Evaluation. The results of the measurements, exemplarily shown for some selected beam patterns in Fig. 4.6a and Fig. 4.6b, reveal that the transmit and receive patterns are significantly different. While the transmit beams have a side-lobe-level by about -12 dB to -9 dB, the receive beams have a side-lobe level by about -6 dB to -3 dB compared to the main lobe. The measurements also reveal that the gain of the main lobe decreases with the steering angle by -5 dB for transmitter and receiver. However, the main-lobes steering direction has a small error towards the targeted direction. The relation of the large beamwidth of about 30° and the small difference of the main lobe directions between



(a) Transmit beams measurements. **(b)** Receive beams measurements.

Figure 4.6: Antenna measurements for several nominal steering angles.

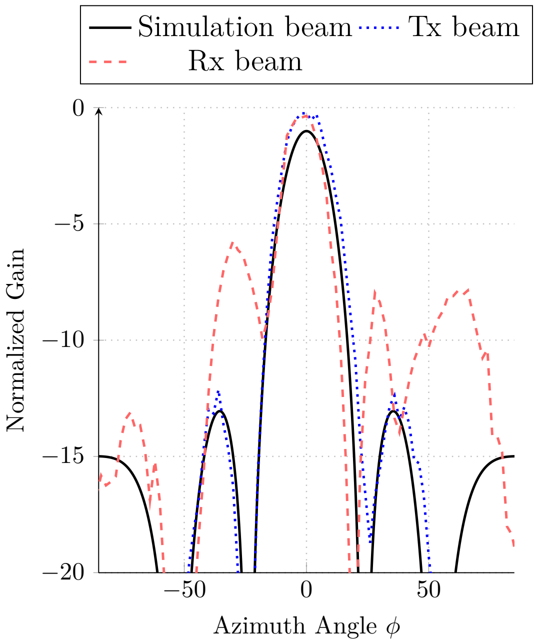
adjacent beams of 5° lead to a large overlapping between adjacent beams. The measurements will help us to interpret the results of the measurement campaign of the beam steering algorithms. Additionally, we compared the measured beams with beams whose patterns are derived by theoretical investigations assuming an ULA of 5 antenna elements with distance $d_s = \frac{\lambda}{2}$. A visual comparison is depicted in Fig. 4.7. It can be seen that the transmit beams are relatively close to the theoretical assumption, while the receiver beams suffer significantly more from large side lobes. It is also observed that the match between theoretical and measured patterns decreases for large steering angles w.r.t. the array normal (e.g. steering to 60° in Fig. 4.7b). The side lobes of the receive patterns for the maximum steering angle have nearly the same gain as the main lobe.

4.3.2 Implementation Strategy

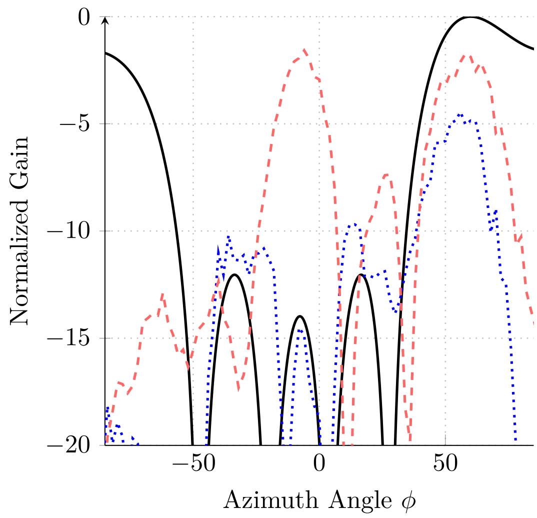
The implementation of beam steering algorithms in a complete system brings special challenges due to the limitation of processing power and time and the bad debugging capabilities which result from the need for real-time processing and closed-loop operation. Therefore, we pursued several steps for the implementation process (see Fig. 4.8 for reference):

1. *Non-real-time Matlab implementation of the algorithm using a simulated channel.* After designing the algorithms, we implemented it to an easy-to-debug Matlab environment. Therefore we used a simulated channel according to the results of Chapter 1. The implementation is done in floating point numbers and non-closed loop.

2. *Implementation in a LabVIEW framework (fixed point) using defined interfaces and simulated channel (beam steering integration environment).* As a second step, we



(a) Beam 12 with nominal steering angle of 0°.



(b) Beam 24 with nominal steering angle of 60°.

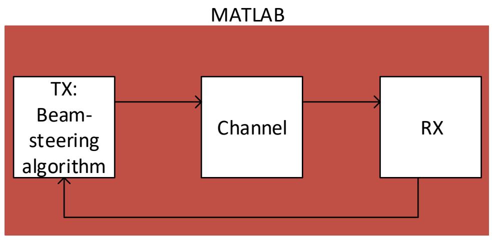
Figure 4.7: Comparison of simulated and measured beams.

implemented the algorithms into a simulation of the MAC protocol of the system in LabVIEW. The measurements were given in fixed-point values, and the interfaces to the protocol were the same as in the real system. We used a simulated channel in LabVIEW and just considered a downlink connection with the assumption of perfect feedback. This step allows us to test the actual implementation for the real-time computer in a non-real-time environment. The beam steering integration environment executes the beam steering protocol related functionality of the MAC layer but abstracts layer one and its control in a channel model. This implementation was tested using simulated channel realizations. This first step allows to debug and test the beam steering protocol. In addition, it allows to test and optimize the execution time of the beam steering algorithm by running the beam steering integration environment on the real-time controller.

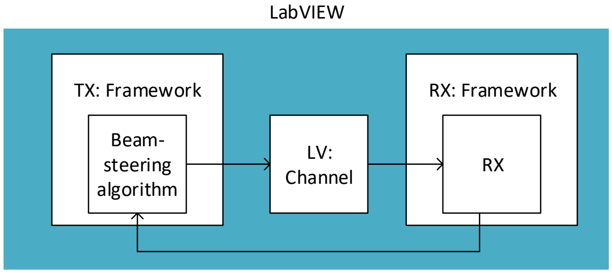
3. *Real-time implementation using a LabVIEW framework, using FPGA PHY implementation and a real-world channel.* The LabVIEW implementation was moved to the real-time system and tested.

4.3.3 Beam-Steering Wrapper

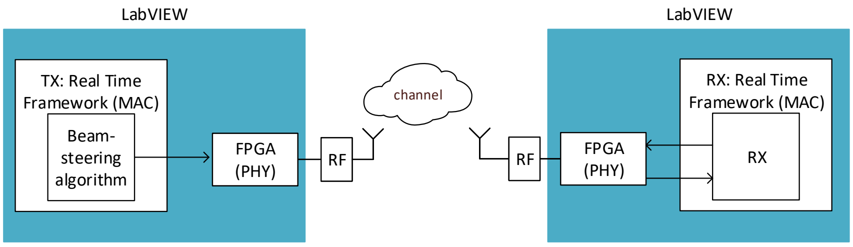
An interface to connect the beam steering algorithms to the rest of the MAC layer implementation was defined in joined work with NI. Due to the high noise figure at the transmitter, we were only able to use power measurements of the channel (instead of SNR measurements). In order to implement the actual algorithm, we designed a wrapper func-



(a) Phase 1: Non real-time Matlab implementation of the algorithm using simulated channel.



(b) Phase 2: Implementation in LabVIEW framework (fixed point) using defined interfaces and simulated channel.



(c) Phase 3: Real-time implementation using LabVIEW framework and using FPGA PHY implementation and real-world channel.

Figure 4.8: Design process of beamforming algorithms for the demonstrator.

tion to encapsulate the actual beam alignment and beam tracking algorithms, to translate and store given inputs for use in the algorithm and to convert the outputs to match the predefined structures and take care about feedback generation. Thus, the beam alignment and beam tracking algorithms are separated from administrative tasks and have to perform the mathematical operation only. The wrapper defines the consecutive steps the algorithm has to accomplish in case of anticipated or unexpected measurement inputs. Additionally, the wrapper specifies the order of the algorithms and the completion of the probing if the alignment is successfully concluded. Due to the above-mentioned difference in the transmit and receive patterns, we were not able to assume channel reciprocity. For this reason, we performed the beam steering algorithm independently for the DL and the UL. As mentioned before, a bidirectional control channel is needed as the UD is not allowed to perform any decisions. It is not ensured that a connection between AP and UD (and vice versa) can be established throughout a radio frame in training mode while the beam steering is active. Thus the control data is transmitted using directional beams at the transmitter and wide beams at the receiver. This control link can be easily established in one radio frame by a one-sided exhaustive search using fewer beam combinations. Thus the control message uses heavy coding to compensate for the larger path-loss. The usage of the wide-beams can be avoided by allowing beam steering processing in the UD or using a secondary link via a different carrier frequency, as it is proposed in NR for the non-standalone case [v1518a].

The structure of the wrapper is sketched in Fig. 4.9. The idea and structure of the LabVIEW implementation for the real-time computer will be presented briefly. At first, the interface is transformed in order to meet the requirements and preferred data structure of the algorithm (A). Then another wrapper for downlink and uplink probing and measurement processing is called in parallel. Note that the downlink and the uplink wrappers are actually the same LabVIEW program (called virtual instrument (VI) in LabVIEW), which can be used in parallel. After the UL and DL wrappers are processed, the Output (e.g., the beam schedule) is again transformed to meet the requirements of the interface. In parallel, the internal memory cluster is written to a feedback node.

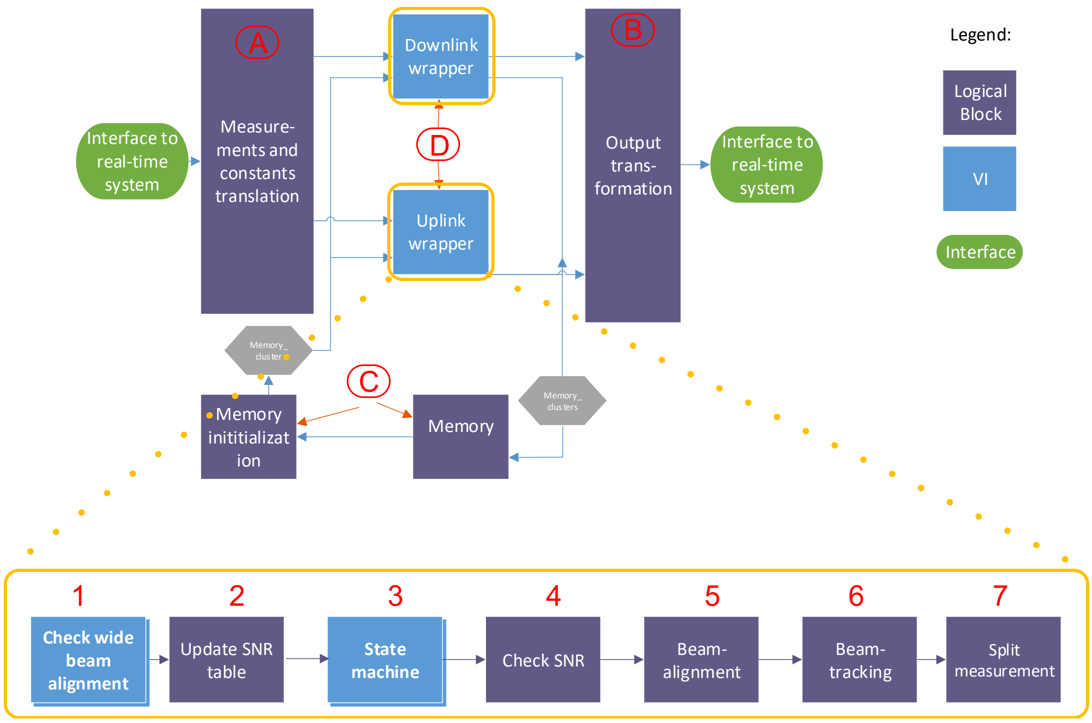


Figure 4.9: Beam-steering wrapper.

UL/DL Wrapper. The UL/DL wrapper is the actual core module, responsible for the beam steering. It interprets measurements, generates schedules for the beam alignment and beam tracking algorithms. It processes the stored parameters and channel measurements and outputs the schedule for the correct radio frame.

1. *Check wide beam alignment.* The program checks the availability of a working control channel and reports it to the next blocks. The scheduling is done by the split-measurement block. Initially, a one-sided full search is applied for DL and UL, and after initialization, the best narrow TX beam and the neighboring ones are probed every radio frame, to increase stability in the connection (Fig. 4.10).

2. *Update SNR table.* This block updates the internal SNR table measurements.

3. *State Machine.* The task of this VI is to decide on the correct state of the algorithm, which has to be processed depending on the history and available measurements. The main purpose is to monitor which are implemented using several consecutive stages (e.g., the gradient-based beam alignment algorithm, see Section 5.6).

4. *Check SNR.* The check SNR code block's purpose is to check whether the required SNR or received power values are successfully received by the AP in order to proceed with the next step of the algorithm. Therefore, the schedule history buffer is used to obtain information about the probed beam pair indices. If not all values are received, a schedule list with missing measurements will be given for rescheduling. The measurement check will also filter non-valid beam indexes.

5. *Beam alignment.* The beam alignment code block is the heart of the wrapper and will execute one of the actual alignment algorithms. The algorithm VI has to be called several times in order to fulfill its task (e.g., a schedule has to be created, and the measurements have to be processed in the simplest case). All relevant status information is given with input and output variables, which are stored in the memory. Detailed information about the beam alignment implementation is given in Sec. 5.6.

6. *Beam tracking.* The beam tracking code block is responsible for applying the tracking algorithms similar as in the beam alignment code block. The tracking will exclusively be executed if the alignment is finished. More details are given in Sec. 6.2.

7. *Split measurements.* The block splits the beam pair indices which are determined to be probed into a part which is probed in the next radio frame and a buffer. Also, the code includes beams for a continuous check of the wide beam connection and the current target beam to the schedule list in order to increase the robustness of the connection. Thus, even if the current beam probing schedule does not allow a stable connection, the current target beam and the wide beam connection will provide sufficient robustness of communications.

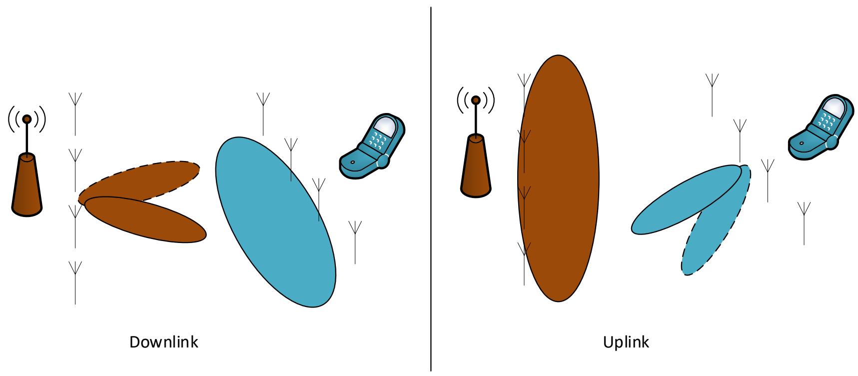


Figure 4.10: Wide beam search scheme.

4.4 Summary

In this chapter, we introduced the mmW real-time demonstrator. We presented the structure of the demonstrator and sketched the goals of the implementation. The demonstrator is able to demonstrate a cellular communication in the mmW band using directional, electrically steerable antennas, a full real-time PHY implementation mainly based on FPGAs and a real-time MAC implementation in a real-time computer. The system supports variable MCS modulation and thus data rates from 147 Mbit/s to 2318 Mbit/s. It uses bidirectional TDD closed loop communication. In the second part of the chapter, we introduced the contribution of the author on the demonstrator precisely. Measurements of the steerable antennas in the anechoic chamber showed that the patterns of the antennas are subject to impairments. Significant deviations compared to theoretical ideal patterns have to be considered in the case of the receive beam patterns. The system was used for demonstrating beam alignment and beam tracking algorithms.

Chapter 5

Beam Alignment for Millimeter Wave Low-Cost Systems

In Chapter 2, we showed that the goal of precoding and combining for a general MIMO system is to increase the achievable data rate and that this problem is equivalent to a search for a beam combination which receives the highest power in the case of an ABF system. In the previous chapter, we showed how a practical system for mmW communications could be designed. Due to implementation impairments and the high cost of high-frequency hardware, it is likely that commercial systems for mmW communications also will use the ABF system with codebook based antenna arrays.

One fundamental question is how to apply a reliable training procedure to find the beam pair that will give the best receive power most efficiently, i.e., using a short time period. This problem is commonly referred to as to beam alignment. Using the theory of the multiple hypothesis tests, we already showed in Chapter 3 how to design the length of the training signal in the case of perfect CSI and proposed a heuristic method to find the length of the signal in the case of no channel knowledge. However, up until now, we assumed to train all possible beam combinations at least once. For outdoor cellular systems, the required antenna gain is high. Consequently, the beamwidth has to be small. Therefore, in order to enable communications in every direction, a large number of beams has to be considered in the codebooks and thus in the beam alignment procedure. The training time increases quadratically as the cardinality of the antennas on both sides of the link increase. Thus the question arises, whether all beam pairs have to be probed, or if it is sufficient to probe just a subset of all possibilities.

In this chapter, we develop two algorithms to solve this problem. After theoretical analysis, we implement one of these beam alignment algorithms in the testbed and demonstrate the effectiveness of a suitable algorithm experimentally.

5.1 Beam Alignment Problem

As before we consider the system with beam switching ULA antennas at the transmitter and at the receiver with codebooks $\tilde{\mathbf{F}}_{\mathrm{rf}}$ and $\tilde{\mathbf{W}}_{\mathrm{rf}}$, respectively, where the steering vectors are assumed to be sorted according to their main lobe direction, and use a SC-FDE. We omit the subscript from now on for better readability. A scheme of the system is shown in Fig. 3.1. Therefore we can use the narrow-band channel model described in Sec. 2.2.

Let us repeat the receive signal of the ABF (2.74) for further use

$$y[n_\mathrm{s}] = \mathbf{w}_l^H \mathbf{H} \mathbf{f}_k s[n_\mathrm{s}] + \mathbf{w}_l^H \mathbf{n}[n_\mathrm{s}]. \tag{5.1}$$

We want to maximize the receive power, thus we give an optimization problem for a given sample time n_s

$$\left(\hat{k}, \hat{l}\right) = \arg \max_{\substack{\mathbf{f}_k \in \tilde{\mathbf{F}} \\ \mathbf{w}_l \in \tilde{\mathbf{W}}}} \underbrace{\left|y(k,l)\right|^2}_{J(k,l)} = \arg \max_{\substack{\mathbf{f}_k(\phi_k^{\mathrm{tx}}) \in \tilde{\mathbf{F}} \\ \mathbf{w}_l(\phi_l^{\mathrm{rx}}) \in \tilde{\mathbf{W}}}} \underbrace{\left|y(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})\right|^2}_{J(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})}. \tag{5.2}$$

As stated in Section 2.5, the solutions in literature for solving this problem for ABF systems with a fixed codebook are beam sweeping techniques that measure beam combination to cover all pointing or steering directions to obtain complete implicit channel knowledge. Besides the most simple exhaustive search, the alignment can also be implemented in a hierarchical manner [Wan09], leveraging wide beams in the beginning. This method was adopted in standards such as IEEE 805.15.ac [IEEb] and IEEE 802.11ad [IEEa] for static indoor communication. As the use of wider beams requires larger spreading factors (length of the training sequences), the necessary search time will still be large compared to other approaches.

We can interpret the receive power $J(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})$ as a function of two variables which, given a channel realization $\mathbf{H}$, is dependent on the steering of the main lobe of the transmitter beam ϕ_k^{tx} and the steering of the main lobe of the receiver beam ϕ_l^{rx}. As we assume to use codebooks with limited numbers of possible steering angles for both the transmitter and the receiver side, the arguments of this function are discrete. However, the function can be understood as a sampled version of a continuous function $J_{\mathrm{cont}}(\phi^{\mathrm{tx}}, \phi^{\mathrm{rx}})$ with continuous steering angles using the same definition for J_{cont} as for J, but letting $K \to \infty$ and $L \to \infty$. Unfortunately, we cannot assume this function to be strictly convex or concave for two reasons. First, several MPCs will result in several maxima of this function. Second, even for one MPC, the shape of the angular beam pattern is non-convex. Thus each combination of beam patterns will also be non-convex. The physical interpretation of this fact becomes evident by considering a transmit main lobe steered towards a receiver side-lobe, which will give a maximum in the power function. For this reason, traditional convex optimization techniques cannot be directly applied to this problem.

In the following, we evaluate two approaches to cope with this problem.

The first approach is based on "black-box-algorithms", a term describing mathematical algorithms which solve optimization problems by considering no apriori information about the function to be maximized or minimized. The particular feature of some of these algorithms is the ability to find the global extremum of non-convex functions due to the heuristic approach.

The second approach is to explore the function coarsely to be sure to search in the region around the global maximum and then to use obtained gradient information to approach the global maximum of $J(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})$.

Both algorithms work adaptively, and thus will require feedback for control of the beams from receiver to transmitter. However, data can only be transmitted *after* the alignment is completed. In this work, we assume a secondary link, which will forward control information even if the alignment phase is not finished. While this can be a lower frequency data link for commercial scenarios, we use a wide-beam connection in the testbed for this purpose.

In the following section, we will start with the first approach and evaluate its performance.

5.2 Black Box Based Algorithm

By using the SNR as performance indicator and the number of beam indices as degrees of freedom, we can convert the problem into a discrete optimization problem with incomplete sampling. Because the channel characteristic is not known in advance, we lack of a specific mathematical description of the optimization function, thus we utilize so called "black-box optimization" algorithms for a solution. Here the receive signal for given sample n_s is a function solely of the steering vectors

$$y(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}}) = \underbrace{\mathbf{w}^H(\phi_l^{\mathrm{rx}})\mathbf{H}\mathbf{f}(\phi_k^{\mathrm{tx}})s + \mathbf{w}^H(\phi_l^{\mathrm{rx}})\mathbf{n}}_{\text{black box}} \quad . \tag{5.3}$$

We convert the maximization problem into a minimization problem by changing the sign of the expression in order to use existing convex optimization algorithms

$$(\phi_{\mathrm{min}}^{\mathrm{tx}}, \phi_{\mathrm{min}}^{\mathrm{rx}}) = \arg\min_{\phi_l^{\mathrm{rx}}, \phi_k^{\mathrm{rx}}} \underbrace{(c_0 - |y(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})|^2)}_{\text{cost function } J_{\mathrm{min}}(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})} \tag{5.4}$$

$$c_0 = \max_{k,l} |y(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})|^2. \tag{5.5}$$

Employing the cost function $J_{\mathrm{min}}(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})$ allows us to take advantage of existing black box optimization algorithms.

Black box algorithms are solving an optimization problem for not analytically known functions by only using observations given by known excitations of the system. This problem is exhaustively covered in mathematical literature [Sch11, Hem10]. According to this literature, subdividing optimization problems into categories will help to find a suitable algorithm for solving the given task. The first distinction is whether derivative information can be obtained and if it is reliable or not. The second categorization applies to the fact, whether it is "cheap" or "expensive" to obtain a single observation/evaluation of the cost function. In our case, the derivative information can be numerically approximated but is usually unreliable [Hem10, p. 12]. Also it is clear that one evaluation of the cost function is expensive, as it implies a change of the current beam indices.

For the first reason, we focus on derivative-free black box algorithms, which can be coarsely subdivided into three categories.

Meta-heuristic search algorithms use a stochastic approach in order to generate sets of new candidate points for observation from already evaluated candidates [Hem10][p.13].

Most well-known black-box optimization techniques as the so-called genetic algorithms and evolutionary algorithms (for example from Goldberg and Holland for the first group and Beyer and Schwefel for the latter) are part of this category. However, they are not suited to address our problem, as they are designed to optimize functions with "cheap observations", they usually need hundreds of evaluations in order to discover an optimum. Additionally, most of them assume a continuous variable space.

The second category of derivative-free black box algorithms, called deterministic sampling, makes use of deterministic approaches to find new candidates of points to evaluate. A prominent representative of this group is the DIRECT algorithm [Hem10][p. 17]. As before, as these algorithms are designed to solve functions with cheap observations, they seem not suitable for our problem.

In contrast to the previous two approaches, metamodel based black box algorithms are designed for expensive cost functions. Most of them use surrogate models designed from previous iterations in order to obtain new candidates for evaluation, which in turn are used for optimizing the model. Detailed information about this topic and a survey about available algorithms can be found in [Sch11], [Hem10, p.18ff].

Among these algorithms, the metamodel based black box algorithms, Wang et al. [WSW04] proposed a mode pursuing sampling method (MPS), which has been successfully applied to many benchmark problems [SWE08]. However, despite its excellent performance, it can not be directly applied to the problem as it is initially designed for the continuous variable space. Sarif et al. [SWE08] adapted the C-MPS algorithm for the discrete domain, obtaining good estimations for the optimum by using moderate numbers of evaluations for some exemplary functions. Consequently, we will utilize this algorithm in order to find the best beam combination between transmitter and receiver.

5.2.1 Mode-Pursuing Sampling Method for Discrete Variable Space

The concept of the mode pursuing sampling method for a discrete variable space (D-MPS, [SWE08]) is to achieve a balance between exploration of possible new optima and exploitation of discovered optima regions. Based on this approach, there are two regions (circles/hyperspheres) of candidates for the variable pair $\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}}$, denoted by $D1$ and $D2$ respectively. Circle $D1$ starts with a small initial radius increasing, if evaluations in this sphere lead to an improvement, while the initial large radius of the second circle area will do the contrary. A linear line spline fitting function is used to build up a model of the function based on past evaluations of the function. Note that the evaluation of the cost function represents a beam change of initiator and responder, sending and receiving the training data, while an iteration corresponds to several m evaluations and one feedback over the secondary link. Table 5.1 gives a list of the used parameters. The algorithm is summarized as follows:

- Step 1: All parameters, especially the radii, have to be initialized to default values (Tab. 1).

- Step 2: For initial points of several values of ϕ_k^{tx} and ϕ_l^{rx}, the function $J(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})$ is evaluated in order to get a baseline for the model. Contrary to [SWE08], this is done by using a coarse initial search which allows to equidistantly explore the angle space of the cost function. The minimum of the found values are regarded as center of the two circle areas from now on.

- Step 3: Performs the basic D-MPS algorithm.

  - Step 3.1: From each of the circular regions of valid arguments ϕ_k^{tx} and ϕ_l^{rx}, $\frac{N_{\mathrm{ch}}}{2}$ arguments are chosen randomly but not evaluated by probing. They now form the set of cheap points.
  - Step 3.2: The model of the two dimensional curve is approximated by a spline interpolation by using the existing expensive and actually evaluated points/beam pairs.
  - Step 3.3: In this step, the spline model helps to determine the most promising beam index candidates (expensive points of the cost function), which will be evaluated, i.e. the beam pairs are probed. The information of which beam to probe will be forwarded from the decision unit to the transceivers by the secondary link.

- Step 4: The evaluated m expensive points from the last step will be added to the list of known points. From this list, the current minimum of the cost function $J(\phi_k^{\mathrm{tx}}, \phi_l^{\mathrm{rx}})$ and the corresponding arguments ϕ_k^{tx} and ϕ_l^{rx} have to be specified. This point will serve as the center for the circles furthermore.

- Step 5: This step actually performs the double circle strategy by changing the radii.

  - If the last step discovered a new minimum: increase the radius of the smaller circle to $1/\alpha$, reduce the radius of the bigger circle by α
  - If there were no improvements in the last n_α iterations: reduce the smaller radius by $1/\alpha$, increase the bigger radius by α
  - Check borders (minimum radius and maximum radius) of the radii and restrict them to the nearest one.

- Step 6: Define new circles by using new radii and center points. Check whether the maximum iteration is reached. Otherwise, go to step 3.

Fig. 5.1 shows a contour plot of an example cost function. Several MPCs cause several minima of the cost function. The evaluations show that the algorithm is able to step from a non-optimal minimum to the desired minimum. That explains the advantage of this algorithm over strict gradient-based algorithms.

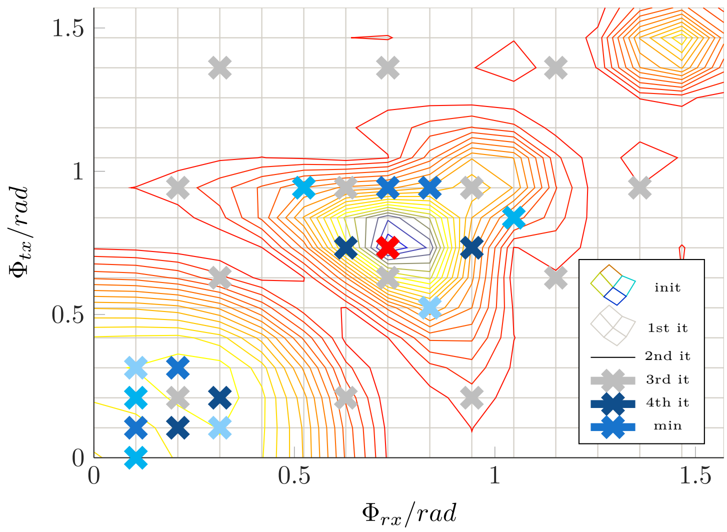


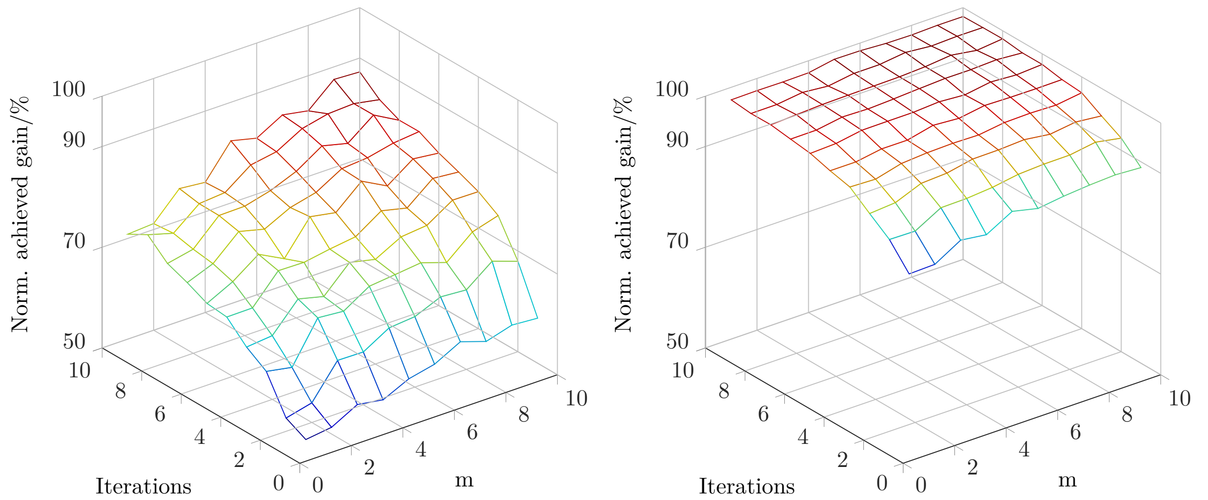
Figure 5.1: Cost function $J(\phi_{\mathrm{tx}}, \phi_{\mathrm{rx}})$ and evaluations (black) and final result (red). It can be seen that due to the double hypersphere, a local extremum was dismissed due to the global optimum. (Source: [KCF16] © 2016 IEEE).

Table 5.1: Parameters set of the algorithm. Standard values are used if not stated differently.

Description	Parameter	Value
cost function	$J(\phi^{\mathrm{tx}}_{k,k_{\mathrm{it}}}, \phi^{\mathrm{rx}}_{l,k_{\mathrm{it}}})$	
model for the cost function constructed by interpolation	$\hat{J}(\phi^{\mathrm{tx}}_{k,k_{\mathrm{it}}}, \phi^{\mathrm{rx}}_{l,k_{\mathrm{it}}})$	
radius of the normed hypersphere S	R_s	
radius of the normed hypersphere R	R_b	
part of the discrete variable set of J	$S(J)$	
number of cheap points per iteration using model	N_{ch}	30
number of initial expensive points/evaluations	$m_{\mathrm{exp,init}}$	24
number of expensive points/evaluations per iteration	m_{exp}	4
number of iterations	n_{it}	4
coefficient for increasing and decreasing the radii	α	
number of beams in the angle space $\phi_{\mathrm{tx}} \in \{0, \pi\}$	B_{tx}	16
number of beams in the angle space $\phi_{\mathrm{rx}} \in \{0, \pi\}$	B_{rx}	16
number of transmit antenna elements	N_{tx}	8
number of receive antenna elements	N_{rx}	8

5.2.2 Simulation Results

We aim to reduce the number of beam changes and thus time, which is needed in order to find the best beam combination. As no existing beam alignment algorithm is known to cope with the restrictions of codebook based ABF except exhaustive search, we will use it as a baseline for our investigations. The performance measure in the absence of noise is the normalized gain, which is defined as the received power using a beam combination found with the proposed algorithm, normalized by the highest achievable received power by using the best beam combination. We take into account the channel described before with a static channel scenario using 8 antenna elements and 16 angular equidistant distributed



(a) 3 initial evaluations. **(b)** 24 initial evaluations.

Figure 5.2: The achieved gain in dependence of the number of iterations and the number of evaluations per iteration (m). Mean over 2000 channel realizations. The gain is normalized w.r.t the gain achieved by the exhaustive search algorithm. $N_{\text{eval}} = 40$.(Source: [KCF16] © 2016 IEEE).

beams at each transmitter and receiver.

At first, the behavior of the algorithm will be tested in simulation when noise is absent. That will provide the opportunity to find a useful parameter set as a starting point for further investigations and additionally will provide insight into the influence of the individual parameters on the result. Fig. 5.2 shows the normalized gain as a function of the number of iterations and the number of evaluations (beam-changes) per iteration. The two subplots differ in the number of initial evaluations, which serve as input for the black box algorithm. Initial evaluations can be seen as a very rough exhaustive search. If using only three equidistant points (5.2a), the achievable performance in terms of the achieved power of a beam pair increases if either the number of iterations or the number of evaluated points per iteration m is increased. In contrast, if 24 initial evaluations are used, the performance does not increase relevantly after four iterations. Near perfect performance of 98 percent of normed gain can be achieved in this case. The total number of evaluations (N_{eval},beam-changes) is calculated as

$$N_{\text{eval}} = N_{\text{init\_eval}} + N_{\text{exp\_points}} \cdot N_{\text{it}}, \tag{5.6}$$

where N_{init_eval} is the number of initial evaluations, $N_{\text{exp\_points}}$ is the number of expensive points per iteration and N_{it} is the number of used iterations.

According to these simulations, the parameters are chosen to be $N_{\text{init\_eval}} = 24$, $N_{\text{exp\_points}} = 4$, $N_{\text{it}} = 4$ and thus a total number of evaluations of $N_{\text{eval}} = 40$, which is at 15.6% of the $B_{\text{tx}} \cdot B_{\text{rx}} = 256$ beam-changes using exhaustive search. It should be noted that 4 feedback signals are needed.

Now we measure the performance of the algorithm including the effect of noise. This will be done in terms of achievable spectral efficiency [bit/s/Hz] compared to the best

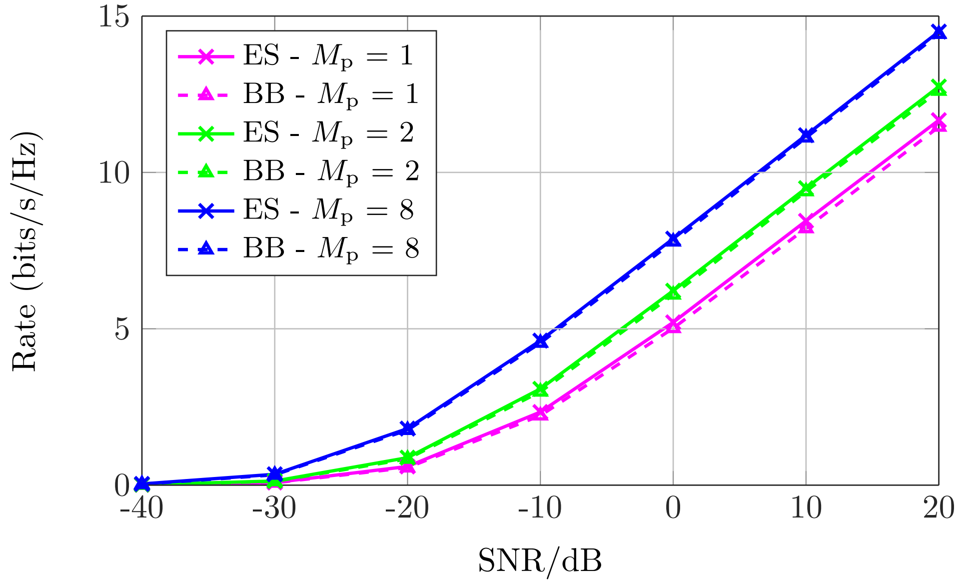


Figure 5.3: Spectral efficiency of exhaustive search (ES) and black-box based beam alignment algorithm (BB) in dependence of the number of multipath components (M_p) for the proposed channel model. Monte Carlo simulation over 2000 channel realizations. (Source: [KCF16] © 2016 IEEE).

achievable spectrum efficiency, defined as

$$R = \log_2 \left(1 + \frac{P_{\text{tx}}}{N_{\text{tx}}} \left(\mathbf{w}^H \mathbf{w} \right)^{-1} \mathbf{w}^H \mathbf{H} \mathbf{f} \mathbf{f}^H \mathbf{H}^H \mathbf{w} \right). \tag{5.7}$$

The values of the $SNR \in \{-40 \text{ dB} ... 20 \text{ dB}\}$ correspond to a distance of $d \in (0.23 \text{ m} ... 165 \text{ m})$ if the RF output power is assumed to be 5 dBm. Figs. 5.3 and 5.4 show the influence of the noise to the capacity. It shows that the algorithm proposed here provides comparable performance in terms of possible spectral efficiency while using only a fraction of the evaluations and the corresponding time. That is not only valid for a line of sight scenario, but also using a channel with several multipath components, which makes the optimization problem more difficult due to the existence of more local maxima of the cost function.

A drawback of this approach is the complexity of the algorithm, which is a result of the need for creating a metamodel in order to find promising regions for the maximization. Therefore, we will solve the problem by using a simpler algorithm in the next section.

5.3 Gradient-Based Beam Alignment

Another possibility to solve the problem (5.2) is to use the gradient information of the function. The proposed gradient-based beam alignment algorithm aims to reduce the number of probings and is split into two phases. While the aim of the initial coarse alignment is to identify the area of the optimization function where the function can be regarded concave in the region of the global maximum, the goal of the fine alignment phase is to increasingly enhance the receive power using different beam combinations, i.e. finding the global maximum of the optimization function. The two stages of the algorithm require

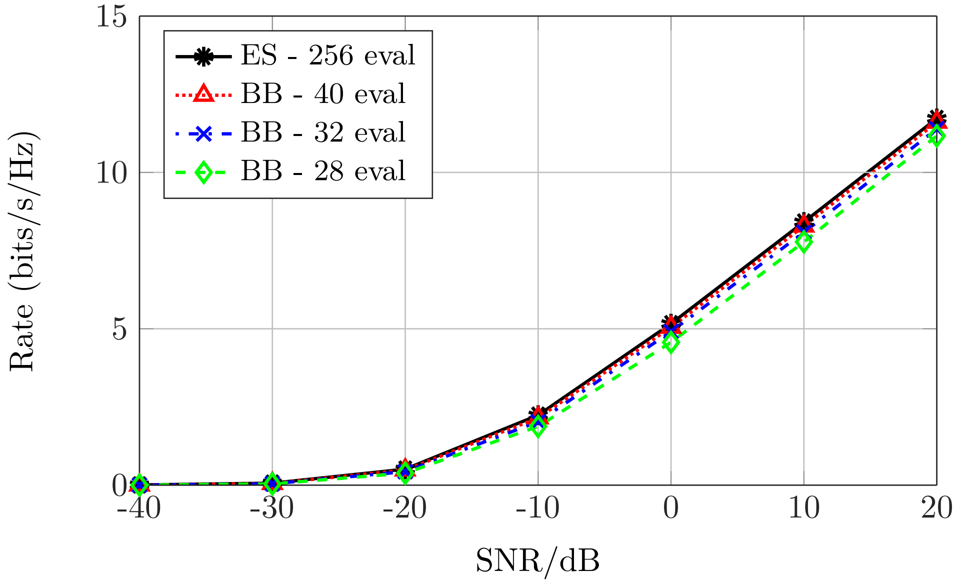


Figure 5.4: Spectral efficiency of exhaustive search (ES) and black-box based beam alignment algorithm (BB) with $N = 50$ and $n_{\mathrm{it}} = 4, 2, 1$ respectively. (Source: [KCF16] © 2016 IEEE).

different codebooks. The initial search requires a codebook which is complete, meaning the incoming wave can be received with at least the half power of the maximum receive power of the beam pattern. The half power beamwidth (HPBW) is commonly approximated by $HPBW = \pi/N$; therefore we need a codebook with cardinality at least equal to the number of antennas N. The fine alignment uses, in contrast, an over-complete codebook with more beams than antennas, where the beams are overlapping. The property of this codebook can be described by the overlap factor (OLF), defined as $OLF = HPBW/\Delta\phi$. The codebook for the initial search can be derived from the codebook of the fine search by choosing every d^{th} beam of the fine search codebook. Note that the beams of the codebook can be distributed equidistantly by either wavenumber or angular domain. We use the latter definition, in order to be compatible with the experimentation setup.

5.3.1 Initial Alignment

First, we apply an initial search by evaluating the beam combinations of every d_{tx} transmit and every d_{rx} receive beam of the respective codebooks, resulting in a codebook $\tilde{\mathbf{F}}_{\mathrm{init}} = \{\mathbf{f}_k, k = 0, d_{\mathrm{tx}}, 2d_{\mathrm{tx}}, ..., K-1\}$ of size $K_{\mathrm{init}} = \lfloor K/d_{\mathrm{tx}} \rfloor$ for the transmitter and codebook $\tilde{\mathbf{W}}_{\mathrm{init}} = \{\mathbf{w}_l, l = 0, d_{\mathrm{rx}}, 2d_{\mathrm{rx}}, ..., L-1\}$ of size $L_{\mathrm{init}} = \lfloor L/d_{\mathrm{rx}} \rfloor$ for the receiver. The result of the initial beam alignment at iteration $n = 0$ is, based on the maximization problem (5.2)

$$\left(\hat{k}_0, \hat{l}_0\right) = \arg \max_{\substack{\mathbf{f}_k \in \tilde{\mathbf{F}}_{\mathrm{init}} \\ \mathbf{w}_l \in \tilde{\mathbf{W}}_{\mathrm{init}}}} J(k, l) \qquad . \tag{5.8}$$

The result of the initial beam alignment gives a beam combination, which is located close to the optimum beam pair, which gives the maximum receive power. The training time of the initial alignment phase can be given by the multiplication of the codebooks sizes with the individual training time T_{signal} of a beam pair

$$T_{\text{init}} = K_{\text{init}} \cdot L_{\text{init}} \cdot T_{\text{signal}}. \tag{5.9}$$

5.3.2 Fine Alignment

The goal of the fine alignment is to find the optimum beam pair, which gives the maximum receive power. Based on the assumption that the current beam combination at iteration $n = 0$ is located close to this optimum, we can use a gradient-based approach to achieve the goal. The step is repeated for several iterations n until the abort criterion is reached. Following the current beam pair combination at iteration n, adjacent beam combinations $k_n \pm 1$ and $l_n \pm 1$ are evaluated. Thus the approximation of the local gradient is given by

$$\tilde{\nabla}\boldsymbol{J}(\hat{k}_n, \hat{l}_n) = \begin{bmatrix} \tilde{\nabla} J_{\text{tx}}(\hat{k}_n, \hat{l}_n) \\ \tilde{\nabla} J_{\text{rx}}(\hat{k}_n, \hat{l}_n) \end{bmatrix} = \begin{bmatrix} \frac{J(\hat{k}_n+1,\hat{l}_n)-J(\hat{k}_n-1,\hat{l}_n)}{2\Delta\phi_{\text{tx}}} \\ \frac{J(\hat{k}_n,\hat{l}_n+1)-J(\hat{k}_n,\hat{l}_n-1)}{2\Delta\phi_{\text{rx}}} \end{bmatrix}. \tag{5.10}$$

The formula of the approximate gradient is given by the difference quotient of each dimension. Motivated by the steepest descend approach we use the gradient information to increase the receive power by choosing the new beam pair candidate of the codebook by reducing the error $\tilde{\phi}_{\text{e}_n}^{\text{tx}}$ and $\tilde{\phi}_{\text{e}_n}^{\text{rx}}$ to the next valid angles ϕ^{tx} and ϕ^{rx} corresponding to a beam index

$$\begin{bmatrix} \phi_{\hat{k}_{n+1}}^{\text{tx}} \\ \phi_{\hat{l}_{n+1}}^{\text{rx}} \end{bmatrix} = \begin{bmatrix} \phi_{\hat{k}_n}^{\text{tx}} \\ \phi_{\hat{l}_n}^{\text{rx}} \end{bmatrix} + \mu_n \tilde{\nabla}\boldsymbol{J}(\hat{k}_n, \hat{l}_n) - \begin{bmatrix} \tilde{\phi}_{\text{e}_n}^{\text{tx}} \\ \tilde{\phi}_{\text{e}_n}^{\text{rx}} \end{bmatrix}, \tag{5.11}$$

where μ_n is the step size for iteration n. This is equivalent to choosing the nearest valid $\phi_{\hat{k}_{n+1}}^{\text{tx}}$ and $\phi_{\hat{l}_{n+1}}^{\text{rx}}$. The adjacent beam pair combinations $\hat{k}_{n+1} \pm 1$ and $\hat{l}_{n+1} \pm 1$ are now evaluated and the fine alignment step is repeated. A stopping criterion is formulated by bracketing a maximum value of the cost function between two smaller values

$$J(\hat{k}_n, \hat{l}_n) > J(\hat{k}_n \pm 1, \hat{l}_n \pm 1). \tag{5.12}$$

A block diagram of the algorithm is presented in Fig. 5.5.

A complete description of the algorithm is given in Alg. 3.

The training time of the fine alignment stage is flexible and is given approximately depending on the number of needed iterations N_{it} and the feedback time T_{FB} of the control channel

$$T_{\text{fine}} = N_{\text{it}} \cdot (5 \cdot T_{\text{signal}} + T_{\text{FB}}). \tag{5.13}$$

In each iteration, 5 new beam pairs are evaluated, which are the new candidate and the neighboring beam pairs. If some beam pairs are already evaluated they are skipped.

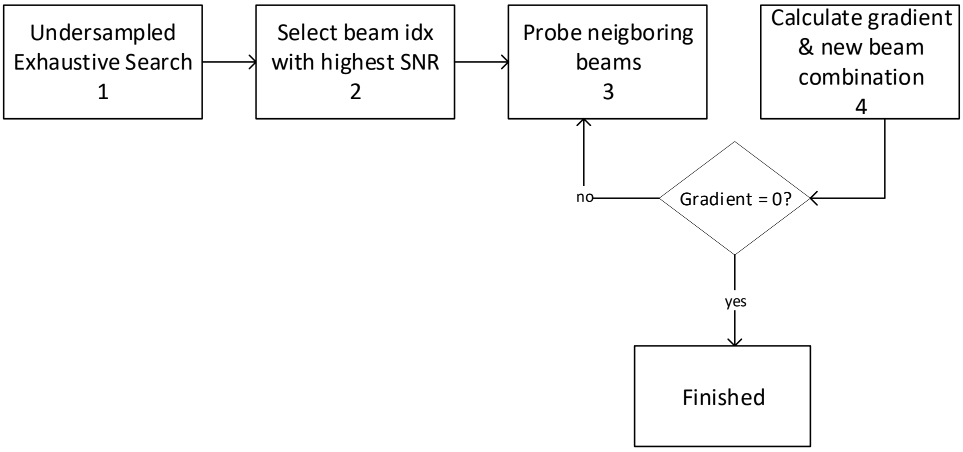


Figure 5.5: Block diagram of gradient beam alignment algorithm.

input : $\bar{a}_{\text{target}}$, $\text{SNR}_{\text{assumed}}$
output: beam combination $\hat{k}_{\text{final}}$, $\hat{l}_{\text{final}}$

A: Initial Alignment;
begin
  evaluate all beam combinations of codebooks $k \in \tilde{\mathbf{F}}_{\text{init}}, l \in \tilde{\mathbf{W}}_{\text{init}}$;
  $(k_0, l_0) = \arg \max\limits_{\substack{k \in \tilde{\mathbf{F}}_{\text{init}} \\ l \in \tilde{\mathbf{W}}_{\text{init}}}} J(k, l)$;
end

B: Fine Alignment for iteration n;
begin
  evaluate adjacent beam combinations $\{J(\hat{k}_n - 1, \hat{l}_n), J(\hat{k}_n + 1, \hat{l}_n), J(\hat{k}_n, \hat{l}_n - 1), J(\hat{k}_n, \hat{l}_n + 1)\}$;
  calculate gradient $\tilde{\nabla} \boldsymbol{J}(k_n, l_n))$;
  find step size parameter μ_{n+1} ;
  obtain $(\hat{k}_n, \hat{l}_n)$ from (5.11);
  stop if $J(\hat{k}_n, \hat{l}_n) > J(\hat{k}_n + i, \hat{l}_n + j) \quad i, j \in \{-1, 1\}$;
end

Algorithm 3: Gradient based beam alignment

5.4 Definition of the Initial Codebook

In this section, we focus on the problem of how to define the codebook for the initial search $\tilde{\mathbf{F}}_{\text{rf,init}}$ and $\tilde{\mathbf{W}}_{\text{rf,init}}$. As the beam width can not change, the entries of the initial have to be part of the full codebook. For this reason, we define the codebooks by taking every $d_{\text{T}}{}^{\text{th}}$ or $d_{\text{R}}{}^{\text{th}}$ entry of the full codebook, i.e. $\tilde{\mathbf{F}}_{\text{rf,init}} = \{\mathbf{f}_k, k = 0, d_{\text{T}}, 2d_{\text{T}}, .., K-1\}$ and $\tilde{\mathbf{W}}_{\text{rf,init}} = \{\mathbf{w}_l, l = 0, d_{\text{R}}, 2d_{\text{R}}, .., L-1\}$. Thus the problem is to find the distances d_{T} and d_{R}, respectively. Without loss of generality, we restrict the problem to the receive dimension in the following discussion by omitting the effect of transmit beamforming, which means that the objective function $J(k', l)$ is dependent only from receive beam index l, by fixing $k = k'$. d_R can be found in the same way.

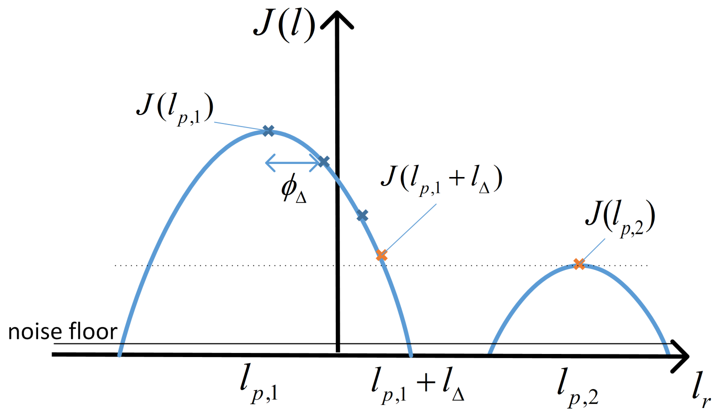


Figure 5.6: One-dimensional objective function with one global and a smaller local peak. (Source: [KCF17] © 2017 IEEE).

We assume that two planar waves with power $P_1 > P_2$ are approaching the receive array with N_{rx} antennas. Then, (5.1) reduces to $y = \sqrt{P_1}\mathbf{w}_l^H\mathbf{a}_1 + \sqrt{P_2}\mathbf{w}_l^H\mathbf{a}_2$ for the synchronization case, and the objective function $J(l) = |y(l)|^2$ can be obtained by evaluating all entries of the full codebook $\tilde{\mathbf{W}}_{\mathrm{rf}}$ with angular difference $\Delta\phi_{\mathrm{R}}$ between main lobe directions of adjacent beams. From this equation, it can also be seen that in the area where the effect of one wave will be dominant, the shape of the objective function follows the shape of the beam pattern. Note that we assume here the full exploration (exhaustive search) of the objective function to get an insight of the typical structure of the function $J(l)$, while the goal is to shrink the number of needed evaluations in the actual beam alignment algorithm to a minimum. We use an over-complete codebook $\tilde{\mathbf{W}}_{\mathrm{rf}}$ with large overlap between adjacent beam patterns. The overlap between adjacent beam patterns can be expressed by using the HPBW of the beam patterns and the angular distance $\Delta\phi$ between adjacent beam patterns. We define the OLF as a measure of the over-completeness of the codebook as

$$OLF = \frac{HPBW}{\Delta\phi}. \tag{5.14}$$

Assuming the beam patterns shape to be identical, the cost function gives a sampled version of the beam pattern for each path. For further investigations, we only focus on main lobes and omit the side lobes for a moment, and assume the angular difference of the two paths $\phi_{\mathrm{A},1}$ and $\phi_{\mathrm{A},2}$ is large enough, so that the beam pattern shapes or slopes do not overlap until the noise level. An example for the one-dimensional cost function is given in Fig. 5.6. The maximal receive power introduced by wave one is given by choosing receive beam pattern $k_{p,1}$ with $J(l_{p,1}) = J_{p,1}$, giving the global maximum of the cost function. The maximal receive power exited by wave two is given by choosing receive beam pattern $k_{p,2}$ with $J(l_{p,2}) = J_{p,2}$, resulting in a local maximum. If the result of the initial search gives a beam pair around the global optimum, the slope of the beam pattern will cause the gradient to find the global maximum. If, in contrast, the initial search result gives a beam around the local optimum, the slope of the beam pattern will guide the gradient only to

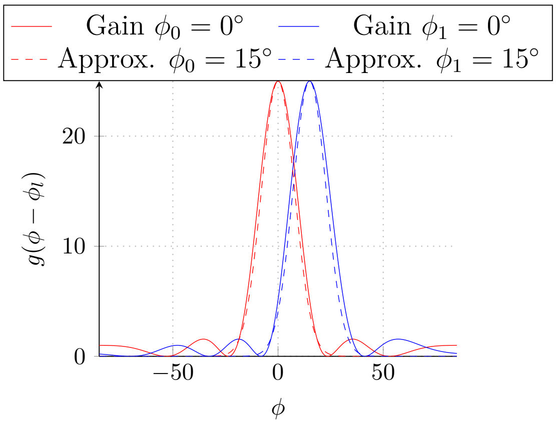


Figure 5.7: Approximation of the beam pattern function. The error of the approximation will increase with rising ϕ_0 as the beamwidth is widened, unrealistic small approximated beamwidth. As this is a disadvantage for our investigation, it can be seen as worst case. (Source: [KCF17] © 2017 IEEE).

the local maximum of the objective function. In order to ensure that the global maximum is found, we have to make sure that the initial search evaluates at least one beam pattern where the shape of the peak caused by wave one is higher than the maximum value of the cost function caused by wave two. The number of beam patterns in this area of the cost function determines the distance d_R we can allow. As described above, the area of the objective function, which is defined by the influence of wave one, is given as a sampled version of the beam pattern. The beam pattern for a linear array in power domain is given by

$$g(\phi - \phi_l) = \frac{\sin^2\left(\frac{N_\mathrm{rx}\pi}{2}\sin\phi - \sin\phi_l\right)}{\sin^2\left(\frac{\pi}{2}\sin\phi - \sin\phi_l\right)} \tag{5.15}$$

$$g(\phi - \phi_l) \overset{(a)}{\approx} A \cdot \mathrm{e}^{\frac{-(\phi-\phi_l)^2}{2\sigma^2}}, \tag{5.16}$$

where (a) follows an approximation of mainlobe of the beam pattern by a Gaussian function (Fig. 5.7). Parameter A can be easily found by investigating the maximum of the pattern function

$$A = \max g_l(\phi) = N_\mathrm{rx}^2. \tag{5.17}$$

The parameter σ can be obtained by evaluating the first zero of the beam pattern for $\phi_l = 0$ with constraint $\phi_0 = 3\sigma$

$$g(\phi_0) = 0 \tag{5.18}$$

$$\phi_0 = \arcsin\frac{2}{N_\mathrm{rx}} \tag{5.19}$$

$$\text{s.t. } \phi_0 = 3\sigma \tag{5.20}$$

$$\sigma = \frac{\arcsin\frac{2}{N_\mathrm{rx}}}{3}. \tag{5.21}$$

We choose $\sigma = 1/3\phi_0$ because the integral over the range of $\phi \in \{-3\sigma, 3\sigma\}$ contains about 99.7% of the area of the Gaussian shaped function.

The condition that the area of the cost function $\mathcal{J}_1 = \{J(l_{p,1} - l_\Delta), ..., J(l_{p,1} + l_\Delta)\}$ has to be larger than the value $J(l_{p,2})$ of the objective function caused by wave two leads to the following relation

$$\forall J \in \mathcal{J}_1 > J(l_{p,2}) \tag{5.22}$$

$$J(l_{p,1} + l_\Delta) > J(l_{p,2}) \tag{5.23}$$

$$\frac{J(l_{p,1})}{N_{\mathrm{rx}}^2} g(\phi_{l_{p,1}} + \phi_{l_\Delta}) \overset{(b)}{>} J(l_{p,2}) \tag{5.24}$$

$$\mathrm{e}^{-\left(\frac{\frac{d}{2}^2 \Delta\phi^2}{2\sigma^2}\right)} \overset{(c)}{>} \frac{J(l_{p,2})}{J(l_{p,1})}, \tag{5.25}$$

with ϕ_Δ is the angular difference between the main lobes and where (a) follows from the fact that the shape of the maximum of the objective function equals the beam shape and (b) follows from the fact that at least one point of the initial search must be contained in $\mathcal{J}$, i.e. $2l_\Delta = d_R\phi_\Delta$. After some calculations, by using the definition of the HPBW and (5.14), d_R is given by

$$d_R < \frac{2\sigma}{\Delta\phi}\sqrt{2\ln\frac{J(l_{p,1})}{J(l_{p,2})}} \tag{5.26}$$

$$d_R < \frac{\sigma N_{\mathrm{rx}}\, OLF}{\pi}\sqrt{8\ln\frac{J(l_{p,1})}{J(l_{p,2})}}. \tag{5.27}$$

These considerations give a closed form expression of the allowed distance d_{R} depending on the power of the two planar waves. As this parameter is also dependent on the design of the full codebook $\tilde{\mathbf{W}}_{\mathrm{rf}}$, it is not surprising that it is also dependent on the OLF. Note that the number of additional waves with smaller power is independent in our considerations.

5.4.1 Effect of Side Lobes

Until now, we only regard the effect of the mainlobe to the objective function J. Side lobes influence the shape of the function by producing smaller local maxima in addition to a dominant maximum by the main lobes. Therefore, from the perspective of the objective function, side lobes have the same effect as incoming waves with small power. Thus we can model the effect of the side lobes by a shifted version of the beam pattern with smaller power. For this reason an important usage of (5.27) comes with the investigations of the effect of side lobes. In this case, the ratio between the global and local maximum of the objective function $J(l_{p,2})/J(l_{p,1})$ is fixed. To find an analytical solution for the first side-maximum of (5.15) is not easy. For this reason, we follow the common approach [Bal05] to approximate the maximum of the beam pattern function in (5.15) by assuming the numerator of the original gain function is maximized

$$\sin\left(\frac{N_{\mathrm{rx}}\pi}{2}\sin\phi\right)^2 = 1 \tag{5.28}$$

$$\frac{N_{\mathrm{rx}}\pi}{2}\sin\phi = \frac{3}{2}\pi \tag{5.29}$$

$$\phi = \arcsin\frac{3}{N_{\mathrm{rx}}} \tag{5.30}$$

$$g_{sidelobe} = \frac{1}{\sin\left(\pi/2\cdot\sin\left(\arcsin(\frac{3}{N_{\mathrm{rx}}})\right)\right)^2} \tag{5.31}$$

Thus the ratio between the mainlobe and first side lobe gain is

$$\frac{J(l_{p,2})}{J(l_{p,1})} = \frac{1}{\sin(\pi/2\cdot 3/N_{\mathrm{rx}})^2\cdot N_{\mathrm{rx}}^2}. \tag{5.32}$$

5.4.2 Expectation Value

If the ratio $\frac{J(l_{p,1})}{J(l_{p,2})}$ is decreased, d_{R} also has to be decreased (5.27). If the ratio becomes too small, the distance has to be $d_{\mathrm{R}} = 1$, meaning the initial search is equal to exhaustive search and no benefit in terms of decreasing the number of evaluations is reached. Nevertheless, the probability of finding the global maximum over different channel realizations can be quite high using a larger d_R. To give an understanding of the correct parametrization in these cases, we use the expectation value of the maximum of the objective function over several channel realizations found by the algorithm (maximum of objective function J_{alg}) compared to the global maximum of the objective function (maximum of the objective function $J_{\max}$). In contrast to the case above, we cannot ensure that the global maximum is reached but we can control the probability that the global optimum is found. Assume again a scenario where two planar waves with power $P_1 > P_2$ but $P_1 \approx P_2$ are hitting one array with N_{rx} antennas. Given a targeted expectation of the ratio over several channel realizations $\mathbb{E}\left\{\frac{J_{\mathrm{alg}}}{J_{\max}}\right\} < 1$, the goal is to choose the distance parameter d_{R} in dependency of the OLF property of the codebook. We again assume no overlap and assume the shape caused by the waves is symmetrical sampled by the beam patterns (Fig. 5.8). The expectation value of J_{alg} can be described using the probability of finding the local or global maximum by

$$\begin{aligned}\mathbb{E}\left\{J_{\mathrm{alg}}\right\} =& P\left(J(l_{p,1})\right)J(l_{p,1}) + P\left(J(l_{p,2})\right)J(l_{p,2})\\ &+ P\left(N_{\mathrm{rx}}\sigma_n^2\right)N_{\mathrm{rx}}\sigma_n^2,\end{aligned} \tag{5.33}$$

where $N_{\mathrm{rx}}\sigma_n^2$ describes the noise level. Finding the local or global maximum is successful if the maximum receive power of a beam initial search is located around the global maximum one or local maximum two respectively, i.e. the power of the receive signal has to be higher than the noise. In the following, we name the set of all beam patterns probed for the initial search $\tilde{\mathbf{W}}_{\mathrm{rf,init}}$ and the resulting set of cost function values $\mathcal{J}_{\mathrm{init}}$. If the beam pattern which gives the local optimum is contained in the initial codebook

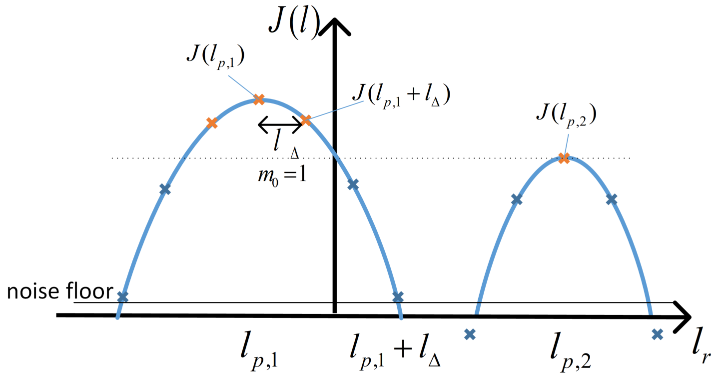


Figure 5.8: One-dimensional objective function with one global and a local peak. 3 adjacent beams at the top of the global maximum are higher than the local maximum. (Source: [KCF17] © 2017 IEEE).

i.e. $(l_{p,2} \in \tilde{\mathbf{W}})_{\text{rf,init}}$ (Fig. 5.8), the probability that the initial search finds a beam pattern around the shape of the global maximum $\mathcal{J}_1 = \{J(l_{p,1}-m), ..., J(l_{p,1}+m)\}$ is given by

$$P\left(J_{\text{alg}} = J_{p,1} | l_{p,2} \in \tilde{\mathbf{W}}_{\text{rf,init}}\right) = \min\left(\frac{2m_0+1}{d_{\text{R}}}, 1\right) \tag{5.34}$$

where m_0 is the maximum index distance to $l_{p,1}$, which gives a receive power value higher than the local peak, i.e. $J(l_{p,1}-m_0) > J(l_{p,2})$. Consequently, $(2m_0+1)$ gives the number of beam patterns which give a receive power higher than $J(l_{p,2})$. Thus the equation describes the probability that at least one of the $(2m_0+1)$ beam patterns is part of $\mathcal{J}_1$ and is limited by one. The probability that $l_{p,2} \in \tilde{\mathbf{W}}_{\text{rf,init}}$ is

$$P\left(l_{p,2} \in \tilde{\mathbf{W}}_{\text{rf,init}}\right) = \frac{1}{d}. \tag{5.35}$$

Following this argumentation, we can calculate the probability that the initial search finds a beam pattern around the global maximum under the assumption that not the beam pattern $l_{p,1}$ but at least one of the adjacent beam patterns is part of the initial codebook $\left[(l_{p,1}-1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right] \vee \left[(l_{p,1}+1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right]$ as

$$P\left(J_{\text{alg}} = J_{p,1} | \left[(l_{p,2} \pm 1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right]\right) = \min\left(\frac{2m_1+1}{d_{\text{R}}}, 1\right). \tag{5.36}$$

The probability that $\left[(l_{p,1}-1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right] \vee \left[(l_{p,1}+1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right]$ is

$$P\left(\left[(l_{p,1}-1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right] \vee \left[(l_{p,1}+1) \in \tilde{\mathbf{W}}_{\text{rf,init}}\right]\right) = \frac{1 + d_{\text{R}} \mid 2n_1}{d_{\text{R}}}. \tag{5.37}$$

where $d_{\text{R}} \mid 2n_1$ is one if d is divisor of $2n_1$, m_i is the distance of peak one to the point, which is higher the current point $J(l_{p,2}+n_i)$, where n_i is the distance to local optimum index $l_{p,2}$. An example is given in Fig. 5.9.

Based on the assumption that the initial search evaluates points of the cost function, which are n points away from local maximum $J(l_{p,2})$, the maximum allowed distance to

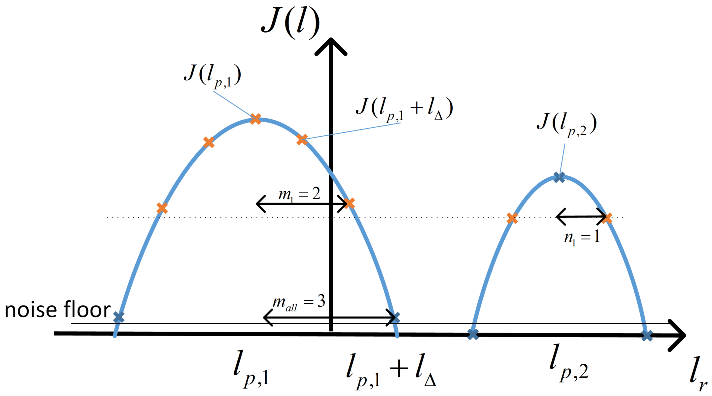


Figure 5.9: One-dimensional objective function with one global and a local peak. 5 adjacent beams at the top the global maximum are higher than the power received by the pattern next to the local maximum. (Source: [KCF17] © 2017 IEEE).

$J(l_{p,1})$, which is m, can be calculated by using the approximation of the beam patterns

$$\frac{J_{p,1}}{N_{\text{rx}}^2} g(m\Delta\phi) > \frac{J_{p,2}}{N_{\text{rx}}^2} g(n\Delta\phi)\,. \tag{5.38}$$

Using approximation in (5.15) we can write

$$\frac{m^2\Delta\phi^2}{2\sigma^2} < \ln\frac{J_{p,1}}{J_{p,2}} + \frac{n^2\Delta\phi^2}{2\sigma^2}\,. \tag{5.39}$$

Then we can isolate the maximal allowed distance m to the peak $J_{p,1}$

$$m < \frac{\sigma}{\Delta\phi}\sqrt{2\ln\frac{J_{p,1}}{J_{p,2}} + \frac{n^2\Delta\phi^2}{\sigma^2}} \tag{5.40}$$

Using the notation of the OLF this equation equals to

$$m < \frac{\sigma N_{\text{rx}} OLF}{\pi}\sqrt{2\ln\frac{J_{p,1}}{J_{p,2}} + \frac{n^2\pi^2}{N_{\text{rx}}^2 OLF^2\sigma^2}} \tag{5.41}$$

$$m < \sqrt{\frac{2\sigma^2 N_{\text{rx}}^2 OLF^2}{\pi^2}\ln\frac{J_{p,1}}{J_{p,2}} + n^2}. \tag{5.42}$$

Thus the probability that the maximum received power of the initial search can be calculated according to (5.33)

$$\begin{aligned} P\left(J_{p,1}\right) = \sum_{\substack{n>n_{all}\\ m(n)}} \min\left(\frac{2m_l+1}{d},1\right)\min\left(\frac{1}{d}+\frac{2n_k \mid d}{d},1\right) \\ + \min\left(\frac{2m_{all}+1}{d},1\right)\min\left(\frac{2n_{all}+1}{d},1\right) \end{aligned}, \tag{5.43}$$

where m_{all} and n_{all} are the maximal allowed distances to $J(l_{p,1})$ and $J(l_{p,2})$, which are defined by the noise level. The maximum allowed distances m_{all} and n_{all} respectively can be defined by searching the intersection of the beam-shape with the noise level

$$\frac{J_{p,1}}{N_{\mathrm{rx}}^2} g(m_{all}) = N_{\mathrm{rx}} \sigma_n^2 \tag{5.44}$$

$$m_{all} = \frac{\sigma}{\Delta\phi} \sqrt{2 \ln \frac{J_{p,1}}{N_{\mathrm{rx}} \sigma_n^2}}. \tag{5.45}$$

Again using the OLF in the terms, the maximum allowed distances to the peak are

$$m_{all} = \frac{\sigma N_{\mathrm{rx}} OLF}{\pi} \sqrt{2 \ln \frac{J_{p,1}}{N_{\mathrm{rx}} \sigma_n^2}} \tag{5.46}$$

$$n_{all} = \frac{\sigma N_{\mathrm{rx}} OLF}{\pi} \sqrt{2 \ln \frac{J_{p,2}}{N_{\mathrm{rx}} \sigma_n^2}}. \tag{5.47}$$

The expectation value can now be calculated according to (5.33) as

$$\begin{aligned} \mathbb{E}\left\{J_{\mathrm{alg}}\right\} =& P\left(J_{p,1}\right) J_{p,1} \\ &+ \Big(1 - P\left(J_{p,1}\right) \\ &- P\left(J_{\mathrm{alg}} = N_{\mathrm{rx}} \sigma_n^2\right)\Big) J_{p,2} \\ &+ P\left(P = N_{\mathrm{rx}} \sigma_n^2\right) N_{\mathrm{rx}} \sigma_n^2 \end{aligned} \tag{5.48}$$

with the probability that the algorithm misses all peaks and J_{alg} equals to noise level is

$$\begin{aligned} &P\left(J_{\mathrm{alg}} = N_{\mathrm{rx}} \sigma_n^2\right) = \\ &(1 - \min \frac{2m_{all} + 1}{d})(1 - \min \frac{2n_{all} + 1}{d}). \end{aligned} \tag{5.49}$$

5.4.3 Step Size Parameter μ

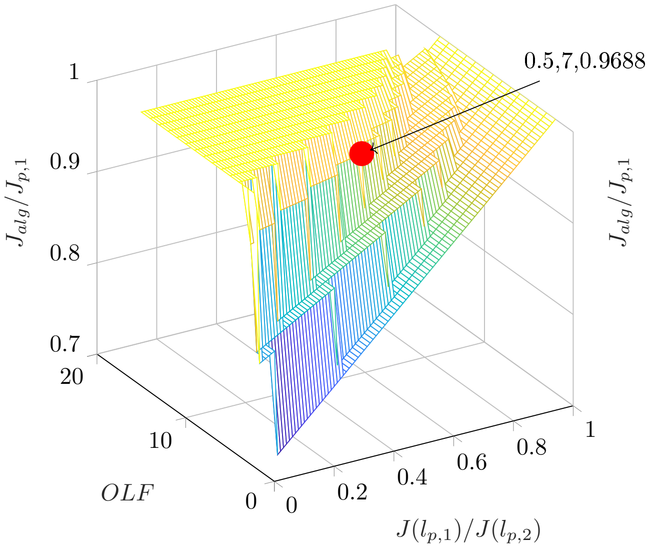
We assume that no derivative information of the function values of the cost function is available in order to determine the optimum step-size parameter μ_n. Instead, we use the assumed concavity of the optimization function in an area around the maximum value of the initial alignment. Thus the function can be approximated by a parabola. In each iteration, we get three adjacent points of the objective function, which can be used to determine the parameters of the approximating parabola. The step size parameter μ_n can be set, by finding the maximum point of the approximation.

Defining the approximation of the optimization function in dimension k by fixing $l = l'$

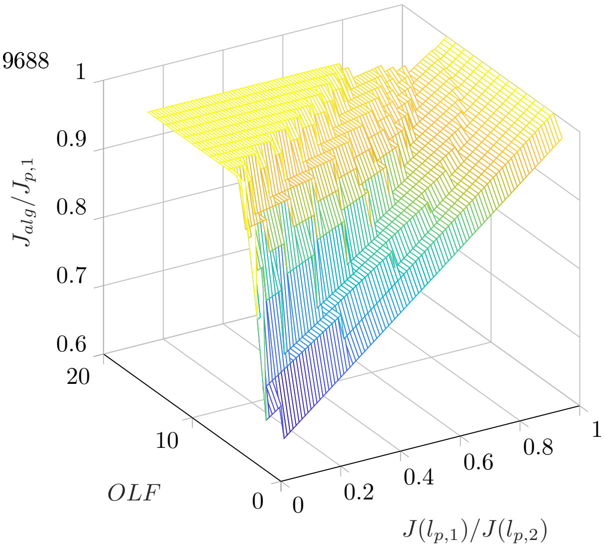
$$J(\hat{k}_n) = a(\phi_{\hat{k}_n}^{\mathrm{tx}})^2 + b\phi_{\hat{k}_n}^{\mathrm{tx}} + c \quad , \tag{5.50}$$

parameters a, b and c can be calculated by using $J(\hat{k}_n - 1, \hat{l}_n)$, $J(\hat{k}_n, \hat{l}_n)$ and $J(\hat{k}_n + 1, \hat{l}_n)$ using standard linear algebra. The argument for the maximum value of $J(\hat{k}_n)$ is given by

$$\phi_{\hat{k}_{\max}}^{\mathrm{tx}} = -\frac{b}{2a}. \tag{5.51}$$



(a) $d_R = 4$.



(b) $d_R = 6$.

Figure 5.10: $\frac{J_{\text{alg}}}{J_{p,1}}$ against OLF and $\frac{J_{p,2}}{J_{p,1}}$. The number of antennas is $N_{\text{rx}} = 5$, distances are $d_R = 4$ and $d_R = 6$. (Source: [KCF17] © 2017 IEEE).

Thus the approximation of the optimal parameter μ_{tx} for the transmitter dimension can be set to

$$\mu_{\text{tx},n+1} = \frac{\phi^{\text{tx}}_{\hat{k}_{n+1}} - \phi^{\text{tx}}_{\hat{k}_n}}{\nabla J_{\text{tx}}(\hat{k}_n, \hat{l}_n)}. \tag{5.52}$$

The procedure is repeated for the receive dimension l, and the total step size parameter μ is given by

$$\mu_{n+1} = \sqrt{\mu_{\text{tx}}^2 + \mu_{\text{rx}}^2}. \tag{5.53}$$

As in any iterative optimization function, a global maximum may be missed by using a step size parameter μ_n, which is too large. That results in bracketing of the maximum by two beam pair candidates. This situation can be determined by the fact that the result of the cost function is falling instead of rising for increasing iteration steps

$$J(\hat{k}_n, \hat{l}_n) < J(\hat{k}_{n-1}, \hat{l}_{n-1}). \tag{5.54}$$

Consequently, in such situations, we reduce the step size parameter

$$\mu_{n+1} = \frac{\mu_n}{2}. \tag{5.55}$$

In contrast, parameter μ_n could be too small to allow the beam pair candidate to change at all. Therefore, we define a lower bound of the step size parameter by using the distances of the mainlobe directions in the angular domain $\Delta\phi_{\text{tx}}$ and $\Delta\phi_{\text{rx}}$

$$\mu_{\min} = \sqrt{\frac{\Delta\phi_{\text{tx}}^2 + \Delta\phi_{\text{rx}}^2}{\nabla J_{\text{tx}}(k_n, l_n)^2 + \nabla J_{\text{rx}}(k_n, l_n)^2}}. \tag{5.56}$$

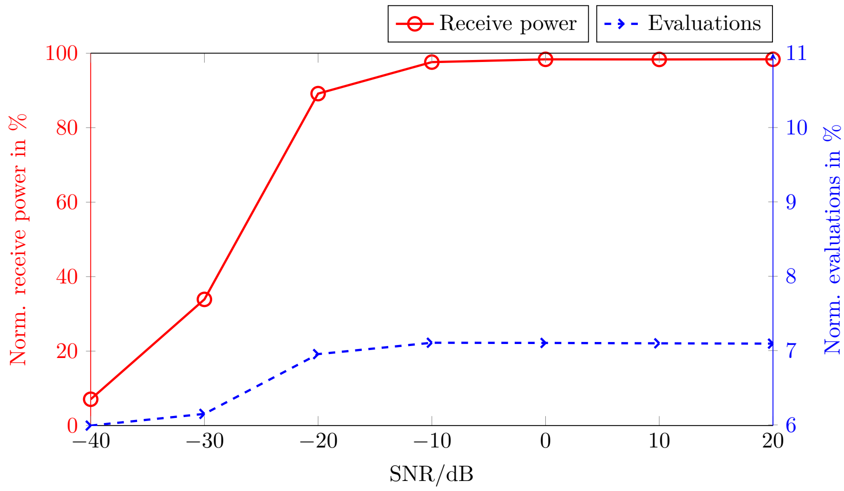


Figure 5.11: Relative receive power over distance. (Source: [KCF17] © 2017 IEEE).

5.5 Simulation Results

We now demonstrate the applicability of the system by simulation. We use the channel model introduced in (2.15) for a small cell scenario with $M_p = 2$ paths with a complex Gaussian distributed amplitude and independently uniformly distributed AoAs ϕ_{A} and AoDs and ϕ_{D}. The transmitter and receiver are assumed to have a linear array of each $N_{\mathrm{tx}} = N_{\mathrm{rx}} = 5$ antenna elements. The defined codebooks with cardinality $|\tilde{\mathbf{F}}_{\mathrm{rf}}| = |\tilde{\mathbf{W}}_{\mathrm{rf}}| = 25$, which are equidistant distributed in angular domain in the range of $\phi \in \{-\pi/3, ..., \pi/3\}$ with angular distance of $\Delta\phi_{\mathrm{tx}} = \Delta\phi_{\mathrm{rx}} = 5°$ have an OLF of $OLF \approx 7$. We need to know the maximum distance $d_R = d_T$ for the initial search for the given OLF. We want to be sure that the effect of the side lobes can be always neglected. Based on (5.27) and (5.32), we can calculate $\frac{J(k_{p,2})}{J(k_{p,1})} = 0.061$ and thus $d_{T\setminus R} < 7$. For the two paths we specify our system as follows. We want to allow a ratio of $\frac{P_{peak2}}{P_{peak1}} = \frac{1}{2}$, and an expectation value of $\mathbb{E}\left\{J_{\mathrm{alg}}/J_{\mathrm{max}}\right\} = 0.97$ for the given OLF. Therefore, we choose $d_{T\setminus R} = 4$ as a larger distance result in a smaller expectation value (Fig. 5.10).

We compared the performance in three ways. First, the ratio J_{gr}/J_{es} compares the received power given by the gradient search and exhaustive search for the same channel realization. Fig. 5.11 shows that the proposed method achieves almost the same receive power as exhaustive search with rising SNR, as a too large noise floor destroys the smooth nature of the objective function, making the gradient information useless. The second performance criterion is the ratio of beam pair evaluation, which is needed for algorithm termination ev_{gr}/ev_{es} (Fig. 5.11). It shows that the gradient needs at most 7.2% of the probing effort of the full exhaustive search. The third performance criterion, the achievable normalized rate, is depicted in Fig. 5.12. Although the gradient algorithm achieves significantly less receive power than the exhaustive search for very bad channel conditions, it does not reflect in the achievable rate, as the channel is anyhow too bad to be used for communications for the given system.

In summary, the results show that the proposed gradient-based algorithms achieve almost the same performance in receive power and rate as ES for values reasonable SNR

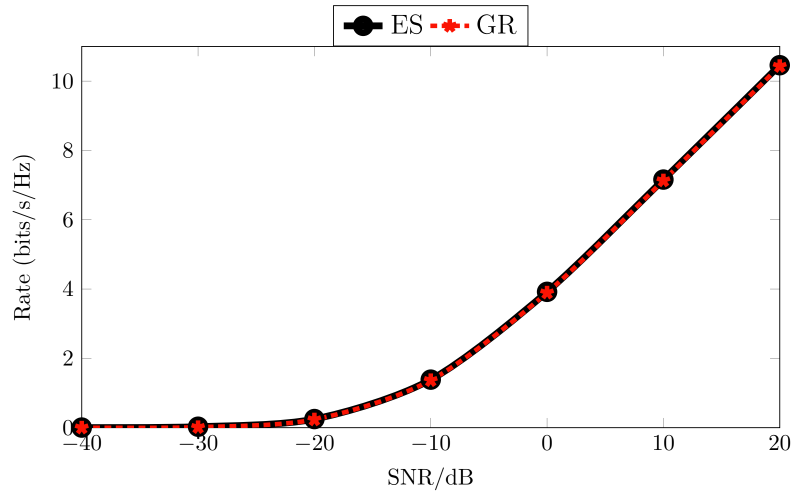


Figure 5.12: Comparison of normed achievable rate for exhaustive search (ES) and gradient based algorithm (GR). (Source: [KCF17] © 2017 IEEE).

>-20 dB, where communication is meaningful. The disadvantage of little smaller average achieved receive power is compensated by the advantage of only using about 7% of the probing compared to the exhaustive search algorithm. However, this comes at the cost of additional feedback.

5.6 Implementation

As the black box based algorithm for solving the beam alignment problem is too complex for being implemented in a real-time computer using LabVIEW, we implemented the promising gradient-based approach for usage in the testbed.

Exhaustive Search Implementation. The exhaustive search algorithm serves as a baseline for performance comparison. Additionally, parts of it are also used for the gradient algorithm in the initial search phase. The basic idea of the exhaustive search is to probe all possible beam pair combinations between AP and UD and search for the one with the best receive power.

Gradient-Based Beam alignment. A scheme of the implementation of the gradient algorithm is presented in Fig. 5.13. We reuse some code blocks from the exhaustive search implementation for the initial search operation here, which covers not the complete set but only a subset of beam combinations. The following steps presented in the block diagram in Fig. 5.13 differ between the theoretical block diagram in Fig. 5.5 and the practical implementation, because the gradient calculation and the scheduling of the adjacent beams of the new candidate have to be performed in one radio frame. Therefore, gradient calculation is shifted to implementation block two but refers to step three. The gradient calculation will be skipped for the first iteration, where the current candidate is determined by the initial search. In the following, we will summarize the individual steps in Fig. 5.13.

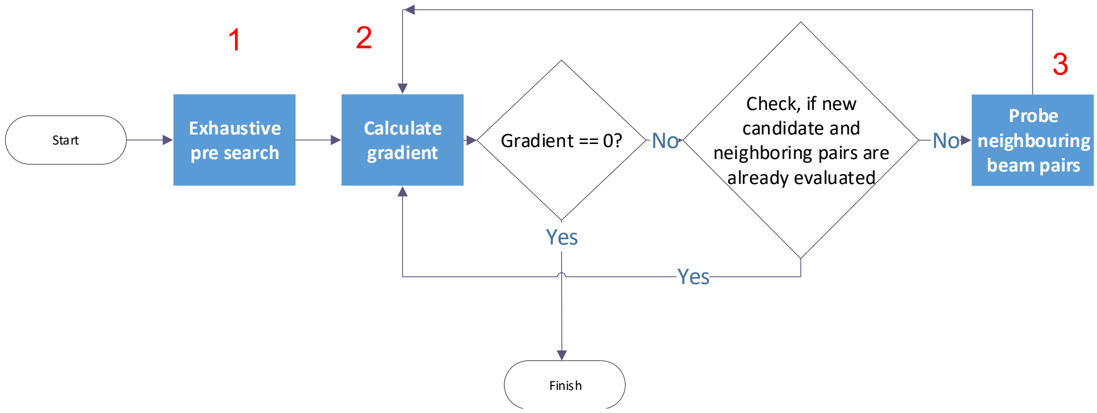


Figure 5.13: Sketch of gradient algorithm implementation blocks. Block 2 is skipped in the first iteration, as the neighboring beams have to be probed first.

1. *Initial search.* To use the ES implementation for the gradient algorithm, a feature has to be built in, to allow "under-sampled exhaustive search", which means that just a part of all beam pair combinations is scheduled. The implementation consists of two cases also called working modes. The "create schedule" working mode creates the beam-schedule to probe the necessary beam pair combination (depending on the number of beams of the codebook and whether it is an under-sampled search or not). After the pre-search is finished, which means that for each stage of the initial search the "create schedule" and the "find best beam" phase have to be executed, the result of the best beam pair of the under-sampled search is used as an input for the "calculate gradient" block.

2. *Calculate gradient.* The "calculate gradient" block is the core of the algorithm. The most important code snippets will be discussed here. First, in order to calculate the gradient, the values of the cost function $J_n(\phi^{\text{tx}}, \phi^{\text{rx}} + \Delta\phi^{\text{rx}}), J_n(\phi^{\text{tx}}, \phi^{\text{rx}} - \Delta\phi^{\text{rx}}), J_n(\phi^{\text{tx}} + \Delta\phi^{\text{tx}}, \phi^{\text{rx}}), J_n(\phi^{\text{tx}} - \Delta\phi^{\text{tx}}, \phi^{\text{rx}})$ are obtained from the internal array. The second step is the termination criterion checking whether the gradient is zero or negative, which is equivalent to the case that the neighboring beam-index combinations receive less power than the current one. If it is ensured that the stopping criterion is not fulfilled, the gradient and the next candidate is calculated. After the new beam pair candidate is determined, it is checked whether it is the same as before. If this is the case, the parameter μ_n is not large enough and will be reset to the minimum valid value. If the output beam indices are not equal to input beam indices either $(\phi^{\text{tx},n} \neq \phi^{\text{tx}}_{n+1}, \phi^{\text{rx}}_n \neq \phi^{\text{rx}}_{n+1})$, the program will be shifted to the next step, if the next value is not evaluated yet. If the value $J_n(\phi^{\text{tx}}_{n+1}, \phi^{\text{rx}}_{n+1})$ is already evaluated, it is checked whether it is smaller than the current value. If this is the case, then the maximum is likely to be jumped over and is in between the beam indices of step n and $(n+1)$. This effect is shown for a one-dimensional optimization curve in Fig. 5.14.

3. *Probe neighboring beam pair.* The task of this VI is to probe neighboring beams of the new determined candidate beam index in order to calculate the gradient in

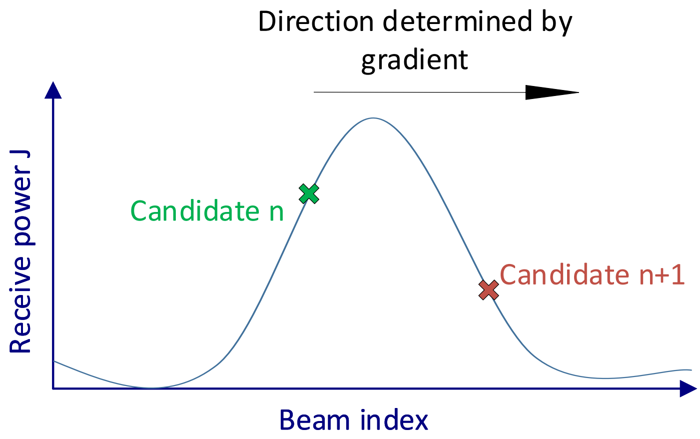
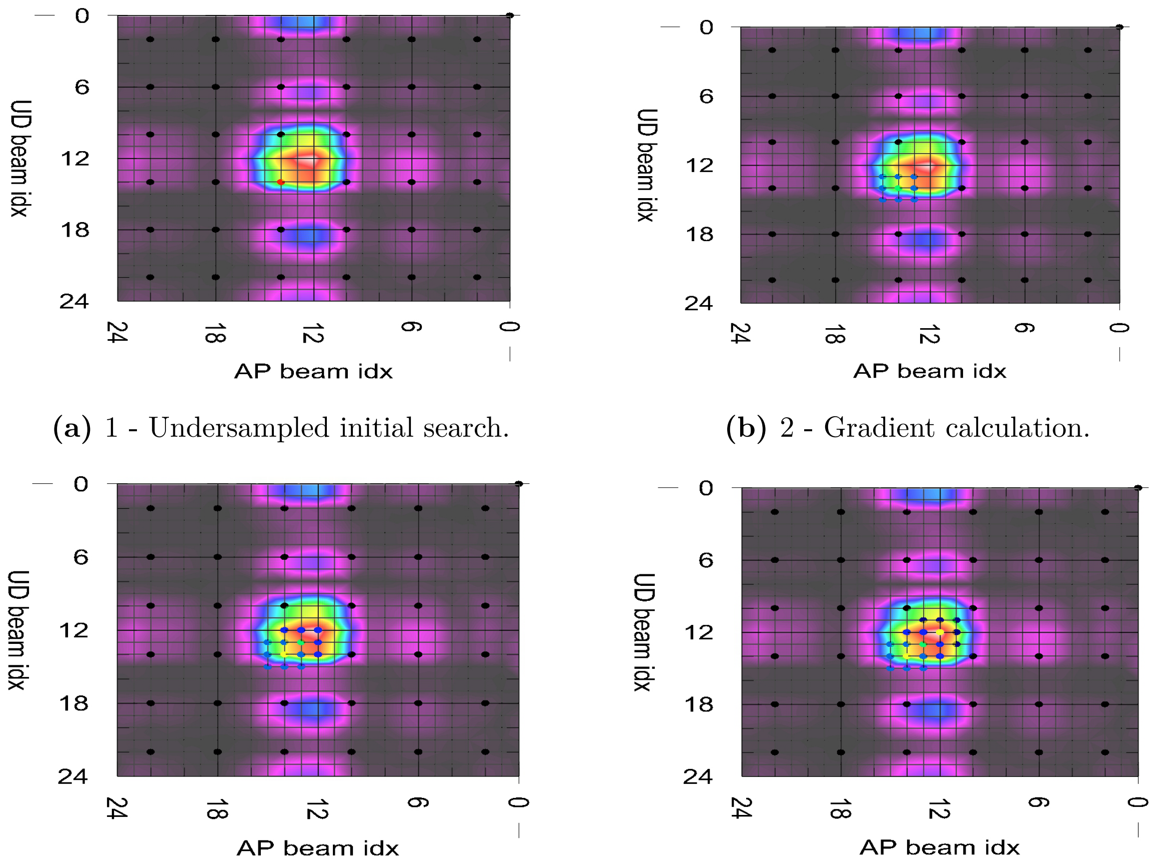


Figure 5.14: Example where a large parameter μ_n causes mismatch.



(a) 1 - Undersampled initial search.

(b) 2 - Gradient calculation.

(c) 3 - Candidate probing and gradient calculation.

(d) 4 - 2nd candidate and termination.

Figure 5.15: Individual steps of the gradient beam alignment method. Screenshot from actual LabVIEW implementation.

the next RF iteration. If the candidate is at the border of the available beams, the candidate beam is used instead for gradient calculation.

5.7 Experimental Validation

In order to test the practicability of the proposed algorithm, we compared simulation and experiments by using the algorithm. Unfortunately, the frame structure does not allow to use adaptive probing intervals for one beam-combination as discussed in chapter 3, therefore we used a training signal of 512 samples in the simulation to adapt to the requirements of the implementation.

5.7.1 Simulation Parameters

The simulation was performed using a Matlab reference implementation. The parameters were chosen to match the requirements of the experimental setup in order to be able to compare the results (Tab. 4.1). The transmitter and receiver are assumed to have a linear array of each $N_{\text{tx}} = N_{\text{rx}} = 5$ antenna elements. We use the channel model introduced in (2.15). We assume to have $L = 2$ paths with a complex Gaussian distributed amplitude and independently uniformly distributed AoA ϕ_{A} and AoD ϕ_{D}. We use simulated beam patterns with angular distance of $\Delta\phi_{\text{tx}} = \Delta\phi_{\text{rx}} = 5°$. Assuming a transmit power of $P_{\text{tx}} = 16$ dBm as in the experiment and thermal noise based on the bandwidth of $B = 750$ MHz, the simulation is performed for different distances and corresponding path losses. Furthermore, a pseudo-noise sequence with 512 samples was used to achieve an estimate of the receive power level in order to be compatible with the proposed frame structure.

5.7.2 Test Results

The experiments were performed on an indoor scenario for different distances, where one wall with windows served as a reflector of a secondary NLOS path additionally to the LOS path (Fig. 5.17). While the main direction of the AP and UD face each other, the direction of the antenna array of the UD was shifted with respect to the normal direction to $-45°, -20°, 0°, 20°, 45°$. That slightly changed the effect of the secondary path to the channel so that the algorithm could be tested with various channel conditions. In contrast, the simulation was performed using random AoAs, AoDs and path strengths.

As before, the performance criterion of the algorithm was chosen to be the received power gained by the beam combination of the gradient-based algorithm compared to the received power achieved by the solution of the exhaustive search $P_{rel} = \frac{P_{gr}}{P_{es}}$. Additionally, the number of beam pair evaluations the algorithm needs in order to finish of the proposed algorithm compared to the exhaustive search $ev_{rel} = \frac{ev_{gr}}{ev_{es}}$ was used as complexity criterion. The experimental setup is depicted in Figs. 5.17 and 5.18.

Fig. 5.16 shows the result in terms of relative receive power and relative number of iterations for the gradient beam alignment procedure of both the experiments and the simulations. The relative receive power for the simulated channel is almost constant about 98% of normalized gain, which is in line with the simulations performed in Fig. 5.11. In contrast, the achieved relative gain of the experiment varies in dependency of the direction of the receiver and the distance. The reason for this that a rotation and movement of the receiver shifts the relative position of the scatterers with respect to the receiver and thus

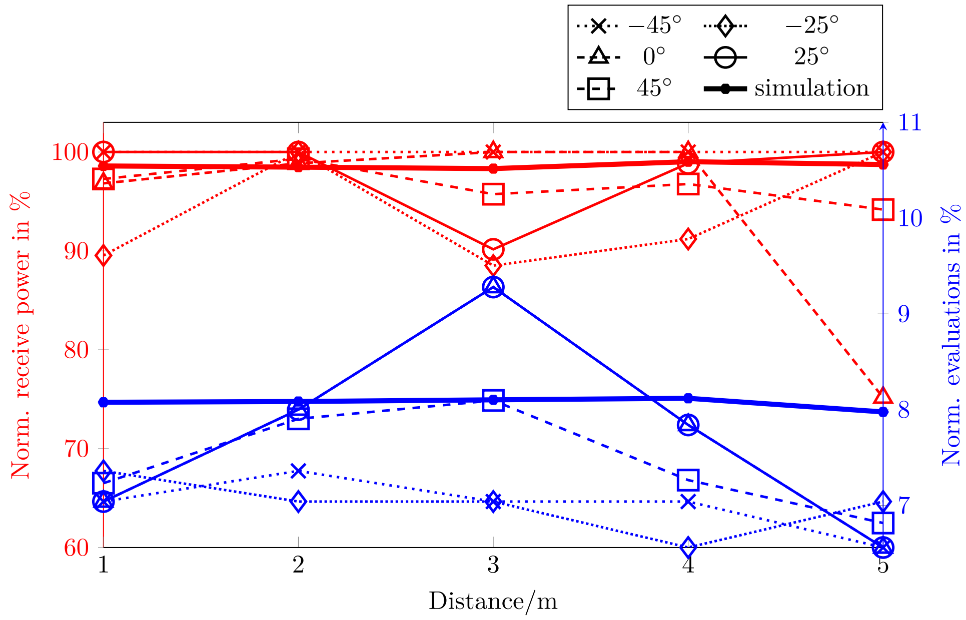


Figure 5.16: Achieved gain and evaluation numbers of proposed beam alignment compared to exhaustive search. (Source: [KRCF18] © 2018 IEEE).

experiences a new channel realization, which could have different properties. For example, a strong secondary path caused by one reflection of the glass may have an impact on the performance of the algorithm, as it may stop at a local optimum. Another issue is that the beam patterns significantly vary from the ideal shape. Especially the increased side lobe level has a significant influence on the gradient-based optimization approach as had been discussed in Section 4.3.1. Still, the relative power in the experiments is mostly higher than 90%.

The number of evaluations for the simulation is around 8% of the evaluations needed by the exhaustive search. The number of evaluations for the beam alignment varies in the experiment but is in the same region (6% to 9% of the exhaustive search evaluations) as the result for the simulation. This result shows that the gradient has to take an approximately constant number of iterations to find a local or global optimum and the algorithm behaves stably.

The experiments could be only performed up to a distance of 5 m, as a control channel was needed for the feedback. This control channel was realized using the mmW link utilizing one receive antenna to avoid a complicated beam alignment procedure. Due to the lack of directionality for the control link, the increase of distance beyond 5 m caused unreliable feedback and synchronization between the transmitter and receiver. This effect, however, is more related to the actual realization of the control feedback channel in the demonstrator than to the design of the algorithm.

Summarizing, the results show the trend that the performance of the experiments and the simulations are comparable and thus the algorithm is applicable in practice. The reason in the deviation can be explained mainly by the non-ideal beam patterns with higher side lobes. Nonetheless, the results prove that the algorithm is able to work even under non-ideal conditions.

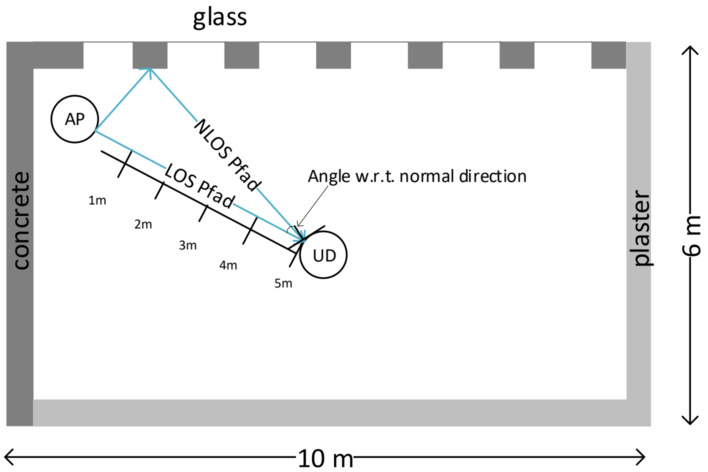


Figure 5.17: Experimentation setup scheme. Apart from LoS connection, the windows reflect a NLoS signal. The distance between the two nodes is varied from 1m to 5m in the experiment.

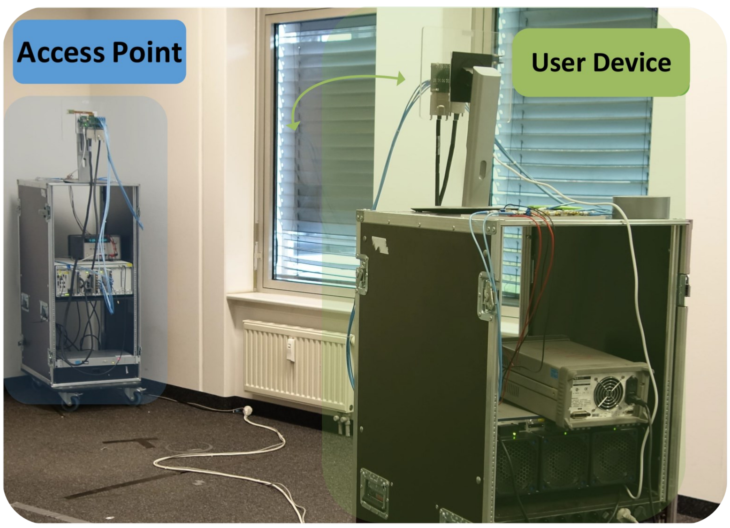


Figure 5.18: Experimentation setup picture. The potential of window reflection is visible. (Source: [KRCF18] © 2018 IEEE).

5.8 Summary

Aligning the beams in a time efficient way is crucial in high-frequency point-to-point communications systems. Using ABF in combination with a fixed codebook for the transmitter of size K and L respectively, the possibilities of creating efficient algorithms is limited. The commonly used ES algorithm uses $K \cdot L$ probing slots, increasing with the square of the size of the array and thus with the cardinality of the codebooks. As large arrays are needed to cope with the large path loss especially for future outdoor cellular applications, we presented two beam alignment algorithms to decrease the number of probing slots needed to find the best beam pair. The first algorithm uses the mathematical concepts of black-box optimization to determine the optimum of the function. It shows that the spectral efficiency of the ES algorithm can be almost approached by using about 15.6% of the training compared to ES and four feedbacks in our example. The performance-to-evaluation ratio can be adapted by adjusting various parameters. However, this algorithm is too complicated to be implemented. The second algorithm is based on finding the appropriate region of the cost function $J(k, l)$ around the global maximum and then uses obtained gradient information to reach the global maximum. Simulations show that, if adequately parametrized, only about 7.2% of pilot slots have to be probed compared to ES, by still achieving similar spectral efficiency. However, the application of this algorithm is limited to a small number of MPCs and a considerable overlap of the beam patterns in order to determine the region of interest of the function reliably. This algorithm was also implemented in the testbed. The theoretical performance could be verified in general also in the experimental test over several channel realizations resulting of the rotation of the array, although the impairments on the beam patterns caused a decrease of the performance in some cases.

Chapter 6

Beam Tracking for Millimeter Wave Systems

In contrast to static communications, the support of mobility is only given if beams of the AP and UD are continuously adapted in such a way that the connection is stable also under mobility. Few works for tracking are already available in the literature. The beam tracking algorithm developed in [PDW17] uses omni-directional transmission and a HBF system scheme. Also [HKG+14] uses HBF. Omni-directional transmission limits the maximal distance of reception by the high path loss, while HBF is complex to realize due to the use of several RF chains. The works [ZGF16], [JMCK17] and [VVH16] employ ABF with one RF chain using a Kalman filter approach. However, also the perfect access to phases of the individual antenna elements is quite complex to realize. A system with a simpler ABF structure with a given codebook is presented in [HPR+15], which uses an algorithm to probe adjacent beams.

Consequently, most mmW beam tracking experiments described in literature rely also on this simple system design. Experiments for the 28 GHz regime have been done in [OOA+15, OIA+16], investigating a HBF system with a simple exhaustive search like tracking algorithm in outdoor scenarios. Notable experimental results for cellular-type application using the 70 GHz band were shown in [IKS+15b, IKO+15, YIS+16, IYK+17] focusing on realistic closed-loop SC communication using one-sided search. While [IKS+15b] describes simple beam alignment and beam tracking algorithms of the system and shows the general applicability of the concept, [IKO+15] explores properties of the channel of an indoor shielded environment by comparing to an anechoic chamber, claiming that it is sufficient to track adjacent beams for dominant LoS connections. Experiments in LoS outdoor environments were conducted in [YIS+16] showing a system which is capable of dealing with higher velocities. Within these papers, the tracking problem is reduced to a dominant LoS channel in simple environments. Experiments in more complex settings were performed in [IYK+17], first showing a result in a NLoS environment. However, the algorithm was not changed due to the good-natured channel, which did not require fast switched beam changes. Throughout the experiments, the tracks investigated were mainly simple lines without curves.

This chapter motivates the need for an extension to a simple codebook based beam tracking algorithm. The investigations show that even in ill-conditioned NLoS channels good performance can be achieved by the proposed scheme, and experiments with a mobile UD show the feasibility of the algorithm in real-world applications and its robustness to impairments.

6.1 Tracking Algorithm

The main optimization criterion of the tracking algorithm is to maximize the level of power or increase the power level compared to the last channel block. Like in [OOA+15] this goal is achieved by searching a set of promising beam pair combinations. In contrast to the paper above, we do not only consider the adjacent beam pairs of the current candidate, but also a regular grid of beam pair combinations. Mathematically speaking, the procedure is

$$(k,l)_{n+1} = \arg \max_{\substack{\mathbf{f}_k \in \tilde{\mathbf{F}}_n \\ \mathbf{w}_l \in \tilde{\mathbf{W}}_n}} |y(k,l)|^2, \tag{6.1}$$

for finding the maximum of all probed beam pairs. For a typical movement, the neighboring beam pairs of the current optimum beam pair are good candidates to search. Additionally, due to the dominant effect of shadowing and the limited diffraction effect at mmW frequencies together with the usage of high directive beams, it is likely that the path corresponding to the best beam pair is blocked or the connection quality is degraded at some time. In order to search for new paths, which may occur to the usage of scattering objects, we additionally probe the other beam pairs in a regular shape. We use every d^{th} transmit and receive beam pair for each channel block, resulting in probing a grid of beam pair combinations. While d stays constant the adjacent transmit and receive beams are used for probing at the following channel block. Thus, all beam pair combinations are probed in d^2 channel blocks. The set of beam candidates at channel block n is given as

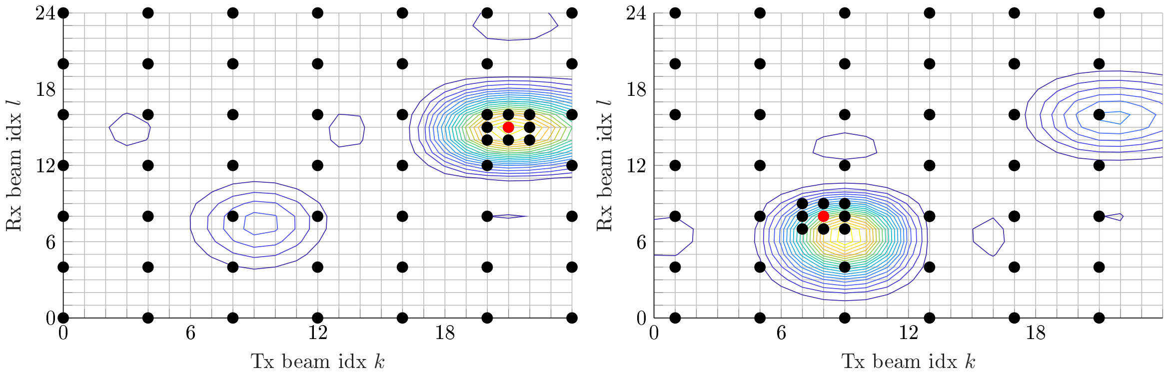
$$\begin{aligned}\tilde{\mathbf{F}}_n = \mathbf{f}_k | k = &\underbrace{\left\{ n_{\text{mod}_\text{f}}, n_{\text{mod}_\text{f}} + d, n_{\text{mod}_\text{f}} + 2d, ..., n_{\text{mod}_\text{f}} + \left(\left\lfloor \frac{K - n_{\text{mod}_\text{f}} - 1}{d} \right\rfloor \right) \cdot d \right\}}_{\text{grid}} \\ &\cup \underbrace{\{k_n - 1, k_n, k_n + 1\}}_{\text{neighboring beams}}\end{aligned} \tag{6.2}$$

$$\begin{aligned}\tilde{\mathbf{W}}_n = \mathbf{w}_l | l = &\left\{ n_{\text{mod}_\text{w}}, n_{\text{mod}_\text{w}} + d, n_{\text{mod}_\text{w}} + 2d, ..., n_{\text{mod}_\text{w}} + \left(\left\lfloor \frac{L - n_{\text{mod}_\text{w}} - 1}{d} \right\rfloor \right) \cdot d \right\} \\ &\cup \{l_n - 1, l_n, l_n + 1\} ,\end{aligned} \tag{6.3}$$

where d defines the coarseness of the evaluated grid and has to be smaller than K and L and $n_{\text{mod}_\text{f}} = n \mod d$ and $n_{\text{mod}_\text{w}} = \lfloor \frac{n}{d} \mod d \rfloor$. We use $d = 4$ like in the last chapter.

In summary, the algorithm enables reliable NLOS communication with mobility. Contour figures of an exemplary channel realization are shown in Fig. 6.1.

In order to prove the applicability of the simple algorithm, we performed simulations by using the time-varying channel model presented in Sec. 2.2. If not stated otherwise, the same parameters as for the experimentation setup are used (Tab. 4.1). Like



(a) Receive power for all Tx and Rx beam combination for a channel realization at time n. Black dots depicts probing. of the algorithm.

(b) Receive power for all Tx and Rx beam combination for a channel realization at time n+1. Thanks to the grid the new path is discovered.

Figure 6.1: Contour plots of receive power for 2 subsequent channel realizations for $d = 4$. (Source: [KCF18] © 2018 IEEE).

in the experimentation setup, we assume a system with $N_{\text{tx}} = N_{\text{rx}} = 5$ antenna elements and $K = L = 25$ beam pattern uniformly angularly distributed in a scan range of $-60°, .., \phi, ..., 60°$. We assume a speed of the UD of $v = 3$km/h and the coarseness of $d = 4$ beams of the grid. The mean distance between AP and UD is 10m. The channel block time is 1 ms. The channel is assumed to have $M_{\text{p}} = 2$ multi-path components. In order to test the efficiency of the algorithm, we compare the data rate of a system, which would always use the best beam pair combination, with the data-rate achieved with the proposed tracking algorithm. Fig. 6.2 shows the results. It can be seen that the deviation between the perfect solution and our approach is marginal while tracking without grid loses about 15% performance due to blocking in this scenario.

6.2 Implementation

The implementation restrictions are mainly the same as discussed in the previous section. However, another effect is that the design and particularly the duration of the radio frame has a significant influence on the tracking capabilities. Additionally, the delay time between scheduling the desired beam indices and actually probing them also affects the overall reaction time, and therefore performance losses are expected. The implementation of the beam tracking algorithms is carried out within the same framework as the implementation of the alignment algorithms. For this reason, only the implementation of the actual beam tracking algorithms has to be discussed here.

The main beam tracking VI contains all relevant functions in order to realize the beam tracking algorithm described above. A sketch of a block diagram and a data flow diagram can be seen in Figs. 6.3 and 6.4, respectively. According to the restrictions and realistic assumptions, the implementation of the algorithm has been extended in some details. At first, the maximum receive power of any probed beam combination (by using

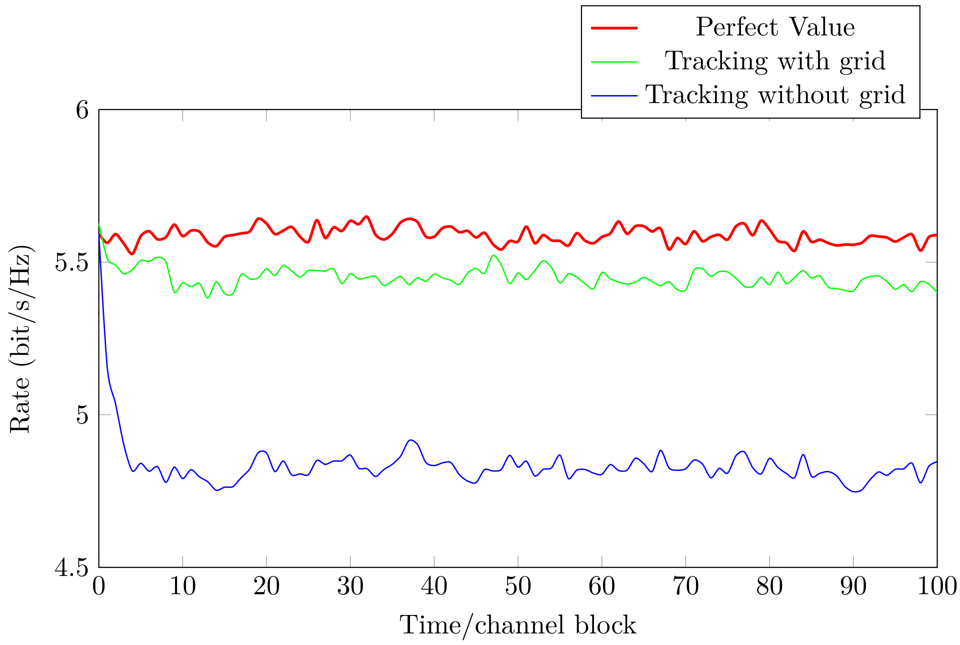


Figure 6.2: Simulation of achievable rate for perfect channel knowledge, probing neighboring beams and neighboring beams with grid (proposed). (Source: [KCF18] © 2018 IEEE).

the newest measurements) is determined. In the initial step, this equals to the outcome of the beam alignment algorithm. The received power is compared to the one obtained in the last radio frame and is changed if the receive power difference exceeds a certain amount. It is also checked whether the received power allows minimal communication or whether the beam alignment has to be restarted. If this condition is fulfilled, a schedule for probing all adjacent beams is created. Additionally, this schedule is combined with a schedule which probes a subset of all possible beam combinations in such a way that all beam combinations are probed after a certain number of radio frames (exhaustive search over several radio frames). That ensures that also hidden paths are discovered.

These schedules are then combined and serve as a basis for the tracking algorithm in future radio frames. These steps are described in detail below.

1. *Tracking check values.* The tracking check VIs purpose is to obtain the maximum SNR of all valid beam combinations which are already measured. Therefore, the arrays are used to store the status of the beam pairs. One array contains the information of which beam indices are already probed at all, another array contains the actual received power values, and a third structure contains the information of the radio frame at which the beam indices are scheduled. In order to use only valid and current measurements, only the indices of the arrays whose index is valid and where the radio frame on which it was scheduled not too old compared to a maximum delay constant are used for finding the maximum receive power.

2. *Schedule adjacent beams.* Schedules neighboring beam indices of the current beam combination index. If the beam indices are out of the range of possible beam indices, they are skipped.

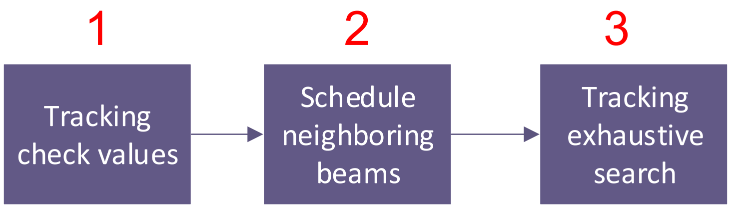


Figure 6.3: Block diagram of the tracking implementation.

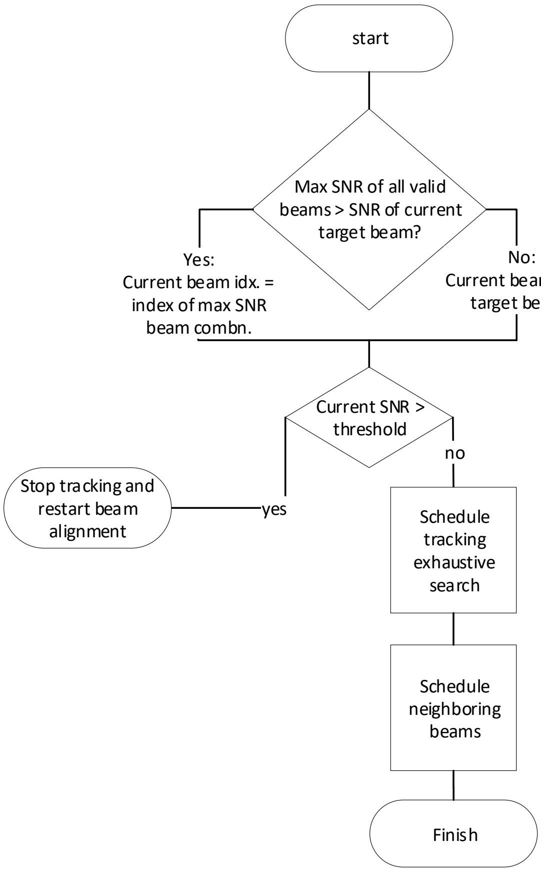


Figure 6.4: Data flow diagram of the tracking implementation.

3. *Tracking exhaustive search.* It has the purpose to schedule all possible beam combinations in a certain time frame. Therefore, an adapted version of the exhaustive search is used where only a subset of beam pairs according to a regular grid is probed. This method ensures that a second multipath component could be detected with high probability in a reasonable time and increases robustness to blockage as reasonable data is always available to exploit major changes in the characteristics of the channel. The density of the points (coarseness of the grid) can be set by a parameter.

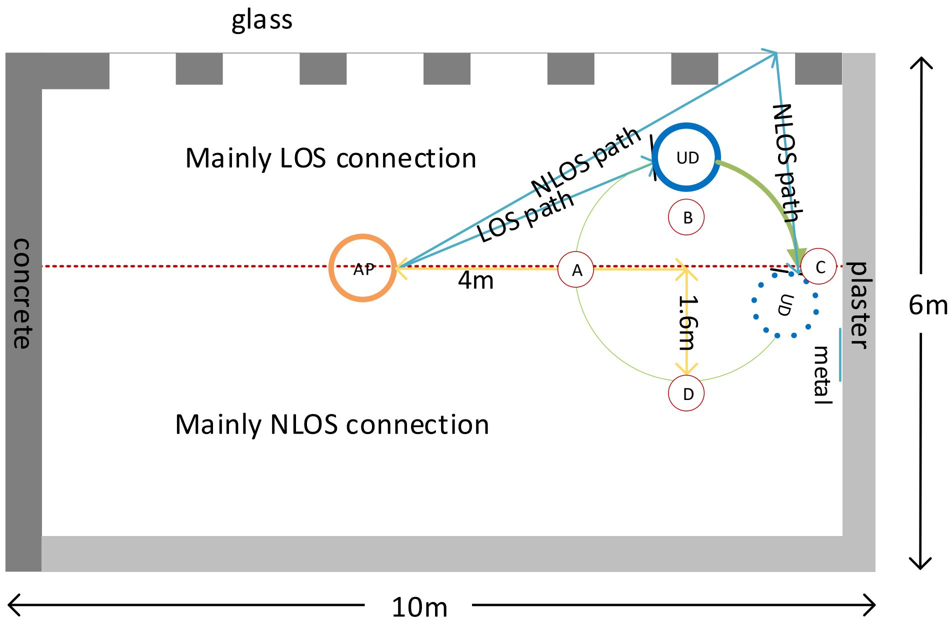


Figure 6.5: Measurement site. A) Begin of course. B) LoS and NLoS path. C) LOS path vanishes. D) Dominant NLoS reflection caused by metal inside plaster wall.

6.3 Experimental Verification

We performed a mobility experiment with static AP and mobile UD using a robot. The robot followed a circular track with a velocity of about 0.8 km/h (Fig. 6.5). As it was an indoor experiment, it was possible to model the effect of a LoS and NLoS connection and also the blockage of paths. The antenna was mounted at the rear of the robot and pointing in the opposite direction than the movement was performed (Fig. 6.6). The track started at the closest point towards the AP (position A at Fig. 6.5).

Ray-tracing Channel Model. In order to compare the measurements with simulations, we created a ray tracing model of the experimentation track. The main parameters are discussed in Sec. 2.2. We remodeled the room for the experiment, modeled the strength for each MPC for each point of the track received by the UD and simulated the received power of the array with five antenna elements, assuming the relative permittivity of granite given in Tab. 2.1 for the walls. The goal is twofold. First, we want to evaluate how precisely the modeling can be considered and whether the beam tracing algorithm misses some points. Second, the model gives insight into the change of the dominant paths over time and the effect of NLOS to the beam tracking algorithm. Fig. 6.7 shows the simulated setup.

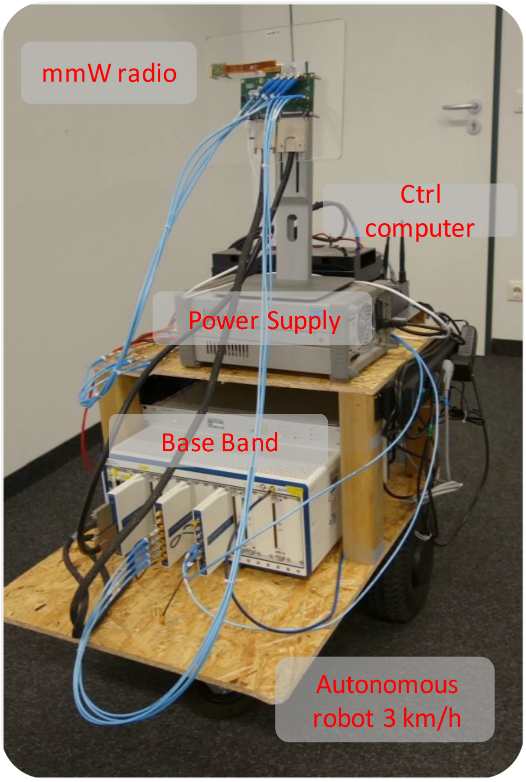


Figure 6.6: UE with robot. (Source: [KCF18] © 2018 IEEE).

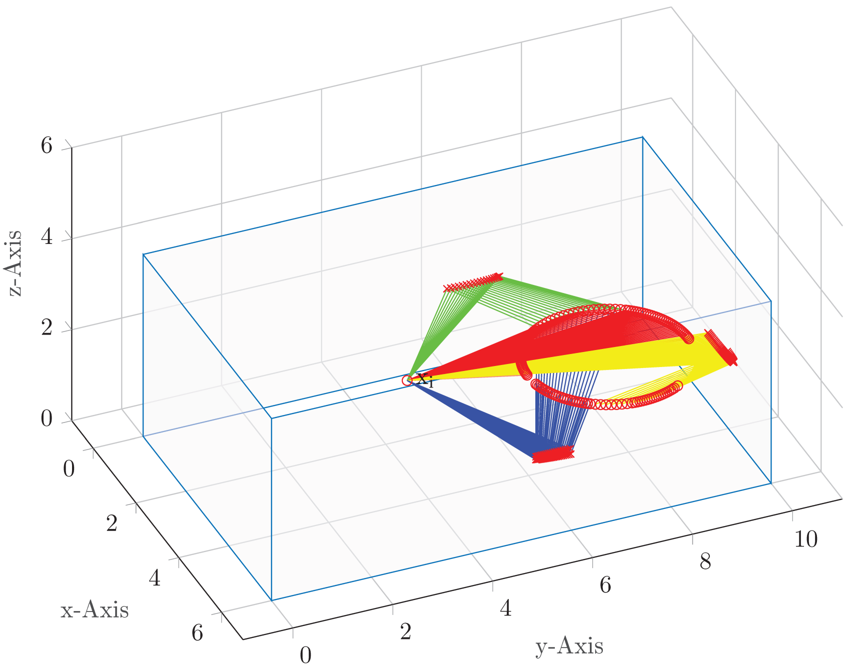


Figure 6.7: Ray tracing simulation. The red dots represent the track of the UE, while the other colors represent different MPCs. (Source: [KCF18] © 2018 IEEE).

Experimental Results. Measurements for the instantaneous transmit and receive beam of both devices and the corresponding DL transmission are depicted in Figs. 6.8-6.10 for both the measurements (beam tracking algorithm) and the ray tracing simulation (exhaustive search). As the automatic gain control (AGC) had to be used in the measurement device, absolute receive power values are not a meaningful performance measure of the quality of the link. Therefore, we use the throughput as performance measure.

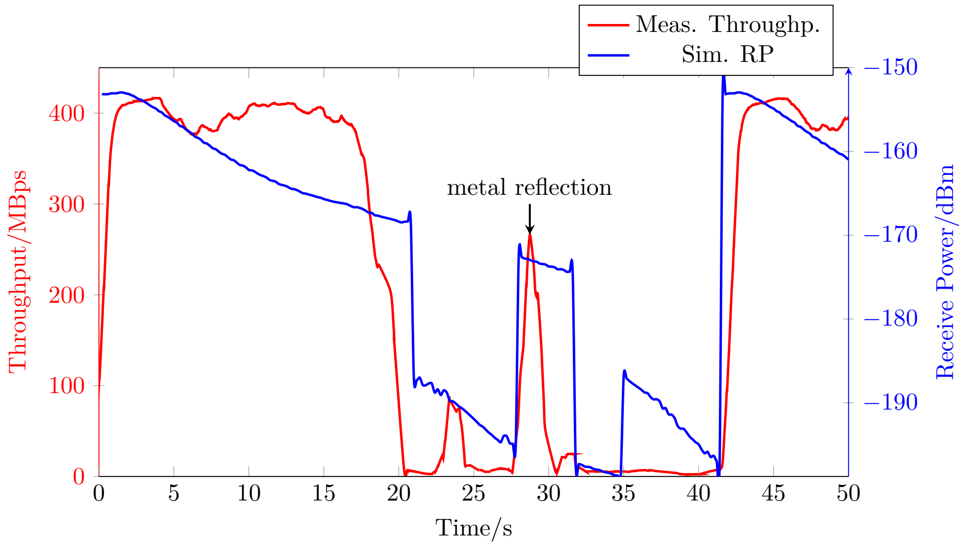


Figure 6.8: Measurements of throughput compared with the gain of the ray racing simulation. (Source: [KCF18] © 2018 IEEE).

The transmit beams have better characteristics than the receive beams due to lower side lobes. Thus the transmit beam number reflects the AoD quite well while larger errors are assumed for the linking between receive beam index and AoA. The large side lobes can cause fast changes for the receive beam index corresponding to the highest receive power. At the beginning of the measurement, the UD has LoS connection to the AP (position A in Fig 6.5). It chooses a beam with maximum deviation to the normal direction of the UD array. When the UD is driving to position C it changes the beam to the other side, meaning the beams are switching from beam 25 ($\phi_A = 60°$) to beam 1 ($\phi_A = -60°$), which can be seen in Fig. 6.10 from 0s to 20s. Meanwhile, the AP beam is just changed slightly, as the angle change from its point of view is not high (Fig. 6.9). At the movement from position A to C the LoS connection ensures high throughput (Fig. 6.8).

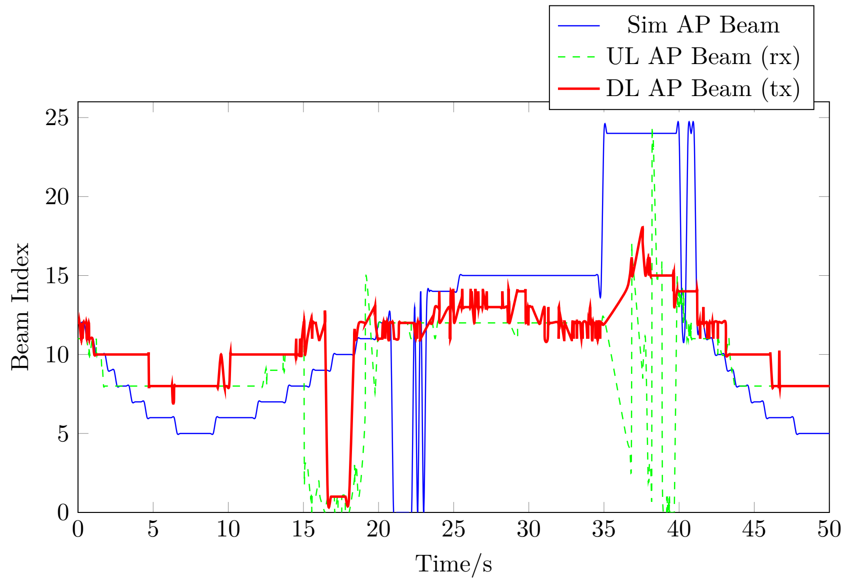


Figure 6.9: Measurements of the chosen AP beam compared with ray racing simulation. (Source: [KCF18] © 2018 IEEE).

After 20s, the robot is going into the NLOS dominated part of the room (position C). Thus a new path has to be considered, and still, a connection can be established. However, the plaster wall only provides very weak reflection, so that only limited throughput can be achieved in the time interval from 20s to 40s, while control signaling is still possible. It is interesting that the communication can automatically detect and exploit a metal part in the wall, where the reflection coefficient and the resulting throughput significantly rises. As the dominant path is reflected mostly from the wall opposing the AP, the same linear decrease in the UD beam number can be seen, which is interfered by other reflections in the latter part of the NLoS path. The throughput rises when entering the LoS region again. It is noted that the path gets blocked, or new paths are created several times in this circular track, even with limited numbers of reflectors. Thus, we may conclude that the grid in the algorithm is mandatory. It is also shown that reliable communication is also possible with mobility under NLoS cases although the data transmission is severely limited there. The result of the experiments also give insight into possible use cases in street canyon environments if the relation between beam-angle and wall-distance remains the same and as not many scatterers are existent in this indoor scenario. The comparison of the beam tracking measurements and the ray tracing model suggest the conclusion that the simple algorithm is sufficient to solve the tracking problem for moderate speeds of the UD even in a NLOS scenario with blockage. Some variations of the simulated and the measured UD beam indices may originate from a higher reflection coefficient of the plaster wall than assumed by the simulation.

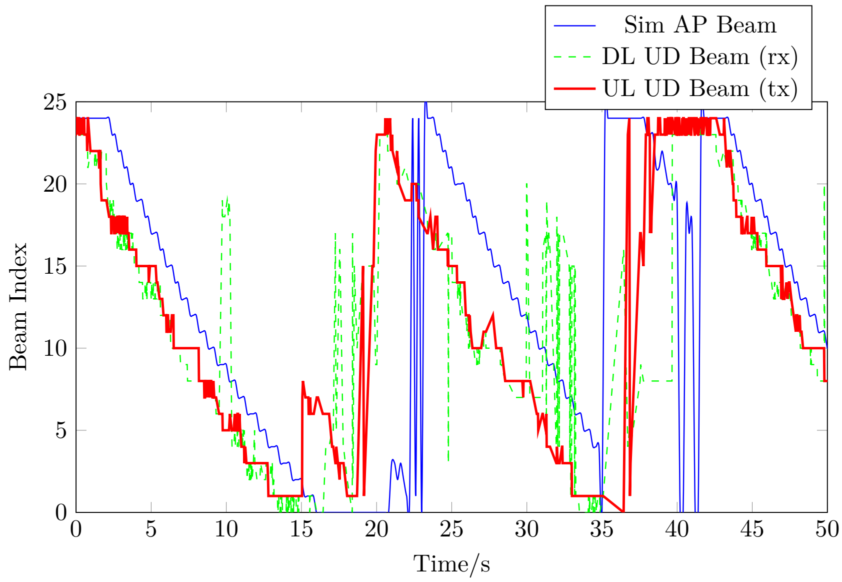


Figure 6.10: Measurements of the chosen UD beam compared with ray racing simulation. (Source: [KCF18] © 2018 IEEE).

6.4 Summary

In this chapter, a simple but effective tracking mechanism motivated by simulation was developed. In addition, it was shown to work experimentally with a real-time mmW system. The experimental validation showed that a connection between the AP and UD can be maintained even under NLoS connections and mobility of the UD while throughput is limited due to the lack of effective scatterers.

Chapter 7

Conclusions and Future Works

In this thesis, we presented and investigated strategies for practical analog beamforming (ABF) systems to cope with a directional transmission. Due to the high path loss, the use of high gain antennas is mandatory in millimeter wave (mmW) cellular systems. We presented methods for optimizing the training length and for effective and robust beam alignment and beam tracking algorithms. Furthermore, we investigated the influence of the training length to the performance of the beam selection. Additionally, we experimentally demonstrated algorithms in static and mobile scenarios. The work was published partly in [KCF18, KRCF18, KCF17, KCF16]. The contributions of this work can be summarized as follows:

- ▷ We started with an investigation on the millimeter wave (mmW) channel and deduced several channel models. We introduced ABF as a promising low-complex architecture for simple mmW applications.
- ▷ Interpreting the beam alignment problem as an M-ary hypothesis test, mathematical expressions using the Q-functions of the probability of error for missing the best beam pair can be derived, given CSI knowledge.
- ▷ Without full channel state information (CSI) knowledge, composite testing allows using maximum likelihood (ML) estimates of the radio channel properties to derive the selection probability. The normalized amplitude was found to be a more suitable performance measure.
- ▷ Based on this theoretical approach, two adaptive algorithms have been proposed to define the training signal length effectively, in order to reach a predefined targeted normalized amplitude. The algorithms' performance was demonstrated by simulation.
- ▷ Black-box algorithms are suitable tools to cope with non-convex optimization problems. Therefore they can be used for the beam alignment problem in the context of codebook based ABF systems. Using a specific black box approach, the mode pursuing sampling method for a discrete variable space (D-MPS), a beam alignment algorithm was proposed. It was shown to be effective by using only 15.6% of the training slots of that of conventional exhaustive search (ES) at the cost of several additional feedbacks.

- ▷ In addition, a gradient-based beam alignment algorithm was proposed for codebook-based ABF systems with low complexity. When the parameters of the algorithms are chosen suitably according to the channel properties, it was found to be effective for angular sparse channels. Experiments could validate the superiority of this algorithm by reducing the probing overhead to 8% of that of ES plus feedback.
- ▷ A low-complexity tracking algorithm was developed and experimentally verified despite serious impairments of the measured beam patterns.
- ▷ The comparison between ray-tracing simulations and experimental measurements shows that the channel can be accurately modeled by using quasi-optical laws and solely taking into account first order reflections.

7.1 Future Works

The experimental part of this work was limited to an ABF system. With the help of advanced hybrid beamforming (HBF) systems, more sophisticated channel estimation and precoding algorithms could be examined for practical applicability. Additionally, the following problems may be worth to be considered in future research:

- ▷ The role of mmW impairments like enhanced phase noise and antenna imperfections are rarely considered in the literature yet. Specifically, the impact of these impairments on beam alignment and beam tracking algorithms is an important topic for future commercial devices. A detailed investigation of this topic was not possible in this thesis due to the complex nature of the real-time demonstrator.
- ▷ The fact that the connection can only be established after successful beam alignment makes it necessary to rely on a secondary, non mmW link for certain applications. The interaction between the mmW link and a traditional cellular link is an interesting and important topic.
- ▷ This thesis limits the experimental validation to indoor environments due to constraints of the testbed. Beam alignment and beam tracking have to be verified in typical outdoor scenarios in the context of cellular applications.

Appendix A

Derivation of the Selection Probability of the M-ary Hypothesis Test

We can write the detection event $P\left(\bar{z}_m = \max_{i\in\{1...M\}\backslash m} \bar{z}_i\right)$ as an intersection of the sub-events $E_i = \bar{z}_m > \bar{z}_i$

$$P_{\mathrm{s}}^m = P\left(\bar{z}_m = \max_{i\in\{1...M\}\backslash m} \bar{z}_i\right) = P\left(\bigcap_{i\in\{1...M\}\backslash m} \underbrace{\bar{z}_m > \bar{z}_i}_{E_i}\right). \tag{A.1}$$

All the events have to be fulfilled at the same time. As each signal $\bar{z}_i$ is Gaussian distributed and thus also the difference of two signals is Gaussian, we can write the detection probability as an integral over a multivariate normal distribution with $M-1$ dimensions with $\bar{\mathbf{z}}_{\mathrm{d}} = [\bar{z}_m - \bar{z}_1, ..., \bar{z}_m - \bar{z}_i, ..., \bar{z}_m - \bar{z}_M]^T$, where we use the differences of the individual signals and constant bounds for the integral

$$\begin{aligned}
&P\left(\bigcap_{i\in\{1...M\}\backslash m} \bar{z}_m - \bar{z}_i > 0\right) = \\
&\int\limits_{\bar{z}_m-\bar{z}_1=0}^{\infty} ... \int\limits_{\bar{z}_m-\bar{z}_M=0}^{\infty} f_{\bar{\mathbf{z}}_{\mathrm{d}}}(\bar{\mathbf{z}}_{\mathrm{d}}) d\bar{\mathbf{z}}_{\mathrm{d}}, \quad f_{\bar{\mathbf{z}}_{\mathrm{d}}}(\bar{\mathbf{z}}_{\mathrm{d}}) \sim \mathcal{N}_M\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \\
&\text{with :} \\
&\boldsymbol{\mu} = \begin{bmatrix} \mu_{m1} \\ ... \\ \mu_{mM} \end{bmatrix} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 2\bar{\sigma}^2 & ... & \bar{\sigma}^2 \\ ... & ... & ... \\ \bar{\sigma}^2 & ... & 2\bar{\sigma}^2 \end{bmatrix}.
\end{aligned} \tag{A.2}$$

However, the sub-events E_i and the difference of the signals $\bar{z}_m - \bar{z}_i$ are not independent and therefore the integral can not be simplified. If we use the individual and independent signals $\bar{z}_i$ as integration variables instead, we have to use the integration variable $\bar{z}_m$ as bound with $\bar{\mathbf{z}} = [\bar{z}_1, , ...\bar{z}_i, ...\bar{z}_M]^T$

$$P\left(\bigcap_{i\in\{1...M\}\backslash m} \bar{z}_m > \bar{z}_i\right) \tag{A.3}$$

$$= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\bar{z}_m} f_{\bar{\mathbf{z}}}(\bar{\mathbf{z}}) d\bar{\mathbf{z}}, \quad f_{\bar{\mathbf{z}}}(\bar{\mathbf{z}}) \sim \mathcal{N}_M\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \tag{A.4}$$

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_m \\ ... \\ \mu_i \end{bmatrix} \quad \boldsymbol{\Sigma} = \bar{\sigma}^2 \mathbf{I} \tag{A.5}$$

In this way, using the decoupling of the multivariate normal distribution into a product of M independent normal distributions, we can use the Q function to simplify the representation. This leads to

$$P_{\mathrm{s}}^m = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} e^{\frac{-(\bar{z}_m-\mu_m)^2}{2\bar{\sigma}^2}} \cdot \tag{A.6}$$

$$\prod_{i\in\{1...M\}\backslash m} \underbrace{\int_{-\infty}^{\bar{z}_m} \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} e^{\frac{-(\bar{z}_m-\mu_i)^2}{2\bar{\sigma}^2}} d\bar{z}_i}_{\left(1-\mathrm{Q}\left(\frac{\bar{z}_m-\mu_i}{\bar{\sigma}}\right)\right)} \, \mathrm{d}\bar{z}_m \tag{A.7}$$

Bibliography

[3GP17] 3GPP. Study on channel model for frequency spectrum above 6 GHz. Technical Specification (TS) 38.900, 3rd Generation Partnership Project (3GPP), 07 2017. Version 14.2.0.

[AALH13] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath. Hybrid precoding for millimeter wave cellular systems with partial channel knowledge. In *2013 Information Theory and Applications Workshop (ITA)*, pages 1–5, Feb 2013.

[AALH14] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath. Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831–846, Oct 2014.

[ABB+07] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyosti, D. Laurenson, G. Matz, A.F. Molisch, C. Oestges, and H. Ozcelik. Survey of Channel and Radio Propagation Models for Wireless MIMO Systems. *EURASIP Journal on Wireless Communications and Networking*, 2007(1):019070, Feb 2007.

[AH16] A. Alkhateeb and R. W. Heath. Gram Schmidt based greedy hybrid precoding for frequency selective millimeter wave MIMO systems. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3396–3400, March 2016.

[ALH15] A. Alkhateeb, G. Leus, and R. W. Heath. Compressed sensing based multi-user millimeter wave systems: How many measurements are needed? In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2909–2913, April 2015.

[ALS+14] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip. Millimeter Wave Channel Modeling and Cellular Capacity Evaluation. *IEEE Journal on Selected Areas in Communications*, 32(6):1164–1179, June 2014.

[ARAS+14] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath. Spatially Sparse Precoding in Millimeter Wave MIMO Systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, March 2014.

[AWW+13] Y. Azar, G. N. Wong, K. Wang, R. Mayzus, J. K. Schulz, H. Zhao, F. Gutierrez, D. Hwang, and T. S. Rappaport. 28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York city. In *Proceedings IEEE International Conference on Communications*, pages 5143–5147, June 2013.

[Bal05] Constantine A. Balanis. *Antenna Theory: Analysis and Design.* Wiley-Interscience, 2005.

[Ban99] Rajeev Bansal. The Far-Field: How Far is Far Enough?, 1999.

[Bea06] N. C. Beaulieu. A useful integral for wireless communication theory and its application to rectangular signaling constellation error rates. *IEEE Transactions on Communications*, 54(5):802–805, May 2006.

[BHR+15] C. N. Barati, S. A. Hosseini, S. Rangan, P. Liu, T. Korakis, S. S. Panwar, and T. S. Rappaport. Directional Cell Discovery in Millimeter Wave Cellular Networks. *IEEE Trans. on Wireless Communications*, 14(12):6664–6678, Dec 2015.

[Cap69] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, Aug 1969.

[CGK+13] M. Cudak, A. Ghosh, T. Kovarik, R. Ratasuk, T. A. Thomas, F. W. Vook, and P. Moorut. Moving Towards Mmwave-Based Beyond-4G (B-4G) Technology. In *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pages 1–5, June 2013.

[CKRF16] H. L. Chiang, T. Kadur, W. Rave, and G. Fettweis. Low-complexity spatial channel estimation and hybrid beamforming for millimeter wave links. In *Proceedings IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, pages 1–7, Sept 2016.

[CRKF17] H. L. Chiang, W. Rave, T. Kadur, and G. Fettweis. A low-complexity beamforming method by orthogonal codebooks for millimeterwave links. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3375–3379, March 2017.

[CRKF18] H. L. Chiang, W. Rave, T. Kadur, and G. Fettweis. Hybrid Beamforming Based on Implicit Channel State Information for Millimeter Wave Links. *IEEE Journal of Selected Topics in Signal Processing*, 12(2):326–339, May 2018.

[Fos98] M.J. Foschini, G.J.and Gans. On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas. *Wireless Personal Communications*, 6(3):311–335, Mar 1998.

[Fri46] H. T. Friis. A Note on a Simple Transmission Formula. *Proceedings of the IRE*, 34(5):254–256, May 1946.

[GTC+14a] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie. Millimeter-Wave Enhanced Local Area Systems: A High-Data-Rate Approach for Future Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 32(6):1152–1163, June 2014.

[GTC+14b] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie. Millimeter-Wave Enhanced Local Area Systems: A High-Data-Rate Approach for Future Wireless Networks. *IEEE Journal on Selected Areas in Communications*, 32(6):1152–1163, June 2014.

[Gus14] Carl Gustafson. *60 GHz wireless propagation channels: characterization, modeling and evaluation*. PhD thesis, Department of Electrical and Information Technology, Lund University, Lund, Sweden, 2014.

[Hem10] Thomas Hemker. *Derivative Free Surrogate Optimization for Mixed-Integer Nonlinear Black Box Problems in Engineering*. TU Darmstadt, May 2010.

[HGPR+16] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed. An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):436–453, April 2016.

[HIXR15] S. Han, C. I, Z. Xu, and C. Rowell. Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Communications Magazine*, 53(1):186–194, January 2015.

[HKG+14] J. He, T. Kim, H. Ghauch, K. Liu, and G. Wang. Millimeter wave MIMO channel tracking systems. In *Proceedings IEEE Global Communication Conference Workshops*, pages 416–421, Dec 2014.

[HPR+15] K. Hosoya, N. Prasad, K. Ramachandran, N. Orihashi, S. Kishimoto, S. Rangarajan, and K. Maruhashi. Multiple Sector ID Capture (MIDC): A Novel Beamforming Technique for 60-GHz Band Multi-Gbps WLAN/PAN Systems. *IEEE Transactions on Antennas and Propagation*, 63(1):81–96, Jan 2015.

[IEEa] IEEE P802.11ad, Part 11. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band, IEEE Computer Society, 2012.

[IEEb] IEEE Std 802.15.3c Part 15.3. Wireless Medium Access Control (MAC) and Physical Layer (PHY) SpeciïňĄcations for High Rate Wireless Personal Area Networks (WPANs) Amendment 2: Millimeterwave-based Alternative Physical Layer Extension, IEEE Computer Society, Oct. 2009.

[IKO+15] Y. Inoue, Y. Kishiyama, Y. Okumura, J. Kepler, and M. Cudak. Experimental evaluation of downlink transmission and beam tracking performance for 5G mmW radio access in indoor shielded environment. In *Proceedings IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, pages 862–866, Aug 2015.

[IKS+15a] Y. Inoue, Y. Kishiyama, S. Suyama, J. Kepler, M. Cudak, and Y. Okumura. Field Experiments on 5G mmW Radio Access with Beam Tracking in Small Cell Environments. In *Proceedings IEEE Global Communication Conference Workshops*, pages 1–6, Dec 2015.

[IKS+15b] Y. Inoue, Y. Kishiyama, S. Suyama, J. Kepler, M. Cudak, and Y. Okumura. Field Experiments on 5G mmW Radio Access with Beam Tracking in Small Cell Environments. In *Proceedings IEEE Global Communication Conference Workshops*, pages 1–6, Dec 2015.

[IYK+17] Y. Inoue, S. Yoshioka, Y. Kishiyama, J. Kepler, M. Cudak, S. Suyama, and Y. Okumura. Field Experimental Trials for 5G Mobile Communication System Using 70 GHz-Band. In *Proceedings IEEE Wireless Communications and Networking Conference Workshops* , pages 1–6, March 2017.

[JD92] Don H. Johnson and Dan E. Dudgeon. *Array Signal Processing: Concepts and Techniques.* Simon & Schuster, Inc., New York, NY, USA, 1992.

[JMCK17] S. Jayaprakasam, X. X. Ma, J. W. Choi, and S. Kim. Robust Beamtracking for mmWave Mobile Communications. *IEEE Communications Letters*, PP(99):1–1, 2017.

[JRB+17] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt, and E. Eberlein. QuaDRiGa- Quasi Deterministic Radio Channel Generator, User Manual and Documentation. Technical report, Fraunhofer Heinrich Hertz Institute, 2017.

[JSDM91] McGarrity J.S., J.J. Soraghan, T.S. Durrani, and S. Mayrargue. The ESPRIT and MUSIC algorithms using the covariance matrix. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 3285–3288 vol.5, Apr 1991.

[Kam08] K.D. Kammeyer. *Nachrichtenübertragung.* Teubner, 4 edition, 2008.

[Kay98] S.M. Kay. *Fundamentals of Statistical Signal Processing: Detection theory.* Prentice Hall Signal Processing Series. Prentice-Hall PTR, 1998.

[KCF16] T. Kadur, H. L. Chiang, and G. Fettweis. Effective Beam Alignment Algorithm for Low Cost Millimeter Wave Communication. In *Proceedings IEEE Vehicular Technology Conference*, pages 1–5, Sept 2016.

[KCF17] T. Kadur, H.-L. Chiang, and G. Fettweis. Efficient Gradient-Based Beam Alignment for Low-Cost mmW Devices. In *Proceedings IEICE SmartCom - International Workshop on Smart Wireless Communications*, May 2017.

[KCF18] T. Kadur, H. Chiang, and G. Fettweis. Experimental Validation of Robust Beam Tracking in a NLoS Indoor Environment. In *Proceedings IEEE International Conference on Telecommunications*, pages 644–648, June 2018.

[KLO+15] Y. Kim, H. Y. Lee, J. Oh, J. Lee, W. Roh, and K. Cheun. Feasibility of Mobile Cellular Communications at Millimeter Wave Frequency. In *Proceedings IEEE Global Communications Conference*, pages 1–6, Dec 2015.

[KMH+] Pekka Kyösti, Juha Meinilä, Lassi Hentilä, Xiongwen Zhao, Tommi Jämsä, Christian Schneider, Milan Narandzic, Marko Milojevic, Aihua Hong, Juha Ylitalo, Veli-Matti Holappa, Mikko Alatossava, Robert J. C. Bultitude, Yvo L. C. de Jong, Terhi Rautiainen, Juha Ebitg, Meinilä, Xiongwen Ebitg, Zhao, Mikko Uoulu, Alatossava, Veli-Matti Uoulu, Holappa, Robert Crc, Bultitude, and Crc Yvo De Jong. WINNER II Channel Models Part I Channel Models.

[KMH+07] Pekka Kyösti, Juha Meinilä, Lassi Hentilä, Xiongwen Zhao, Tommi Jämsä, Christian Schneider, Milan Narandzić, Marko Milojević, Aihua Hong, Juha Ylitalo, Veli-Matti Holappa, Mikko Alatossava, Robert Bultitude, Yvo de Jong, and Terhi Rautiainen. WINNER II Channel Models. Technical report, EC FP6, September 2007.

[KRCF18] T. Kadur, W. Rave, H. Chiang, and G. Fettweis. Experimental validation of a robust beam alignment algorithm in an indoor environment. In *Proceedings IEEE Wireless Communications and Networking Conference*, pages 1–6, April 2018.

[KRF19] T. Kadur, W. Rave, and G. Fettweis. Application of Multiple Hypothesis Testing for Beam Selection. In *Proceedings IEEE International Conference on Communications*, June 2019.

[LLH94] B. Langen, G. Lober, and W. Herzig. Reflection and transmission behaviour of building materials at 60 GHz. In *Proceedings of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Wireless Networks*, volume 2, pages 505–509 vol.2, Sep. 1994.

[Lyo04] Richard G. Lyons. *Understanding Digital Signal Processing (2Nd Edition)*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2004.

[MC78] J.L. Melsa and D.L. Cohn. *Decision and estimation theory*. McGraw-Hill series in electrical engineering: Communications and information theory. McGraw-Hill, 1978.

[MiW17] MiWaveS. MmWave Access and Backhaul Link Tests and Presentation of Final Demonstrator , Deliverable D6.5, 2017.

[Mol05] Andreas Molisch. *Wireless Communications*. Wiley-IEEE Press, 2005.

[MRRA+15] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath. Channel estimation and hybrid combining for mmWave: Phase shifters or switches? In *Proceedings Information Theory and Applications Workshop*, pages 90–97, Feb 2015.

[MRRGP+16] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath. Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches? *IEEE Access*, 4:247–267, 2016.

[MV12] N. Maleki and A. Vosoughi. Channel-aware M-ary distributed detection: Optimal and suboptimal fusion rules. In *Proceedings IEEE Statistical Signal Processing Workshop (SSP)*, pages 644–647, Aug 2012.

[OIA+16] T. Obara, Y. Inoue, Y. Aoki, S. Suyama, J. Lee, and Y. Okumurav. Experiment of 28 GHz Band 5G super wideband transmission using beamforming and beam tracking in high mobility environment. In *Proceedings IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, pages 1–5, Sept 2016.

[OOA+15] T. Obara, T. Okuyama, Y. Aoki, S. Suyama, J. Lee, and Y. Okumura. Indoor and outdoor experimental trials in 28-GHz band for 5G wireless communication systems. In *Proceedings IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, pages 846–850, Aug 2015.

[Pay17] Sohail Payami. *Hybrid beamforming for massive MIMO systems*. PhD thesis, University of Surrey, UK, 2017.

[PDW17] J. Palacios, D. De Donno, and J. Widmer. Tracking mm-Wave channel dynamics: Fast beam training strategies under mobility. In *Proceedings IEEE Conference on Computer Communications*, pages 1–9, May 2017.

[PH13] Gamantyo Hendrantoro Puji Handayani. A Model of Double-Directional Indoor Channels for Multiterminal Communications. *International Journal of Antennas and Propagation*, 2013:18, 2013.

[PHAH97] P. B. Papazian, G. A. Hufford, R. J. Achatz, and R. Hoffman. Study of the local multipoint distribution service radio channel. *IEEE Transactions on Broadcasting*, 43(2):175–184, June 1997.

[PM96] John G. Proakis and Dimitris G. Manolakis. *Digital Signal Processing (3rd Ed.): Principles, Algorithms, and Applications.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[PYP08] Y. Pinhasi, A. Yahalom, and S. Petnev. Propagation of ultra wide-band signals in lossy dispersive media. In *Proceedings of IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems*, pages 1–10, May 2008.

[RHDM14] Theodore S. Rappaport, Robert W. Heath, Robert C. Daniels, and James N. Murdock. *Millimeter Wave Wireless Communications.* Prentice Hall, 2014.

[RMG11] T. S. Rappaport, J. N. Murdock, and F. Gutierrez. State of the Art in 60-GHz Integrated Circuits and Systems for Wireless Communications. *Proceedings of the IEEE*, 99(8):1390–1436, Aug 2011.

[RMRGPH15] C. Rusu, R. Méndez-Rial, N. González-Prelcicy, and R. W. Heath. Low complexity hybrid sparse precoding and combining in millimeter wave MIMO systems. In *Proceedings IEEE International Conference on Communications*, pages 1340–1345, June 2015.

[RS11] V. Raghavan and A. M. Sayeed. Sublinear Capacity Scaling Laws for Sparse MIMO Channels. *IEEE Transactions on Information Theory*, 57(1):345–364, Jan 2011.

[RSM+13] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter Wave Mobile Communications for 5G Cellular: It Will Work! *IEEE Access*, 1:335–349, 2013.

[Sch11] Tom Schaul. *Studies in Continuous Black-box Optimization.* Dissertation, Technical University Munich, 2011.

[SJJS00] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst. Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel. *IEEE Journal on Selected Areas in Communications*, 18(3):347–360, March 2000.

[SMB01] M. Steinbauer, A. F. Molisch, and E. Bonek. The double-directional radio channel. *IEEE Antennas and Propagation Magazine*, 43(4):51–63, Aug 2001.

[Smu94] Peter F. M. Smulders. Deterministic modelling of indoor radio propagation at 40–60 Ghz. *Wireless Personal Communications*, 1(2):127–135, Jun 1994.

[SSS98] D. L. Sengupta, T. K. Sarkar, and D. Sen. Centennial of the semiconductor diode detector. *Proceedings of the IEEE*, 86(1):235–243, Jan 1998.

[SV87] A. A. M. Saleh and R. Valenzuela. A Statistical Model for Indoor Multipath Propagation. *IEEE Journal on Selected Areas in Communications*, 5(2):128–137, February 1987.

[SWA+13] M. Samimi, K. Wang, Y. Azar, G. N. Wong, R. Mayzus, H. Zhao, J. K. Schulz, S. Sun, F. Gutierrez, and T. S. Rappaport. 28 GHz Angle of Arrival and Angle of Departure Analysis for Outdoor Cellular Communications Using Steerable Beam Antennas in New York City. In *IEEE Vehicular Technology Conference*, pages 1–6, June 2013.

[SWE08] Behnam Sharif, G. Gary Wang, and Tarek Y. ElMekkawy. Mode Pursuing Sampling Method for Discrete Variable Optimization on Expensive Black-Box Functions. *Journal of Mechanical Design*, 130, 2008.

[SY15] F. Sohrabi and W. Yu. Hybrid digital and analog beamforming design for large-scale MIMO systems. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2929–2933, April 2015.

[Tel99] I. Emre Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, 10:585–595, 1999.

[TNMR14] Timothy A. Thomas, Huan Cong Nguyen, George R. MacCartney, and Theodore S. Rappaport. 3D mmWave Channel Model Proposal. In *Proceedings IEEE Vehicular Technology Conference*, 2014.

[v1417a] 3GPP TR 38.901 v14.3.0. Study on channel model for frequencies from 0.5 to 100 GHz. Technical report, 3GPP, 2017.

[v1417b] 3GPP TS 136.101 v14.3.0. Physical channels and modulation. Technical report, 3GPP, 2017.

[v1518a] 3GPP TS 38.101-3 v15.3.0. Physical channels and modulation. Technical report, 3GPP, 2018.

[v1518b] 3GPP TS 38.211 v15.3.0. Physical channels and modulation. Technical report, 3GPP, 2018.

[VT02] H.L. Van Trees. *Optimum Array Processing.* Number Part 4. Wiley-Interscience, 2002.

[VVH16] V. Va, H. Vikalo, and R. W. Heath. Beam tracking for mobile millimeter wave communication systems. In *Proceedings IEEE Global Conference on Signal and Information Processing*, pages 743–747, Dec 2016.

[Wan09] J. Wang. Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems. *IEEE Journal on Selected Areas in Communications*, 27(8):1390–1399, October 2009.

[WJ02] J. W. Wallace and M. A. Jensen. Modeling the indoor MIMO wireless channel. *IEEE Transactions on Antennas and Propagation*, 50(5):591–599, May 2002.

[WSW04] Liqun Wang, Songqing Shan, and G Gary Wang. Mode-pursuing sampling method for global optimization on expensive black-box functions. *Engineering Optimization*, 36(4):419–438, 2004.

[YIS+16] S. Yoshioka, Y. Inoue, S. Suyama, Y. Kishiyama, Y. Okumura, J. Kepler, and M. Cudak. Field experimental evaluation of beamtracking and latency performance for 5G mmWave radio access in outdoor mobile environment. In *Proceedings IEEE Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, pages 1–6, Sept 2016.

[YKZ+16] Y. Yu, K. Kang, C. Zhao, Q. Zheng, H. Liu, S. He, Y. Ban, L. Sun, and W. Hong. A 60-GHz 19.8-mW Current-Reuse Active Phase Shifter With Tunable Current-Splitting Technique in 90-nm CMOS. *IEEE Transactions on Microwave Theory and Techniques*, 64(5):1572–1584, May 2016.

[ZGF16] C. Zhang, D. Guo, and P. Fan. Tracking angles of departure and arrival in a mobile millimeter wave channel. In *Proceedings IEEE International Conference on Communications*, pages 1–6, May 2016.

[ZMS+13] H. Zhao, R. Mayzus, S. Sun, M. Samimi, J. K. Schulz, Y. Azar, K. Wang, G. N. Wong, F. Gutierrez, and T. S. Rappaport. 28 GHz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in New York city. In *Proceedings IEEE International Conference on Communications*, pages 5163–5167, June 2013.

[ZVM10] H. Zhang, S. Venkateswaran, and U. Madhow. Channel Modeling and MIMO Capacity for Outdoor Millimeter Wave Links. In *2010 IEEE Wireless Communication and Networking Conference*, pages 1–6, April 2010.