TECHNISCHE UNIVERSITÄT DRESDEN

Index Modulation Techniques for Energy-efficient Transmission in Large-scale MIMO Systems

Merve Sefunç

von der Fakultät Elektrotechnik und Informationstechnik
der Technischen Universität Dresden

zur Erlangung des akademischen Grades

DOKTORINGENIEUR
(Dr.-Ing.)

genehmigte Dissertation

Vorsitzende:	Prof. Dr.-Ing. habil. Frank Ellinger
Gutachter:	Prof. Dr.-Ing. Eduard A. Jorswieck
	Dr. Marco Di Renzo
Tag der Einreichung:	27. 06. 2019
Tag der Verteidigung:	09. 10. 2019

To my loved ones

Acknowledgments

I would like to express my sincere gratitude to my supervisor Professor Eduard A. Jorswieck for his guidance throughout the way that led me to finalize this dissertation. Under his supervision, I enjoyed the freedom to explore the research direction that is most fulfilling. Each discussion with him was a source of inspiration to form a comprehensive and objective critique.

I am elated to be part of the 5Gwireless Project which was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant. I am particularly thankful to Doctor Marco Di Renzo and all the parties of the project for organizing various events throughout the years that provided the necessary environment for fruitful collaborations.

I would like to thank Doctor Marco Di Renzo once again for serving as the reviewer of this dissertation, Professor Dirk Plettemeier for taking the part as the examiner, and Professor Frank Ellinger for being the head of my promotion committee. I am deeply indebted to all committee members for providing insightful comments, and for the challenging and fruitful discussions.

I would also like to extend my deepest gratitude to Professor Alessio Zappone who has been an anchor on my professional path with his profound knowledge as an enviable researcher and a supportive friend. I cannot begin to express my thanks to Doctor Christian Scheunert, an unrivaled person, who helped me to overcome many technical problems, provided me organizational support, proofread my thesis and helped me through the stressful times. My gratitude extends to Hrjehor Mark who helped me to edit my dissertation into its final form with his endless energy. I am thankful for successfully collaborating with Ahmed Raafat, Yoann Corre, Mohammed Zahid Aslam, Professor Adrian Agustin and Professor Josep Vidal. I am sincerely thankful to my colleagues at TU Dresden for insightful discussions, administrative support, and especially grateful to Carsten R. Janda, Doctor Pin-Hsun Lin and Doctor Martin Mittelbach for reading parts of my dissertation and for their fruitful comments. I wish to express my appreciation to Andrew Lonnstrom for proofreading my dissertation.

I would like to express my deepest gratitude to my family, who always believes in me and nurtures me. I am grateful to İbrahim Erkut Yüzgeçcioğlu for walking besides me through the years and supporting me in challenging times.

Abstract

This thesis exploits index modulation techniques to design energy- and spectrum-efficient system models to operate in future wireless networks. In this respect, index modulation techniques are studied considering two different media: mapping the information onto the frequency indices of multicarrier systems, and onto the antenna array indices of a platform that comprises multiple antennas.

The index modulation techniques in wideband communication scenarios considering orthogonal and generalized frequency division multiplexing systems are studied first. Single cell multiuser networks are considered while developing the system models that exploit the index modulation on the subcarriers of the multicarrier systems. Instead of actively modulating all the subcarriers, a subset is selected according to the index modulation bits. As a result, there are subcarriers that remain idle during the data transmission phase and the activation pattern of the subcarriers convey additional information.

The transceivers for the orthogonal and generalized frequency division multiplexing systems with index modulation are both designed considering the uplink and downlink transmission phases with a linear combiner and precoder in order to reduce the system complexity. In the developed system models, channel state information is required only at the base station. The linear combiner is designed adopting minimum mean square error method to mitigate the inter-user-interference. The proposed system models offer a flexible design as the parameters are independent of each other. The parameters can be adjusted to design the system in favor of the energy efficiency, spectrum efficiency, peak-to-average power ratio, or error performance.

Then, the index modulation techniques are studied for large-scale multiple-input multiple-output systems that operate in millimeter wave bands. In order to overcome the drawbacks of transmission in millimeter wave frequencies, channel properties should be taken in to account while envisaging the wireless communication network. The large-scale multiple-input multiple-output systems increase the degrees of freedom in the spatial domain. This feature can be exploited to focus the transmit power directly onto the intended receiver terminal to cope with the severe path-loss. However, scaling up the number of hardware elements results in excessive power consumption. Hybrid architectures provide a remedy by shifting a part of the signal processing to the analog domain. In this way, the number of bulky and high power consuming hardware elements can be reduced. However, there will be a performance degradation as a consequence of renouncing the fully digital signal processing. Index modulation techniques can be combined with the hybrid system architecture to compensate the loss in spectrum efficiency to further increase the data rates.

A user terminal architecture is designed that employs analog beamforming together with spatial modulation where a part of the information bits is mapped onto the indices of the antenna arrays. The system is comprised a switching stage that allocates the user terminal antennas on the phase shifter groups to minimize the spatial correlation, and a phase shifting stage that maximizes the beamforming gain to combat the path-loss. A computationally efficient optimization algorithm is developed to configure the system. The flexibility of the architecture enables optimization of the hybrid transceiver at any signal-to-noise ratio values.

A base station is designed in which hybrid beamforming together with spatial modulation is employed. The analog beamformer is designed to point the transmit beam only in the direction of the intended user terminal to mitigate leakage of the transmit power to other directions. The analog beamformer to transmit the signal is chosen based on the spatial modulation bits. The digital precoder is designed to eliminate the inter-user-interference by exploiting the zero-forcing method. The base station computes the hybrid beamformers and the digital combiners, and only feeds back the digital combiners of each antenna array-user pair to the related user terminals. Thus, a low complexity user architecture is sufficient to achieve a higher performance. The developed optimization framework for the energy efficiency jointly optimizes the number of served users and the total transmit power by utilizing the derived upper bound of the achievable rate. The proposed transceiver architectures provide a more energy-efficient system model compared to the hybrid systems in which the spatial modulation technique is not exploited.

This thesis develops low-complexity system models that operate in narrowband and wideband channel environments to meet the energy and spectrum efficiency demands of future wireless networks. It is corroborated in the thesis that adopting index modulation techniques both in the systems improves the system performance in various aspects.

Contents

Abbreviations

4G	4th generation
5G	5th generation
AA	antenna array
ABFSM	analog beamforming with spatial modulation
ADC	analog to digital converter
AoA	angle of arrival
AoD	angle of departure
BB	base band
BER	bit error rate
BPSK	binary phase shift keying
BS	base station
CP	cyclic prefix
CSI	channel state information
DAC	digital to analog converter
DL	downlink
EE	energy efficiency
ESIM	enhanced subcarrier index modulation
FFT	fast Fourier transform
GFDM	generalized frequency division multiplexing
GSFIM	generalized space-frequency division multiplexing
HBF	hybrid beamforming
HBFSM	hybrid beamforming with spatial modulation
ICI	inter-carrier-interference
IFFT	inverse fast Fourier transform
i.i.d.	independent and identically distributed
IM	index modulation
IoT	Internet of Things
ISI	inter-symbol-interference
IUI	inter-user-interference
JDD	joint detection and demodulation
LB	lower bound
LNA	low noise amplifier

LoS	line-of-sight
LTE	long term evolution
MIMO	multiple-input multiple-output
ML	maximum-likelihood
MMSE	minimum mean square error
mmWave	millimeter wave
MSE	mean square error
MU	multiuser
NLoS	non-line-of-sight
OFDM	orthogonal frequency division multiplexing
P/S	parallel to serial
PA	power amplifier
PAPR	peak-to-average power ratio
PDF	probability density function
PS	phase shifter
QAM	quadrature amplitude modulation
QoS	quality of service
RF	radio frequency
S/P	serial to parallel
SDN	software-defined networking
SE	spectrum efficiency
SFIM	space and frequency index modulation
SIM	subcarrier index modulation
SM	spatial modulation
SNR	signal-to-noise ratio
SRRC	square-root-raised-cosine
SSK	space shift keying
SW	switch
TDD	time division duplex
THz	terahertz
UAV	unmanned aerial vehicles
UB	upper bound
UL	uplink
ULA	uniform linear arrays
UT	user terminal
WiFi	wireless fidelity
ZF	zero-forcing

Notation

$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
$x \propto y$	value x is proportional to value y
$\binom{k}{n}$	combination of n things taken k at a time without repetition
$\lvert x \rvert$	absolute value of a real or complex number x
$\lfloor x \rfloor$	greatest integer less than or equal to x
$\Re(x)$	real part of a complex number x
$\Im(x)$	imaginary part of a complex number x
$\{x_k\}$	set with elements x_k
$\lvert\lvert \mathbf{x} \rvert\rvert$	ℓ_2-norm of vector $\mathbf{x}$
$\mathrm{diag}(\mathbf{x})$	diagonal matrix with elements of vector $\mathbf{x}$ on the diagonal
$\mathrm{tr}(\mathbf{x})$	diagonal matrix with elements of vector $\mathbf{x}$
$\mathbf{X} = (x_{ij})$	matrix $\mathbf{X}$ with i rows and j columns, consisting of elements x_{ij}
$\mathbf{X}^{\mathrm{T}}$	transpose of a matrix or vector $\mathbf{X}$
$\mathbf{X}^{\mathrm{H}}$	Hermiteian of a matrix or vector $\mathbf{X}$
$\mathbf{X}^{\dagger}$	$(\mathbf{X}^{\mathrm{H}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{H}}$ if $\mathbf{X}$ has full column rank or $\mathbf{X}^{\mathrm{H}}(\mathbf{X}\mathbf{X}^{\mathrm{H}})^{-1}$ if $\mathbf{X}$ has full row rank
$\mathbf{I}n$	identity matrix of size n
$\lvert\mathbf{X}\rvert$	determinant of a square matrix $\mathbf{X}$
$\mathrm{qr}(\mathbf{X})$	QR decomposition of matrix $\mathbf{X}$
$\mathrm{E}(x)$	expectation of a random variable x
$\Pr(x = k)$	probability of event that the random variable x equals k
$\mathrm{var}(x)$	variance of a random variable x
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$\mathcal{CN}(\mu, \sigma^2)$	circularly symmetric complex Gaussian distribution with mean μ and variance σ^2
$\mathcal{U}(a, b)$	uniform distribution over interval $[a, b]$
$h(x)$	differential entropy of x
$h(x \mid z)$	conditional differential entropy of x given z
$I(x; y)$	mutual information between x and y
$I(x; y \mid z)$	conditional mutual information between x and y given z

Chapter 1

Introduction

1.1 Motivation

Future 5th generation (5G) and beyond networks will have to serve an unprecedented amount of wireless devices, which is forecasted to reach tens of billions by 2030 [1]. Numerous use cases, e.g., enhanced mobile broadband (eMBB), ultra-reliable low latency communications (URLLC), and massive machine type communications (mMTC), will be implemented with the deployment of the 5G networks to provide reliable service [2]. The resulting network will have a heterogeneous architecture that consists of diverse technologies. Small and macro cells will operate in cooperation, long term evolution (LTE) and wireless fidelity (WiFi) will coexist at the same unlicensed frequency bands, and device-to-device communications will be enabled. In order to serve such a massive amount of connected devices, which comprises not only mobile terminals but also smart appliances, autonomous cars, Internet of Things (IoT), and so on, a 1000x data rate increase is needed. It is apparent that such a huge data rate increase cannot be reached by simply scaling up the transmit power, due to sustainable growth, economical, and environmental concerns. Thus, energy efficiency (EE) has been acknowledged as a key requirement of future networks in addition to spectrum efficiency (SE) [3].

There are several candidates to fulfill the SE and EE requirements of future wireless networks among which large-scale multiple-input multiple-output (MIMO) systems [4], millimeter wave (mmWave) communications [5], and flexible waveform designs [6] particularly have been put forward for the upcoming 5G networks. Moving towards beyond 5G networks, terahertz (THz) communications will be utilized that enables more accurate positioning, various health applications, short-range communications [7]. To provide a better coverage and enable wireless communications at the remote areas, unmanned aerial vehicles (UAV) will be exploited that can potentially facilitate wireless broadcast and support high rate transmissions [8]. As a result of network densification and heterogeneity, software-defined networking (SDN) [9] and network slicing [10] will be adopted to accommodate various technologies. The former provides a dynamic architecture that configures the network to improve the performance and monitoring while the latter will allow to partition the network resources for the simultaneous operation of diversified technologies.

Besides all the aforementioned technologies, index modulation (IM) techniques have attracted significant attention thanks to their innovative way of conveying information compared to traditional communication systems. Specifically, the attractive feature of IM techniques is mapping the information onto the indices of different media. Because of this feature, IM techniques provide an additional dimension to the conventional modulation schemes that map the information onto the phase, amplitude or frequency of a sinusoidal carrier signal. Unlike these conventional schemes, IM techniques are able to map the information onto the antenna/antenna array (AA) indices that are used to transmit/receive the signal, on/off status of radio frequency (RF) mirrors, modulation types, signal powers, time/frequency indices and also a combination of those in addition to the conventional modulation schemes.

This thesis exploits the IM techniques in order to design energy- and spectrum-efficient system models to operate in future wireless networks. In this respect, IM techniques are studied considering two different media: mapping the information onto the frequency indices of multicarrier systems and onto the AA indices of a platform that comprises multiple antennas.

1.2 Overview and Contribution

Multicarrier waveforms are widely utilized due to their high SE nature where the transmission of multiple streams at the same time instant is enabled. The orthogonal frequency division multiplexing (OFDM) multicarrier communication system is already implemented for various technologies including 4th generation (4G) LTE [11]. However, the demand for SE in 5G wireless networks is much larger than in current 4G networks since 5G networks will involve not only the cellular communications, but also several different vertical sectors such as sensors for smart cities, device-to-device communications, self driving vehicles. One of the physical layer techniques that has emerged recently is the IM techniques. Hence, there is a need for alternative signal processing techniques that will increase the SE and EE of the network.

Modulation and Multiplexing. In order to satisfy these requirements, IM is introduced for OFDM waveforms in [12] which enables transmitting additional bits by modulating them onto the subcarrier indices. As opposed to the conventional OFDM system, there are subcarriers that remains idle during the transmission. The number and index of the subcarriers that are modulated by a conventional modulation scheme are established based on the IM part of the information bits. In this particular system model, the IM part of the information bits are grouped as ones and zeroes, then the group with larger cardinality is chosen to be modulated. Hence, there is a need for control subcarriers to explicitly signal the type of the majority bit value to the receiver. In order to prevent the need of excess subcarriers that can also be exploited to transfer information bits, an enhanced subcarrier index modulation (ESIM)-OFDM system has been proposed [13]. This study proposes a procedure to avoid the control subcarriers where the available subcarriers are divided into smaller groups. The number of the subcarriers in each group is associated with the modulation order of the symbol. Thus, this method requires higher order modulations to reach the same SE as that of the conventional OFDM system.

The dependency to the modulation order on the IM part of the information bits limits the SE and the flexibility of the resulting system model. Hence, OFDM-IM system is proposed in [14] where the number of active subcarriers are predefined regardless of the modulation order and the indices of the subcarriers are chosen according to the incoming bits. The results show that the OFDM-IM technique improves the error performance of the system compared to the conventional OFDM and ESIM-OFDM systems. Since a grouping architecture is adopted in order to exploit OFDM-IM system, the resulting performance can be further improved by employing an interleaving method [15, 16]. The idea is to locate the subcarriers that belong to the same group apart from each other while forming the final OFDM-IM block to be transmitted. Thus, the subcarriers of the adjacent subchannels belong to the separate groups which brings frequency diversity gain to OFDM-IM.

In order to further investigate the performance enhancement by exploiting IM technique on the frequency domain, a closed-form lower bound to the achievable rate of the OFDM-IM systems with interleaved grouping is derived [16]. The study denotes that the lower bound can be exploited to predict the optimal subcarrier activation strategy which maximizes the performance gain of OFDM-IM over conventional OFDM systems. Besides the achievable rate, bit error rate (BER) is also an important performance metric that indicates the robustness of the developed system model. As it has been proven by extensive simulations, the OFDM-IM system improves the BER performance compared to the conventional OFDM. Moreover, a tight closed-form approximation of the BER is derived introducing the expression for the number of bit errors occurring in both the index domain and the complex domain, in the presence of both imperfect and perfect detection of active subcarrier indices [17]. Then the analysis is extended to the case that the channel state information (CSI) is uncertain which enables to investigate the impacts of CSI uncertainty on the diversity gain and error floor of the BER [18]. Addition of a third dimension to the constellation diagram on frequency domain increases the diversity and brings various advantages to the system as studied with the aforementioned works. However, it also increases the load of the receiver since IM part of the information bits should be detected successfully. In most of the system models that have been proposed, there is a small number of subcarrier activation pattern which is out of use. A methodology is proposed to encode all possible subcarrier activation patterns which also lowers the false detection caused by having excess activation pattern [19].

Various methods to exploit IM technique on the subcarriers of an OFDM system is studied and the performance metrics are analyzed. Furthermore, OFDM-IM system is extended to MIMO scenario in order to exploit also the spatial multiplexing and achieve higher SE values [20]. Low complexity transceiver architectures are crucial in MIMO systems since the complexity and the required computation power are increased with the number of antennas. Hence, a low complexity transceiver structure is developed and performance of various detection schemes are studied. As it has been reported before, IM techniques can be combined to increase the diversity and improve the performance as with generalized space-frequency division multiplexing (GSFIM) system [21, 22]. Therein, both spatial domain and frequency domain are exploited to encode bits through indexing which increases the complexity of the resulting system model. Hence, an encoding procedure is proposed to exploit the low complexity computation of combinatorics in combinatorial number system together with a detection algorithm based on a multi-stage message passing approach.

Although there are similarities on the system design of a point-to-point and multiuser (MU)-MIMO scenarios from the spatial multiplexing perspective, inter-user-interference (IUI) of the latter case brings additional challenges to tackle in the system design. Therefore, a transceiver architecture is designed for the subcarrier index modulation (SIM)-OFDM in MU networks where an iterative detector based on the generalized approximate message passing algorithm is developed to tackle the IUI [23]. Furthermore, the study is extended to the large-scale MIMO systems with imperfect CSI [24]. Furthermore, surveys on IM techniques are recently published in [25, 26] which provide a broader look into various applications and performance evaluations in the subject.

The OFDM system enables to transmit information on multiple orthogonal subcarriers at a time instant. The technology is already adopted by many applications as it increases the SE significantly. Although the orthogonality reduces the inter-carrier-interference (ICI), it results in large peak-to-average power ratio (PAPR) and out-of-band emission. To tackle the disadvantages of the OFDM system and provide a more generic design, generalized frequency division multiplexing (GFDM) system is proposed [27]. In contrast to the OFDM system, GFDM allows one to transmit information by an arbitrarily filtered block that comprises multiple subsymbols in the time domain and multiple subcarriers in the frequency domain. The generation of an GFDM block is not straightforward as for the OFDM case, hence a low complexity matrix model that provides a practical method to generate the GFDM block is proposed [28]. Moreover, the performance of the GFDM system is compared with the LTE standard as far as spectral properties and implementation complexity aspects are concerned. Since linear signal processing methods lower the complexity of the resulting system, the matrix model to generate the GFDM block is further improved to enable design of linear receiver structures and the BER is exploited as the performance metric to compare various receivers [29]. Moreover, possible applications of GFDM scheme for various 5G network scenarios are further studied in [30].

Similar to the OFDM system, IM techniques can be exploited to increase the performance of the GFDM [31]. Although BER of the GFDM-IM system outperforms the conventional GFDM and OFDM-IM systems at mid- to high-signal-to-noise ratio (SNR) regions, more erroneous detection of the subcarrier activation pattern results in performance degradation in low SNR region. Hence, a message passing-based detector can be adopted in order to improve the BER performance of the GFDM-IM system [32]. Besides enhancing the SE and improving the BER performance, utilizing IM techniques also reduces the out-of-band emission of the GFDM systems [33]. Moreover, the spatial domain can be exploited in addition to the frequency domain to further improve the SE where the information bits are mapped both onto the indices of the transmit antennas and the subcarriers along with the conventionally modulated symbols [34]. The space and frequency index modulation (SFIM)-GFDM system outperforms the spatial modulation (SM)-GFDM system regarding the BER [35] where the IM technique is exploited only in the spatial domain. A generic GFDM system with flexible IM techniques that enables switching between different IM techniques to adapt the channel conditions can be designed [36]. Thanks to the flexibility of the resulting system architecture, a multilayer transmission scheme can be exploited by effectively utilizing the space, frequency, and time dimensions to adjust the BER performance and SE.

Contribution. Inspired by the aforementioned studies, this thesis first discuses a system model for a MU MIMO network that utilizes the OFDM-IM system [37]. Thereafter, the work is extended for the GFDM-IM [38] that brings additional challenges to overcome to the system design. In order to study the performance of the proposed system models, uncoded BER is taken into account as a performance metric for both systems. Since MU networks are considered, interference among the users both in the uplink (UL) and downlink (DL) transmission phases is the fundamental challenge for the design of the system. With the additional complexity brought by utilizing the IM techniques, it necessitates to tackle the interference among the users by linear combiner and precoder architectures without degrading the resulting performance. To enable the detection of the information conveyed by both IM and conventional modulation parts, a more versatile yet lower complexity algorithm is essential. Therefore, this thesis proposes linear combiner and precoder designs by exploiting the spatial multiplexing of MIMO system and describes low complexity detection algorithms that are applicable for practical scenarios.

Energy-efficient mmWave Communications. Although the multicarrier modulation system together with IM techniques improves the performance of the resulting system model, the shortage of bandwidth in low frequencies motivates the migration to higher frequencies in order to satisfy the vast increase in data rates [39, 40]. Availability of spectrum in the mmWave band has been considered by the research community as a way to design the wireless communication networks to meet the expected large traffic demands [41].

A shortcoming of the mmWave band is the severe path-loss and sensitivity to blockage of mmWave propagation bring additional challenges to the network [42]. In spite of the unfavorable characteristics of the channel, large number of antennas can be packed in small areas due to the small wave length at high frequencies and asymptotically achieve the capacity of the channel [4]. Hence, large-scale MIMO technology became one of the promising candidates to combat these challenges and satisfy the demand for high data rates [43]. By increasing the number of antennas, the performance of the system can be improved as a trade-off between the antenna gain and the additional degrees of freedom

[44]. Considering a system with a massive number of antennas, it is possible to serve multiple user terminal (UT) in the same time and frequency resources simultaneously [45]. The usage of a large number of antennas helps focusing the energy in the form of narrow beams, so that the quality of communication is improved and the radiated power is steered to the intended direction [46]. Moreover, linear precoding schemes can be employed to drastically reduce the complexity of the system [47].

Fully digital architectures, such as the block diagonalization algorithm in [48], are proven to achieve high performance. However, implementation of a fully digital architecture at a terminal with a large number of antennas is not practical due to space and energy consumption constraints [49]. Adopting analog beamforming techniques decreases the number of power hungry hardware elements while benefiting the directivity of the channel [50]. Although analog beamforming techniques help to combat the severe path-loss of the mmWave bands, control over the precoding matrix is limited in the analog domain. Therefore, hybrid architectures should be designed with a reduced number of RF chains by employing analog beamforming together with digital precoding [51, 52]. There has been extensive research on hybrid architectures, such as a two-stage hybrid beamformer/combiner algorithm [53], where the minimum number of RF chains to realize the performance with a fully-connected structure is studied. An asymptotically optimal hybrid beamforming design that exploits the singular vectors of the channel to generate the analog beamformer and combiner is proposed in [54, 55]. Then, a codebook-based analog precoding design is proposed where the CSI is obtained via a beam sweep procedure [56]. Furthermore, an algorithm is proposed that configures the hybrid precoders at the transmitter side and the analog combiner at the receive end with a small training and feedback overhead [57]. There, the authors have shown that the hybrid beamforming system achieves higher data rate values compared to the analog beamforming and approaches the performance of fully digital beamforming. A different perspective is adopted to lower the complexity of the hybrid beamforming design problem by firstly obtaining the optimum number of RF chains, then calculating the precoders [58]. Because of utilizing large number of antennas, considering per-antenna power constraint results in a more realistic scenario while developing the system architecture [59]. Another hybrid beamforming system is proposed in [60] that enables multiple data streams to the UT by jointly designing the analog beamformer and digital precoder in one stage. In addition, a minimum mean square error (MMSE) based multiuser hybrid beamformer is designed by leveraging the sparsity of the mmWave channel [61]. However, the fully connected hybrid MIMO architecture comprises many analog devices (phase shifters, power splitters and combiners). Thus, when a moderate number of RF chains is deployed, the fully connected hybrid MIMO consumes more power than the fully digital MIMO [62]. A more energy-efficient system can be designed by adopting a partially-connected architecture [63]. However, abandoning the fully connected hybrid architecture results in degradation of the SE even with the most sophisticated algorithms for the design of the partially-connected hybrid precoders/combiners [62, 64].

In order to improve the SE without compromising the simplicity and low power consumption, SM techniques, where a part of the information bits is mapped in the spatial domain such as onto the antenna or AA indices, can be adopted [65, 66, 67]. System models can be designed such that the loss on the SE caused by lowering the number of RF chains to reduce the total power consumption is compensated by adopting SM techniques. As proposed in [68], a point-to-point system with generalized SM employing analog beamforming achieves better SE than a system that comprises more RF chains. Therein, a single RF chain per data stream is activated to transmit the conventionally modulated symbols, and analog beamforming is employed by the subarray selected by spatially modulated bits. The implementation of SM in MU MIMO networks is studied in [69] and a precoding method is proposed in order to cancel IUI. A closer look into the transmitter design of SM together with analog beamforming is provided in [70]. Furthermore, the application of SM with hybrid beamforming for a mmWave railway communication system is studied in [71]. There, the receive antenna arrays at the front and at the end of the train are regarded as virtual users served by two different data streams transmitted by the base station (BS). An analog beamformer and combiner are employed at the transmitter and receiver, respectively, even though the system is not specifically designed for a multiuser scenario.

Contribution. By considering the tendency of exploiting the empty spectrum on mmWave bands, this thesis focuses on developing system models that operate on higher frequencies as the next step [72, 73]. The proposed system models utilize the properties of the signal and channel characteristics brought by operating on the mmWave band. Firstly, a point-to-point large-scale MIMO system that operates in mmWave narrowband channel environment [74] is developed. In order to cope with the poor scattering nature of mmWave channels and the severe path-loss of mmWave propagation, a large number of antennas is employed at the BS and at the UT to provide a large beamforming gain. An analog beamforming aided SM scheme for the UL transmission phase is proposed to maximize the EE at the UT for a given SE requirements. The extension of the work to the MU networks is straightforward, since a fully digital BS architecture is adopted for the UL transmission phase. Then, a system model for an MU large-scale MIMO network that operates in mmWave narrowband channel environment is proposed [72]. There, a hybrid BS architecture is considered that exploits the SM technique by mapping a part of the information onto the indices of the AAs. In this scheme, BS employs an analog beamformer to direct the beam to the intended user and a digital precoder to cancel IUI while users employ digital combiners. EE is one of the fundamental requirements of the future networks and the basis of the proposed system models. Therefore, an optimization framework for bit-per-Joule EE maximization is proposed for the developed system model [75, 76].

The proposed hybrid beamforming systems that adopt SM techniques are designed for narrowband channels in this thesis. An extension to the wideband channels in which IM technique is adopted on the subcarriers of the multicarrier system in addition to the SM is possible to further increase the diversity of the system. This can be achieved by treating each subchannel of the multicarrier system as a narrowband channel. In this case, analog and digital precoders/combiners should be designed independently for each subchannel. However, this would result in a much complex system model and increase the number of hardware elements for the signal processing on the analog domain. Hence, a hybrid system can be designed with an analog beamfomer for the entire spectrum and digital beamformers for each subchannels [77, 78].

1.3 Outline

Increasing the SE of the network has been the biggest concern of the research community since the data exchange went beyond voice transmission in various communication scenarios. Deployment of MIMO technologies and taking advantage of multicarrier systems have been the traditional approaches to meet the demands of today's networks. However, the number of connected devices is exponentially growing as well as the heterogeneity of the communication scenarios. Hence, concerns on the power consumption of the network continuously increased and ultimately EE became a key requirement for the future wireless networks. Traditional MIMO and OFDM systems are insufficient to satisfy these ever-growing demands. Thus, this thesis discusses the implementation of IM techniques to tackle the design of energy- and spectrum-efficient wireless communication systems. Motivated by this background, this thesis will discuss innovative multiple antenna and multicarrier transmission schemes.

Chapter 2 elaborates on the fundamentals of the technologies that are utilized in this thesis. First, the principles of the multicarrier systems are unveiled and the basic structure of the OFDM and GFDM transceivers are explained. Then, the necessity of the large-scale MIMO systems are explicated, advantages and disadvantages brought by deploying large number of antennas are further discussed. Lastly, the fundamental principles of the IM techniques that are adopted to develop the system models are elucidated. Furthermore, some practical examples are presented to unveil the numerous possibilities of system designs when IM techniques are considered.

Chapter 3 focuses on multicarrier systems that exploit IM techniques in the frequency domain, namely OFDM and GFDM. The former enjoys the orthogonality of the subcarriers while the latter further increases the SE by transmitting the data block in multiple time instances. The MU communication and IM bring additional complexity to the system due to the multi-stream signal processing, IUI and the extension of the modulation of the data to the frequency domain. Thus, linear signal processing methods are adopted and the receiver is designed to decode the data by a low complexity detection algorithm without sacrificing the performance. An IM system is considered in which the number of available and active subcarriers are predefined, and the subcarrier activation pattern conveys information. The proposed system model is designed for a MU network where both BS and UTs comprise multiple antennas. Hence, precoder and combiner designs at the BS have the utmost importance to mitigate the IUI and process the signal to prepare for the joint detection of the information bits conveyed by IM and conventional modulation parts. To tackle this problem, MMSE-based linear precoder and combiner are designed for the DL and UL transmission phases, respectively. The receiver is designed to detect the IM bits by comparing the received power on the subcarriers, then decode the conventionally modulated symbols carried by the detected active subcarriers.

Chapter 4 studies transmission in the mmWave bands, specifically studying the use of SM technique in this context. It is aimed to develop system models where the communication at the high frequency bands is turned into advantage by exploiting large number of antennas and adopting hybrid architectures. The SM technique is a special case of IM techniques where the information bits are mapped in the spatial domain such as indices of antennas or AAs. By combining these technologies, Chapter 4 discusses large-scale MIMO systems that exploit IM techniques in the spatial domain. There is an available spectrum in mmWave bands which is mostly empty. Since the lower frequency bands are excessively crowded, mmWave bands are highly attractive to serve large amount of devices that will be part of future wireless networks. However, the poor scattering nature and the severe path-loss of the mmWave bands bring additional challenges to overcome in designing the system. In order to combat the disadvantages brought by transmission in high frequencies, large-scale MIMO systems can be adopted. Due to the degrees of freedom of deploying a large number of antennas, linear signal processing methods and beamforming technologies can be exploited. Although large-scale MIMO systems offer a favorable solution for the SE requirements, scaling up the number of hardware elements in the system increases the power consumption drastically. Therefore, a more efficient design is necessary instead of increasing the number of hardware elements, so that the EE requirements can be met as well. In this respect, Chapter 4 proposes transceiver architectures that exploit hybrid beamforming techniques together with SM scheme. First, a low power consuming UT architecture is designed where the objective is lowering the number of bulky and high power consuming hardware elements. The UT comprises a large number of antennas that are grouped to form multiple AAs considering the channel characteristics. Performance of the network highly depends on the channel statistics since the mmWave bands have a limited scattering environment, hence the system design is more challenging. Therefore, an algorithm is developed to design the system that maximizes the EE of the UT for a given SE. In the resulting system model, SM bits are mapped onto the indices of AAs which are activated to transmit a conventionally modulated symbol by employing analog beamforming. Since a fully digital large-scale BS architecture is considered, a zero-forcing (ZF) combiner is employed prior to detection of the received signal. Then, the focus is shifted to the design of a low power consuming BS architecture that transmits multiple data streams to fully digital UTs. The BS comprises a large number of antennas that are grouped to form AAs. Moreover, SM bits are mapped onto the indices of activated AAs for each user's data stream. Since a MU network is considered, BS employs a ZF-based digital precoder to eliminate the IUI. Furthermore, analog beamforming is employed to direct the signal to the intended user. Thanks to the hybrid beamforming at the BS, one AA can be utilized to transmit data to multiple UTs simultaneously, hence the SM technique is fully exploited. At last, an EE optimization framework is developed in order to jointly optimize the number of UTs to be served and the total transmission power. The problem is formulated as the maximization of the EE with respect to the transmit power and number of active users, enforcing constraints on maximum power, maximum and minimum number of active users, and a minimum achievable guaranteed rate. The optimization is carried out considering both the standard linear power consumption model, as well as the case in which the transmit amplifier operates in the non-linear region. In both cases, globally optimal optimization algorithms with affordable complexity are developed.

Finally, Chapter 5 concludes the thesis and discusses the open problems for further research in these directions.

Chapter 2

Preliminaries and Fundamentals

This chapter elaborates the basic principles of the technologies that will be exploited to design system models in Chapter 3 and Chapter 4. Among the many candidate technologies for future wireless networks, multicarrier systems are already deployed for various technologies. Section 2.1 describes the elementary transceiver architecture and signal processing methods for OFDM and GFDM systems. Since implementations has already started, large-scale MIMO systems will take part in the network to provide better service and to cover larger areas. Therefore, the discussion of large-scale MIMO systems follows in Section 2.2 by addressing the advantages and disadvantages of scaling up the number of hardware elements at the terminals. Conventional signal processing methods are not adequate as the number of connected devices increases drastically. IM techniques are proposed to further increase the SE and EE of the multicarrier and large-scale MIMO systems. Section 2.3 elucidates the principles of IM techniques and discusses the additional diversity brought to the system. Lastly, Section 2.4 introduces a generic single cell MU network architecture and unveils the assumptions to develop the system models in the following chapters.

2.1 Multicarrier Systems

In communication systems, information is carried with the help of sinusoidal signals. These carrier signals can be modulated in phase, amplitude, or frequency in order to convey information. In general, this methodology is called M-ary modulation where M denotes the modulation order and enables to transmit $\log_2 M$ bits per symbol. The modulated symbols are transmitted over a wireless channel in a certain frequency band. The resulting input-output relationship when a single M-ary modulated symbol x is transmitted can be expressed as

$$y = hx + n, \tag{2.1}$$

where y is received signal, h is the channel impulse response with arbitrary distribution, n is the additive noise with $\mathcal{CN}(0, \sigma^2)$ distribution. When a single carrier communication system is considered, in order to achieve high data rates,

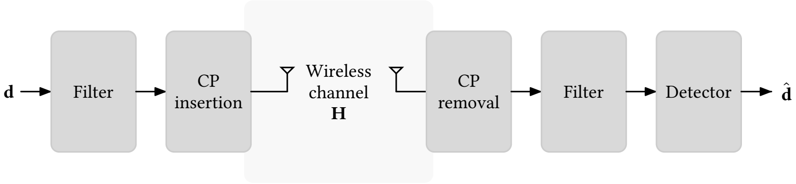


Figure 2.1: Block diagram of a multicarrier transceiver.

higher order modulation order should be adopted. However, conventional modulation schemes may suffer from inter-symbol-interference (ISI) and frequency selective fading. The ISI occurs when the received symbol is distorted by the subsequent symbols caused by the multipath fading environment which is the usual case in real-world scenarios. Frequency selective fading occurs when the channel impulse response varies over the bandwidth that the transmitted signal occupies. The frequency band in which the channel impulse response is constant is called coherence bandwidth and the signal experiences flat fading over the coherence bandwidth. In order to achieve higher data rates without suffering ISI and frequency selective fading, multicarrier waveforms can be adopted in a way that the available bandwidth is divided into subchannels which are shorter than the coherence bandwidth, and a subcarrier that is modulated with a lower order modulation scheme carries the information. Each subcarrier spans a lower bandwidth and a higher time interval can be utilized. This results in communication suffering only from flat fading over multiple narrowband channels.

Orthogonal Frequency Division Multiplexing. OFDM is one of many multicarrier waveforms that is implemented in several modern systems. By dividing the available transmission bandwidth OFDM enables transmitting a large number of orthogonal, narrowband subcarriers in parallel. A transceiver block diagram of a basic OFDM system is depicted in Figure 2.1. Let $\mathbf{d} \in \mathbb{C}^{N_{\text{tot}} \times 1}$ be the set of symbols to be transmitted where N_{tot} is the number of total subcarriers. For brevity, signal processing blocks, i.e., serial to parallel (S/P) and parallel to serial (P/S) converters, up to the modulation and demodulation of the symbols is not demonstrated in Figure 2.1. The OFDM block is generated by utilizing a square filter which is equivalent to the inverse fast Fourier transform (IFFT) in practical implementation. Hence, the transmit symbols are processed by an IFFT in the filter block, i.e.

$$\mathbf{x} = \mathbf{A}^{\mathrm{H}}\mathbf{d}, \tag{2.2}$$

where $\mathbf{A} = (a_{k,l}) \in \mathbb{C}^{N_{\mathrm{tot}} \times N_{\mathrm{tot}}}$ is a fast Fourier transform (FFT) matrix with elements $a_{k,l} = \frac{1}{\sqrt{N_{\mathrm{tot}}}} \exp\left(j2\pi \frac{kl}{N_{\mathrm{tot}}}\right)$ for all $k, l \in \{0, 1, \ldots, N_{\mathrm{tot}} - 1\}$. In order to guarantee the orthogonality of the signal passing through a multipath channel, a cyclic prefix (CP) is inserted at the end of the OFDM block which is equivalent to copying the first N_{CP} samples of the OFDM block to its end. The resulting received signal is

$$\tilde{\mathbf{r}} = \tilde{\mathbf{H}}\tilde{\mathbf{x}} + \mathbf{n}. \tag{2.3}$$

Herein, $\tilde{\mathbf{H}} \in \mathbb{C}^{N_{\mathrm{tot}}+N_{\mathrm{CP}} \times N_{\mathrm{tot}}}$ is the circular convolution matrix generated from the channel impulse vector $\mathbf{h} \in \mathbb{C}^{C \times 1}$ with arbitrary distribution. Since the channel impulse response comprises multiple paths, the effect of a signal block $\tilde{\mathbf{x}} \in \mathbb{C}^{N_{\mathrm{tot}}+N_{\mathrm{CP}} \times 1}$ passing through the channel is the convolution of the channel impulse vector and the transmitted signal. This operation can be performed in matrix form by exploiting the circular convolution matrix $\tilde{\mathbf{H}}$ which simplifies the transmission model as in (2.3). The additive noise vector $\mathbf{n} \in \mathbb{C}^{N_{\mathrm{tot}}+N_{\mathrm{CP}} \times 1}$ comprises independent and identically distributed (i.i.d.) elements that follow circularly symmetric complex Gaussian distribution with zero mean and σ^2 variance. After the reception of the OFDM block, the CP is removed and FFT is employed

$$\mathbf{y} = \mathbf{A}\mathbf{r}. \tag{2.4}$$

Prior to the demodulation of the symbols, equalization is performed by the help of matched filtering $\mathbf{W} = \mathbf{H}^{\mathrm{H}}$ or ZF filtering $\mathbf{W} = \mathbf{H}^{\dagger}$ or linear MMSE filtering $\mathbf{W} = (\mathbf{H}^{\mathrm{H}}\mathbf{H} + \sigma^2\mathbf{I}_{N_{\mathrm{tot}}})^{-1}\mathbf{H}^{\mathrm{H}}$ where $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{tot}} \times N_{\mathrm{tot}}}$ is the circular convolution matrix of the channel after removal of the CP. The number of subcarriers N_{tot} at an OFDM block is generally large in order to achieve high SE. Thus, the channel matrix $\mathbf{H}$ and the filtering matrix $\mathbf{A}$ also have high dimensions. In order to design a system with low complexity, only linear processing methods are considered.

Generalized Frequency Division Multiplexing. Next, the properties of a GFDM system are elucidated which is a non-orthogonal multicarrier system. The GFDM system can be considered as a generalization of the OFDM that enables transmitting the data in multiple time and frequency resources. A GFDM block comprises L subsymbols in the time domain and N_{tot} subcarriers in the frequency domain. In contrast to the OFDM system, the GFDM block can be generated with an arbitrary filter and the CP inserted only once for multiple subsymbols, which provides larger SE than the OFDM system. The transceiver block diagram of a basic GFDM system is again as depicted in Figure 2.1 with slightly different operations in the blocks. Let $\mathbf{d} \in \mathbb{C}^{Q\times 1}$ be the set of symbols to be transmitted where $Q = LN_{\text{tot}}$ is the total number of samples. The GFDM block is generated by processing the data vector $\mathbf{d}$ with the transmit matrix $\mathbf{A} \in \mathbb{C}^{Q\times Q}$ that is computed as proposed in [29]

$$\mathbf{x} = \mathbf{A}\mathbf{d}. \tag{2.5}$$

Herein, $\mathbf{A}$ has the structure of

$$\mathbf{A} = [\mathbf{g}_{0,0}, \dots, \mathbf{g}_{N_{\text{tot}}-1,0}, \mathbf{g}_{0,1}, \dots, \mathbf{g}_{N_{\text{tot}}-1,1}, \mathbf{g}_{0,L-1}, \dots, \mathbf{g}_{N_{\text{tot}}-1,L-1}] \tag{2.6}$$

$$\text{where} \quad g_{n,l} = g\big((q - nL)_{\text{mod}_Q}\big)\ \exp\left(j2\pi \tfrac{nq}{N_{\text{tot}}}\right) \tag{2.7}$$

is the circularly shifted transmit filter and $\mathbf{g}_{n,l} \in \mathbb{C}^{Q\times 1}$ is the vector consisting of the filter samples

$$\mathbf{g}_{n,l} = [g_{n,l}(0), g_{n,l}(1), \dots, g_{n,l}(Q-1)]^{\mathrm{T}}. \tag{2.8}$$

Each individual sample of the GFDM block can be represented as

$$\mathbf{x}(q) = \sum_{n=0}^{N_{\text{tot}}-1} \sum_{l=0}^{L-1} d_{n,l}\, g_{n,l}(q), \tag{2.9}$$

with sampling index $q \in \{0, 1, \dots, Q-1\}$. After processing the symbols to be transmitted, a CP of length N_{CP} is inserted at the end of the GFDM block. The remaining procedure is the same as for OFDM systems.

The aforementioned multicarrier waveforms enable improving the SE by transmitting multiple samples per channel use. However, these systems also come with some disadvantages. OFDM systems are easy to implement with an FFT algorithm. Since the pulse shaping filter is rectangular, it results in out-of-band emission due to the side lobs in the frequency domain. OFDM systems provide orthogonality that makes subcarriers and symbols independent of each other, but this also requires perfect synchronization, and equalization plays an important role in the performance of the resulting system. In order to prevent ISI, the insertion of the CP is necessary, which results in a degradation of the SE. In addition, OFDM systems suffer from high PAPR which is due to the presence of a large number of independently modulated subcarriers. Hence, the peak power value on the individual subcarrier can be larger than the average of the whole OFDM block.

In GFDM systems, pulse shaping can be performed with any arbitrary filter which results in lower out-of-band emission. These systems have non-orthogonal subcarriers, hence the time and frequency synchronization is not as crucial as that of OFDM systems. However, non-orthogonality causes self-created ISI and ICI whenever the nulls of the subcarriers do not coincide with the peak of remaining subcarriers. Only a single CP per GFDM block is inserted for total LN_{tot} subcarriers. In contrast, L CP per OFDM block with N_{tot} subcarriers should be inserted to transmit the same amount of information as for the GFDM. Hence, GFDM systems increase the SE compared to OFDM systems. Although OFDM and GFDM systems provide spectrum-efficient communication, in order to satisfy the EE requirements of future wireless networks and to further increase the SE of the system, it is convenient to adopt IM techniques on top of these multicarrier systems.

2.2 Large-scale Multiple Input Multiple Output Systems

In recent years, MIMO systems have been implemented in several wireless communication networks since they enable transmitting multiple data streams at the same time and in the same frequency band, thus increasing the SE. MIMO communications have already been employed for both point-to-point and MU scenarios. Implementing multiple antennas brings additional degrees of freedom which can be exploited, e.g., to increase SE due to multiple simultaneous data streams, to decrease interference since the transmitted beam can be directed to the intended UT, to benefit from the enhanced reliability due to the diversity of the transmitted and received signals over the channel.

With recent advances in technology, it is possible to pack a large number of antennas at the transceivers. There are more powerful processors to control and process the signals that belong to many antennas. In addition, there is ongoing research on designing more robust and less energy consuming hardware. Therefore, hundreds of antennas can be deployed to design large-scale MIMO systems.

A generic block diagram of a large-scale MIMO system is depicted in Figure 2.2. In this system model, the number of transmit or receive antennas N_{T} and N_{R}

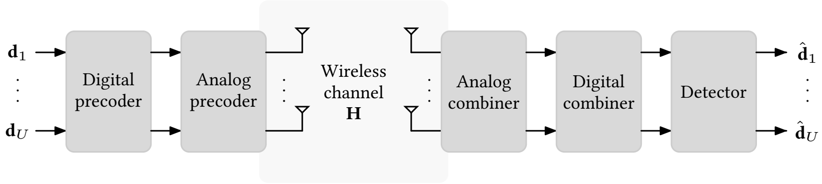


Figure 2.2: Block diagram of a large-scale MIMO transceiver.

are much larger than the number of data streams U to be transmitted per channel use. Since the large number of antennas provides sufficient diversity linear processing methods can be utilized even under poor scattering channel environments. Hence, large-scale MIMO systems facilitate designing the system with lower complexity.

In the block diagram depicted in Figure 2.2, the transmitter and receiver comprise digital and analog precoders and combiners. It is possible to perform signal processing operations entirely in the baseband to have a better control over the entries of the digital precoding and combining matrices. However, this increases the system complexity and the power consumption drastically when large-scale MIMO systems are considered due to the large power consumption of, e.g., the RF chain, the digital to analog converter (DAC), and the analog to digital converter (ADC). Therefore, analog precoders and combiners are used in the system design where the phase of the signal is controlled in the RF domain. As a result of utilizing the analog blocks in the system model, it is possible to implement a lower number of power hungry devices, e.g., power amplifier (PA), RF chain, DAC, ADC, low noise amplifier (LNA), than the total number of antennas in the system. Adopting linear processing, the received signal of a system that is working in a narrowband channel environment is

$$\mathbf{r} = \mathbf{HFPd} + \mathbf{n}, \tag{2.10}$$

where the transmitted signal in this case is $\mathbf{x} = \mathbf{FPd}$. The post-processed signal after hybrid combining is

$$\mathbf{y} = \mathbf{WDr}, \tag{2.11}$$

where $\mathbf{P}$, $\mathbf{W} \in \mathbb{C}^{U \times U}$ are the digital precoder and combiner as well as $\mathbf{F} \in \mathbb{C}^{N_\mathrm{T} \times U}$ and $\mathbf{D} \in \mathbb{C}^{U \times N_\mathrm{R}}$ are the analog precoder and combiner, respectively. $\mathbf{H} \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{T}}$ is the channel matrix with arbitrary distribution, $\mathbf{d} \in \mathbb{C}^{U \times 1}$ is the input vector comprising U symbols to be transmitted, and $\mathbf{n} \in \mathbb{C}^{N_\mathrm{R} \times 1}$ is the circularly symmetric complex Gaussian noise vector with i.i.d. $\mathcal{CN}(0, \sigma^2)$ distributed elements.

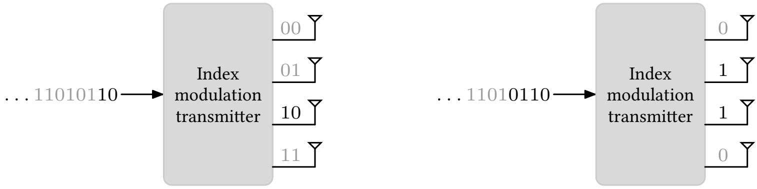
The design of the precoders and combiners can differ according to the targeted system model and the resulting network architecture. There is an extensive amount of work on the design of hybrid systems both in narrowband and wideband channel environments for point-to-point and MU networks. The surveys on the hybrid beamforming techniques in [79, 52] and the references therein provide a broader look for further reading.

Large-scale MIMO systems are useful particularly in mmWave bands to combat the poor scattering nature of the channel and the severe path-loss. Large AAs can be utilized in order to direct the beam to the receiver and to concentrate the transmit power with minimal leakage. This feature not only increases the received SNR, but also mitigates the IUI in MU networks. Due to the small wavelength of high frequencies, antennas occupy less space, hence large arrays can be implemented even for portable devices. However, bulky and high power consuming hardware elements such as RF chains, DACs, and ADCs complicate the design of large-scale MIMO systems. The aforementioned hybrid beamforming schemes enable lowering the number of these hardware elements by moving a part of the signal processing to the analog domain. However, lowering the number of these elements also limits the number of data streams to be transmitted simultaneously. Therefore, more advanced signal processing techniques are necessary to compensate the loss in SE. In this respect, IM techniques can provide a remedy by exploiting the excessive number of antennas to transmit information without additional power consumption.

2.3 Index Modulation Techniques

In traditional communication systems, information is transmitted by modulating the phase, amplitude, or frequency of a sinusoidal carrier signal. In contrast, IM techniques enable transmitting of information by mapping the bits onto different entities of the system. Then, this additional diversity is exploited in order to transmit information. Moreover, IM techniques can be adopted as stand alone or together with conventional modulation schemes.

Index Modulation in the Spatial Domain. In order to deliberate the diversity brought by IM techniques, a transmitter consisting of multiple antennas is chosen to exemplify exploiting the spatial domain. Consider a transmitter



(a) Single antenna activation. **(b)** Multi antenna activation.

Figure 2.3: Index modulation in the spatial domain.

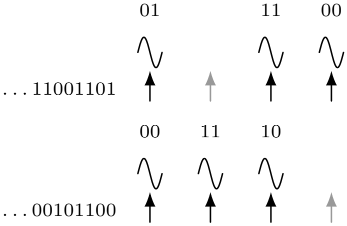


Figure 2.4: Index modulation in the frequency domain.

architecture that comprises $N_\mathrm{T} = 4$ transmit antennas as depicted in Figure 2.3a and Figure 2.3b. There are various possibilities in order to transmit information by adopting IM techniques with such a system architecture.

Example I: Activating a Single Antenna. Assuming a narrowband channel environment, only a single carrier signal is transmitted per channel use. Although there exist $N_\mathrm{T} = 4$ transmit antennas, only one of them is activated during the transmission and the index of the activated antenna conveys information as shown in Figure 2.3a. This kind of IM technique is called SM and enables to transmit $\log_2 N_\mathrm{T}$ bits of information. In addition to the SM, the carrier signal can be modulated with a conventional M-ary modulation scheme. Hence, $\log_2 N_\mathrm{T} + \log_2 M$ bits of information can be transmitted per channel use. Note that the spatially-modulated bits are transmitted inherently with the index of the activated antenna and do not require additional transmit power. Thus, adopting SM increases the SE without any additional energy cost, which leads to improving the EE, as well, compared to the traditional modulation schemes.

Example II: Activating Multiple Antennas. Considering the same narrowband system model, the on/off status of the transmit antennas can be altered based on the incoming bit stream as depicted in Figure 2.3b. In this case, a sinusoidal carrier signal is transmitted from the antenna if the bit mapped onto that antenna is 1 and the antenna remains idle otherwise. This kind of IM technique is called space shift keying (SSK) and enables transmitting N_T bits of information. In addition to the SSK scheme, the carrier signal can be modulated with a conventional M-ary modulation scheme. Hence, $N_\mathrm{T} + \log_2 M$ bits of information can be transmitted per channel use. In this case, each antenna transmits the same conventionally modulated signal, hence the probability of correct detection of the modulated signal at the receiver side is increased due to the diversity of the received signal. Another variation of the SSK scheme is to define transmit power levels for bits 0 and 1 and activate all the antennas during transmission. The power level of the transmitted carrier signal differs based on the bit mapped onto the specific antenna. With this variation, the number of paths over which the modulated signal is received is multiplied, hence the probability of correct detection at the receive side is further increased.

These are only two of many different variations of the IM idea to map information onto the indices of the transmit antennas. Another advantage of the IM technique is that it allows one to design the system model with a lower number of RF chains than that of conventional systems. Since this enables a diversity gain in the spatial domain, a system that consumes lower power than the counterparts can be designed without compromising the SE.

Index Modulation in the Frequency Domain. Next, a wideband channel application of the IM technique is considered. Without loss of generality, a system with $N = 4$ available carriers is considered as depicted in Figure 2.4. In contrast to the conventional wideband system in which all the available carriers are modulated, the activation pattern of the carriers is selected based on the incoming bit stream. In this specific scenario, $K = 3$ out of $N = 4$ carriers are modulated and the activation pattern of the carriers convey additional information to the conventional modulation. As a result, $\lfloor \log_2 \binom{N}{K} \rfloor$ bits of information are transmitted in addition to the $K \log_2 M$ conventionally modulated bits. Note that the wideband application of the IM can be combined with the systems in Figure 2.3a and Figure 2.3b. Therefore, $\log_2 N_\mathrm{T} + \lfloor \log_2 \binom{N}{K} \rfloor + K \log_2 M$ bits for Example I and $N_\mathrm{T} + \lfloor \log_2 \binom{N}{K} \rfloor + K \log_2 M$ bits for Example II can be transmitted per channel use, respectively.

Although IM techniques provide advantages, e.g., increased SE and EE, CSI is generally crucial in order to detect the bits that are transmitted by one of the IM techniques. The resulting system model is therefore complex and more sophisticated detection algorithms are required.

This thesis deals with a system model that can operate both in narrowband and wideband channels as well as in low and high frequency bands and that can satisfy the SE and EE demands of future wireless networks. Hence, this thesis discusses transceiver architectures that employ IM techniques with linear precoders and combiners and reduced complexity detection schemes. This section has unveiled the IM techniques that will be adopted to design a large-scale MIMO system to cope with the growing SE and EE demands in wireless networks. The surveys on the SM in [80] and IM in [25, 26] and the references therein provide a broader look into the various applications for further reading.

2.4 Single Cell Multiuser Networks

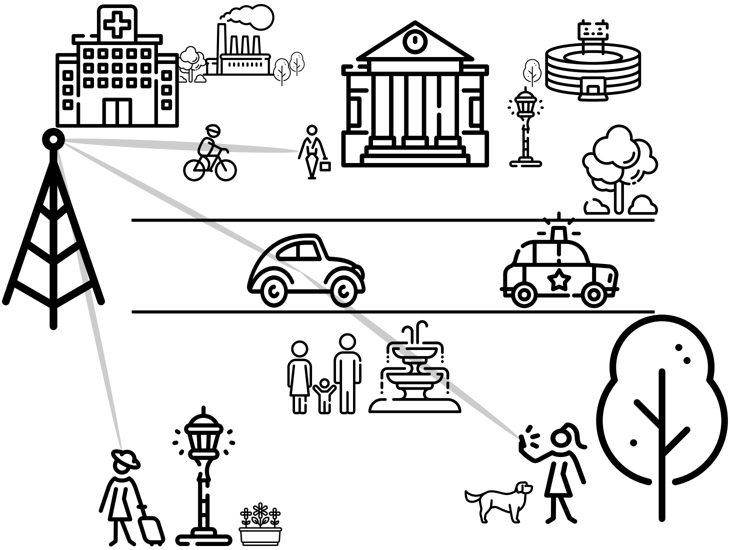


Figure 2.5: Illustration of a generic network architecture for a single cell multiuser wireless communications scenario.

An illustration for a wireless communications network with a single BS and multiple UTs is depicted in Figure 2.5. It is assumed that both the BS and the UTs are equipped with multiple antennas in order to exploit the spatial multiplexing. The technologies that are utilized at each terminal and the signal processing blocks are elaborated in Chapter 3 and Chapter 4. Therein, case specific transceiver architectures are explained in detailed and the performance of the resulting systems are analyzed.

Considering the UL data transmission phase of the MU scenario, the received signal at the BS can be expressed as

$$\mathbf{r} = \sum_{u=1}^{U} \sqrt{P_u}\mathbf{H}_u\mathbf{x}_u + \mathbf{n}, \tag{2.12}$$

where P_u is the transmit power budget of the user u, $\mathbf{H}_u \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{T}}$ is the channel between the BS and the user, $\mathbf{x}_u$ is the transmitted signal and $\mathbf{n}$ is the additive noise. Due to the interference, received signal $\mathbf{x}_u$ is distorted by the signals transmitted from the remaining UTs. Similarly, the received signal at user u considering the DL data transmission phase is

$$\mathbf{r}_u = \sum_{i=1}^{U} \sqrt{P_i}\mathbf{H}_u\mathbf{x}_i + \mathbf{n}. \tag{2.13}$$

where $\mathbf{x}_{i\neq u}$ refers to the transmitted signal from the interfering UTs. As a consequence of the IUI, the probability of the erroneous detection increases. Signal processing methods to mitigate the IUI should be adopted to increase the communication quality. Therefore, CSI should be acquired at least at the BS to enable designing precoders and combiners according to the channel statistics between the BS and UTs.

A time division duplex (TDD) transmission is adopted in which the BS acquires the CSI during the UL training phase. According to the scenario considered in Chapter 3 and Chapter 4 the UTs either do not require full CSI but partial information on the channel statistics. In the case that the UTs require parts of the CSI on the channel statistics, the required information is fed back during the DL training phase. In the proposed system models, IM techniques are exploited in order to increase the SE of the network by considering EE requirements as well. The systems are designed to employ linear precoders and combiners to obtain a low complexity architecture. Detection algorithms are designed to jointly decode the IM and conventionally modulated parts of the bit stream.

Chapter 3 will develop a system model that operates on a wideband channel environment and employs IM techniques on the subcarriers of multicarrier systems in order to increase the SE and EE of the overall network. The transmission model will be developed based on (2.3) for OFDM-IM and GFDM-IM systems considering the MU communications. When a wireless communication system operates in the low frequency band, the transmitted signal travels through a rich scattering channel environment since the wavelength of the electromagnetic wave is large which results in lower absorption and more reflection and diffraction while passing through the channel. A common way of simulating such channel environments is a Rayleigh fading model which can be represented as an i.i.d. complex Gaussian distribution with $\mathcal{CN}(0,1)$. Hence, a multipath Rayleigh fading channel is adopted while studying the UL and DL transmission performance of the proposed system model in Chapter 3.

A system model that employs IM in the spatial domain and operates in the mmWave band will be developed in Chapter 4. The transmission model will be developed based on (2.10) for hybrid beamforming with spatial modulation (HBFSM) systems considering the MU communications. In order to study the UL and DL transmission performances of the system, a narrowband channel environment is adopted which follows a geometry-based model [81, 82]. The CSI can be acquired by a pilot-based channel estimation using the polynomial expansion scheme proposed in [83] or a low complexity adaptive compressed-sensing-based algorithm proposed in [84] which offer low complexity methods. The number of scattering clusters is limited in mmWave band due to the severe path-loss of the waves traveling at high frequencies. In order to take this effect into consideration in the system performance evaluation, the channel matrix between the BS and a UT is modeled as the well-known geometry-based model

$$\mathbf{H} = \sqrt{\frac{N_\mathrm{T} N_\mathrm{R}}{C}} \sum_{c=1}^{C} \alpha_c \, \mathbf{a}_\mathrm{R}\left(\theta_c\right) \mathbf{a}_\mathrm{T}\left(\phi_c\right)^\mathrm{H} . \tag{2.14}$$

Herein, C is the number of scattering clusters of the channel $\mathbf{H} \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{R}}$ between the BS and the UT where N_T and N_R denote the number of transmit and receive antennas, respectively. Furthermore, α_c is the gain of the c-th path that is modeled as $\mathcal{CN}(0,1)$, while $\theta_c \sim \mathcal{U}[0, 2\pi]$ and $\phi_c \sim \mathcal{U}[0, 2\pi]$ represent the azimuth angle of departure (AoD) from the BS and angle of arrival (AoA) at the UT. By assuming uniform linear arrays (ULA), the transmit and receive array response vectors of the c-th path $\mathbf{a}_{\mathrm{T}}(\phi_c)$ and $\mathbf{a}_{\mathrm{R}}(\theta_c)$ are generated as

$$\mathbf{a}(\varphi) = \frac{1}{\sqrt{N}} \left[1, \mathrm{e}^{jkd\sin(\varphi)}, \ldots, \mathrm{e}^{j(N-1)kd\sin(\varphi)}\right]^{\mathrm{T}}, \tag{2.15}$$

where φ is the angle of the considered path and N is the number of elements in the array. Moreover, $k = \frac{2\pi}{\lambda}$ where λ is the signal wavelength and $d = \frac{\lambda}{2}$ is the inter-element spacing.

The channel model in (2.14) can be decomposed as

$$\mathbf{H} = \mathbf{A}_{\mathrm{R}}\mathbf{D}\mathbf{A}_{\mathrm{T}}^{\mathrm{H}}, \tag{2.16}$$

where matrices $\mathbf{A}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{R}} \times C}$ and $\mathbf{A}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{T}} \times C}$ consist of the antenna array response vectors of all the paths

$$\mathbf{A}_{\mathrm{R}} = [\mathbf{a}_{\mathrm{R}}(\theta_1), \mathbf{a}_{\mathrm{R}}(\theta_2), \ldots, \mathbf{a}_{\mathrm{R}}(\theta_C)], \tag{2.17a}$$

$$\mathbf{A}_{\mathrm{T}} = [\mathbf{a}_{\mathrm{T}}(\phi_1), \mathbf{a}_{\mathrm{T}}(\phi_2), \ldots, \mathbf{a}_{\mathrm{T}}(\phi_C)]. \tag{2.17b}$$

Finally, the diagonal matrix $\mathbf{D} \in \mathbb{C}^{C \times C}$ has the complex path gains at the diagonal entries $\sqrt{N_{\mathrm{T}} N_{\mathrm{R}}/C}[\alpha_1, \alpha_2, \ldots, \alpha_C]$.

IM techniques provide additional diversity which can be exploited by mapping the information bits to the various media. These bits are transmitted inherently without requiring transmission power. Although this feature enables increasing the SE of the network by maintaining the EE of the system, it requires more advanced signal processing and detection algorithms. In this respect, this thesis focuses on developing system models that employ linear precoders and combiners by adopting ZF and MMSE receivers, which allow processing the signal to enable spatial multiplexing, to mitigate the IUI and to prepare the signal for detection. Since the BS has more computational power than the UTs, the system performs a large part of the signal processing at the BS side.

The following sections present in detail the system models for each specific case. The problem of developing system models that adopt IM techniques in wideband channels with multicarrier transmission is addressed in Chapter 3. Then, large-scale MIMO systems that operate in narrowband channel environments are developed in Chapter 4 where IM is exploited in the spatial domain.

Chapter 3

Multicarrier Systems with Index Modulation

This chapter studies the IM techniques in wideband communication scenarios. Since the demand on data rates increases drastically, multicarrier systems will continue to be deployed due to their high SE nature. In order to meet with ever-increasing demands on the wireless communication networks, OFDM and GFDM multicarrier systems that exploit IM techniques are proposed for MU scenarios.

Although the advantages brought by the conventional OFDM system were sufficient for LTE networks, the SE and EE need to be further improved to satisfy the excessive requirements of future wireless networks. Therefore, Section 3.1 develops a system model that exploits both the conventional OFDM system and IM technique. Bringing an IM technique into the OFDM system architecture enables forming a three dimensional constellation diagram where the third dimension depends on the index activation pattern of the subcarriers. As it is already elucidated in Section 2.3, the source of information conveyed by the index activation pattern is transmitted without any additional power consumption for the conventionally modulated symbol. The system model for the OFDM-IM is developed next and improvement on the performance is corroborated by comparison with existing systems.

Furthermore, a GFDM system offers a non-orthogonal multicarrier communication that can be considered as a generalized form of the conventional OFDM system. Although the GFDM system architecture is more complex in comparison to the OFDM system it provides larger SE in return. A GFDM symbol consists of multiple time and frequency slots and requires a single CP insertion for the entire data block. Hence, the SE is improved since the need of the redundancy is much lower than in the conventional OFDM system. As discussed for the OFDM system, IM techniques can be exploited for a GFDM system as well to further increase the SE and EE of the network. In this respect, Section 3.2 develops a GFDM-IM system model for MU networks. In order to corroborate the improvement by exploiting the GFDM-IM system, the performance is compared with the developed OFDM-IM system and the conventional GFDM system.

3.1 Orthogonal Frequency Division Multiplexing

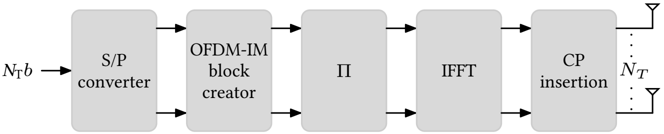
This section introduces the system model for an MU network that adopts an OFDM-IM system for wireless transmission. The general specifications of the OFDM-IM system are elaborated in Section 3.1.1 and the transceiver architectures are developed for the UL and the DL transmission phases. Furthermore, Section 3.1.2 studies the performance of the proposed system model with reference to the achieved SE and PAPR.

3.1.1 System Model

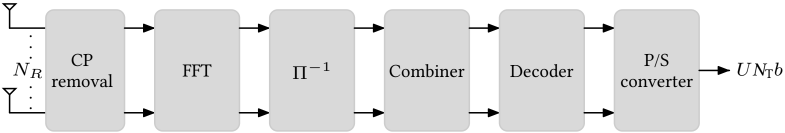
A single cell MU network architecture was introduced in Section 2.4. This section focuses on the detailed design of each terminal in the network by considering a wideband communications system. In order to combat the frequency selectivity and to achieve high data rates the available spectrum can be divided into N_{tot} narrowband subchannels. In this way, the symbols transmitted through the subchannels experience almost flat fading. A system that adopts a conventional OFDM system would modulate all the available subcarriers per channel use and consequently, $N_{\text{tot}} \log_2 M$ bits would be transmitted per channel use, where M is the modulation order. In contrast, a subset of the available subcarriers is chosen to be active and modulated with M-ary symbols when the OFDM-IM system is considered. Thus, a part of the incoming bit stream is mapped onto the index activation pattern of the subcarriers. Note that the bits that are mapped onto the subcarrier indices are transmitted inherently with the knowledge of the index activation pattern that requires no additional transmission power.

Assume that K_{tot} out of N_{tot} available subcarriers are chosen to be active according to the incoming bit stream per channel use. M-ary modulated symbols are transmitted on these K_{tot} subcarriers while $N_{\text{tot}} - K_{\text{tot}}$ subcarriers remain idle. Thus, the size of the set that contains the index activation pattern is $\binom{N_{\text{tot}}}{K_{\text{tot}}}$ which is a large number for realistic values of N_{tot}. Therefore, a grouping architecture is adopted in order to reduce the complexity of data detection on the receiver side. Hence, the subcarriers are divided into G groups in order to have a feasible receiver structure. The resulting number of available and active subcarriers in each group are $N = N_{\text{tot}}/G$ and $K = K_{\text{tot}}/G$. Consequently, $b = Gp$ bits are transmitted per channel use where $p = p_1 + p_2$ is such that $p_1 = \lfloor \log_2 \binom{N}{K} \rfloor$ bits belong to the IM part and $p_2 = K \log_2 M$ bits are transmitted as M-ary modulated symbols.

In the following subsections, the system model of OFDM-IM for both UL and



(a) User terminal.



(b) Base station.

Figure 3.1: OFDM-IM block diagram for the uplink transmission phase.

DL transmission in MU networks is introduced. The BS design to deal with the IUI for these systems is explained in detail.

Uplink Transmission

The transmitter and receiver block diagrams for the uplink transmission are depicted in Figure 3.1a and Figure 3.1b, respectively [37]. The MU network comprises U UTs equipped with N_T transmit antennas and a BS equipped with N_R receive antennas. Since a low complexity system model is targeted, the condition that the BS has sufficient degrees of freedom should be satisfied. Hence, it is assumed that $N_\mathrm{R} \geq UN_\mathrm{T}$ in order to design a linear combiner which operates with high performance. Both the users and the BS have fully digital architectures. It is assumed that the BS has perfect CSI and the users are not aware of the channel statistics.

The bit stream for each user consists of $N_\mathrm{T}b$ bits to be transmitted to the BS. The S/P converter block splits the incoming bit stream for each antenna and b bits are assigned per antenna to be transmitted per channel use. The OFDM-IM block creator first starts by dividing these bits into $G = N_\mathrm{tot}/N$ groups such that $b = Gp$. Note that the number of groups G and the number of available subcarriers at each group N can be designed according to the system requirements, e.g. high SE, low PAPR, or a simple receiver architecture. Furthermore, p bits are divided into two parts according to the modulation order M and the number of active subcarriers K. The combination of the K

active subcarriers that will be modulated for the transmission over the channel is selected from the look-up table according to p_1 bits

$$\mathbf{i}_{tu}^{g} = [i_{tu}^{g}(1), i_{tu}^{g}(2), \dots, i_{tu}^{g}(K)]^{\mathrm{T}}, \tag{3.1}$$

where $g \in \{1, 2, \dots, G\}$, $t \in \{1, 2, \dots, N_{\mathrm{T}}\}$ and $u \in \{1, 2, \dots, U\}$. Herein, $i_{tu}^{g}(k) \in \{1, 2, \dots, N\}$ is the k-th selected subcarrier index of the g-th group to transmit from the t-th transmit antenna of the u-th user. Note that, when p_1 bits are not a power of 2, the look-up table will be truncated and the decoding of the active subcarrier indices must be adapted accordingly. Furthermore, the remaining p_2 bits are used to generate M-ary modulated data to transmit on the selected subcarriers

$$\mathbf{s}_{tu}^{g} = [s_{tu}^{g}(1), s_{tu}^{g}(2), \dots, s_{tu}^{g}(K)]^{\mathrm{T}}. \tag{3.2}$$

Herein, $s_{tu}^{g}(k)$ is the modulated symbol to be assigned to the k-th selected subcarrier index of the g-th group to transmit from the t-th transmit antenna of the u-th user. The resulting subcarrier indices and the modulated symbols for all groups are collected in $\mathbf{i}_{tu} \in \mathbb{C}^{K_{\mathrm{tot}} \times 1}$ and $\mathbf{s}_{tu} \in \mathbb{C}^{K_{\mathrm{tot}} \times 1}$ vectors to form the OFDM-IM block per antenna

$$\mathbf{i}_{tu} = [(\mathbf{i}_{tu}^{1})^{\mathrm{T}}, (\mathbf{i}_{tu}^{2})^{\mathrm{T}}, \dots, (\mathbf{i}_{tu}^{G})^{\mathrm{T}}]^{\mathrm{T}}, \tag{3.3}$$

$$\mathbf{s}_{tu} = [(\mathbf{s}_{tu}^{1})^{\mathrm{T}}, (\mathbf{s}_{tu}^{2})^{\mathrm{T}}, \dots, (\mathbf{s}_{tu}^{G})^{\mathrm{T}}]^{\mathrm{T}}. \tag{3.4}$$

Once the selection procedure is completed, the modulated data is assigned to the active subcarriers. First, the frequency domain OFDM-IM symbol of the g-th group $\mathbf{x}_{tu}^{g} = [x_{tu}^{g}(1), x_{tu}^{g}(2), \dots, x_{tu}^{g}(N)]^{\mathrm{T}}$ is generated where $x_{tu}^{g}(n)$ is the signal to be transmitted on the n-th subcarrier. Thereafter, in order to ensure that the symbols at each subcarrier are transmitted through uncorrelated channels, interleaved grouping is performed to build the final OFDM-IM block

$$\mathbf{x}_{tu}[1, \dots, g, \dots, g+G, \dots, g+(N-1)G, \dots, N_{\mathrm{tot}}]^{\mathrm{T}} = \mathbf{x}_{tu}^{g}[1, 2, \dots, N]^{\mathrm{T}} \tag{3.5}$$

where $g \in \{1, 2, \dots, G\}$, $t \in \{1, 2, \dots, N_{\mathrm{T}}\}$ and $u \in \{1, 2, \dots, U\}$. The interleaver block Π performs this operation and locates the subcarriers that belong to the same group separated G subchannels from each other. Note that $G(N-K)$ entries of $\mathbf{x}_{tu}$ are zero which is the number of subcarriers that are selected to stay idle during the data transmission. After reorganizing the frequency domain symbol $\mathbf{x}_{tu}$, an N_{tot}-point IFFT operation is performed. Thereafter, a

CP of length N_{CP} is inserted to eliminate the ISI. The resulting OFDM-IM block with length $N_{\text{tot}} + N_{\text{CP}}$ is transmitted from each transmit antenna over a C-tap frequency-selective Rayleigh fading channel for each user. The input-output relationship in the frequency domain per subcarrier is

$$\mathbf{r}_n = \sqrt{P}\sum_{u=1}^{U} \mathbf{H}_{nu}\mathbf{x}_{nu} + \mathbf{n}_n, \tag{3.6}$$

where $n \in \{1, 2, \ldots, N\}$ is the index of the subcarrier. Therein, P is the total transmit power budget that is equally allocated for each user as $P = P_u/U$, $\mathbf{x}_{nu} \in \mathbb{C}^{N_{\text{T}}\times 1}$ is the transmitted OFDM-IM block of the u-th user, $\mathbf{H}_{nu} \in \mathbb{C}^{N_{\text{R}}\times N_{\text{T}}}$ is the corresponding effective channel matrix with $\mathcal{CN}(0,1)$ distributed elements, and $\mathbf{n}_n \in \mathbb{C}^{N_{\text{R}}\times 1}$ is a noise vector with i.i.d. $\mathcal{CN}(0,\sigma^2)$ distributed elements.

After the reception of the OFDM-IM block the BS first removes the CP. Afterwards, an N_{tot}-point FFT is performed to transform the signal back to the frequency domain. In order to prepare the signal for combining and demodulation the OFDM-IM block with length N_{tot} is divided into G groups by the deinterleaving block Π^{-1} where each group comprises N subcarriers. Thereafter, MMSE combining is performed to eliminate the IUI and to successfully reconstruct the transmitted symbol. The resulting post-processed signal at each group is

$$\mathbf{y}_n^g = \mathbf{W}_n^g \mathbf{r}_n^g = \mathbf{W}_n^g \mathbf{H}_n^g \mathbf{x}_n^g + \mathbf{W}_n^g \mathbf{n}_n^g, \tag{3.7}$$

where

$$\mathbf{W}_n^g = ((\mathbf{H}_n^g)^{\text{H}}\mathbf{H}_n^g + \sigma^2 \mathbf{I}_{UN_{\text{T}}}/P)^{-1}(\mathbf{H}_n^g)^{\text{H}} \tag{3.8}$$

is the MMSE combiner of the g-th group and the n-th subcarrier such that the channel matrix $\mathbf{H}_n^g = [\mathbf{H}_{n1}^g, \mathbf{H}_{n2}^g, \ldots, \mathbf{H}_{nU}^g] \in \mathbb{C}^{N_{\text{R}}\times UN_{\text{T}}}$ is rearranged and consists of all single user channel matrices. The $UN_{\text{T}} \times 1$ dimensional signal vector $\mathbf{x}_n^g = [(\mathbf{x}_{n1}^g)^{\text{T}}, (\mathbf{x}_{n2}^g)^{\text{T}}, \ldots, (\mathbf{x}_{nU}^g)^{\text{T}}]^{\text{T}}$ contains the transmitted data of all users.

After employing the MMSE combiner to each subcarrier, the resulting $\mathbf{y}_n^g \in \mathbb{C}^{UN_{\text{T}}\times 1}$ symbols for each group are rearranged as

$$\mathbf{Y}^g = [\mathbf{Y}_1^g, \mathbf{Y}_2^g, \ldots, \mathbf{Y}_U^g], \tag{3.9}$$

where $\mathbf{Y}_u^g \in \mathbb{C}^{N\times N_{\text{T}}}$ is the post-processed signal of the u-th user. From the post-processed matrix $\mathbf{Y}^g$ it is easy to detect the active indices. Assume $\in \mathbb{Z}^{2^{p_1}\times K}$

is the matrix that contains all possible index combinations of the truncated look-up table. An example of is given for $N = 4$ and $K = 2$

$$= \begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 3 & 4 & 3 \end{bmatrix}^T . \tag{3.10}$$

By adopting such a look-up table, the decision metric is calculated for each combination

$$\mathbf{m}(j) = \sum_{k=1}^{K} |\mathbf{y}_{tu}^{g}((j,k))|, \tag{3.11}$$

where $j \in \{1, 2, \ldots, 2^{p_1}\}$, $g \in \{1, 2, \ldots, G\}$, $t \in \{1, 2, \ldots, N_\mathrm{T}\}$, and $u \in \{1, 2, \ldots, U\}$. Furthermore, the active subcarrier combination is determined by searching for the maximum entry of $\mathbf{m}$

$$\hat{j} = \arg\max_{j} \mathbf{m}(j). \tag{3.12}$$

Finally, $\hat{j}$ is inserted into the look-up table and the subcarrier activation pattern is detected as

$$\hat{\mathbf{i}}_{tu}^{g} = (\hat{j}, :). \tag{3.13}$$

After detection of the active subcarrier indices the symbols on these subcarriers are collected in $\hat{\mathbf{x}}_{tu}^{g} = \mathbf{y}_{tu}^{g}(\hat{\mathbf{i}}_{tu}^{g})$. Therein, $\hat{\mathbf{i}}_{tu}^{g}$ and $\hat{\mathbf{x}}_{tu}^{g}$ are $\mathbb{C}^{K \times 1}$ vectors that contain indices of K active subcarriers and M-ary modulated data transmitted from the t-th transmit antenna of the u-th user, respectively. After collecting the signals on the detected subcarriers, M-ary demodulation is performed in order to detect the symbols $\hat{\mathbf{s}}_{tu}^{g}$. The resulting detected indices and M-ary modulated data of all groups from all users are $[\hat{\mathbf{I}}_1, \ldots, \hat{\mathbf{I}}_U]$ and $[\hat{\mathbf{S}}_1, \ldots, \hat{\mathbf{S}}_U]$, respectively, where $\hat{\mathbf{I}}_u$ and $\hat{\mathbf{S}}_u$ are $\in \mathbb{C}^{K_\mathrm{tot} \times N_\mathrm{T}}$ for $u \in \{1, 2, \ldots, U\}$.

The performance of the OFDM-IM system is presented for the UL transmission phase in Figure 3.2 [37]. In order to study the performance of the developed system model the uncoded BER is considered as the performance metric. The OFDM-IM system is proposed in order to increase the amount of information bits that is transmitted per channel use without increasing the total transmit power. In this way, it is aimed to meet the EE and SE requirements of future wireless networks by adopting progressive signal processing techniques. As already elaborated throughout this section, a part of the information bits is mapped onto the activation pattern of the subcarriers that carry the conventionally modulated

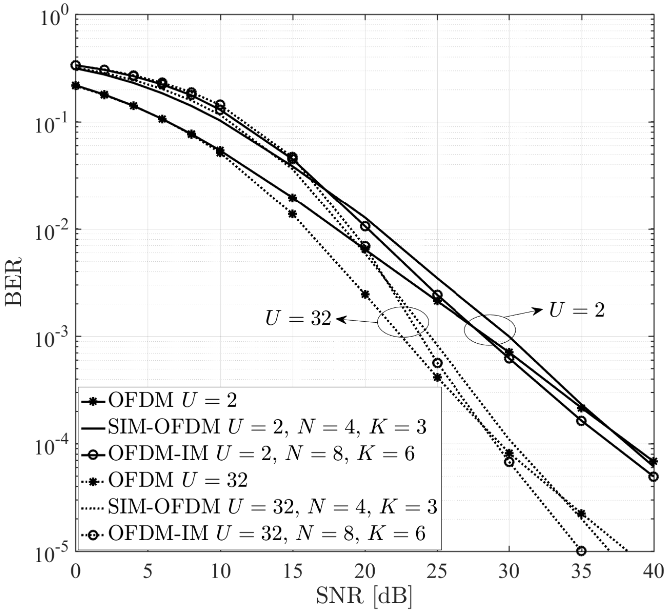


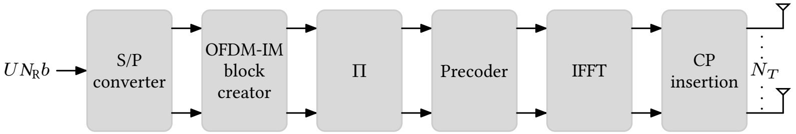
Figure 3.2: Uncoded bit error rate performance comparison of the uplink transmission phase for the OFDM-IM, SIM-OFDM and OFDM systems in networks consisting of 2 and 32 users, respectively ($N_{\mathrm{T}} = 1$, $N_{\mathrm{R}} = UN_{\mathrm{T}}$, $N_{\mathrm{tot}} = 128$ with 4-QAM).

symbols. In contrast, all available subcarriers are actively modulated to transmit the symbols with the conventional OFDM system. Although the conventional OFDM system has a lower complexity than the OFDM-IM since there is only one source of information to encode and decode, it also results in a large PAPR as elucidated in Section 3.1.2. An analog system model was proposed for the MU scenario in [23] where SIM is adopted to map the information bits to the subcarrier activation pattern. However, the number of available subcarriers per group is fixed (due to the modulation order $G = M$) and a single subcarrier remains idle during the transmission. The resulting system achieves the same SE as the conventional OFDM system.

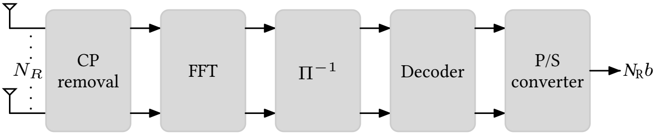
In order to study the performance of the proposed system model the uncoded BER of the aforementioned system models are compared with the OFDM-IM for the MU scenario. A wideband channel with $N_{\text{tot}} = 128$ subcarriers is considered to transmit information to the BS. The channel between the BS and the users is assumed to be a $C = 8$-tap frequency selective Rayleigh fading channel. Therefore, a CP of length $N_{\text{CP}} = 16$ is inserted at the end of each block in order to eliminate the ISI. It is assumed that the BS has perfect CSI which is acquired during the training phase prior to the data transmission. Furthermore, a 4-quadrature amplitude modulation (QAM) scheme is adopted to modulate the active subcarriers. Figure 3.2 depicts the UL BER performance of the system with $U = 2$ and $U = 32$ users, respectively. Thus, the effect of the IUI to the performance can be studied as well. It is assumed that each user has a single transmit antenna and that the number of antennas at the BS equals the number of users $N_{\text{R}} = UN_{\text{T}}$ so that the linear MMSE receiver operates with high performance. Since the number of active subcarriers is determined by the modulation order in the SIM-OFDM system, it is assumed that $N = 4$, $K = 3$ for the simulations. On the other hand, the numbers of total and active subcarriers in a group can be determined independently for the OFDM-IM system. Therefore, these parameters are selected as $N = 8$ and $K = 6$ for the OFDM-IM case to unveil the advantages of the flexibility of the developed system model.

Figure 3.2 elucidates that the OFDM-IM system achieves better performance than the conventional OFDM and SIM-OFDM in both the 2- and the 32-user scenarios. Since only GK subcarriers carry information on the spectrum, the average distance, i.e., subcarrier spacing, between the M-ary symbols on the spectrum is larger than in the other two systems. Furthermore, the interleaved grouping provides additional protection to the modulated symbols against correlated channels. Note also that the selection of the available subcarriers and the active subcarriers in a group is independent from the modulation order. On the contrary, in SIM-OFDM system, the number of subcarriers in a group is defined as M that is also the modulation order. Moreover, only one subcarrier among the available subcarriers in a group remains idle during the data transmission. This property of the system limits the system design while in OFDM-IM N and K can be chosen freely according to the system requirements. This feature of OFDM-IM provides a flexible design and allows a trade-off between SE and PAPR.

The relatively poor performance of the OFDM-IM in a low SNR regime can be explained with the higher erroneous detection of the subcarrier indices. When the active subcarrier indices are not detected correctly at the receiver side decoding of the modulated symbols on these detected subcarriers will also be erroneous. Another disadvantage brought by the IM part of the system



(a) Base station.



(b) User terminal.

Figure 3.3: OFDM-IM block diagram for the downlink transmission phase.

is the additional complexity of the decoding process. There are intermediate steps between combining and M-ary demodulation in order to detect the active subcarrier indices. Despite its additional complexity the OFDM-IM system has a better error performance than the conventional OFDM and the SIM-OFDM systems. Furthermore, OFDM-IM provides a flexible system model that enables designing the system according to the requirements such as high SE, low PAPR and low BER.

Downlink Transmission

The transmitter and the receiver block diagrams for the DL transmission are depicted in Figure 3.3a and Figure 3.3b, respectively [37]. This time, the BS transmits U data streams simultaneously to the users equipped with N_R receive antennas. Moreover, the BS is equipped with N_T transmit antennas such that $N_\mathrm{T} \geq UN_\mathrm{R}$ in order to design a linear precoder that operates with high performance. The OFDM-IM symbol $\mathbf{x}_{ru}^{g}$ for each group is formed in the same manner with the UL transmission phase where $g \in \{1, 2, \dots, G\}$, $r \in \{1, 2, \dots, N_\mathrm{R}\}$, and $u \in \{1, 2, \dots, U\}$. Thereafter, the OFDM-IM block per antenna and group $\mathbf{x}_{ru} \in \mathbb{C}^{N_\mathrm{tot} \times 1}$ is built by performing interleaved grouping in order to transmit the subsymbol through uncorrelated channels.

In the uplink transmission phase, only the BS has the CSI while the users are oblivious of the channel statistics. On the contrary, the users have partial information about the channel statistics in the DL transmission phase in order to successfully decode the received signal. By taking advantage of the CSI, an MMSE-based precoder is employed to eliminate the IUI. The precoded frequency domain OFDM-IM symbol is

$$\bar{\mathbf{x}}_n = \beta_n \sum_{u=1}^{U} \mathbf{P}_{nu}\mathbf{x}_{nu}, \tag{3.14}$$

where $n \in \{1, 2, \ldots N\}$ is the index of the subcarrier. Therein, $\mathbf{x}_{nu} \in \mathbb{C}^{N_\mathrm{R}\times 1}$ is the signal on the n-th subcarrier to be transmitted to the u-th user. $\mathbf{P}_{nu} \in \mathbb{C}^{N_\mathrm{T}\times N_\mathrm{R}}$ is the precoding matrix designed to mitigate the IUI such that

$$\mathbf{P}_n = \left(\mathbf{H}_n^\mathrm{H}\mathbf{H}_n + \frac{\sigma^2\mathbf{I}_{UN_\mathrm{T}}}{P_u}\right)^{-1}\mathbf{H}_n^\mathrm{H}. \tag{3.15}$$

Herein, $\mathbf{P}_n = [\mathbf{P}_{n1}, \mathbf{P}_{n2}, \ldots, \mathbf{P}_{nU}]$ is the precoding matrix for all users where $u \in \{1, 2, \ldots, U\}$. Furthermore, β_n is the normalization coefficient per subcarrier in order to satisfy the total transmit power constraint

$$\beta_n = \sqrt{\frac{P}{\mathrm{tr}\{\mathbf{P}_n\mathbf{P}_n^\mathrm{H}\}}}. \tag{3.16}$$

After precoding, in order to separate the signals of all users, the OFDM-IM symbol is transformed into the time domain by employing an IFFT. The time domain signal is transmitted through a C-tap frequency-selective Rayleigh fading channel following the insertion of the CP to eliminate ISI. The received signal at the u-th user after removing the CP and interleaved regrouping is

$$\mathbf{r}_{nu}^g = \mathbf{H}_{nu}^g\bar{\mathbf{x}}_n^g + \mathbf{n}_{nu}^g, \tag{3.17}$$

where $\bar{\mathbf{x}}_n^g \in \mathbb{C}^{N_\mathrm{T}\times 1}$ is the transmitted symbol, $\mathbf{H}_{nu}^g \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{T}}$ is the Rayleigh fading channel matrix between the u-th user and the BS where the elements follow $\mathcal{CN}(0, 1)$ distribution and $\mathbf{n}_{nu}^g \in \mathbb{C}^{N_\mathrm{R}\times 1}$ is the noise vector with i.i.d. elements that follow circularly symmetric complex Gaussian distribution with zero mean and σ^2 variance, respectively.

Since the precoding has been performed in order to eliminate the IUI, the user only needs to acquire the normalization coefficient β_n to calculate $\mathbf{y}_{nu}^g = \mathbf{r}_{nu}^g/\beta_n$. The user then proceeds to decode the received OFDM-IM symbol. Note that the proposed design reduces the user training burden significantly since

β_n is a scalar. In the case that users require complete CSI, the channel matrix $\mathbf{H}_n \in \mathbb{C}^{N_\mathrm{T} \times N_\mathrm{R}}$ per subcarrier n would be required where $n \in \{1, 2, \ldots, N_\mathrm{tot}\}$, which results in a long training period for each user and thus reduces the spectrum efficiency of the system.

Once the user obtained the post-processed signal $\mathbf{y}^g_{nu}$, the active subcarrier indices $\hat{\mathbf{i}}^g_{ru}$ and the M-ary modulated symbols $\hat{\mathbf{s}}^g_{ru}$ on these subcarriers are decoded in the same manner as in the UL transmission phase.

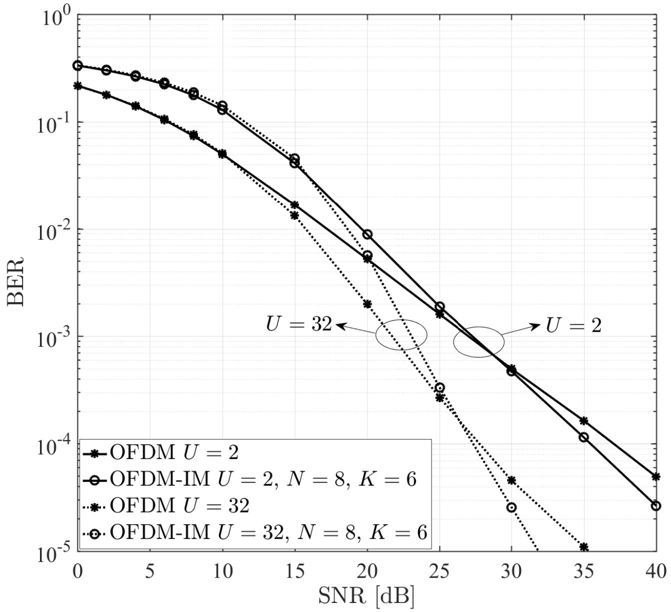


Figure 3.4: Uncoded bit error rate performance comparison of the downlink transmission phase for the OFDM-IM and OFDM systems in networks consisting of 2 and 32 users, respectively ($N_\mathrm{T} = UN_\mathrm{R}$, $N_\mathrm{R} = 1$, $N_\mathrm{tot} = 128$, $N = 8$, $K = 6$ with 4-QAM).

As for the UL transmission phase, the performance of the OFDM-IM system is studied by considering the uncoded BER as the performance metric. For the DL transmission phase, the only comparable system model is the conventional OFDM since the SIM-OFDM system was studied for the UL transmission phase in [23]. Figure 3.4 elucidates the error performance of the developed system model compared to the conventional OFDM system for the MU scenario [37]. Similar to the UL transmission phase, a wideband channel with $N_{\text{tot}} = 128$ subcarriers is considered to transmit information to the UTs. The channel between the BS and the users is assumed to be a $C = 8$-tap frequency selective Rayleigh fading channel. Hence, a CP of length $N_{\text{CP}} = 16$ is inserted at the end of each block in order to eliminate the ISI. It is assumed that the BS acquires the CSI perfectly during the training phase and computes the MMSE precoders at the begining of each coherence time. Furthermore, a 4-QAM is adopted to modulate the active subcarriers. During the DL transmission phase, the BS transmits parallel data streams to the single antenna users in the network. Hence, the number of transmit antennas at the BS equals the number of users $N_{\text{T}} = UN_{\text{R}}$ which also enables designing high-performance MMSE precoders. The number of total and active subcarriers in a group is determined as $N = 8$ and $K = 6$ for an OFDM-IM system, respectively. As shown in Figure 3.4 that the DL performance of the OFDM-IM system is similar to that of the UL scenario. The MMSE-precoder successfully mitigates the IUI and the OFDM-IM outperforms the conventional OFDM in a mid- to high-SNR regime.

3.1.2 Performance Analysis

Next, the performance of the OFDM-IM system is studied from the SE and PAPR perspective. In order to unveil the advantages brought by the proposed system model the performance is compared with that of conventional OFDM and SIM-OFDM [23] systems.

Spectrum Efficiency

The upper bound of the achievable SE of an OFDM-IM system is

$$SE_{\text{OFDM-IM}} = \frac{G(\lfloor \log_2 \binom{N}{K} \rfloor + K \log_2 M)}{N_{\text{tot}} + N_{\text{CP}}}. \tag{3.18}$$

Herein, N_{tot} is the total number of subcarriers allocated for the communication, N_{CP} is the length of the CP which depends on the channel statistics, and M is the modulation order. The parameters G, N, and K are the numbers of groups, available subcarriers, and active subcarriers per group, respectively. The parameters G, N, and K can be configured according to the system requirements.

On the other hand, the upper bound of the achievable SE of a SIM-OFDM system is

$$SE_{\text{SIM-OFDM}} = \frac{G(\log_2 M + K \log_2 M)}{N_{\text{tot}} + N_{\text{CP}}}. \tag{3.19}$$

Since the number of active subcarriers K in each group is related to the modulation order in the SIM-OFDM system according to $K = M - 1$, the SE can be expressed as

$$SE_{\text{SIM-OFDM}} = \frac{N_{\text{tot}} \log_2 M}{N_{\text{tot}} + N_{\text{CP}}}. \tag{3.20}$$

Furthermore, the upper bound of the achievable SE of a conventional OFDM system is

$$SE_{\text{OFDM}} = \frac{N_{\text{tot}} \log_2 M}{N_{\text{tot}} + N_{\text{CP}}}. \tag{3.21}$$

As can be seen from (3.20) and (3.21), the upper bounds of the SE of the SIM-OFDM and the conventional OFDM systems are equal. On the contrary, the parameters N and K can be configured independently from the modulation order for the OFDM-IM system. The comparison of the SEs of OFDM-IM, SIM-OFDM, and conventional OFDM is given for various parameters. A wideband channel is divided into $N_{\text{tot}} = 128$ subchannels to transmit multiple streams through a frequency selective channel. It is assumed that each stream has the CP length $N_{\text{CP}} = 16$. As it can be seen from Table 3.1, the SEs of the conventional OFDM and the SIM-OFDM systems are equal. The major benefit of the SIM-OFDM system is the lower PAPR value which is explained in the next section. However, the OFDM-IM system can be configured as required, which can result in better SE compared to alternative systems.

Peak-to-Average Power Ratio

In a multicarrier system, there is a large number of independently modulated symbols. Therefore, the power levels of each subcarrier are generally different from each other. The PAPR is defined as the ratio of the maximum power of a sample in a given OFDM transmit symbol to the average power of that OFDM symbol. The large PAPR values result in a degradation of the PA efficiency, thus the PA is forced to operate in the non-linear region. Consequently, there is research on reducing the PAPR and different methods have been proposed over the years [85, 86]. Nevertheless, the PAPR value is proportional to the number of active subcarriers [85, 23], and employing IM techniques provides lower PAPR

Table 3.1: Spectrum-efficiency values of the OFDM-IM, SIM-OFDM, and OFDM systems ($N_{\text{tot}} = 128$, $N_{\text{CP}} = 16$).

System model	M	N	K	SE
OFDM-IM	4	16	13	1.94
	8	32	28	2.75
SIM-OFDM	4	4	3	1.88
	8	8	7	2.66
OFDM	4	-	-	1.88
	8	-	-	2.66

inherently. Since the number of activated subcarriers is smaller than the number of available subcarriers, the OFDM systems that adopt IM attain smaller PAPR values compared to the conventional OFDM and SIM-OFDM systems

$$PAPR_{\text{OFDM-IM}} \leq PAPR_{\text{SIM-OFDM}} < PAPR_{\text{OFDM}}, \tag{3.22}$$

where the PAPR values for the considered multicarrier systems can be expressed for comparison as

$$PAPR_{\text{OFDM-IM}} \propto 10 \log_{10} K_{\text{tot}}^{\text{OFDM-IM}}, \tag{3.23}$$

$$PAPR_{\text{SIM-OFDM}} \propto 10 \log_{10} K_{\text{tot}}^{\text{SIM-OFDM}}, \tag{3.24}$$

$$PAPR_{\text{OFDM}} \propto 10 \log_{10} N_{\text{tot}}. \tag{3.25}$$

3.2 Generalized Frequency Division Multiplexing

In this section the design problem of a GFDM-IM system in an MU network is addressed. Since the GFDM system is a generalization of OFDM the developed system model of GFDM-IM is similar to that of OFDM-IM. However, the general specifications of the GFDM-IM system are elaborated in Section 3.2.1 for clarity, and differences of the transceiver architectures for the UL and the DL transmission phases are emphasized. The performance of the proposed system model with reference to the achieved SE is studied in Section 3.2.2.

3.2.1 System Model

This section focuses on the detailed design of each terminal in an MU network by adopting a GFDM system. The considered GFDM block is filtered arbitrarily and consists of L subsymbols and N_{tot} subcarriers. The transmission of a conventional GFDM block enables transmitting $Q = LN_{\text{tot}}$ symbols per channel use which results in transmitting $Q \log_2 M$ bits where M is the modulation order. Additionally, the IM technique can be exploited similar to the OFDM-IM system in order to transmit a part of the incoming bit stream with the activation pattern of the subcarriers on each subsymbol. As a result, not all available subcarriers at a subsymbol are modulated but K_{tot} out of N_{tot} subcarriers are chosen to carry M-ary modulated signals according to the incoming bit stream. The remaining $L(N_{\text{tot}} - K_{\text{tot}})$ subcarriers stay idle during the transmission and the location of the active subcarriers carries additional information besides the conventionally modulated data.

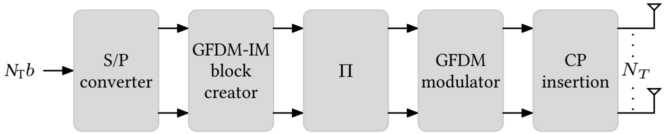
As explained in Section 3.1, when the number of available subcarriers N_{tot} is large the size of the set that contains the index activation pattern is large as well. Thus, a grouping architecture is adopted in order to design a feasible receiver. Let G be the number of groups for each subsymbol, then the number of available and active subcarriers in each group is $N = N_{\text{tot}}/G$ and $K = K_{\text{tot}}/G$, respectively. The resulting total number of bits transmitted from one of the transmit antennas with a GFDM-IM block is $b = LGp$ where $p = p_1 + p_2$ is the total number of bits per group with $p_1 = \lfloor \log_2 \binom{N}{K} \rfloor$ for the IM part and $p_2 = K \log_2 M$ for the M-ary modulated symbols.

In the following subsections, the system model of GFDM-IM for both UL and DL transmission in MU networks is introduced. The BS design to deal with the IUI for these systems is explained in detail.

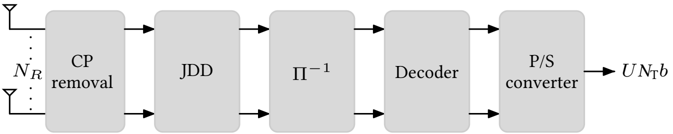
Uplink Transmission

The transmitter and the receiver block diagrams for the uplink transmission are depicted in Figure 3.5a and Figure 3.5b, respectively [38]. In the considered system model, there are U users each equipped with N_{T} transmit antennas and the BS which is equipped with N_{R} receive antennas where $N_{\text{R}} \geq UN_{\text{T}}$. It is assumed that the BS has perfect CSI and that the users are not aware of the channel statistics.

The bit stream at each UT consists of $N_{\text{T}}b$ bits transmitted to the BS during the UL transmission phase. The S/P converter block splits the incoming bit stream



(a) User terminal.



(b) Base station.

Figure 3.5: GFDM-IM block diagram for the uplink transmission phase.

into b bits per antenna. The GFDM-IM block creator first starts by dividing b bits into L groups for each subsymbol. In order to configure the system for the GFDM-IM system with grouping architecture the available N_tot subcarriers of each subsymbol are divided into G groups where $N = N_\mathrm{tot}/G$. Thereafter, according to the IM bits, K subcarriers are activated by using a predefined look-up table as in (3.10). Recall that, according to the number of bits conveyed by the IM part of the system, the look-up table might be truncated. The activation pattern of the K subcarriers that will be modulated for the transmission is selected from the look-up table according to the p_1 bits

$$\mathbf{i}_{ltu}^{g} = [i_{ltu}^{g}(1), i_{ltu}^{g}(2), \dots, i_{ltu}^{g}(K)]^\mathrm{T}, \tag{3.26}$$

where $g \in \{1, 2, \dots, G\}$, $l \in \{1, 2, \dots, L\}$, $t \in \{1, 2, \dots, N_\mathrm{T}\}$, and $u \in \{1, 2, \dots, U\}$. Herein, $i_{ltu}^{g}(k) \in \{1, 2, \dots, N\}$ is the k-th selected subcarrier index of the g-th group of the l-th subsymbol to transmit from the t-th transmit antenna of the u-th user. In the case that p_1 is not an order of 2 and the look-up table Φ is truncated; the decoding of the indices of the active subcarriers must be adapted accordingly. Furthermore, the remaining p_2 bits are used to modulate K M-ary symbols

$$\mathbf{s}_{ltu}^{g} = [s_{ltu}^{g}(1), s_{ltu}^{g}(2), \dots, s_{ltu}^{g}(K)]^\mathrm{T}. \tag{3.27}$$

Herein, $s_{ltu}^{g}(k)$ is the modulated symbol for the subcarrier $i_{ltu}^{g}(k)$ of the g-th group of the l-th subsymbol to transmit from the t-th transmit antenna of the

u-th user. The resulting subcarrier indices and the modulated symbols for all groups are collected in the vectors $\mathbf{i}_{ltu} \in \mathbb{C}^{K_{\text{tot}} \times 1}$ and $\mathbf{s}_{ltu} \in \mathbb{C}^{K_{\text{tot}} \times 1}$ for each subsymbol

$$\mathbf{i}_{ltu} = [(\mathbf{i}^1_{ltu})^{\mathrm{T}}, (\mathbf{i}^2_{ltu})^{\mathrm{T}}, \ldots, (\mathbf{i}^G_{ltu})^{\mathrm{T}}]^{\mathrm{T}}, \tag{3.28}$$

$$\mathbf{s}_{ltu} = [(\mathbf{s}^1_{ltu})^{\mathrm{T}}, (\mathbf{s}^2_{ltu})^{\mathrm{T}}, \ldots, (\mathbf{s}^G_{ltu})^{\mathrm{T}}]^{\mathrm{T}}. \tag{3.29}$$

The modulated data is assigned to the active subcarriers and the data vector $\mathbf{d}^g_{ltu} \in \mathbb{C}^{N \times 1}$ for the g-th group and the l-th subsymbol is generated as

$$\mathbf{d}^g_{ltu} = [d^g_{ltu}(1), d^g_{ltu}(2), \ldots, d^g_{ltu}(N)]^{\mathrm{T}}. \tag{3.30}$$

Note that $N - K$ subcarriers on the vector $\mathbf{d}^g_{ltu}$ remain idle during the transmission. Thereafter, the data vector of each subsymbol is generated by combining all groups $\mathbf{d}_{ltu} = [(\mathbf{d}^1_{ltu})^{\mathrm{T}}, (\mathbf{d}^2_{ltu})^{\mathrm{T}}, \ldots, (\mathbf{d}^G_{ltu})^{\mathrm{T}}]^{\mathrm{T}}$. The resulting data block $\mathbf{d}_{tu} \in \mathbb{C}^{Q \times 1}$ is generated by combining the data of all subsymbols

$$\mathbf{d}_{tu} = [\mathbf{d}^{\mathrm{T}}_{1tu}, \mathbf{d}^{\mathrm{T}}_{2tu}, \ldots, \mathbf{d}^{\mathrm{T}}_{Ltu}]^{\mathrm{T}}. \tag{3.31}$$

Finally, a block interleaver is employed in a similar manner as in (3.5) to transmit each subcarrier in a group through uncorrelated channels, and the GFDM-IM data block of the u-th user to be transmitted from the t-th transmit antenna $\bar{\mathbf{d}}_{tu} \in \mathbb{C}^{Q \times 1}$ is generated.

Afterwards, the remaining procedure is the same as that in the conventional GFDM system. The resulting GFDM symbol is filtered prior to the transmission

$$\mathbf{x}_{tu} = \mathbf{A}\bar{\mathbf{d}}_{tu}. \tag{3.32}$$

Herein, $\mathbf{A}$ is the $Q \times Q$ transmit matrix generated as proposed in [29] with the following structure

$$\mathbf{A} = [\mathbf{g}_{0,0}, \ldots, \mathbf{g}_{N_{\text{tot}}-1,0}, \mathbf{g}_{0,1}, \ldots, \mathbf{g}_{N_{\text{tot}}-1,1}, \mathbf{g}_{0,L-1}, \ldots, \mathbf{g}_{N_{\text{tot}}-1,L-1}] \tag{3.33}$$

$$\text{where} \quad g_{n,l} = g((q - nL)_{mod_Q}) \exp\left(j2\pi \frac{nq}{N_{\text{tot}}}\right) \tag{3.34}$$

is the circularly shifted transmit filter and $\mathbf{g}_{n,l} \in \mathbb{C}^{Q \times 1}$ is the vector consisting of the filter samples

$$\mathbf{g}_{n,l} = [g_{n,l}(0), g_{n,l}(1), \ldots, g_{n,l}(Q-1)]^{\mathrm{T}}. \tag{3.35}$$

Assuming perfect channel equalization, the CP insertion and removal is omitted for ease of representation, the resulting data to be transmitted from all antennas of the u-th user can be written in a vector form as $\mathbf{x}_u = [\mathbf{x}_{1u}^{\mathrm{T}}, \mathbf{x}_{2u}^{\mathrm{T}}, \dots, \mathbf{x}_{Tu}^{\mathrm{T}}]^{\mathrm{T}}$. Finally, the GFDM-IM block is transmitted from each user over a C-tap frequency-selective Rayleigh fading channel. The received signal at the BS is

$$\mathbf{r} = \sqrt{P} \sum_{u=1}^{U} \mathbf{H}_u \mathbf{x}_u + \mathbf{n}. \tag{3.36}$$

$\mathbf{n}_n \in \mathbb{C}^{QN_{\mathrm{R}} \times 1}$ is i.i.d. circularly symmetric complex Gaussian with $\mathcal{CN}(0, \sigma^2)$ distributed elements, while $\mathbf{H}_u \in \mathbb{C}^{QN_{\mathrm{R}} \times QN_{\mathrm{T}}}$ has the following structure

$$\mathbf{H}_u = \begin{bmatrix} \mathbf{H}_{11u} & \mathbf{H}_{12u} & \dots & \mathbf{H}_{1Tu} \\ & \vdots & \ddots & \vdots & \\ \mathbf{H}_{R1u} & \mathbf{H}_{R2u} & \dots & \mathbf{H}_{RTu} \end{bmatrix}, \tag{3.37}$$

where $\mathbf{H}_{rtu} \in \mathbb{C}^{Q \times Q}$ is the circular convolution matrix generated from the channel impulse response coefficients $\mathbf{h}_{tu} = [h_{tu1}, h_{tu2}, \dots, h_{tuC}]$ where C is the number of channel taps and h_{tuc} are circularly symmetric complex Gaussian random variables distributed as $\mathcal{CN}(0, 1/C)$.

MMSE combining is employed that comprises the matrix $\mathbf{A}$ as well in order to perform joint detection and demodulation (JDD) for GFDM [34]. Hence, the MMSE combiner consists of the channel convolution matrix and the GFDM modulator matrix is designed as

$$\mathbf{W} = (\mathbf{B}^{\mathrm{H}}\mathbf{B} + \sigma^2 \mathbf{I}_{QN_{\mathrm{T}}U}/P)^{-1}\mathbf{B}^{\mathrm{H}} \tag{3.38}$$

where $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_U]$ with

$$\mathbf{B}_u = \begin{bmatrix} \mathbf{H}_{11u}\mathbf{A} & \mathbf{H}_{12u}\mathbf{A} & \dots & \mathbf{H}_{1Tu}\mathbf{A} \\ & \vdots & \ddots & \vdots & \\ \mathbf{H}_{R1u}\mathbf{A} & \mathbf{H}_{R2u}\mathbf{A} & \dots & \mathbf{H}_{RTu}\mathbf{A} \end{bmatrix}. \tag{3.39}$$

The resulting estimated vector $\tilde{\mathbf{d}} = \mathbf{W}\mathbf{r} \in \mathbb{C}^{QN_{\mathrm{T}}U \times 1}$ contains the transmitted data from all users. Furthermore, the estimated data vector of each user is divided into L groups for each subsymbol

$$\tilde{\mathbf{d}}_{tu} = [\tilde{\mathbf{d}}_{1tu}^{\mathrm{T}}, \tilde{\mathbf{d}}_{2tu}^{\mathrm{T}}, \dots, \tilde{\mathbf{d}}_{Ltu}^{\mathrm{T}}]^{\mathrm{T}}. \tag{3.40}$$

Thereafter, the data vector of each subsymbol $\tilde{\mathbf{d}}_{tu}$ is further divided into G groups in order to detect the active subcarriers and demodulate the M-ary symbols carried by these subcarriers.

$$\tilde{\mathbf{d}}_{ltu} = [(\tilde{\mathbf{d}}^1_{ltu})^{\mathrm{T}}, (\tilde{\mathbf{d}}^2_{ltu})^{\mathrm{T}}, \ldots, (\tilde{\mathbf{d}}^G_{ltu})^{\mathrm{T}}]^{\mathrm{T}}, \tag{3.41}$$

where $l \in \{1, 2, \ldots, L\}$, $t \in \{1, 2, \ldots, N_{\mathrm{T}}\}$, and $u \in \{1, 2, \ldots, U\}$.

In order to detect the active subcarriers in the received data $\tilde{\mathbf{d}}^g_{ltu} \in \mathbb{C}^{N\times 1}$, the decision metric $\mathbf{m}$ is computed for each possible combination of the active subcarriers by using the look-up table

$$\mathbf{m}(j) = \sum_{k=1}^{K} |\tilde{\mathbf{d}}^g_{ltu}((j,k))|, \tag{3.42}$$

where $j \in \{1, 2, \ldots, 2^{b_1}\}$ and $g \in \{1, 2, \ldots, G\}$. Once the decision metric $\mathbf{m}$ is computed for all possibilities, the maximum entry of this metric, which also indicates the active subcarrier combination, is found as $\hat{j} = \arg\max_j \mathbf{m}(j)$. Finally, $\hat{j}$ is inserted to and the active subcarriers are detected as follows

$$\hat{\mathbf{i}}^g_{ltu} = (\hat{j}, :). \tag{3.43}$$

Afterwards, the transmitted symbols on these subcarriers can be easily demodulated by a classical M-ary demodulation. Finally, $\hat{\mathbf{i}}^g_{ltu} \in \mathbb{C}^{K\times 1}$ contains the detected active subcarrier indices and $\hat{\mathbf{s}}^g_{ltu} \in \mathbb{C}^{K\times 1}$ contains the detected M-ary modulated symbols of the g-th group of the l-th subsymbol transmitted from the t-th antenna of the u-th user. The resulting detected indices and M-ary modulated data of all users are $\hat{\mathbf{I}} = [\hat{\mathbf{I}}_1, \ldots, \hat{\mathbf{I}}_U]$ and $\hat{\mathbf{S}} = [\hat{\mathbf{S}}_1, \ldots, \hat{\mathbf{S}}_U]$, respectively, where $\hat{\mathbf{I}}_u$ and $\hat{\mathbf{S}}_u$ are $\in \mathbb{C}^{LK_{\mathrm{tot}}\times N_{\mathrm{T}}}$ for $u \in \{1, 2, \ldots, U\}$.

Furthermore, the performance of the GFDM-IM system is studied for the UL transmission phase. Since there is no other work on the system that adopts GFDM-IM at MU networks the performance of the developed system model is compared with a conventional GFDM system and the OFDM-IM system proposed in Section 3.1. The design of a system model that operates at MU networks introduces additional challenges to overcome the IUI. In addition to this, generating a GFDM block is already more complex than generating a conventional OFDM block since an arbitrary filter is utilized and a GFDM block comprises multiple time and frequency samples. Therefore, a linear MMSE receiver is exploited in order to process the GFDM-IM block to compensate the complexity of the resulting system model.

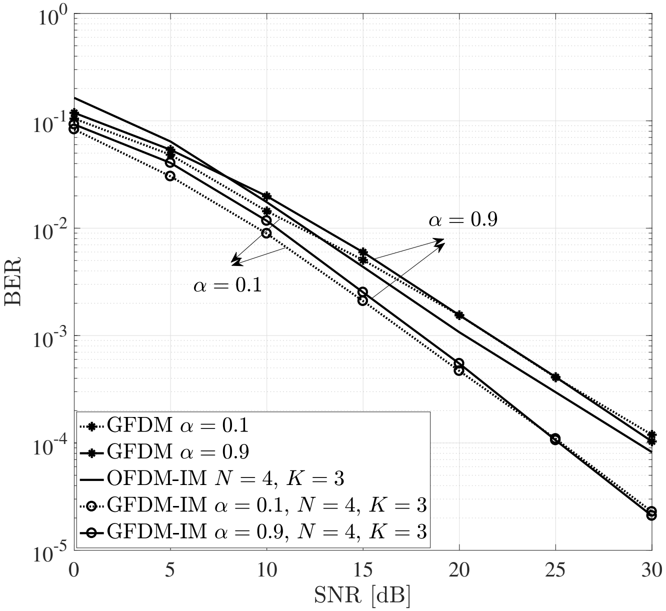


Figure 3.6: Uncoded bit error rate performance comparison of the uplink transmission phase for the GFDM-IM, OFDM-IM, and GFDM systems in a 2-user network ($N_T = 1$, $N_R = 2$, $L = 5$, $N_{tot} = 128$, $N = 4$, $K = 3$ with 4-QAM).

The performance of the proposed system is studied by considering the uncoded BER as the performance metric. The channel between the BS and the users is assumed to be a $C = 10$-tap frequency selective Rayleigh fading channel. Thus, a CP of length $N_{CP} = 32$ is inserted at the end of each block to eliminate the ISI. It is assumed that the BS acquires the CSI perfectly during the training phase, hence the linear receiver operates without performance degradation. Moreover, 2- and 4-user networks are considered in order to study the performance of the system. Each user is equipped with a single antenna $N_T = 1$ and the number of receive antennas at the BS equals the number of users $N_R = UN_T$. During a single channel use, a number of $L = 5$ subsymbols and $N_{tot} = 128$ subcarriers are exploited. The number of available subcarriers in a group for the GFDM-IM and OFDM-IM systems is determined as $N = 4$. Among these subcarriers

$K = 3$ of them are activated at each transmission. Additionally, a 4-QAM scheme is adopted to modulate the active subcarriers. The resulting samples are processed with a square-root-raised-cosine (SRRC) filter to generate the GFDM block where the shape of the filter is adjusted with the roll-off factor α. Note that a smaller roll-off factor sharpens the edges of the filter which results in a square filter when $\alpha = 0$. Hence, setting the roll-off factor to $\alpha = 0$ makes the subcarriers of the resulting GFDM block orthogonal. In order to observe the effect on the filtering choice to the performance two roll-off factors are considered: $\alpha = 0.1$ and $\alpha = 0.9$.

Figure 3.6 elucidates the error performance of the GFDM-IM system compared to the conventional GFDM and the OFDM-IM systems for a 2-user network. It has been shown in Section 3.1 that exploiting IM techniques on the subcarriers of the OFDM system improves the error performance of the system. The studies on the GFDM system in the literature unveiled that GFDM achieves better error performance than the OFDM system. Besides, the OFDM system can be considered as a variation of the GFDM in which the number of subsymbols is $L = 1$ and the samples are filtered with a square filter. The simulation results reveal that the error performance of the GFDM-IM system is better than in both the conventional GFDM and the OFDM-IM systems. This performance improvement is the result of the MMSE-based JDD and the improved distance of the symbols in the frequency domain caused by the IM technique. Note that increasing the roll-off factor results in a slightly worse performance up to the mid-SNR region.

The uncoded BER performance of the proposed system is studied for a 4-user network in Figure 3.7. The simulation results show that the interference among the users is successfully mitigated. The behavior of the considered systems are similar to the 2-user case, hence the designed receiver successfully eliminates the IUI. Moreover, the increased diversity by increasing the number of BS antennas results in a performance improvement for all considered system models.

Downlink Transmission

The block diagrams of the transmitter and receiver in the DL transmission phase are depicted in Figure 3.8a and Figure 3.8b, respectively [38]. During the DL transmission phase, the BS transmits U data streams simultaneously to the users. In this system model, the BS is equipped with N_T transmit antennas while each user is equipped with N_R receive antennas where $N_\mathrm{T} \geq U N_\mathrm{R}$. Moreover, it is assumed that BS has a perfect CSI and that the users need only partial

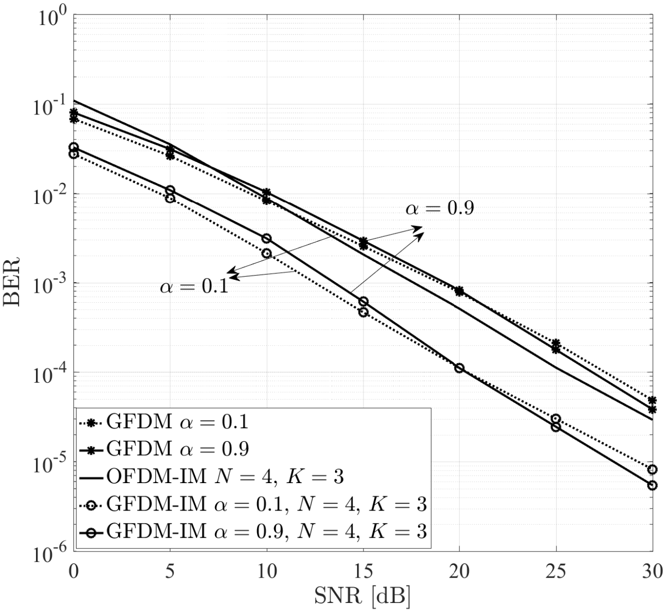


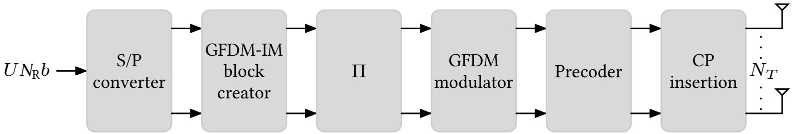
Figure 3.7: Uncoded bit error rate performance comparison of the uplink transmission phase for the GFDM-IM, OFDM-IM, and GFDM systems in a 4-user network ($N_\mathrm{T} = 1$, $N_\mathrm{R} = 4$, $L = 5$, $N_\mathrm{tot} = 128$, $N = 4$, $K = 3$ with 4-QAM).

information on the channel statistics to successfully decode the received signal.

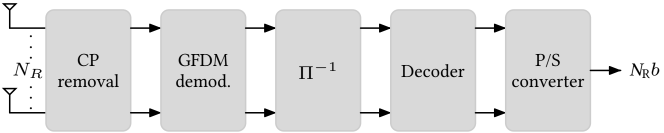
For each user u at the transmit chain of the BS arrive $N_\mathrm{R}b$ bits. The GFDM-IM block $\mathbf{x}_u \in \mathbb{C}^{QN_\mathrm{R}\times 1}$ is formed in the same manner as for the UL transmission phase where $u \in \{1, 2, \dots, U\}$. Furthermore, the GFDM-IM block is precoded by an MMSE-based precoder prior to the transmission in order to mitigate the IUI. The resulting received signal at the u-th user is

$$\mathbf{r}_u = \beta \sum_{i=1}^{U} \mathbf{H}_u \mathbf{P}_i \mathbf{x}_i + \mathbf{n}, \tag{3.44}$$

where $\mathbf{H}_u \in \mathbb{C}^{QN_\mathrm{R}\times QN_\mathrm{T}}$ has the same structure as in (3.37). Hence, the circular convolution matrix of the channel for all users can be generated as



(a) Base station.



(b) User terminal.

Figure 3.8: GFDM-IM block diagram for the downlink transmission phase.

$\mathbf{H} = [\mathbf{H}_1^{\mathrm{T}}, \mathbf{H}_2^{\mathrm{T}}, \dots, \mathbf{H}_U^{\mathrm{T}}]^{\mathrm{T}}$ where $\mathbf{H} \in \mathbb{C}^{UQN_{\mathrm{R}} \times QN_{\mathrm{T}}}$. Then, the MMSE precoder $\mathbf{P} \in \mathbb{C}^{QN_{\mathrm{T}} \times UQN_{\mathrm{R}}}$ is calculated as follows

$$\mathbf{P} = \left(\mathbf{H}^{\mathrm{H}} \mathbf{H} + \frac{\sigma^2 \mathbf{I}_{UQN_{\mathrm{R}}}}{P_u} \right)^{-1} \mathbf{H}^{\mathrm{H}}, \tag{3.45}$$

where $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_U]$. Furthermore, β is the normalization coefficient in order to satisfy the total transmit power constraint

$$\beta = \sqrt{\frac{P}{\mathrm{tr}\{\mathbf{P}\mathbf{P}^{\mathrm{H}}\}}}. \tag{3.46}$$

Finally, $\mathbf{n} \in \mathbb{C}^{QN_{\mathrm{R}} \times 1}$ is the circularly symmetric complex Gaussian noise vector where the elements follow a i.i.d. $\mathcal{CN}(0, \sigma^2)$ distribution.

After the reception of the GFDM-IM block the user first removes the CP. Afterwards, the GFDM demodulation is performed and the estimated vector $\tilde{\mathbf{d}} \in \mathbb{C}^{QN_{\mathrm{R}} \times 1}$ is calculated

$$\tilde{\mathbf{d}}_u = \mathbf{A}^{\dagger} \mathbf{r}_u / \beta. \tag{3.47}$$

Note that the normalization coefficient β is the only parameter that is needed to fed back to the users during the training phase. After the GFDM demodulation each user employs deinterleaving and regroups the signal in order to detect the active subcarrier indices and the modulated symbols.

Once the user has the post-processed signal $\tilde{\mathbf{d}}^g_{lru}$ the active subcarrier indices $\hat{\mathbf{i}}^g_{lru}$ and the M-ary modulated symbols $\hat{\mathbf{s}}^g_{lru}$ on these subcarriers are decoded in the same manner as in the UL transmission phase.

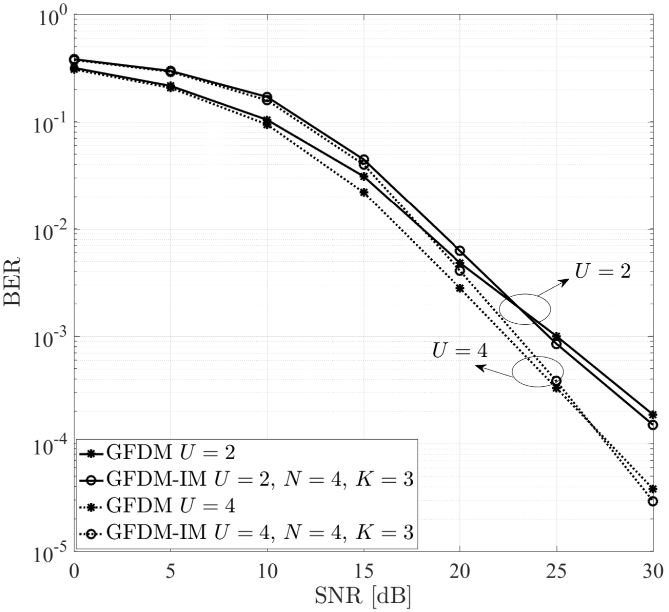


Figure 3.9: Uncoded bit error rate performance comparison of the downlink transmission phase for the GFDM-IM and GFDM systems in the networks consisting of 2 and 4 users, respectively ($N_\text{T} = UN_\text{R}$, $N_\text{R} = 1$, $\alpha = 0.9$, $L = 5$, $N_\text{tot} = 128$, $N = 4$, $K = 3$ with 4-QAM).

Next, the performance of the developed GFDM-IM system for MU networks is studied in Figure 3.9 for 2- and 4-user networks. Similar to the previous cases, the uncoded BER is considered as the performance metric. It has been already shown for the UL transmission phase that the GFDM-IM system achieves a better error performance than the OFDM-IM system. Hence, only the error performance of the conventional GFDM is taken into account for the DL transmission phase to reinforce the advantages of the GFDM-IM system. For the simulations a wideband channel with $N_\text{tot} = 128$ subcarriers and a data block of $L = 5$

subsymbols is considered. The channel between the BS and the users is assumed to be a $C = 10$-tap frequency selective Rayleigh fading channel. A CP with length $N_{\mathrm{CP}} = 10$ is inserted at the end of each data block to eliminate the ISI. In order to modulate the active subcarriers a 4-QAM scheme is adopted. Moreover, 2- and 4-user networks are considered for the DL transmission phase where each UT is equipped with a single antenna $N_{\mathrm{R}} = 1$. Thus, the number of antennas at the BS is equal to the number of users in the network $N_{\mathrm{T}} = UN_{\mathrm{R}}$. The number of available subcarriers for each group is set to $N = 4$, while the number of active subcarriers is $K = 3$ for the GFDM-IM system. In order to generate the final GFDM block, an SRRC filter with roll-off factor $\alpha = 0.9$ is utilized. Figure 3.9 demonstrates that the GFDM-IM system outperforms the conventional GFDM system starting from the mid-SNR region. As for the previous system models that exploit IM techniques erroneous detection of the active subcarrier indices results in performance degradation in the low-SNR regime.

3.2.2 Performance Analysis

The GFDM system already provides a better SE and a higher flexibility on the type of filter that can be used for modulation and demodulation for the GFDM block. As in the OFDM-IM case, IM can be employed to the GFDM system in order to achieve better spectrum- and energy-efficiency and to satisfy requirements of future networks. Next, the performance of the GFDM-IM is studied in terms of the achieved SE value, comparing the performance with a conventional GFDM system and the OFDM-IM system proposed in Section 3.1.

Spectrum Efficiency

The upper bound of the achievable SE of the GFDM-IM system is

$$SE_{\mathrm{OFDM\text{-}IM}} = \frac{LG(\lfloor \log_2 \binom{N}{K} \rfloor + K \log_2 M)}{LN_{\mathrm{tot}} + N_{\mathrm{CP}}}. \tag{3.48}$$

Recall that L is the number of subsymbols, N_{tot} is the total number of subcarriers allocated for the communication, N_{CP} is the length of the CP depending on the channel statistics, and M is the modulation order. Furthermore, the parameters G, N, and K are the numbers of groups, available subcarriers and active subcarriers per group, respectively. The parameters G, N, and K can be configured according to the system requirements as for the OFDM-IM case.

On the other hand, the upper bound of the achievable SE of the OFDM-IM system was given in (3.18) and for the conventional GFDM system it is

$$SE_{\text{GFDM}} = \frac{LN_{\text{tot}} \log_2 M}{LN_{\text{tot}} + N_{\text{CP}}}. \tag{3.49}$$

The comparison of the SE of GFDM-IM, OFDM-IM, and conventional GFDM is given in Table 3.2 for various parameters. It is assumed for the GFDM-IM and GFDM systems that the transmission consists of $L = 5$ subsymbols. Moreover, the wideband channel per subsymbol is divided into $N_{\text{tot}} = 128$ subchannels to transmit multiple streams through a frequency selective channel. It is assumed that the symbols are modulated with 4-QAM. As shown in Table 3.2, the SE of the GFDM-IM system is better than that of the conventional GFDM system for the given parameters. Although the SE of the conventional GFDM and the OFDM-IM systems are similar when $N_{\text{CP}} = 16$ the conventional GFDM system achieves better SE when the length of the CP is increased to $N_{\text{CP}} = 32$.

Table 3.2: Spectrum-efficiency values of the GFDM-IM, OFDM-IM, and GFDM systems ($N_{\text{tot}} = 128$ with 4-QAM).

System model	N_{CP}	L	N	K	SE
GFDM-IM	16	5	16	13	2.14
	32	5	16	13	2.08
OFDM-IM	16	-	16	13	1.94
	32	-	16	13	1.75
GFDM	16	5	-	-	1.95
	32	5	-	-	1.90

3.3 Summary

This chapter focused on the IM techniques for multicarrier systems. Firstly, an MU system model was developed that exploits the IM technique on the subcarriers of the OFDM system. Instead of actively modulating all available subcarriers of the OFDM, a subset is selected according to the IM part of the information bits. As a result, there are subcarriers that remain idle during the data transmission phase and the activation pattern of those subcarriers convey additional information. In order to lower the complexity of the encoding and decoding of the OFDM-IM system, a grouping architecture was exploited. Although grouping the subcarriers results in a lower data rate transmitted by the IM part it reduces the complexity of the receiver significantly. Moreover, an interleaving method was adopted while constructing the final data block that ensures that the subcarriers which belong to the same group are not allocated to adjacent channels. Thus, the subcarriers of the same group are not effected by the correlation between the subchannels. Then, the GFDM-IM system was developed in a similar manner as the OFDM-IM system. For the GFDM-IM system, IM technique was utilized for the subcarriers of each subsymbol by adopting interleaved grouping. Then, the final GFDM-IM block was generated comprising all subcarriers of the subsymbols.

The transceivers for the OFDM-IM and GFDM-IM systems were designed both considering UL and DL transmission phases with linear combiner and precoder in order to reduce the system complexity. In the developed system models the BS acquires the CSI, computes the combiner for the UL transmission phase and computes the precoder for the DL transmission phase. The linear MMSE method was exploited to eliminate the IUI which results in a low complexity system model. Moreover, the SE and PAPR expressions were given as the parameters of the number of groups, available subcarriers, and active subcarriers. The proposed system models offer flexible design since these parameters are independent of each other. A multicarrier system can be designed by adopting either a OFDM-IM or GFDM-IM system according to the requirements. Lower PAPR values can be achieved by activating less subcarriers. The SE can be improved by adjusting the number of groups, available subcarriers, and active subcarriers. The EE can be improved by transmitting more index modulated bits while keeping the transmit power per conventionally modulated bits fixed. Moreover, the error performance can be enhanced by using the same total transmit power as in conventional OFDM and GFDM systems. This results in higher transmit power per active subcarrier in OFDM-IM and GFDM-IM systems.

Chapter 4

Hybrid Beamforming with Spatial Modulation

This chapter studies the IM techniques in large-scale MIMO systems that operate in mmWave bands. In order to overcome the drawbacks of mmWave bands, channel properties should be taken into account while envisaging the wireless communication network. As already elaborated in Chapter 2, large-scale MIMO systems increase the degrees of freedom in the spatial domain. This feature can be exploited to focus the transmit power directly onto the intended receiver terminal to cope with the severe path-loss. However, scaling up the number of hardware elements results in an increased power consumption. Hybrid architectures provide a remedy by shifting a part of the signal processing to the analog domain. In this way the number of bulky and high power consuming hardware elements can be reduced. Needless to say, there will be a performance degradation as a consequence of renouncing the fully digital signal processing. Consequently, IM techniques can be combined with the hybrid system architecture to compensate the loss in SE and to further increase the data rate.

Since the UTs are generally portable devices space and power (battery) constraints limit the number of hardware elements that can be implemented and the computation power of the terminal. In this respect, Section 4.1 develops a system model in which a hybrid UT architecture that employs SM is designed. An algorithm is developed to configure the analog circuitry to maximize the EE for a given SE by considering the channel statistics. The performance of the resulting system model is analyzed by extensive simulations both for stochastic and deterministic channel environments.

Section 4.2 focuses on developing a spectrum- and energy-efficient BS architecture. A hybrid beamformer is designed where the number of high-power-consuming hardware elements is limited by the number of UTs to be served. SM is exploited by mapping a part of the information bits onto the indices of the AAs. The developed algorithm for EE maximization optimizes the transmit power and the number of UTs to be served, subject to the maximum power constraints, maximum and minimum number of UTs, and minimum achievable rate guarantee. Simulation results corroborate the performance improvement by the proposed system model and the global optimality of the developed algorithm.

4.1 Uplink Transmission

This section develops an energy-efficient large-scale UT architecture that exploits the SM technique to communicate with a fully digital BS in mmWave bands. Implementing a fully digital architecture at the UT that comprises a large number of antennas is challenging due to the space and battery life constraints. To overcome this issue a UT architecture that employs analog beamforming with spatial modulation (ABFSM) is introduced in Section 4.1.1. Then, a closed form expression for the achievable rate of the system considering Gaussian input for the conventionally modulated part of the transmitted signal is derived in Section 4.1.2. Moreover, an optimization algorithm is developed in Section 4.1.3 to design the system parameters to maximize the EE of the resulting system. The system performance is evaluated by adopting a stochastic channel model and a deterministic channel model that is predicted utilizing the ray-based Volcano technology by SIRADEL [87].

4.1.1 System Model

A large-scale MIMO system model is considered in which the focus is on the design of a low power consuming UT architecture that employs ABFSM. In the proposed system model the UT comprises N_{UT} transmit antennas and the BS comprises N_{R} receive antennas. The communication protocol is assumed to be TDD where acquiring the CSI is required only at the BS. After acquiring the CSI the BS runs the optimization algorithm proposed in Section 4.1.3 in order to determine the system parameters detailed in the remainder of this section. Thereafter, the BS provides the UT with the outcome of the optimization algorithm during the DL training phase. Since the information required by the UT is limited the DL training phase results in a low training overhead. The BS computes the ZF combiner to process the received signal during the UL data transmission phase.

Figure 4.1 depicts an energy-efficient UT architecture for the UL transmission phase that comprises a single RF chain and a large number of power-efficient devices such as RF switches and phase shifters [74]. In order to employ analog beamforming an architecture that consists of two stages is considered: a phase shifting stage to provide a large transmit beamforming gain and a switching stage to group the antennas of an array considering the correlation among the channels of the antennas. Hence, the resulting channel matrices between the AAs and the BS have full rank.

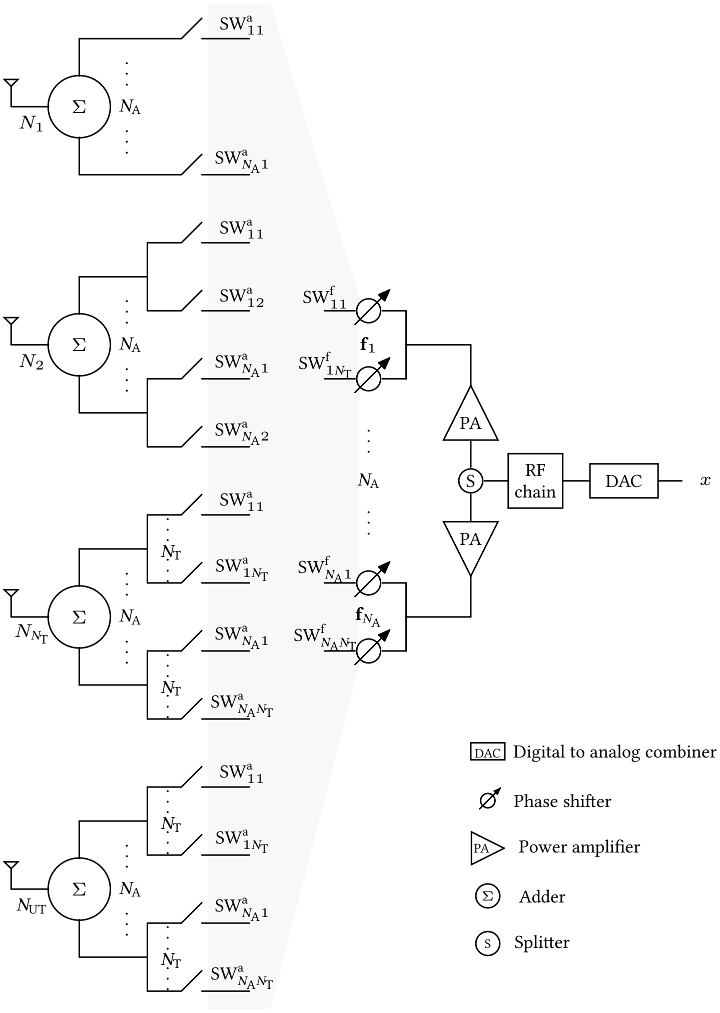


Figure 4.1: User terminal block diagram of the ABFSM scheme for the uplink transmission phase.

Phase Shifting Stage. During the UL transmission phase the UT exploits N_A AAs to transmit the conventionally modulated signal. Each AA consists of N_T

phase shifters that help to form the analog beamforming vector $\mathbf{f}_a$. Each phase shifter has to be connected to a single antenna only. In the specific AA, each phase shifter is connected to a distinct antenna, however, among the different arrays there can be common antennas. The number of AAs N_A and the number of phase shifters per array N_T are determined to maximize the EE of the system at the beginning of each coherence time by employing the low complexity optimization algorithm described in Section 4.1.3. The resulting UT comprises $N_\mathrm{A}N_\mathrm{T}$ phase shifters that contribute to the communication.

Switching Stage. The optimization algorithm establishes the connections of the UT circuitry illustrated in Figure 4.1. The connection of the switches at the phase shifters side is labeled as $\mathrm{SW}^\mathrm{f}_{i,j}$ and at the antennas side is labeled as $\mathrm{SW}^\mathrm{a}_{i,j}$ where i denotes the AA index that the phase shifter belongs to and j denotes the antenna index that the phase shifter is connected to. In the hardware circuitry, $\mathrm{SW}^\mathrm{f}_{i,j}$ and $\mathrm{SW}^\mathrm{a}_{i,j}$ with same indices are connected, however this connection is omitted in Figure 4.1 to preserve the clarity of the illustration. Maximum N_A switches are closed for each transmit antenna during the channel coherence time. The switching architecture between the transmit antennas and the phase shifters is as follows:

- The transmit antenna N_1 can be connected to the first phase shifters of each AA. The connections are established between $\mathrm{SW}^\mathrm{f}_{i,1}$ and $\mathrm{SW}^\mathrm{a}_{i,1}$ where $i \in \{1, 2, \ldots, N_\mathrm{A}\}$.
- The transmit antenna N_2 can be connected to the first two phase shifters of each AA. The connection possibilities of the transmit antenna N_2 depend on the connections of the transmit antenna N_1. As an example, if the transmit antennas N_1 and N_2 contribute to the analog beamformer $\mathbf{f}_1$ then the connection from $\mathrm{SW}^\mathrm{f}_{1,1}$ to $\mathrm{SW}^\mathrm{a}_{1,1}$ and the connection from $\mathrm{SW}^\mathrm{f}_{1,2}$ to $\mathrm{SW}^\mathrm{a}_{1,2}$ contribute to the signal transmission, however if the transmit antenna N_1 is not connected to $\mathbf{f}_1$ and transmit antenna N_2 contributes to the analog beamformer $\mathbf{f}_1$ then the connection from $\mathrm{SW}^\mathrm{f}_{1,1}$ to $\mathrm{SW}^\mathrm{a}_{1,1}$ to transmit antenna N_2 contributes to the signal transmission. The connections are established between $\mathrm{SW}^\mathrm{f}_{i,j}$ and $\mathrm{SW}^\mathrm{a}_{i,j}$ where $i \in \{1, 2, \ldots, N_\mathrm{A}\}$ and $j \in \{1, 2\}$.
- The same procedure follows up to transmit antenna $(N_\mathrm{T} - 1)$ and the connections are established between $\mathrm{SW}^\mathrm{f}_{i,j}$ and $\mathrm{SW}^\mathrm{a}_{i,j}$ where $i \in \{1, 2, \ldots, N_\mathrm{A}\}$ and $j \in \{1, 2 \ldots, N_\mathrm{T} - 1\}$.
- The connections for the remaining antennas are established between $\mathrm{SW}^\mathrm{f}_{i,j}$ and $\mathrm{SW}^\mathrm{a}_{i,j}$ where $i \in \{1, 2, \ldots, N_\mathrm{A}\}$ and $j \in \{1, 2 \ldots, N_\mathrm{T}\}$.

Since the same antenna can be shared by multiple analog beamformers each antenna can be fed by maximum N_A signals. Each antenna is connected to an adder block that is used to sum the signals coming from different phase shifters to the designated antennas through the switches to perform analog beamforming. Only the antennas that are selected as active by the optimization algorithm contribute to the communication and the rest remains idle.

Hardware Power Consumption Model. The power consumption of the considered UT circuitry is modeled as

$$P_\mathrm{c} = N_\mathrm{A} N_\mathrm{T}(P_\mathrm{SW} + P_\mathrm{PS}) + N_\mathrm{A} P_\mathrm{PA} + P_\mathrm{RF} + P_\mathrm{DAC} + P_\mathrm{BB}. \tag{4.1}$$

Therein, P_PA refers to the power consumption of the PA taken as the reference P_ref in the hardware circuitry. The power consumption of the UT devices at 28 GHz is modeled as in [62, 88] for the hardware power dissipation

$$\begin{aligned} P_\mathrm{PA} &= P_\mathrm{PS} = P_\mathrm{ref},\ P_\mathrm{SW} = 0.25 P_\mathrm{ref}, \\ P_\mathrm{RF} &= 2P_\mathrm{ref},\ P_\mathrm{DAC} = P_\mathrm{BB} = 10 P_\mathrm{ref}, \end{aligned} \tag{4.2}$$

where the power consumption of the hardware elements, such as phase shifter (PS), DAC, switch (SW), RF chain, and base band (BB) computation, are defined by using the reference power consumption value P_ref.

At the receive end of the channel, the BS employs a ZF combiner prior to the detection of the spatially and M-ary modulated bits. In this way the interference between the signals from different AAs is eliminated. In order to employ the ZF combiner successfully the channel between the BS and UT should be full rank. Therefore, an analog switching stage is utilized to perform antenna selection and grouping that reduces the correlation among the antennas of the UT and ensures the rank requirement. As a result, N_A analog beamformers are employed where each AA consists of N_T antennas in order to boost the transmit beamforming gain. Since the antennas at each group are chosen in order to combat the spatial correlation among the UT antennas the received power at the BS is maximized. This is achieved by selecting the number of antennas and the antenna indices per AA jointly by considering the channel statistics.

The incoming data stream prior to the transmission comprises two parts. The former is modulated according to a conventional M-ary modulation scheme, the latter is mapped onto the indices of the AAs. During the UL transmission phase, N_A AAs transmit the same M-ary modulated symbol from the activated antennas which provides diversity of the received signal at the BS. Each AA transmits the modulated symbol either with a high power level α_H that indicates the spatial bit 1 or with a low power level α_L that indicates the spatial bit 0. As a result, N_A spatially modulated bits and $\log_2 M$ conventionally modulated bits are transmitted per channel use. By exploiting different power levels for the bits 1 and 0 instead of deactivating the antennas that represent bit 0 the complete message set with 2^{N_A} combinations can be transmitted based on the incoming bit stream. If the system was designed in a way that the antennas that represent 0 are deactivated then transmitting N_A 0s would require special treatment since no M-ary modulated signal could be transmitted in this case. The resulting received signal at the BS is

$$\mathbf{r} = \sqrt{\beta P}\mathbf{H}\mathbf{F}x + \mathbf{n}. \tag{4.3}$$

Therein, β is the coefficient that ensures the constant average transmit power at the UT, and P is the transmit power budget of the UT. It is assumed that the channel $\mathbf{H} \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{UT}}$ between the UT and the BS follows a geometry-based channel model given in (2.14). Furthermore, x is the M-ary modulated symbol and $\mathbf{n} \in \mathbb{C}^{N_\mathrm{R} \times 1}$ is the noise vector with i.i.d. circularly symmetric complex Gaussian elements $\mathcal{CN}(0,\, \sigma^2)$.

4.1 Definition. $\mathbf{A}_\mathrm{SW} \in \mathbb{R}^{N_\mathrm{UT} \times N_\mathrm{A} N_\mathrm{T}}$ is the *switching matrix* that satisfies

$$\mathbf{A}_\mathrm{SW}^{(i,j)} \in \{0{,}1\}, \tag{4.4a}$$

$$\|\mathbf{A}_\mathrm{SW}^{(i,(a-1)N_\mathrm{T}+1:aN_\mathrm{T})}\|_0 \in \{0, 1\}, \tag{4.4b}$$

where $i \in \{1, 2, \ldots, N_\mathrm{UT}\}$, $j \in \{1, 2, \ldots, N_\mathrm{A} N_\mathrm{T}\}$ and $a \in \{1, 2, \ldots, N_\mathrm{A}\}$.

The switching matrix $\mathbf{A}_\mathrm{SW}$ connects the phase shifters to the UT antennas in such a way that $\mathbf{A}_\mathrm{SW}(i, j) = 1$ when the j-th phase shifter represented by the j-th column of $\mathbf{A}_\mathrm{SW}$ is connected to the i-th UT antenna represented by the i-th row of $\mathbf{A}_\mathrm{SW}$ and hence, $\|\mathbf{A}_\mathrm{SW}^{(j)}\|_0 = 1$ where $j \in \{1, 2, \ldots, N_\mathrm{A} N_\mathrm{T}\}$. Moreover, (4.4b) holds from the fact that the phase shifter inside a specific beamformer is connected to a distinct UT antenna.

4.2 Definition. Let $\mathbf{f}_a \in \mathbb{C}^{N_\mathrm{T} \times 1}$ be the *analog beamformer* of the a-th AA where $a \in \{1, 2, \ldots, N_\mathrm{A}\}$. Then, $\mathbf{A}_\mathrm{PS} \in \mathbb{C}^{N_\mathrm{A} N_\mathrm{T} \times N_\mathrm{T}}$ is the phase shifting matrix having the analog beamformers of all AAs in its diagonal

$$\mathbf{A}_\mathrm{PS} = \mathrm{diag}(\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{N_\mathrm{A}}). \tag{4.5}$$

Moreover, $\|\mathbf{f}_a\|^2 = 1$ has unit power where each element has constant amplitude and satisfies $|\mathbf{f}_a^{(i)}| = |\mathbf{f}_a^{(j)}|$ for all $i, j \in \{1, 2, \ldots, N_\mathrm{T}\}$.

Hence, the beamforming matrix $\mathbf{F} \in \mathbb{C}^{N_\mathrm{UT} \times N_\mathrm{A}}$ is decomposed as

$$\mathbf{F} = \mathbf{A}_\mathrm{SW}\mathbf{A}_\mathrm{PS}\mathbf{t}, \tag{4.6}$$

where $\mathbf{t} = (1 - 2\alpha)\mathbf{s} + \alpha\mathbf{1}_{N_\mathrm{A}}$ is utilized to map the spatial symbol $\mathbf{s} \in \mathbb{R}^{N_\mathrm{A} \times 1}$ to high and low transmit power levels $\alpha_\mathrm{H} = 1 - \alpha$ and $\alpha_\mathrm{L} = \alpha$, respectively, with $0 \le \alpha \le 1/2$. Moreover, the power level in order to transmit the M-ary modulated symbol employing the a-th analog beamformer $\mathbf{f}_a$ depends on the N_A bits from the incoming data stream where $a \in \{1, 2, \ldots, N_\mathrm{A}\}$ and $\Pr(\mathbf{s} = \mathbf{s}_j) = 1/2^{N_\mathrm{A}}$ are such that $\mathbf{s}_j$ is the j-th spatial symbol out of 2^{N_A} combinations. Finally, the power coefficient β is computed as

$$\beta = \frac{1}{\mathrm{E}(\|\mathbf{A}_\mathrm{SW}\mathbf{A}_\mathrm{PS}\mathbf{t}x\|^2)} = \frac{1}{\mathrm{tr}(\mathbf{A}_\mathrm{SW}\mathbf{A}_\mathrm{PS}\mathbf{R}_\mathrm{ss}\mathbf{A}_\mathrm{PS}^\mathrm{H}\mathbf{A}_\mathrm{SW}^\mathrm{H})}, \tag{4.7}$$

$$\mathbf{R}_\mathrm{ss} = \frac{1}{2^{N_\mathrm{A}}}\sum_{j=1}^{2^{N_\mathrm{A}}} \mathbf{t}_\mathrm{j}\mathbf{t}_\mathrm{j}^\mathrm{H}. \tag{4.8}$$

Herein, $\mathbf{t}_\mathrm{j}$ indicates the mapping of the j-th spatial symbol out of 2^{N_A} combinations. After the reception of the signal at the BS ZF combining is employed to enable the detection of the spatially and M-ary modulated symbols

$$\mathbf{y} = \mathbf{W}\mathbf{r}, \tag{4.9}$$

where the ZF combining matrix is computed as $\mathbf{W} = \left(\mathbf{H}_\mathrm{eff}^\mathrm{H}\mathbf{H}_\mathrm{eff}\right)^{-1}\mathbf{H}_\mathrm{eff}^\mathrm{H}$, where $\mathbf{H}_\mathrm{eff} \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{A}}$ is the effective channel matrix

$$\mathbf{H}_\mathrm{eff} = \mathbf{H}\mathbf{A}_\mathrm{SW}\mathbf{A}_\mathrm{PS} = [\mathbf{H}_1\mathbf{f}_1, \mathbf{H}_2\mathbf{f}_2, \ldots, \mathbf{H}_{N_\mathrm{A}}\mathbf{f}_{N_\mathrm{A}}] \tag{4.10}$$

and $\mathbf{H}_a = \mathbf{H}\mathbf{A}_\mathrm{SW}^{((a-1)N_\mathrm{T}+1:aN_\mathrm{T})} \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{T}}$ is the subchannel matrix of the a-th AA. The resulting post-processed signal $\mathbf{y} \in \mathbb{C}^{N_\mathrm{A} \times 1}$ comprises the spatially and M-ary modulated symbols. Hence, the a-th entry of $\mathbf{y}$ has the following values

$$\mathbf{y}(a) = \begin{cases} \sqrt{\beta P}\alpha_\mathrm{H}x + n_a^{'} & \text{if} \quad \mathbf{s}(a) = 1, \\ \sqrt{\beta P}\alpha_\mathrm{L}x + n_a^{'} & \text{if} \quad \mathbf{s}(a) = 0, \end{cases} \tag{4.11}$$

where $n_a^{'} \in \mathcal{CN}(0,\, \sigma_a^{'2})$ is the a-th entry of the post-processed noise variable $\mathbf{n}^{'} = \mathbf{W}\mathbf{n}$ with variance $\sigma_a^{'2} = \left(\mathbf{W}\mathbf{W}^{\mathrm{H}}\right)^{(a,a)} \sigma^2$.

Since a fully digital BS is assumed, a maximum-likelihood (ML) detector can be utilized to jointly detect the spatially and M-ary modulated symbols. The joint ML detection problem can be expressed as

$$\begin{aligned} [\hat{\mathbf{s}}, \hat{x}] &= \max_{\mathbf{s}_j, x_m} f\left(\mathbf{y} | \mathbf{s}_j, x_m\right) \qquad (4.12) \\ &= \min_{\mathbf{s}_j, x_m} \left\| \mathbf{R}_{\mathrm{nn}}^{-\frac{1}{2}} \left(\mathbf{y} - \sqrt{\beta}\mathbf{t}_j x_m\right) \right\|^2 , \end{aligned}$$

where $\mathbf{R}_{\mathrm{nn}} = \left(\sigma^2 \mathbf{W}\mathbf{W}^{\mathrm{H}}\right)$, $\mathbf{s}_j$ is the j-th spatial symbol out of $2^{N_{\mathrm{A}}}$ possibilities, and x_m is the m-th modulation symbol from the M-ary constellation diagram. The ML detector is basically an exhaustive search over all $M 2^{N_{\mathrm{A}}}$ possibilities. Note that although it minimizes the error probability for a given signal it also results in a more complex receiver architecture.

4.1.2 Achievable Rate Analysis

SE is adopted as a performance metric to evaluate the considered system model. Therefore, a closed form expression of the achievable rate of the ABFSM scheme is derived by considering a Gaussian symbol x by means of information theoretical methods. By applying the mutual information chain rule, the achievable rate is computed as

$$I\left(\mathbf{s}, x; \mathbf{y}\right) = I\left(\mathbf{s}; \mathbf{y}\right) + I\left(x; \mathbf{y} | \mathbf{s}\right), \qquad (4.13a)$$

$$I\left(\mathbf{s}; \mathbf{y}\right) = h\left(\mathbf{y}\right) - h\left(\mathbf{y} \middle| \mathbf{s}\right) = I_{\mathrm{S}}, \qquad (4.13b)$$

$$I\left(x; \mathbf{y} | \mathbf{s}\right) = h(\mathbf{y}|\mathbf{s}) - h(\mathbf{y}|\mathbf{s}, x) = I_{\mathrm{M}}. \qquad (4.13c)$$

Assuming a Gaussian-distributed input signal x the differential entropy of the post-processed signal $h(\mathbf{y})$ is computed as

$$h(\mathbf{y}) = -\int f\left(\mathbf{y}\right) \log_2 f(\mathbf{y}) d\mathbf{y}, \qquad (4.14)$$

where $f(\mathbf{y})$ is a complex Gaussian mixture random vector

$$f(\mathbf{y}) = \sum_{j=1}^{2^{N_{\mathrm{A}}}} \Pr\left(\mathbf{s}_j\right) f\left(\mathbf{y} \middle| \mathbf{s}_j\right) = \frac{1}{2^{N_{\mathrm{A}}}} \sum_{j=1}^{2^{N_{\mathrm{A}}}} \frac{1}{\pi^{N_{\mathrm{A}}} \left|{}_j\right|} \exp\left(-\mathbf{y}^{\mathrm{H}}{}_j^{-1}\mathbf{y}\right) \qquad (4.15)$$

$${}_{j} = \beta P \mathbf{t}_{\mathrm{j}}^{\mathrm{H}} \mathbf{t}_{\mathrm{j}} + \mathbf{R}_{\mathrm{nn}}.$$

The closed form expression of the differential entropy of a Gaussian mixture $h(\mathbf{y})$ is unknown as explained in Section 4.2 and approximations are derived in Proposition 4.12 and Proposition 4.13 for discrete inputs. Here, an approximation of the mutual information of the post-processed signal $\mathbf{y}$ is derived considering Gaussian input. The conditional differential entropy chain rule on $h(\mathbf{y})$ is applied to prove a closed form expression for the conditional probability density function $f(y_k|y_{k-1},\dots,y_1)$. Finally, the computational complexity is reduced by simplifying the integral in (4.14) to the sum of double integrals as follows

$$h(\mathbf{y}) = h(y_1) + \sum_{k=2}^{N_{\mathrm{A}}} h(y_k|y_{k-1},\dots,y_1) \tag{4.16}$$

with the conditional probability density function (PDF)

$$f(y_k|y_{k-1},\dots,y_1) = \frac{1}{2^k} \sum_{i=1}^{2^k} f\left(y_k \middle| y_{k-1},\dots,y_1,\mathbf{t}_{\mathrm{j}}(1:k)\right). \tag{4.17}$$

4.3 Lemma. *The random variable* $\left(y_k \middle| y_{k-1},\dots,y_1,\mathbf{t}_j(1:k)\right)$ *in* (4.17) *is complex Gaussian distributed with* $\mathcal{CN}\left(0,\sigma_K^2 + P_{j,K}\right)$.

Proof. The proof is given in Appendix B.1. ■

Therefore, the k-th differential entropy $h(y_k|y_{k-1},\dots,y_1)$ is determined by evaluating the double integral numerically regardless of the size of $\mathbf{y}$, thus the computational complexity is reduced significantly. Moreover, the differential entropy $h(\mathbf{y}|\mathbf{s})$ can be expressed in closed form as

$$\begin{aligned} h(\mathbf{y}|\mathbf{s}) &= \sum_{j=1}^{2^{N_{\mathrm{A}}}} \Pr\left(\mathbf{s}_j\right) h(\mathbf{y}|\mathbf{s} = \mathbf{s}_j) = \frac{1}{2^{N_{\mathrm{A}}}} \sum_{j=1}^{2^{N_{\mathrm{A}}}} \log_2(\pi \mathrm{e})^{N_{\mathrm{A}}} \left| \beta P \mathbf{t}_{\mathrm{j}} \mathbf{t}_{\mathrm{j}}^{\mathrm{H}} + \mathbf{R}_{\mathrm{nn}} \right| \\ &= \frac{1}{2^{N_{\mathrm{A}}}} \sum_{j=1}^{2^{N_{\mathrm{A}}}} \left(N_{\mathrm{A}} \log_2(\pi \mathrm{e}) + \log_2 \left| \mathbf{R}_{\mathrm{nn}} \right| \left(1 + \beta P \mathbf{t}_{\mathrm{j}}^{\mathrm{H}} \mathbf{R}_{\mathrm{nn}} \mathbf{t}_{\mathrm{j}}\right)\right). \end{aligned} \tag{4.18}$$

By utilizing the low computational complexity method for evaluating $h(\mathbf{y})$ and the closed form expression of $h(\mathbf{y}|\mathbf{s})$ the spatial rate I_{S} in (4.13b) can be computed efficiently. Moreover, the differential entropy $h(\mathbf{y}|\mathbf{s},x)$ can be expressed as

$$h(\mathbf{y}|\mathbf{s},x) = h(\mathbf{n}^{'}) = \log_2(\pi \mathrm{e})^{N_{\mathrm{A}}} \left| \mathbf{R}_{\mathrm{nn}} \right|. \tag{4.19}$$

Algorithm 1: System parameter optimization of generalized SM

```
**Input**: $\mathbf{H}$ and $N_{\mathrm{A,max}}$
**Output**: SE
Compute qr($\mathbf{H}$) with permutation vector $\mathbf{e}$ such that $\mathbf{H}(:, \mathbf{e}) = \mathbf{QR}$;
    then $\mathbf{R}_{1,1} \geq \mathbf{R}_{2,2} \geq \cdots \geq \mathbf{R}_{N_{\mathrm{UT}},N_{\mathrm{UT}}}$ holds
**for** $n_{\mathrm{A}} = 1 : N_{\mathrm{A,max}}$ **do**
    $\mathbf{H}_{\mathrm{eff}} = \left[\mathbf{H}^{(:,\mathbf{e}_1)}, \ldots, \mathbf{H}^{\left(:,\mathbf{e}_{n_{\mathrm{A}}}\right)}\right]$
    $\mathcal{S}(a) = \max\limits_{0 \leq \alpha \leq 1/2} (I_{\mathrm{S}} + I_{\mathrm{M}})$
**end for**
**return** SE $= \max(\mathcal{S})$
```

According to (4.18) and (4.19), the modulation rate I_{M} in (4.13c) can be expressed in closed form as

$$I_{\mathrm{M}} = \frac{1}{2^{N_{\mathrm{A}}}} \sum_{j=1}^{2^{N_{\mathrm{A}}}} \log_2 \left(1 + \beta P \mathbf{t}_{\mathrm{j}}^{\mathrm{H}} \mathbf{R}_{\mathrm{nn}}^{-1} \mathbf{t}_{\mathrm{j}}\right). \tag{4.20}$$

Finally, the achievable rate of the ABFSM scheme considering Gaussian input is expressed as the sum of the achievable rate of the spatial modulation part and the conventional modulation part $I(\mathbf{s}, x; \mathbf{y}) = I_{\mathrm{S}} + I_{\mathrm{M}}$.

4.1.3 Optimization of the System Parameters

This section proposes an optimization algorithm to maximize the EE of the system under SE constraints. Furthermore, the optimization algorithm is utilized to jointly design the switching matrix $\mathbf{A}_{\mathrm{SW}}$, the phase shifting matrix $\mathbf{A}_{\mathrm{PS}}$, and the power level α for the SM bits. In the proposed UT architecture an analog phase shifting stage is adopted to achieve a high beamforming gain. Although the phase shifters are low-power-consuming elements compared to the RF chains, DACs, etc., utilizing a large number of phase shifters would also increase the power consumption and could degrade the EE of the system. Thus, the system is developed to reach the maximum EE such that the SE of the resulting system is equal to or larger than the SE of a system that employs only the switching stage, namely generalized SM.

Algorithm 1 optimizes the SE of the generalized SM in order to compare it with the proposed system model. The resulting SE value is utilized as an

Algorithm 2: System parameter optimization of ABFSM

```
Input: $\mathbf{H} = \mathbf{A}_\mathrm{R}\mathbf{D}\mathbf{A}_\mathrm{T}^\mathrm{H}$, SE
Output: $\mathbf{A}_\mathrm{SW}^*$, $\mathbf{A}_\mathrm{PS}^*$ and $\alpha^*$
$\theta_\mathrm{max} = \arg\max_{l} |\mathbf{D}_{l,l}|, l \in \{1, 2, \dots, L\}$
for $n_\mathrm{T} = 1 : N_\mathrm{UT}$ do
    $\mathbf{H}_a = \mathbf{H}\mathbf{A}_{\mathrm{SW},a}, a \in \{1, 2, \dots, N_{N_\mathrm{A}}\}$ where $N_{N_\mathrm{A}} = \binom{N_\mathrm{UT}}{n_\mathrm{T}}$
    $\mathbf{f}_a = \mathbf{A}_{\mathrm{T},a}(:, \theta_\mathrm{max}), a \in \{1, 2, \dots, N_{N_\mathrm{A}}\}$
    $\mathbf{H}_\mathrm{eff} = [\mathbf{H}_1\mathbf{f}_1, \dots, \mathbf{H}_{N_{N_\mathrm{A}}}\mathbf{f}_{N_{N_\mathrm{A}}}]$
    Compute qr($\mathbf{H}_\mathrm{eff}$) with permutation vector $\mathbf{e}$ such that $\mathbf{H}_\mathrm{eff}(:, \mathbf{e}) = \mathbf{QR}$; then $\mathbf{R}_{1,1} \geq \mathbf{R}_{2,2} \geq \dots \geq \mathbf{R}_{N_\mathrm{UT},N_\mathrm{UT}}$ holds
    for $n_\mathrm{A} = 1 : N_\mathrm{A,max}$ do
        $\mathbf{A}_\mathrm{SW}(n_\mathrm{A}, n_\mathrm{T}) = \left[\mathbf{A}_{\mathrm{SW},\mathbf{e}_1}, \dots, \mathbf{A}_{\mathrm{SW},\mathbf{e}_{n_\mathrm{A}}}\right]$
        $\mathbf{A}_\mathrm{PS}(n_\mathrm{A}, n_\mathrm{T}) = \mathrm{diag}(\mathbf{f}_{\mathbf{e}_1}, \dots, \mathbf{f}_{\mathbf{e}_{n_\mathrm{A}}})$
        $\mathbf{H}_\mathrm{eff} = [\mathbf{H}_{\mathbf{e}_1}\mathbf{f}_{\mathbf{e}_1}, \dots, \mathbf{H}_{\mathbf{e}_{n_\mathrm{A}}}\mathbf{f}_{\mathbf{e}_{n_\mathrm{A}}}]$
        $\mathcal{S}(n_\mathrm{A}, n_\mathrm{T}) = \max_{0 \leq \alpha \leq 1/2} (I_\mathrm{S} + I_\mathrm{M})$
        $\mathcal{E}(n_\mathrm{A}, n_\mathrm{T}) = \mathcal{S}(n_\mathrm{A}, n_\mathrm{T}) / P_\mathrm{c}$
        $\Omega(n_\mathrm{A}, n_\mathrm{T}) = \alpha$
    end for
end for
Find $N_\mathrm{A}^*, N_\mathrm{T}^*$ by solving (4.21) for $\mathcal{E}$, $\mathcal{S}$ and $\mathrm{SE}_\mathrm{gSM}$
return $\mathbf{A}_\mathrm{SW}^* = \mathbf{A}_\mathrm{SW}(N_\mathrm{A}^*, N_\mathrm{T}^*)$, $\mathbf{A}_\mathrm{PS}^* = \mathbf{A}_\mathrm{PS}(N_\mathrm{A}^*, N_\mathrm{T}^*)$, $\alpha^* = \Omega(N_\mathrm{A}^*, N_\mathrm{T}^*)$
```

input to Algorithm 2 to optimize the parameters of the developed system model. QR decomposition is employed [89] to obtain the channel quality of each UT antenna. $\mathbf{R}$ in step (3) of Algorithm 1 is the upper triangular matrix of the channel $\mathbf{H}$ where the diagonal entries are associated with the strength of the path between the UT and BS antennas. The permutation vector $\mathbf{e}$ is composed by exploiting the matrix $\mathbf{R}$ such that $\mathbf{e}$ comprises the antenna indices starting from the strongest path. $N_\mathrm{A,max}$ represents the maximum number of AAs that is upper bounded by the number of paths C of the channel. Next, the best subchannel matrix that maximizes the SE is selected.

Moreover, the proposed system model is optimized for an arbitrary number of phase shifters n_T for an AA in Algorithm 2. The aim of Algorithm 2 is to find the optimum number of AAs N_A^* such that each AA consists of those antennas that have the lowest correlation among the arrays. There are $N_{N_{\mathrm{A}}} = \binom{N_{\mathrm{UT}}}{n_{\mathrm{T}}}$ possible ways to connect the n_{T} phase shifters of a group to the N_{UT} antennas of the UT which leads to $N_{N_{\mathrm{A}}}$ possibilities for the groups of phase shifters. Furthermore, the a-th analog beamformer $\mathbf{f}_a$ is designed to transmit the beam in the direction of the strongest path. Next, the effective channel matrix $\mathbf{H}_{\mathrm{eff}}$ is generated that includes all possible analog beamformers for $N_{N_{\mathrm{A}}}$ possible connections. With the help of the QR decomposition [89] the least correlated groups are sorted in one step. Thereafter, the SE and EE are evaluated with optimized power levels accounting for the number of AAs ($n_{\mathrm{A}} = 1 : N_{\mathrm{A,max}}$). The procedure is repeated for every possible number of active antennas in a group ($n_{\mathrm{T}} = 1 : N_{\mathrm{UT}}$) until a $N_{\mathrm{UT}} \times N_{\mathrm{A,max}}$ grid of the SE and EE is constructed. Finally, the BS selects the operating points from the grid to maximize the EE such that the SE is larger than that of the generalized SM system without the phase shifting stage, namely

$$
\begin{aligned}
(N_{\mathrm{A}}^*, N_{\mathrm{T}}^*) = & \max_{n_{\mathrm{A}}, n_{\mathrm{T}}} \ \mathrm{EE} && (4.21\mathrm{a})\\
& \text{s.t. } \mathrm{SE} \geq \mathrm{SE}_{\mathrm{gSM}}, && (4.21\mathrm{b})\\
& \quad \mathrm{EE} \in \mathcal{E}, \ \mathrm{SE} \in \mathcal{S}, && (4.21\mathrm{c})\\
& \quad n_{\mathrm{A}} \in \{1, 2, \ldots, N_{\mathrm{A,max}}\}, && (4.21\mathrm{d})\\
& \quad n_{\mathrm{T}} \in \{1, 2, \ldots, N_{\mathrm{UT}}\}. && (4.21\mathrm{e})
\end{aligned}
$$

The optimization of the power level in step (13) of Algorithm 2 leads to a nonlinear objective function in one variable with one linear constraint that can be evaluated numerically.

In the exhaustive search-based design, step (8) of Algorithm 2 should be replaced with an exhaustive search for the selection of N_{A} groups out of $\binom{N_{\mathrm{UT}}}{n_{\mathrm{T}}}$ possible combinations for all $n_{\mathrm{T}} \in \{1, 2, \ldots, N_{\mathrm{UT}}\}$ and $n_{\mathrm{A}} \in \{1, 2, \ldots, N_{\mathrm{A,max}}\}$. The number of grid points of the proposed algorithm $N_{\mathrm{Alg\,2}}$ and the exhaustive search N_{ES} can be computed as

$$N_{\mathrm{Alg\,2}} = N_{\mathrm{UT}} N_{\mathrm{A,max}}, \tag{4.22}$$

$$N_{\mathrm{ES}} = 1 + \sum_{n_{\mathrm{T}}=1}^{N_{\mathrm{UT}}-1} \sum_{n_{\mathrm{A}}=1}^{N_{\mathrm{A,max}}} \binom{\binom{N_{\mathrm{UT}}}{n_{\mathrm{T}}}}{n_{\mathrm{A}}}, \ \forall \binom{N_{\mathrm{UT}}}{n_{\mathrm{T}}} \geq N_{\mathrm{A}}. \tag{4.23}$$

As an illustrative example, consider $N_{\mathrm{UT}} = 16$ and $N_{\mathrm{A,max}} = 4$. The corresponding numbers of grid points are $N_{\mathrm{Alg\,2}} = 64$ and $N_{\mathrm{ES}} = 2.9 \times 10^{15}$. Therefore, the proposed algorithm reduces the computational complexity significantly with a slight performance degradation as given in Table 4.1.

Table 4.1: Energy efficiency of the ABFSM scheme considering Algorithm 2 and the exhaustive search algorithm ($N_{\mathrm{R}} = 128$, $N_{\mathrm{UT}} = 4$, $C = 3$).

SNR [dB]	-8	-4	0	4	8	12
$\mathrm{EE}_{\mathrm{Alg\ 2}}$ [bits/Hz/J]	8.54	10.08	11.55	12.93	14.25	15.51
$\mathrm{EE}_{\mathrm{ES}}$ [bits/Hz/J]	8.55	10.10	11.58	13.00	14.27	15.55

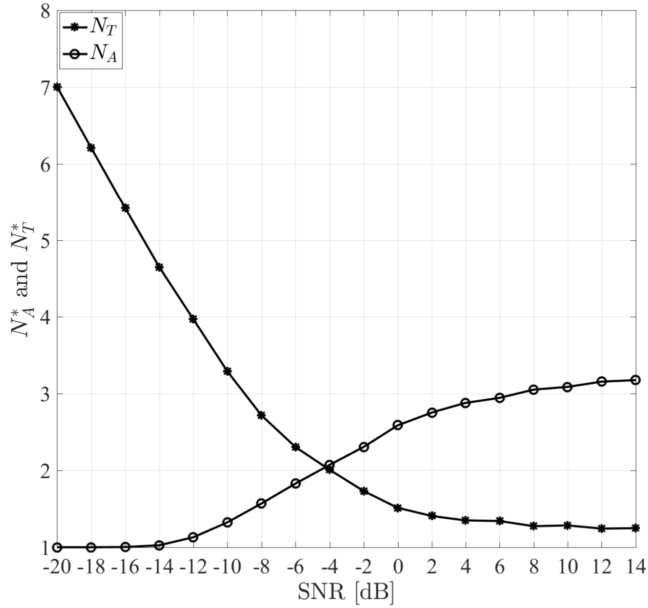


Figure 4.2: Optimum number of AAs and of antennas per array for the ABFSM scheme ($N_{\mathrm{R}} = 128$, $N_{\mathrm{UT}} = 16$, $C = 4$).

Performance Evaluation in a Stochastic Channel Environment. The behavior of Algorithm 2 is evaluated in Figure 4.2 in terms of optimized number of AAs N_A^* and optimized number of antennas per array N_T^* [74]. For the performance evaluation a stochastic channel model is adopted with $C = 4$ scattering clusters. The gain of the path associated to each scattering cluster follows an i.i.d. circularly symmetric complex Gaussian distribution with $\mathcal{CN}(0, 1)$. The AoD and AoA of each path follow a uniform distribution with $\phi \sim \mathcal{U}[-\pi/2, \pi/2]$ and $\theta \sim \mathcal{U}[-\pi/6, \pi/6]$, respectively. The performance of the system is averaged over 1000 channel realizations. A BS with $N_\mathrm{R} = 128$ antennas and a UT with $N_\mathrm{UT} = 16$ transmit antennas where the maximum number of AAs is set to $N_\mathrm{A,max} = 4$. Figure 4.2 shows that, since the aim is at maximizing the EE of the system, the optimum number of phase shifters (transmit antennas) N_T^* for an array is smaller than the total number of transmit antennas $N_\mathrm{UT} = 16$. Therefore, the increase in the number of AAs is necessarily associated with a decrease in the number of phase shifters per AA. Since a high beamforming gain is essential in the low-SNR region the number of phase shifters per AA is larger and thus, the number of arrays is small. Increasing the SNR reduces the required beamforming gain. As a result, the number of phase shifters per AA decreases. Thus, the number of AA increases to attain a high spatial multiplexing gain.

Figure 4.3 compares the SE of the proposed ABFSM system with the generalized SM [74]. At the low-SNR region, the proposed design achieves superior SE as the phase shifting stage of the ABFSM scheme provides a high beamforming gain and compensates the severe path-loss. When the high-SNR region is considered each group may contain one or two phase shifters as depicted in Figure 4.2. Since the small number of antennas at an AA is not sufficient to provide a high beamforming gain the SE of the generalized SM approaches the SE of ABFSM.

In [74] the EE and SE trade-off is studied in Figure 4.4 for the ABFSM system compared to the generalized SM. In the low-SNR region, the proposed design achieves superior SE and EE due to the high gain of the phase shifting stage. In the high-SNR region the generalized SM scheme achieves slightly better EE as the number of AAs is getting smaller (down to one) and the SE gap of the two systems decreases.

Performance Evaluation in a Deterministic Channel Environment. Up to this point the performance of the ABFSM scheme considering a stochastic channel environment was evaluated. Next, the performance of the developed

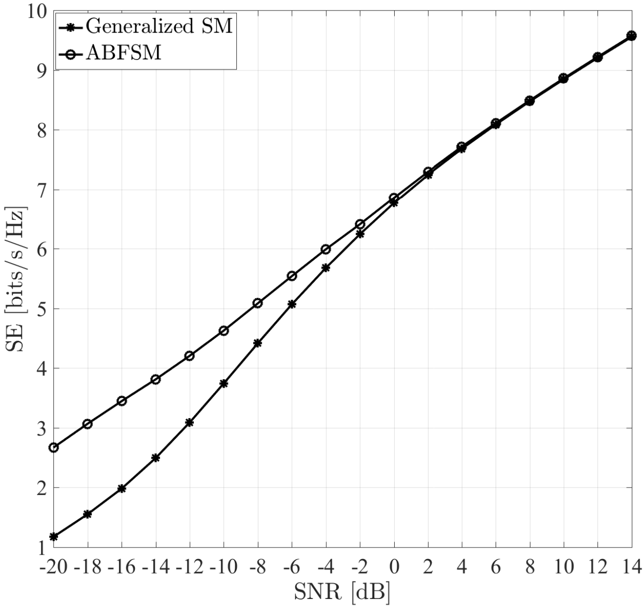


Figure 4.3: Spectrum efficiency of the ABFSM scheme compared to the generalized SM scheme ($N_\mathrm{R} = 128$, $N_\mathrm{UT} = 16$, $C = 4$).

system model is studied considering the poor scattering environment of the mmWave propagation in a deterministic channel environment. A realistic user distribution is considered and deterministic channels per user are generated with the aim of evaluating the performance of the proposed system in a typical small cell scenario at 28 GHz. Furthermore, the performance of the proposed scheme is compared under stochastic and deterministic channel environments. In order to generate the deterministic channel, several outdoor small-cell mmWave-channel samples have been predicted from the ray-based propagation model VolcanoUrban [87]. Those samples are the result of physical interactions between the electromagnetic wave and the real representation of a dense urban environment, more precisely, a district in New York Manhattan. In addition to the line-of-sight (LoS), the indirect paths are a combination of reflections

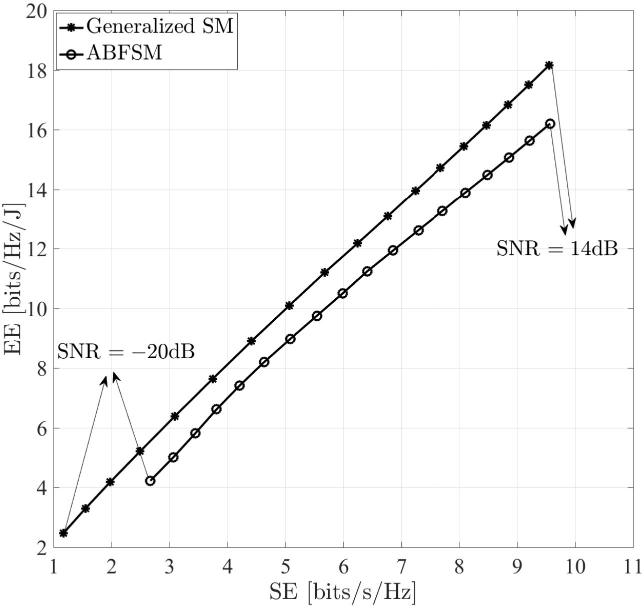


Figure 4.4: Energy- and spectrum-efficiency trade-off of the ABFSM scheme compared to the generalized SM scheme ($N_\mathrm{R} = 128$, $N_\mathrm{UT} = 16$, $C = 4$).

(up to three) and diffraction (up to one) on the building facades. A small cell is positioned at 8 meters above the ground, at a typical location for a lamppost. Three sectors are installed at the small cell. Each sector is feeding a linear AA with boresight direction oriented towards azimuth 0°, 120°, and 240° as depicted in Figure 4.5 [74]. Each linear AA is composed of 128 vertically-polarized antenna elements, which are uniformly distributed in the horizontal plane, at a frequency of 28 GHz, and with half-wavelength separation. The length of each sector AA is 68.6 centimeters. All antenna elements have the same radiation pattern with 60° half power horizontal beamwidth.

The users are assumed to be pedestrians that are randomly distributed on the surrounding pavements within a maximum range of 220 meters from the small cell BS. The user equipment is located at 1.5 meter above the ground. Its

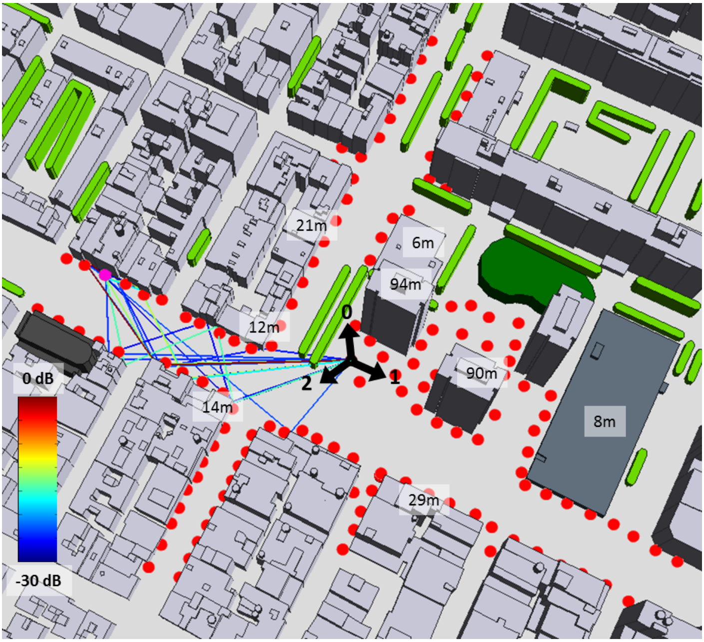


Figure 4.5: Top view of a realistic user distribution (red dots) served by a large-scale MIMO base station consisting of three sectors inside a small mmWave cell in Manhattan, New York City, where the farthest user is located 220 meters away from the base station.

antenna is a ULA with 16 vertically polarized isotropic elements positioned in the horizontal plane. The channel samples are produced from 142 different user positions. Users are positioned either in a wide or a narrow street or even in a small square. A few of them are in an non-line-of-sight (NLoS) situation.

Finally, a total number of 180 channel samples is created: 50, 69, and 61 for Sector 0, 1, and 2, respectively. The number of NLoS samples is 37. Each sample is obtained as follows: the multiple paths are computed using the ray-based model from the central point of the small cell array towards the central point of the user array. The physical description of the departure and the arrival angles of each path allows extrapolating the phase shifts along the arrays. The multi-path field sum is calculated for each couple of transmit/receive elements. The whole approach has the advantage that the field variations in the MIMO channel matrix are related to physical phenomena which intrinsically leads to realistic correlations.

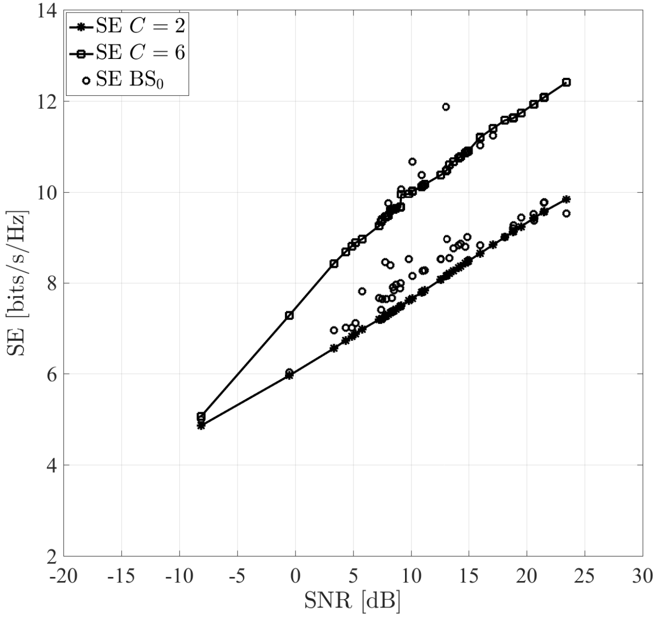


Figure 4.6: Uplink spectrum efficiency of the ABFSM scheme for the Sector 0 of the BS.

The SE of the proposed system model is evaluated on the stochastic channel model with $C = 2$ and $C = 6$ scattering clusters. The deterministic channel model for the scenario depicted in Figure 4.5 is assuming the same path-loss for both models. For the simulation setup it is considered that the noise level $\sigma^2 = -84$ dBm, the transmit power $P = 20$ dBm, and the carrier frequency $f_c = 28$ GHz. Figure 4.6 shows the SE comparison between the stochastic channel model and Sector 0 of the deterministic channel prediction [74]. Sector 0 has the lowest scattering environment due to the LoS users and the vegetation. For that reason, only 4% of the users in this sector achieve an SE larger than the SE of the system operating in a stochastic channel environment with $C = 6$ scattering clusters. Most of the users (94%) in the deterministic channel environment

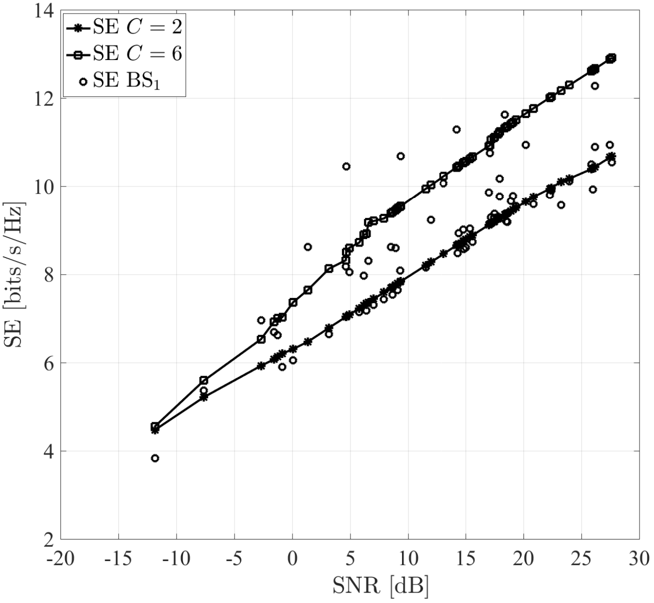


Figure 4.7: Uplink spectrum efficiency of the ABFSM scheme for the Sector 1 of the BS.

achieve an SE between the stochastic channel with $C = 2$ and $C = 6$ scattering clusters. Since Sector 1 has more NLoS users, 8% of the users in this sector experience an SE larger than the stochastic channel with $C = 6$ scattering clusters as shown in Figure 4.7 [74]. Moreover, 87% of the users have an SE between the stochastic channel with $C = 2$ and $C = 6$ scattering clusters. Figure 4.8 depicts the SE performance of Sector 2 [74]. In this sector users are farther away than in the other two cases, therefore, the users in Sector 2 experience a larger path-loss and a lower SE. In addition, most of the users ($\approx 92\%$) have an SE between the stochastic channel with $C = 2$ and $C = 6$ scattering clusters. This experiment shows that the proposed system design attains similar performance under both the theoretical stochastic channel model in (4.25) and the realistic channel model. The considered channel model can be adopted for performance evaluations of the developed system models. Moreover,

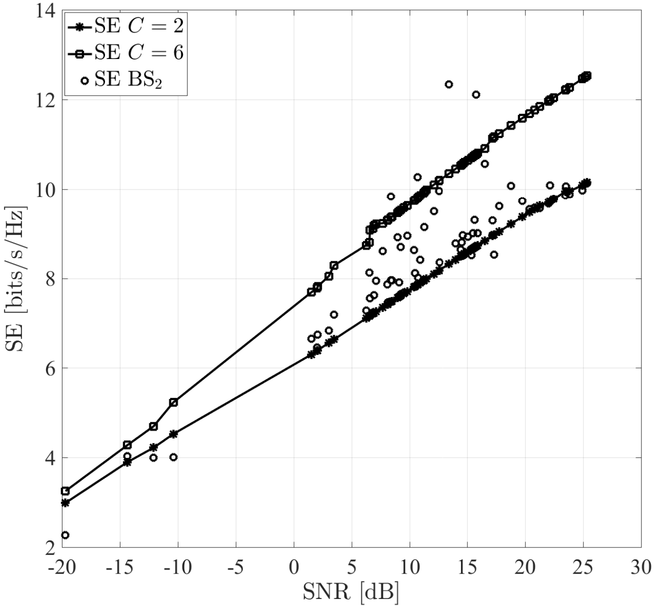
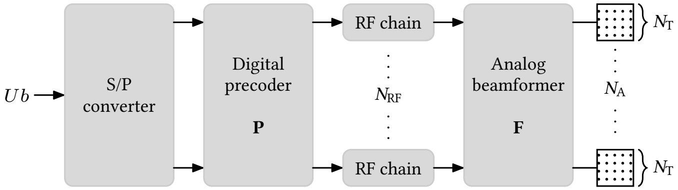


Figure 4.8: Uplink spectrum efficiency of the ABFSM scheme for the Sector 2 of the BS.

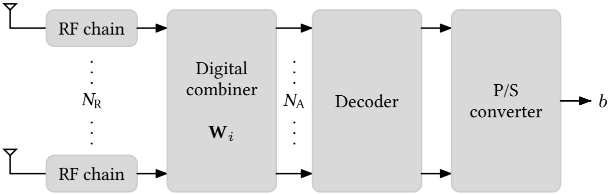
the system performance evaluation of the mmWave channels at 28 GHz, based on a stochastic channel model, gives a realistic assessment if the number of scattering clusters is between $C = 2$ and $C = 6$.

4.2 Downlink Transmission

This section studies an energy-efficient large-scale BS architecture that operates in an MU network. The transceiver is designed by considering the DL transmission phase to operate in a mmWave channel environment in a similar manner as in Section 4.1. In order to cope with the poor scattering nature of the mmWave channels the BS will employ analog beamformers to attain a high beamforming gain and a digital combiner to mitigate the IUI as discussed in Section 4.2.1. The BS utilizes HBFSM in order to achieve high SE without altering



(a) Base station.



(b) User terminal.

Figure 4.9: HBFSM block diagram for the downlink transmission phase.

the total power consumption. Since an MU network with HBFSM system is considered it is assumed that the UTs have a fully digital architecture to focus on a low complexity BS design. In addition, a system model that employs hybrid beamforming (HBF) to transmit only the conventionally modulated signals is presented in Appendix C to demonstrate the performance enhancement of the proposed system model. Moreover, Section 4.2.2 breaks down the design of the hybrid beamformer at the BS and the digital combiner at the UT while the quantization error is further analyzed as a design parameter of the analog beamformer in Section 4.2.3. Next, an approximation of the achievable rate of the developed system model is derived in Section 4.2.4 considering discrete input for the conventionally modulated part of the transmitted signals. Finally, in Section 4.2.5 an optimization framework for EE maximization is developed.

4.2.1 System Model

This section addresses the design of an MU large-scale MIMO system with one transmitter and multiple receivers where the transmitter and receiver block diagrams for the DL transmission phase are depicted in Figure 4.9a and Figure 4.9b, respectively [72]. The transmitter employs hybrid beamforming and the receivers employ digital combining. In the proposed system model the BS is equipped with N_{A} ULA that are fed by N_{RF} RF chains. Each AA consists of N_{T} transmit antennas. It is assumed without loss of generality that N_{T} is fixed for all AAs and there are no common transmit antennas between arrays. Note that the number of AAs is independent from the number of RF chains. It is assumed that the BS has perfect CSI that is acquired during the UL training phase prior to data transmission. Moreover, users require only partial information that is acquired during the DL training phase.

The BS serves U users with N_{R} receive antennas. It is likely that the number of antennas at the UT is small enough to deploy one RF chain per antenna and that every user is served with only one data stream. Consequently, there is no need to deploy analog beamformers at the receiver side as conventional MIMO beamforming techniques suffice. The number of data streams the BS has to process is equal to the number of users U. Given the considered beamforming scheme, the number of data streams (users) U that will be simultaneously transmitted is a lower bound on the number of RF chains $N_{\mathrm{RF}} \geq U$.

Additionally, according to the use of the SM, the incoming bit stream b intended for a specific user is divided into two parts. The first part is modulated according to SM principles by choosing one of the AAs which enables transmitting $b_1 = \log_2 N_{\mathrm{A}}$ bits. The second part is modulated according to a conventional M-ary modulation scheme which enables the transmission of $b_2 = \log_2 M$ bits. Based on the incoming bits for each bit stream one of the AAs is chosen and the M-ary modulated data is transmitted by using the selected AA. Hence, the index of the active AA conveys information without consuming additional power. It is also possible that one AA is activated to transmit information to more than one user per channel use. This can be only realised by using analog beamformers. However, it enables exploiting SM in the MU scenario without any limitation.

A number of U bit streams per channel use arrive at the transmit chain of the BS. According to the SM bits the AA indices are determined for every user. Thereafter, the BS constructs the digital precoder and selects the analog beamformer. U parallel signals are transmitted to the users. The resulting received signal at the i-th user can be expressed as

$$\mathbf{r}_i = \sqrt{P} \sum_{j=1}^{U} \mathbf{H}_{a_j,i} \mathbf{f}_{a_j,j} \mathbf{p}_{a_j,j} \mathbf{x} + \mathbf{n}_i. \tag{4.24}$$

$\mathbf{H}_{a_j,i} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$ is the channel between the a_j-th AA and i-th user accounting for C scattering clusters where a_j is the selected AA index to transmit data to the j-th user. The channel follows a geometry-based model as in (2.14)

$$\mathbf{H}_{a_j,i} = \sqrt{\frac{N_\mathrm{T} N_\mathrm{R}}{C}} \sum_{c=1}^{C} \alpha_{a_j,i_c} \mathbf{a}_{\mathrm{R}_{a_j,i}}(\theta_c) \mathbf{a}^{\mathrm{H}}_{\mathrm{T}_{a_j,i}}(\phi_c). \tag{4.25}$$

$\mathbf{a}_{\mathrm{R}_{a_j,i}}(\theta_c)$ and $\mathbf{a}_{\mathrm{T}_{a_j,i}}(\phi_c)$ are the receive and transmit AA response vector of the c-th path

$$\mathbf{a}_{\mathrm{R}_{a_j,i}}(\theta_c) = \sqrt{\frac{1}{N_\mathrm{R}}}[1, \mathrm{e}^{j\pi \sin\theta_c}, \dots, \mathrm{e}^{j\pi(N_\mathrm{R}-1)\sin\theta_c}]^T, \tag{4.26}$$

$$\mathbf{a}_{\mathrm{T}_{a_j,i}}(\phi_c) = \sqrt{\frac{1}{N_\mathrm{T}}}[1, \mathrm{e}^{j\pi \sin\phi_c}, \dots, \mathrm{e}^{j\pi(N_\mathrm{T}-1)\sin\phi_c}]^T, \tag{4.27}$$

where $c \in \{1, 2, \dots, C\}$. θ_c and ϕ_c are the AoA and AoD of the path and follow the uniform distribution $\mathcal{U}(0, 2\pi]$. Finally, α_{a_j,i_c} represents the channel gain of the c-th path that follows i.i.d. circularly symmetric complex Gaussian distribution with $\mathcal{CN}(0, 1)$. It is assumed that the antennas are placed on the array with half-wavelength separation. Furthermore, $\mathbf{p}_{a_j,j} \in \mathbb{C}^{1\times U}$ and $\mathbf{f}_{a_j,j} \in \mathbb{C}^{N_\mathrm{T}\times 1}$ are the digital precoding and analog beamforming vectors for the j-th user when the a_j-th AA array is chosen for transmission. Equal power allocation in large-scale MIMO systems where the number of BS antennas is larger than the number of receive antennas $N_\mathrm{T} \gg N_\mathrm{R}$ results in a negligible performance degradation compared to the optimized solution [90]. Thus, the total transmit power P of the BS is equally allocated to the users $P_i = P/U$. The selection procedure for the precoders and beamformers is detailed in Section 4.2.2. Moreover, $\mathbf{x} \in \mathbb{C}^{U\times 1}$ is the data vector that consists of the M-ary modulated symbols for U users. $\mathbf{n}_i \in \mathbb{C}^{N_\mathrm{R}\times 1}$ is the circularly symmetric complex Gaussian noise vector where the elements follow an i.i.d. $\mathcal{CN}(0, \sigma^2)$ distribution.

The receiver block diagram of a single user is depicted in Figure 4.9b. It is assumed that the users have a fully digital architecture which results in the same number of RF chains and receive antennas. The received signal is passed onto the combiner with the help of RF chains and digital combining is performed to successfully reconstruct the transmitted symbol

$$\begin{aligned} y_i &= \mathbf{w}^{\mathrm{H}}_{a_i,i} \mathbf{r}_i \\ &= \sqrt{P} \mathbf{w}^{\mathrm{H}}_{a_i,i} \sum_{j=1}^{U} \mathbf{H}_{a_j,i} \mathbf{f}_{a_j,j} \mathbf{p}_{a_j,j} \mathbf{x} + \mathbf{w}^{\mathrm{H}}_{a_i,i} \mathbf{n}_i, \end{aligned} \tag{4.28}$$

where $\mathbf{w}_{a_i,i}$ is the digital combiner of the i-th user when the a_i-th AA is selected based on the SM bits. The digital precoder $\mathbf{p}_{a_j,j}$ is selected based on the ZF method in order to eliminate the interference from other users. Thus, the resulting post-processed symbol at each individual user with the matching digital combiner is equivalent to

$$y_i = \sqrt{P_i}\mathbf{w}_{a_i,i}^{\mathrm{H}}\mathbf{H}_{a_i,i}\mathbf{f}_{a_i,i}x_i + \mathbf{w}_{a_i,i}^{\mathrm{H}}\mathbf{n}_i. \tag{4.29}$$

Since the information of the activated AA to transmit the signal to the user is unknown at the user side the users are oblivious of the knowledge of the matching combiner. Therefore, the decoder should be designed accordingly and the user i needs to have the combiners of each AA $\{\mathbf{w}_{1,i}, \mathbf{w}_{2,i}, \ldots, \mathbf{w}_{N_{\mathrm{A}},i}\}$. In order to decode the received symbol a low complexity ML detector is designed to jointly estimate the AA index and M-ary modulated symbol

$$[\hat{a}_i, \hat{x}_i] = \arg\min_{a,m} \left| \frac{1}{\beta}\mathbf{w}_{a,i}^{\mathrm{H}}\mathbf{r}_i - x_m \right|^2. \tag{4.30}$$

Therein, $\mathbf{w}_{a,i} \in \mathbb{C}^{N_{\mathrm{R}}\times 1}$ is the digital combiner from the a-th AA to the i-th user where $a \in \{1, 2, \ldots, N_{\mathrm{A}}\}$. Moreover, β is the normalization coefficient to ensure a constant transmit power at the BS. Finally, x_m is the m-th symbol of the M-ary constellation diagram. Since the ZF method is utilized for the design of the digital precoder $\mathbf{p}_{a_j,j}$ at the BS and the digital combiner $\mathbf{w}_{a_i,i}$ at the user the interference signals from other users fall into the null-space of the transmitted signal to the i-th user as long as $i \neq j$. In a similar manner, the post-processed signal of the i-th user $\mathbf{w}_{a,i}^{\mathrm{H}}\mathbf{r}_i$ falls into the null-space of the transmitted signal to the i-th user as long as $a_i \neq a$. Hence, the proposed detector architecture allows the user to decode the received signal without the knowledge of the channel $\mathbf{H}_{a_j,i}$, analog beamformer $\mathbf{f}_{a_j,j}$, and digital precoder $\mathbf{p}_{a_j,j}$ but only with the knowledge of its own digital combiner vectors $\mathbf{w}_{a,i}$ where $a \in \{1, 2, \ldots, N_{\mathrm{A}}\}$ and the normalization coefficient β. The construction of the digital combining vectors is further detailed in Section 4.2.2.

4.2.2 Precoder, Beamformer, and Combiner Design

Two approaches are taken into account to design the analog beamforming vector. The first method works by choosing the beamforming vector among the transmit AA response vectors $\mathbf{a}_{\mathrm{T}}(\phi)$ in (4.27) which requires infinite resolution phase shifters in order to generate the beamforming vector of the strongest path with exact AoD ϕ. The second method is exploiting a predefined beamsteering codebook with quantized angles between $(0, 2\pi]$. Although the first approach provides the optimum solution for large arrays [91] it is practically impossible to implement a phase shifter with infinite resolution. The second approach ensures an implementable design but induces a performance degradation due to the quantization of the angle of the strongest path.

4.4 Definition. $\mathcal{F}$ is the *codebook* consisting of N codewords

$$\mathcal{F} = \{\mathbf{f}_n \in \mathbb{C}^{N_\mathrm{T}\times 1} : \mathbf{f}_n^\mathrm{H}\mathbf{f}_n = 1,\, n \in \{1, 2, \ldots, N\}\}, \tag{4.31}$$

such that the codeword $\mathbf{f}_n$ is constructed as

$$\mathbf{f}_n = \sqrt{\frac{1}{N_\mathrm{T}}}[1, \exp\left(j\pi \sin(\phi_n)\right), \ldots, \exp\left(j\pi(N_\mathrm{T}-1)\sin(\phi_n)\right)]^\mathrm{T}. \tag{4.32}$$

Herein, ϕ_n denotes the direction of the n-th analog beamformer.

4.5 Definition. $\mathcal{F}_\mathrm{A}$ is the *infinite resolution codebook* that follows the properties introduced in Definition 4.4 consisting of the transmit AA response vectors

$$\mathbf{f}_n = \mathbf{a}_{\mathrm{T}_n},\, n \in \{1, 2, \ldots, N\}. \tag{4.33}$$

Therein, the angle ϕ_n that is utilized to construct the codeword as in (4.32) denotes the AoD of the n-th path. The size of the codebook is equal to the number of multipaths of the channel $N = C$.

4.6 Definition. $\mathcal{F}_\mathrm{B}$ is the *quantized codebook* that follows the properties introduced in Definition 4.4. The codewords $\mathbf{f}_n$ are constructed by utilizing the quantized angles

$$\phi_n = \frac{2\pi n}{N},\, n \in \{0, 2, \ldots, N-1\}. \tag{4.34}$$

Moreover, the size of the codebook is $N = 2^B$ where B is the resolution of the phase shifters.

4.7 Remark. The construction of the codebook $\mathcal{F}_\mathrm{A}$ requires the CSI since the angle ϕ_n denotes the AoD of the n-th path. Therefore, the codebook should be updated at the beginning of each channel coherence time. In addition, the phase shifters should have infinite resolution in order to steer the beam in the direction of any arbitrary angle. In contrast, the codebook $\mathcal{F}_{\mathrm{B}}$ is constructed once based on the resolution of the phase shifters. However, the resolution of the phase shifter B effects the resulting performance of the overall system due to the quantization error.

Given the codebook in Definition 4.4, the analog beamformer $\mathbf{f}_{a,i}$ per user per AA is chosen in order to maximize the SNR of the link between the selected AA and the intended user where $a \in \{1, 2, \dots, N_{\mathrm{A}}\}$ and $i \in \{1, 2, \dots, U\}$. The analog beamformer design problem is then to find $\mathbf{f}_{a,i}$ that solves

$$\mathbf{f}_{a,i} = \underset{\mathbf{f}_n \in \mathcal{F}}{\arg\max} \, ||\mathbf{H}_{a,i}\mathbf{f}_n||^2. \tag{4.35}$$

Herein, $\mathbf{f}_n$ is the n-th codeword where $n \in \{1, 2, \dots, N\}$. Furthermore, $\mathbf{H}_{a,i}$ is the channel between the a-th AA and i-th user.

The problem in (4.35) is solved independently for each AA–user pair. Therefore, the search space of the optimization problem is $\mathcal{F}$ with cardinality $|\mathcal{F}| = N$ which generally is not a large number. Hence, the exhaustive search method is adopted to provide the optimum solution. Note that the optimum analog beamformers for each AA–user pair are found at the beginning of each coherence time of the channel and are stored at the BS.

The resulting analog beamforming matrix $\mathbf{F} \in \mathbb{C}^{N_{\mathrm{T}} \times U}$ of the present channel use for all users according to the selected AA based on the incoming bit stream is

$$\mathbf{F} = [\mathbf{f}_{a_1,1}, \mathbf{f}_{a_2,2}, \dots, \mathbf{f}_{a_U,U}]. \tag{4.36}$$

In order to generate the digital combiner the BS constructs the matrix $\mathbf{H}_i = [\mathbf{H}_{1,i}\mathbf{f}_{1,i}, \mathbf{H}_{2,i}\mathbf{f}_{2,i}, \dots, \mathbf{H}_{N_{\mathrm{A}},i}\mathbf{f}_{N_{\mathrm{A}},i}]$ for each user. $\mathbf{H}_i$ is then used to calculate the digital combiner of the i-th user to successfully reconstruct the transmitted symbol

$$\mathbf{W}_i = \left[\mathbf{w}_{1,i}, \mathbf{w}_{2,i}, \dots, \mathbf{w}_{N_{\mathrm{A}},i}\right] = (\mathbf{H}_i^{\dagger})^{\mathrm{H}}, \tag{4.37}$$

where $i \in \{1, 2, \dots, U\}$.

4.8 Remark. During the DL training phase the BS feeds only back the $\mathbf{H}_i$ matrix to the user i to enable digital combining which has $N_{\mathrm{A}} \times N_{\mathrm{R}}$ dimension. In the case that all users would have to calculate their own digital combiner the full channel matrix $[\mathbf{H}_{1,i}, \mathbf{H}_{2,i}, \dots, \mathbf{H}_{N_{\mathrm{A}},i}]$ would be needed with $N_{\mathrm{A}}N_{\mathrm{T}} \times N_{\mathrm{R}}$ dimensions, which would result in a longer training period.

After determining the analog beamformers and receive combiners the transmitter generates the effective channel matrix $\mathbf{H}_{\text{eff}}$ in order to calculate the digital precoder to eliminate the IUI

$$\mathbf{H}_{\text{eff}} = \begin{bmatrix} \mathbf{w}_{a_1,1}^{\text{H}} \mathbf{H}_{a_1,1} \mathbf{f}_{a_1,1} & \cdots & \mathbf{w}_{a_1,1}^{\text{H}} \mathbf{H}_{a_U,1} \mathbf{f}_{a_U,U} \\ \vdots & \ddots & \vdots \\ \mathbf{w}_{a_U,U}^{\text{H}} \mathbf{H}_{a_1,U} \mathbf{f}_{a_1,1} & \cdots & \mathbf{w}_{a_U,U}^{\text{H}} \mathbf{H}_{a_U,U} \mathbf{f}_{a_U,U} \end{bmatrix}. \tag{4.38}$$

Finally, the precoding matrix $\mathbf{P} \in \mathbb{C}^{U \times U}$ for all users is determined as

$$\mathbf{P} = [\mathbf{p}_{a_1,1}^{\text{T}} \cdots \mathbf{p}_{a_U,U}^{\text{T}}]^{\text{T}} = \beta \mathbf{H}_{\text{eff}}^{\dagger}, \tag{4.39}$$

where $\mathbf{p}_{a_i,i} \in \mathbb{C}^{1 \times U}$ is the precoder of the i-th user when the a_i-th AA is chosen based on the spatially-modulated bits. The average transmit power constraint P is satisfied by the coefficient β

$$\beta = \text{E}\left(\sqrt{\frac{U}{\text{tr}\left(\mathbf{H}_{\text{eff}}^{\dagger} (\mathbf{H}_{\text{eff}}^{\dagger})^{\text{H}} \right)}} \right), \tag{4.40}$$

where the expectation is over channel realizations. Since the complex channel coefficients are the key parameters for spatially-modulated bits an instantaneous power constraint is not applicable for systems that employ the SM scheme [69].

Figure 4.10 depicts the performance improvement at a large-scale MU-MIMO network when the HBFSM scheme is considered [72]. The performance of the HBFSM scheme is compared with the conventional SM scheme for the MU scenario [69]. In the considered conventional SM system each AA is dedicated to a single user. The SM bits are mapped onto a single antenna at each AA per user and the M-ary modulated symbol is transmitted through the selected antenna. A ZF precoder with average power constraint is employed in order to eliminate the IUI.

In the simulation setup an MU network is considered with $U = 2$ users. It is assumed that each user has $N_{\text{R}} = 1$ receive antennas. 2 spatially-modulated bits are transmitted per user per channel use by either of the considered schemes. A system that employs conventional SM with only digital precoding at the BS has $N_{\text{T}}^{\text{SM}} = 4$ transmit antennas at each AA and the same number of AAs as users $N_{\text{A}} = U$. In contrast, a system that employs the HBFSM scheme has $N_{\text{A}} = 4$ AAs consisting of $N_{\text{T}} = 8$ and $N_{\text{T}} = 64$ transmit antennas at each array,

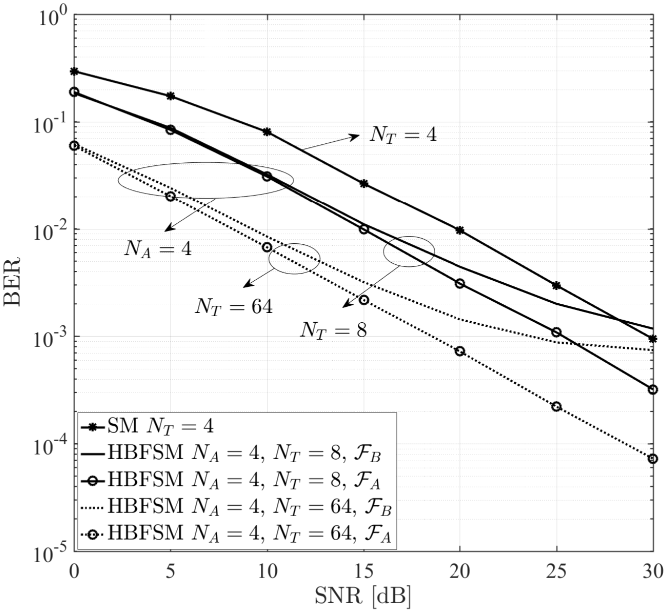


Figure 4.10: Uncoded BER performance comparison of the conventional SM and HBFSM schemes ($U = 2$, $N_\mathrm{R} = 1$, $B = 8$ with 4-QAM).

respectively. Therefore, the total number of bits transmitted through SM per channel use by both schemes is $\log_2 N_\mathrm{T}^\mathrm{SM} = \log_2 N_\mathrm{A} = 2$. A mmWave channel is adopted with $C = 10$ scattering clusters. The cardinality of the codebook is $|\mathcal{F}_\mathrm{A}| = 10$ when the codebook in Definition 4.5 is considered. The phase shifter resolution for the codebook $\mathcal{F}_\mathrm{B}$ is set to $B = 8$ bits, hence the cardinality of the codebook is $|\mathcal{F}_\mathrm{B}| = 256$ when the codebook in Definition 4.6 is considered.

As shown in Figure 4.10, the error performance is better than in the conventional SM scheme up to the high-SNR region although the HBFSM scheme has an error floor when the codebook $\mathcal{F}_\mathrm{B}$ is utilized. When the codebook $\mathcal{F}_\mathrm{A}$ is utilized, HBFSM outperforms conventional SM regardless the SNR value. An AA consisting of only 8 transmit antennas already provides a 5 dB gain over the conventional SM system. Since exploiting a larger number of antennas increases the beamforming gain and enables narrowing the beam, a 12 dB gain can be enjoyed by utilizing AAs equipped with 64 transmit antennas. Although there is an intersection point between HBFSM and conventional SM by increasing the codebook size and number of transmit antennas at an array, the performance of HBFSM can be further improved. Note that each AA is served by a single RF chain. Hence, increasing the number of transmit antennas at an array does not come with a high cost while significantly increasing the performance. The simulation results show that the number of antennas at an array is one of the key parameters that affects the performance. Because of the array gain, the larger the number of transmit antennas the better the beam can be directed to the intended user.

Since the phase shifters have a limited resolution there will be a performance degradation caused by the quantization error. In the next section this quantization error is studied and an analytical expression is derived.

4.2.3 Characterization of the Quantization Error

For large-scale AAs the right and left singular vectors of the channel are the transmit and receive AA response vectors, respectively [91]. The optimization problem is formulated to find the optimal analog beamformer for the channel given in (4.25) by adopting the codebook in Definition 4.5

$$\mathbf{f}_{\mathrm{A}} = \underset{\mathbf{f}_{\mathrm{A},n} \in \mathcal{F}_{\mathrm{A}}}{\arg\max} \, ||\mathbf{H}\mathbf{f}_{\mathrm{A},n}||^2. \tag{4.41}$$

$\mathbf{f}_{\mathrm{A},n}$ is the n-th codeword from the codebook where $n \in \{1, 2, \ldots, N\}$ with $N = C$ being the number of scattering clusters of the channel and the total number of codewords in the codebook. $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$ is the channel matrix and $\mathbf{f}_{\mathrm{A},n} \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ is the n-th codeword of the codebook $\mathcal{F}_{\mathrm{A}}$. In order to achieve the performance of the optimal beamformer infinite resolution phase shifters should be implemented. Since this is not possible for practical systems beamsteering codebooks with quantized angles are commonly considered [92]. However, selecting the analog beamformer among beamsteering codebooks leads to a quantization error. In the sequel, the quantization error between the codewords among the codebooks $\mathcal{F}_{\mathrm{A}}$ and $\mathcal{F}_{\mathrm{B}}$ is derived.

4.9 Definition. Let $\mathbf{f}_a$ and $\mathbf{f}_b$ be the codewords belonging to the codebooks $\mathcal{F}_{\mathrm{A}}$ and $\mathcal{F}_{\mathrm{B}}$, respectively. The *chordal distance* between $\mathcal{F}_{\mathrm{A}}$ and $\mathcal{F}_{\mathrm{B}}$ is defined as

$$d_c(\mathbf{f}_a, \mathbf{f}_b) = \sin[\theta(\mathbf{f}_a, \mathbf{f}_b)], \tag{4.42}$$

where $\theta(\mathbf{f}_a, \mathbf{f}_b) = \arccos|\mathbf{f}_a^{\mathrm{H}}\mathbf{f}_b|$ is the principle angle.

By exploiting the chordal distance d_c that is introduced in Definition 4.9, the quantization error can be upper and lower bounded as in the following proposition.

4.10 Proposition. *Assume that* $\mathbf{f}_A$ *and* $\mathbf{f}_B$ *are the solution of the problem in* (4.35) *for codebooks* $\mathcal{F}_A$ *and* $\mathcal{F}_B$*, respectively. Let* d_c *be the chordal distance between input vectors then the quantization error can be characterized as*

$$d_{c,min} \leq \max_{\mathbf{f}_A \in G(N_T,1)} d_c^2(\mathbf{f}_A, \mathbf{f}_B) \leq d_{c,max}, \tag{4.43a}$$

$$d_{c,min} = 1 - \mathrm{tr}\left(\mathbf{f}_A\mathbf{f}_A^{\mathrm{H}}\mathbf{f}_{B,n}\mathbf{f}_{B,n}^{\mathrm{H}}\right), \tag{4.43b}$$

$$d_{c,max} = 2^{-\frac{B}{N_T-1}}. \tag{4.43c}$$

Therein, $G(N_T, 1)$ *is the Grassmann manifold which is defined as the set of all lines in an* N_T*-dimensional Euclidean space.*

Proof. Let $\mathcal{F}_\mathrm{B}$ be a codebook with B bits resolution phase shifters where the codebook size is $N = 2^B$. Then, the problem in (4.35) can be reformulated as

$$\begin{aligned}\mathbf{f}_\mathrm{B} &= \arg\min_{\mathbf{f}_{\mathrm{B},n}\in\mathcal{F}_\mathrm{B}} d_c^2(\mathbf{f}_\mathrm{A}, \mathbf{f}_{\mathrm{B},n}) \\ &= \arg\min_{\mathbf{f}_{\mathrm{B},n}\in\mathcal{F}_\mathrm{B}} 1 - \mathrm{tr}\left(\mathbf{f}_\mathrm{A}\mathbf{f}_\mathrm{A}^{\mathrm{H}}\mathbf{f}_{\mathrm{B},n}\mathbf{f}_{\mathrm{B},n}^{\mathrm{H}}\right),\end{aligned} \tag{4.44}$$

where (4.44) is an example of a Grassmannian quantization on the Grassmann manifold $G(N_\mathrm{T}, 1)$ [93, 94, 95]. The codebook $\mathcal{F}_\mathrm{B}$ can be considered as a Grassmannian subspace sphere-packing codebook since the elements of $\mathcal{F}_\mathrm{B}$ are distributed over a sphere with radius $||\mathbf{f}_{\mathrm{B},n}^{\mathrm{H}}\mathbf{f}_{\mathrm{B},n}||^2 = 1$, for all $n \in \{1, 2, \ldots, N\}$. The distance between the optimal beamformer and the selected beamformer from the quantized codebook is defined as $d_c(\mathbf{f}_\mathrm{A}, \mathbf{f}_\mathrm{B})$. Based on a Grassmannian subspace sphere-packing codebook the deterministic subspace quantization distortion in (4.44) is defined by

$$d_{c,\mathrm{min}} = 1 - \mathrm{tr}\left(\mathbf{f}_\mathrm{A}\mathbf{f}_\mathrm{A}^{\mathrm{H}}\mathbf{f}_{\mathrm{B},n}\mathbf{f}_{\mathrm{B},n}^{\mathrm{H}}\right). \tag{4.45}$$

Moreover, the maximum value of $d_c^2(\mathbf{f}_\mathrm{A}, \mathbf{f}_{\mathrm{B},n})$ can be upper bounded by adopting the result in [94, Theorem 4]

$$d_{c,\mathrm{max}} = v^{-\frac{1}{N_\mathrm{T}-1}} 2^{-\frac{B}{N_\mathrm{T}-1}}. \tag{4.46}$$

As the result of the derivations by considering the specific case at hand, the metric ball volume coefficient is $v = 1$ as computed in Appendix A. Hence, the upper bound of the quantization error becomes $d_{c,\max} = 2^{-\frac{B}{N_\mathrm{T}-1}}$. ■

4.11 Corollary. *The quantization error in* (4.43) *vanishes asymptotically in* B *since* $\lim_{B\to\infty} 2^{-\frac{B}{N_T-1}} = 0$.

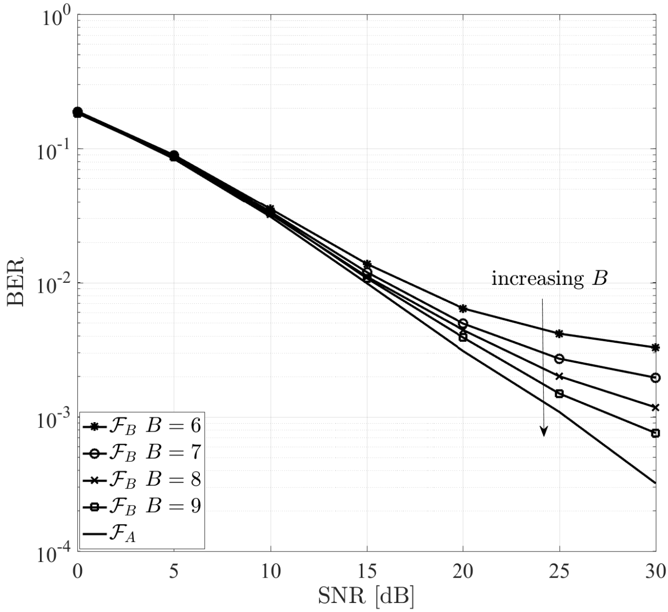


Figure 4.11: Uncoded BER performance of the HBFSM scheme for various codebook sizes ($U = 2$, $N_\mathrm{A} = 4$, $N_\mathrm{T} = 8$, $N_\mathrm{R} = 1$ with 4-QAM).

In Figure 4.11 the uncoded BER performance of the system is depicted for the cases that the analog beamformer vector is chosen among the transmit AA response vectors $\mathcal{F}_\mathrm{A}$ and the beamsteering codebook $\mathcal{F}_\mathrm{B}$ [72]. The simulation results indicate that the quantization error of the codebook $\mathcal{F}_\mathrm{B}$ leads to an error floor for low codebook sizes. Since the digital combiner at the receiver side is generated by using the analog beamforming vector too, an erroneous beamforming vector cumulatively increases the resulting error probability. Contrary, when the analog beamforming vector is chosen from codebook $\mathcal{F}_\mathrm{A}$ the performance is better and there is no error floor in the system. Although the codebook $\mathcal{F}_\mathrm{A}$ provides the optimal beamformer there is no phase shifter with infinite resolution. Hence, the codebook $\mathcal{F}_\mathrm{B}$ with B-bit resolution would be utilized in practical systems. When the resolution of the phase shifters is changed from 6 bits to 9 bits the performance is increased significantly. The performance of the beamsteering codebook will eventually converge to the optimal beamformer with increasing resolution.

4.2.4 Achievable Rate

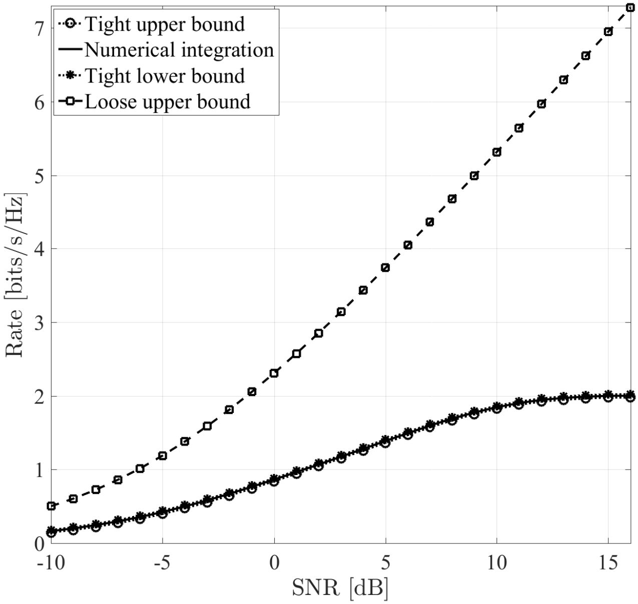


Figure 4.12: Achievable rate approximation of the HBFSM scheme for a single user ($N_\mathrm{A} = 2, N_\mathrm{T} = 8, N_\mathrm{R} = 1$ with BPSK).

The mutual information between the transmitted symbol $\bar{x}_i$ and the post-processed symbol y_i of the i-th user is

$$I(y_i; x_i) = h(y_i) - h(y_i|\bar{x}_i) \ , \tag{4.47a}$$

$$= h(y_i) - h(z_i) \ . \tag{4.47b}$$

$\bar{x}_i$ is the information-bearing symbol and (4.47b) holds due to the fact that $\bar{x}_i$ consists of M-ary and spatially-modulated symbols where the SM part possesses channel statistics as well and z_i is the noise term after receive combining with circularly symmetric complex Gaussian distribution $\mathcal{CN}(0, \sigma^2)$. There is no closed-form solution for the differential entropy of a Gaussian mixture random variable. An approximation is derived in Proposition 4.12.

4.12 Proposition. *Bounds on the achievable rate* $\tilde{R}_i$ *of the* i*-th user can be derived for the post-processed symbol* y_i *with Gaussian mixture distribution*

$$h_l(y_i) - h(z_i) \le \tilde{R}_i \le h_u(y_i) - h(z_i), \tag{4.48a}$$

$$h_l(y_i) = (\gamma + \alpha_{N,N'}) \log_2 e + \log_2 \sigma, \tag{4.48b}$$

$$h_u(y_i) = (\gamma + \beta_{N,N'}) \log_2 e + \log_2 \sigma, \tag{4.48c}$$

$$h(z_i) = \log_2(\pi e \sigma^2) \ . \tag{4.48d}$$

Details on these bounds are given in Appendix B.2.

Proof. First, the PDF of the received signal is derived

$$f_{Y_i}(y_i) = \sum_{m=1}^{M} \sum_{j=1}^{N_\mathrm{A}} f_{Y_i}(y_i|a = j, x = x_m) \mathrm{Pr}(a = j, x = x_m) \ . \tag{4.49}$$

Since the selection of the AA indices and M-ary symbols are independent events the joint probability in (4.49) can be separated as

$$\mathrm{Pr}(a = j, x = x_m) = \mathrm{Pr}(a = j)\mathrm{Pr}(x = x_m) = \frac{1}{N_\mathrm{A}} \frac{1}{M} \ , \tag{4.50}$$

where random variables a and x follow uniform distributions $\mathcal{U}(1, N_\mathrm{A})$ and $\mathcal{U}(1, M)$, respectively. For a given AA index the M-ary symbol pair $(a = j, x = x_m)$, the randomness lies in the noise term that follows a complex Gaussian distribution. Therefore, the PDF of the received signal y_i is the composition of the complex Gaussian PDFs with mean values $(\sqrt{P_i}\mathbf{w}_{j,i}^\mathrm{H}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m)$ where $j \in$

$\{1, 2, \ldots, N_\mathrm{A}\}$ and $m \in \{1, 2, \ldots, M\}$. The resulting PDF follows a Gaussian mixture distribution

$$f_{Y_i}(y_i) = \frac{1}{MN_\mathrm{A}} \sum_{m=1}^{M} \sum_{j=1}^{N_\mathrm{A}} \frac{1}{\pi\sigma} \exp\left\{ -\frac{|y_i - \sqrt{P_i}\mathbf{w}_{j,i}^\mathrm{H}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m|^2}{\sigma^2} \right\}, \tag{4.51}$$

where $\mathbf{H}_{j,i}$, $\mathbf{f}_{j,i}$ and $\mathbf{w}_{j,i}$ are the channel matrix, analog beamformer and receive combiner for the j, i AA–user pair, respectively. Herein, x_m denotes the m-th symbol from the M-ary constellation diagram and P_i is the transmit power of the i-th user. By using the approximation derived for Gaussian mixtures in [96] the differential entropy of y_i follows as in (4.48). ■

Although the upper and lower bounds in (4.48) are tight they result in an expression that is not tractable especially for the EE optimization problem that is discussed in Section 4.2.5. Therefore, a looser upper bound is derived in Proposition 4.13 which is more convenient as far as system optimization is concerned.

4.13 Proposition. *For all $i \in \{1, 2, \ldots, U\}$ the achievable rate $\tilde{R}_i$ of the i-th user assuming discrete-symbol input is upper-bounded as*

$$\begin{aligned}
\tilde{R}_i \leq \log_2 \pi e \Bigg[& \sigma^2 + \frac{P}{UMN_A} \sum_{m=1}^{M} \sum_{j=1}^{N_A} \left| \mathbf{w}_{j,i}^\mathrm{H}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m \right|^2 \\
& - \frac{P}{UM^2N_A^2} \left(\sum_{m=1}^{M} \sum_{j=1}^{N_A} \Re\left\{ \mathbf{w}_{j,i}^\mathrm{H}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m \right\} \right)^2 \\
& - \frac{P}{UM^2N_A^2} \left(\sum_{m=1}^{M} \sum_{j=1}^{N_A} \Im\left\{ \mathbf{w}_{j,i}^\mathrm{H}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m \right\} \right)^2 \Bigg] \\
- \log_2 \pi e \sigma^2 = \bar{R}_i \ ,
\end{aligned} \tag{4.52}$$

where $\bar{R}_i$ denotes the upper bound of the achievable rate of the i-th user.

Proof. See Appendix B.3. ■

Figure 4.12 depicts the achievable rate of the HBFSM scheme. In this figure a BS with $N_\mathrm{A} = 2$ AAs consisting of $N_\mathrm{T} = 8$ transmit antennas is considered.

A single-user network is assumed where the UT has a single receive antenna. The BS transmits a binary phase shift keying (BPSK) symbol from the selected AA based on the SM bits per channel use. The differential entropy of the post-processed signal $h(y) = -\int_{-\infty}^{\infty} f_Y(y) \log f_Y(y) dy$ is computed by using the numerical integration method. Then the achievable rate is calculated as

$$\tilde{R} = h(y) - \log_2(\pi e \sigma^2) \ . \tag{4.53}$$

The lower bound (LB) and upper bound (UB) of the tight approximation are computed by using (4.48) and the looser UB is computed by using (4.52). As shown in Figure 4.12 the tight bounds of Proposition 4.12 are almost the same as the exact achievable rate. However, the gap between the approximation by Proposition 4.13 is getting larger with an increasing SNR. The reason is that this approximation is derived by the principle of maximum differential entropy, i.e., the differential entropy of a random variable that follows a Gaussian distribution provides the maximum attainable differential entropy [97]. Therefore, the differential entropy of a random variable with Gaussian mixture distribution cannot exceed that of a random variable with Gaussian distribution with the same variance value.

4.14 Remark. The bound in (4.52) achieves the best available trade-off between accuracy and mathematical tractability. Indeed, as it will be corroborated later in this section, it provides a higher accuracy than the bound from [71] that does not account for the discrete nature of the modulation symbols. On the other hand, the tight approximation in (4.48) is quite cumbersome and is not suitable for optimization purposes.

4.2.5 Energy Efficiency Optimization

This section develops an optimization framework for bit-per-Joule EE maximization in mmWave large-scale MIMO systems with the HBFSM scheme. The EE maximization problem is formulated subject to maximum power constraints, minimum and maximum number of active users, and minimum per-user rate constraints. The optimization is carried out considering both the traditional power consumption model that assumes a linear operation of the transmit amplifier, as well as the scenario in which the non-linearity of the transmit amplifier cannot be neglected, leading to a more realistic power consumption model.

The approximation by Proposition 4.13 is adopted as the achievable rate expression of the HBFSM system which provides a tractable expression to solve the optimization problem. The right-hand-side of (4.52) can be expressed as

$$\bar{R}_i = \log_2\left(1 + \frac{P}{U} a_i\right) ,$$

$$a_i = \sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} \frac{|b_{j,m,i}|^2}{MN_\mathrm{A}} - \left(\sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} \frac{\Re\{b_{j,m,i}\}}{MN_\mathrm{A}}\right)^2 - \left(\sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} \frac{\Im\{b_{j,m,i}\}}{MN_\mathrm{A}}\right)^2 ,$$

$$b_{j,m,i} = \frac{\mathbf{w}_{j,i}^{\mathrm{H}} \mathbf{H}_{j,i} \mathbf{f}_{j,i} x_m}{\sigma} . \tag{4.54}$$

The variable a_i is guaranteed to be non-negative. Indeed, defining the complex number $X_i = \sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} b_{j,m,i}$ and recalling that $MN_\mathrm{A} \geq 1$ it holds that

$$\sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} |b_{j,m,i}|^2 - \frac{1}{MN_\mathrm{A}}\left[\left(\sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} \Re\{b_{j,m,i}\}\right)^2 + \left(\sum_{m=1}^{M}\sum_{j=1}^{N_\mathrm{A}} \Im\{b_{j,m,i}\}\right)^2\right]$$
$$\geq |X_i|^2 - \Re\{X_i\}^2 - \Im\{X_i\}^2 = 0 . \tag{4.55}$$

Thus, the ergodic achievable sum-rate can be upper bounded as

$$R_s = \mathrm{E}\left(\sum_{i=1}^{U} \bar{R}_i\right) = U\mathrm{E}\left(\log_2\left(1 + \frac{P}{U} a_i\right)\right) , \tag{4.56}$$

wherein it has been assumed that all $\{a_i\}_{i=1}^{U}$ are identically distributed, motivated by the assumption of identical distribution of the fading channels experienced by the U users. Therefore, the ergodic achievable sum-rate can be expressed as $R_s = UR$, where R is the ergodic achievable rate of any user.

Power Consumption Model

This section introduces the power consumption models for the EE maximization problem. In particular, two different models will be considered for the total power consumed for the DL transmission, namely

$$P_\mathrm{t}(U,P) = \eta P + UP_\mathrm{c} , \tag{4.57}$$

$$P_\mathrm{t}(U,P) = \eta_\mathrm{nl} P^{1/n} + UP_\mathrm{c,nl} . \tag{4.58}$$

Equation (4.57) represents the standard power consumption model which accounts for the total power P radiated by the BS to serve all users, with $\eta > 0$ being related to the amplifier efficiency, plus the static hardware power P_c, that is dissipated to operate the communication of a single user in the network. It should be stressed that the expression of P_c could be further detailed to highlight the dependence of P_c on the specifications of the particular hardware blocks that are deployed at each network node, e.g., DACs, ADCs, filters, cooling equipment. However, the specific components that contribute to the value of P_c are inessential as far as the maximization of the EE with respect to P and U is concerned. Indeed, the results that follow hold for any $P_c > 0$ assuming that P_c does not depend on the optimization variables P and U.

The peculiar feature of the model in (4.57) is that it assumes a linear operation for the transmit amplifier. The model in (4.58) extends that in (4.57) by accounting for the presence of non-linear terms at the output of the transmit amplifier. As for the parameters η_{nl} and $P_{\text{c,nl}}$, they have the same interpretation as in (4.57), but are denoted by different symbols because in general their values will be different from those in (4.57). Nevertheless, it is still assumed that both $P_{\text{c,nl}}$ and η_{nl} are non-negative numbers. The model in (4.58) has been previously introduced in [98] and the special case $n = 2$ has been treated in [99, 100]. However, none of these previous works has considered the model in (4.58) in the context of EE maximization. This appears to be a challenging scenario due to the fact that (4.58) is a non-convex function of the transmit power, which complicates the direct use of fractional optimization methods.

In this context the system EE is defined for a given bandwidth B as

$$\text{EE} = B\frac{UR(U,P)}{P_t(U,P)} \,, \tag{4.59}$$

which is to be maximized by determining the optimal (U, P) by considering the linear and non-linear power models, respectively. Section 4.2.5 addresses the maximization of the system EE assuming the linear power model in (4.57) and the case of the non-linear power model in (4.58). The motivation to treat these two cases separately is explained later when the linear power model in (4.57) enables a deeper analytical insight into the optimal resource allocation structure.

Energy Efficiency Optimization with the Linear Power Model

Assume that the power consumption model in (4.57) holds and that the goal is to determine the optimal transmit power P and number of users U that maximize the system EE. This problem is first formulated in the sequel, then the optimal (U, P) is derived by studying the problems of optimal P for fixed U and the optimal U for fixed P, respectively.

Problem Formulation. The problem of maximizing the EE of the considered HBFSM system with respect to the transmit power P and the number of users U can be stated as the optimization program:

$$\max_{U,P} \frac{UR(U,P)}{P+UP_\mathrm{c}} \tag{4.60a}$$

$$\text{s.t. } 0 \le P \le P_\mathrm{max} \tag{4.60b}$$

$$U \in \{U_\mathrm{min}, U_\mathrm{min}+1, \ldots, U_\mathrm{max}\} \tag{4.60c}$$

$$R(U,P) \ge R_\mathrm{min} \ , \tag{4.60d}$$

wherein, without loss of generality, η has been included in P_c, and the resulting multiplicative constant B/η has been neglected since it is not essential for the optimization of P and U. Constraint (4.60b) enforces the physical limitations on the transmit power, constraint (4.60c) ensures that the number of active RF chains/users is between a minimum desired value and the maximum feasible value, while constraint (4.60d) sets a minimum quality of service (QoS) in terms of per-user minimum acceptable achievable rate.

The challenge to solve problem (4.60) lies mainly in the presence of the integer variable U. This prevents the use of standard gradient-based methods, and in its fractional nature, which makes the problem non-convex even in the continuous variable P. The available tools for the optimization of non-convex fractional functions are able to operate with polynomial complexity. Then, the numerator and denominator of the fraction to be maximized are concave and convex functions in the optimization variables. All constraint functions are convex [101]. Even allowing the initial integer variable U to be continuous, the joint concavity of $UR(U,P)$ with respect to (U,P) is still not clear.

In order to solve problem (4.60) the first step is to study the problem with respect to one variable at a time, finding the optimal P for any fixed U and the optimal integer U for any fixed P. Then, joint optimization of U and P will be derived which is proved to be globally optimal solution of problem (4.60).

Optimal P for Fixed U. If U is fixed constraint (4.60c) becomes a single-ratio fractional problem, whose solution can be characterized in closed-form.

4.15 Proposition. *For any given U* (4.60a) *has a unique stationary point P_s and there exists a unique power level P_{min} such that* (4.60d) *is fulfilled with equality. Problem* (4.60) *is feasible if and only if $P_{min} \leq P_{max}$. In this case it admits the unique solution*

$$P^* = \max(P_{min}, \min(P_{max}, P_s)) \ . \tag{4.61}$$

Proof. The objective (4.60a) is non-negative. Moreover, the argument of the statistical expectation at the numerator of (4.60a) tends to zero for $P \to 0$ while it grows logarithmically for $P \to \infty$. Since this is true for every realization of the random variable to be averaged we can conclude that the whole fraction (4.60a) tends to zero for both $P \to 0$ and $P \to \infty$. Thus, for any given U (4.60a) has to have a finite global maximizer with respect to P.

Next, it is observed that with respect to only P the function $R(U, P)$ is strictly concave, because the argument of the statistical expectation defining $R(U, P)$ is strictly concave for any realization of the random variable that is averaged, and because the statistical average is a linear operator, thus preserving the concavity property of its argument. Therefore, with respect to P constraint (4.60a) has a strictly concave numerator and an affine denominator, thus being strictly pseudo-concave. The strict pseudo-concavity of constraint (4.60a) coupled with the fact that it admits a finite global maximizer implies that (4.60a) has only one stationary point P_s, which is also its global maximizer. Moreover, constraint (4.60a) is increasing for $P \leq P_s$ and decreasing for $P \geq P_s$.

Finally, for fixed U the uniqueness of $P_{\min}$ follows from the fact that $R(U, P)$ is strictly increasing in P and $R(U, 0) = 0$. Thus, the equation $R(U, P) = R_{\min}$ has a unique solution with respect to P, say $P_{\min}$. Therefore, problem (4.60) is feasible if and only if $P_{\min} \leq P_{\max}$, and in this case, putting together previous arguments, (4.61) follows. ■

Optimal U for Fixed P. If P is fixed constraint (4.60b) becomes an integer fractional problem. Finding the optimal U appears more challenging than optimizing P, since no low-complexity algorithm exists for non-linear, integer problems, and exhaustive searches are typically used as solution methods, which might be time consuming, especially for large $U_{\max}$ and low $U_{\min}$. Instead, this section determines the optimal integer U in closed-form.

4.16 Lemma. *Consider the following function of real variables:*

$$g : (y, P) \in \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow g(y, P) = yR(y, P) \ . \tag{4.62}$$

The function g is jointly concave in y and P. Moreover, it is strictly increasing and concave in y for fixed P.

Proof. The function g can be seen to be the perspective of the strictly concave function

$$f(P) = \mathrm{E}(\log_2(1 + \alpha P)) \ . \tag{4.63}$$

Then, the joint concavity of g immediately follows from the fact that the perspective function preserves the concavity property [102, Chapter 3].

The second part of the Lemma 4.16 follows upon directly computing the first- and second-order derivatives of g with respect to y, namely

$$\frac{\partial g}{\partial y} = \mathrm{E}\left(\log_2\left(1 + \alpha\frac{P}{y}\right) - \frac{\alpha P/y}{1 + \alpha P/y}\right) \tag{4.64}$$

$$\frac{\partial^2 g}{\partial y^2} = -\mathrm{E}\left(\frac{\alpha 2P^2}{(y + \alpha P)^2}\right) \ , \tag{4.65}$$

and observing that (4.64) is strictly positive[1] and (4.65) is strictly negative, for all values of parameter $y > 0$. ■

4.17 Remark. It can be seen that the function g considered in Lemma 4.16 is formally equivalent to the rate function $UR(U, P)$ in problem (4.60) with the difference that the first argument of the function takes value over the positive real line instead of being an integer value. However, *this does not imply any relaxation to problem* (4.60). The function g has been introduced only because it is instrumental in finding the optimal *integer U* in problem (4.60). Indeed, Proposition 4.18 shows how by exploiting the properties of the function $g(y, P)$ it is possible to derive the optimal integer U that solves (4.60) for any fixed P.

4.18 Proposition. *For any given P the function $f(y) = g(y, P)/(P + yP_c)$ has a unique maximizer y_s, and there exists a unique value y_{min} such that $g(y_{min}, P) = R_{min}$. Then, the optimal integer U for problem* (4.60) *is computed as*

$$U^* = \underset{\{\lceil \bar{y}^* \rceil, \lfloor \bar{y}^* \rfloor\}}{\arg\max} \ EE(U, P) \ , \tag{4.66}$$

---

[1] Recall that $(1 + t)\log(1 + t) > t$ holds for all $t > 0$.

with y^* given by

$$y^* = \max(\max(U_{min}, y_{min}), \min(U_{max}, y_s)) \; . \tag{4.67}$$

Before deriving the proof, it is useful to observe that Proposition 4.18 says that for any given P the optimal number of users U is found by first computing the continuous quantity in (4.67) and then checking which of the two integer values $\lceil \bar{y}^* \rceil$ and $\lfloor \bar{y}^* \rfloor$ yields the larger value of the objective (4.60a).

Proof. The first step of the proof is to show that the function f has a unique maximizer and that the equation $g(y, P) = R_{\text{min}}$ has a unique solution. Both points follow exploiting Lemma 4.16 and by following similar steps as in the proof of Proposition 4.15. Specifically, by virtue of Lemma 4.16 it follows that $g(y, P)$ is strictly concave in y, and thus $f(y)$ is a strictly pseudo-concave function, being the ratio between a strictly concave function and an affine function. Moreover, it is easy to see that $f(y)$ is non-negative and tends to zero for both $y \to 0$ and $y \to \infty$. As a result, for any given P, $f(y)$ is a unimodal function which has a finite global maximizer y_{s}. Moreover, f is increasing for $y \leq y_{\text{s}}$ and decreasing for $y > y_{\text{s}}$. On the other hand, it is also easy to check that $R(y, P)$ is strictly decreasing for $y \geq 0$, and since $\lim_{y \to 0} R(y, P) = +\infty$ the equation $R(y, P) = R_{\text{min}}$ has a unique solution for any $R_{\text{min}} \geq 0$. From these properties it follows that the maximizer of $f(y)$ over the set $y \in [U_{\text{min}}, U_{\text{max}}]$, $R(y, P) \geq R_{\text{min}}$ is given by (4.67).

At this point, the key step in the proof is to observe that for any fixed P objective (4.60a) can be seen as the sampling of the function $f(y)$ with sampling step 1. Then, since $f(y)$ is increasing for $y \leq y_s$ and decreasing for $y > y_s$ it must hold that

$$\text{EE}(\lfloor y^* \rfloor, P) \geq \text{EE}(U, P) \; , \tag{4.68}$$

for any $U \leq y^*$, and

$$\text{EE}(\lceil y^* \rceil, P) \geq \text{EE}(U, P) \; , \tag{4.69}$$

for any $U \geq y^*$. In other words, given the unimodality of the function f and its maximizer y^* the optimal integer $U \in \{U_{\text{min}}, \dots, U_{\text{max}}\}$ must be either the lower or the upper integer of y^*. Hence, (4.66) follows. ■

Before concluding this part, it should be stressed again that, although a continuous version of (4.60a), i.e., $f(x)$, has been exploited in the proof of Proposition 4.18, this entails no loss of optimality and the final solution in (4.66) is the global optimal, *integer* solution of (4.60) with respect to U.

Joint Optimization of U and P. After optimizing the transmit power and number of users separately, now the joint optimization of both U and P will be derived.

To restate problem 4.60 the new variable $z = P/U$ is defined

$$\max_{U,z} \frac{\mathrm{E}(\log_2(1+\alpha z))}{z+P_\mathrm{c}} \tag{4.70a}$$

$$\text{s.t. } 0 \leq z \leq \frac{P_\mathrm{max}}{U} \tag{4.70b}$$

$$U \in \{U_\mathrm{min}, U_\mathrm{min}+1, \dots, U_\mathrm{max}\} \tag{4.70c}$$

$$\mathrm{E}(\log_2(1+\alpha z)) \geq R_\mathrm{min} \ . \tag{4.70d}$$

Inspecting (4.70), the following remarks can be made.

4.19 Remark. The objective (4.70a) depends only on z, and is formally equivalent to the objective (4.60a) regarded only as a function of the transmit power P. As a consequence, the same arguments used in Proposition 4.15 can be used to show that (4.70a) is a unimodal function with a unique stationary point z_s, which coincides with the function global maximizer. Thus, z_s can be found as the unique solution of the equation

$$\frac{\partial}{\partial z}\left(\frac{\mathrm{E}(\log_2(1+\alpha z))}{z+P_\mathrm{c}}\right) = 0 \ . \tag{4.71}$$

4.20 Remark. Similarly, the constraint function in (4.70d) depends only on z and is formally equivalent to (4.60d) regarded only as a function of the transmit power P. Thus, there exists a unique value z_min such that

$$\mathrm{E}(\log_2(1+\alpha z_\mathrm{min})) = R_\mathrm{min} \ . \tag{4.72}$$

According to Remark 4.19 and Remark 4.20 for any given U problem (4.70) has a unique optimal solution z^* given by:

$$z^* = \max(z_\mathrm{min}, \min(z_s, P_\mathrm{max}/U)) \ . \tag{4.73}$$

Since $z = P/U$ the optimal P^* for any given U is computed in closed form as

$$P_U^* = \max(U z_\mathrm{min}, \min(U z_\mathrm{s}, P_\mathrm{max})) \ . \tag{4.74}$$

A first important observation that follows from (4.74) is that the quantities z_min and z_s do not depend on U, being the unique stationary point of (4.70a) and the unique value such that (4.70d) is fulfilled with equality. This implies that (4.70) can be efficiently solved by simply computing (4.74) for any $U \in \{U_{\min}, \ldots, U_{\max}\}$ and then selecting the best pair (U, P_U^*). A procedure to globally solve (4.70) can be formulated as in Algorithm 3.

Algorithm 3: Global optimization of (U, P) with power model (4.57)

```
Compute $z_s$ by solving (4.71);
Compute $z_{\min}$ by solving (4.72);
Compute $P_{U_{\min}}^*$ as in (4.74);
$\mathrm{EE}_{U_{\min}} = \mathrm{EE}(U_{\min}, P_{U_{\min}}^*)$;
$(\widetilde{U}, \widetilde{P}) = (U_{\min}, P_{U_{\min}}^*)$;
for $U = U_{\min} + 1 : U_{\max}$ do
    Compute $P_U^*$ as in (4.74);
    $\mathrm{EE}_U = \mathrm{EE}(U, P_U^*)$;
    if $\mathrm{EE}_U > \mathrm{EE}_{U-1}$ then
        $(\widetilde{U}, \widetilde{P}) = (U, P_U^*)$;
    end if
end for
```

It is to be stressed again that, although Algorithm 3 involves searching over all user set $U \in \{U_{\min}, \ldots, U_{\max}\}$, its complexity is extremely limited since $z_{\min}$ and z_s are to be computed *only once* before the loop over U starts. Once they are available, computing (4.74) is straightforward for any U since it requires only simple algebraic operations.

Finally, note that problem (4.60) might have more than one solution because (4.60a) depends only on the variable z and there might exist two (or more) feasible pairs, say $(U_1, P_{U_1}^*)$ and $(U_2, P_{U_2}^*)$, such that

$$z_s = \frac{P_{U_1}^*}{U_1} = \frac{P_{U_2}^*}{U_2} \,. \tag{4.75}$$

Energy Efficiency Optimization with the Non-linear Power Model

Assume the power consumption model in (4.58) holds and that again the goal is to determine the optimal transmit power P and number of users U that maximize the system EE. This problem is first formulated in the sequel, then the optimal pair (U, P) is determined.

Problem Formulation. Including the parameter η_{nl} into $P_{\text{c,nl}}$ and neglecting the resulting constant term B/η_{nl} the EE maximization problem is formulated with the non-linear power model as

$$\max_{U,P} \frac{U\mathrm{E}\big(\log_2(1+\frac{P\alpha}{U})\big)}{P^{1/n}+UP_{\text{c,nl}}} \tag{4.76a}$$

$$\text{s.t. } 0 \le P \le P_{\max} \tag{4.76b}$$

$$U \in \{U_{\min}, U_{\min}+1, \ldots, U_{\max}\} \tag{4.76c}$$

$$\mathrm{E}\left(\log_2\left(1+\frac{P\alpha}{U}\right)\right) \ge R_{\min}\ . \tag{4.76d}$$

To begin with, the optimal U for fixed P can be determined by the same method described in Section 4.2.5 since the denominator of the EE is still linear in U. Instead, the optimization of the transmit power P is more challenging because, unlike the case considered in Section 4.2.5, the power consumption model is a non-convex function of P for $n > 1$. Thus, even if the numerator of (4.76a) can be proved to be concave following the same argument as in Section 4.2.5, this is not sufficient to conclude that (4.76a) is a pseudo-concave function in P. The remainder of this section is intended to overcome this challenge.

Optimal P for Fixed U Assume U is fixed in (4.76). Then, applying the change of variable $x = P^{1/n}$ allows reformulating (4.76) as follows

$$\max_{x} \frac{\mathrm{E}\big(U\log_2\big(1+\frac{x^n\alpha}{U}\big)\big)}{x+UP_{\text{c,nl}}} \tag{4.77a}$$

$$\text{s.t. } 0 \le x \le P_{\max}^{1/n} \tag{4.77b}$$

$$\mathrm{E}\left(\log_2\left(1+\frac{x^n\alpha}{U}\right)\right) \ge R_{\min}\ . \tag{4.77c}$$

As it will be described next, this variable change makes it possible to prove a similar result for problem (4.77) as in Proposition 4.15, even though it is not possible to directly follow the same argument as in Proposition 4.15, since the numerator of (4.77a) is not concave in x, which prevents directly concluding that (4.77a) is a pseudo-concave function. Nevertheless, the pseudo-concavity of (4.77a) can still be proved as shown next.

4.21 Definition. A function $f : \mathbb{R}^+ \to \mathbb{R}^+$ is called *(strictly) S-shaped* if there exists a point x_0 such that f is (strictly) convex for $0 < x < x_0$ and (strictly) concave for $x > x_0$.

Next, the pseudo-concavity of (4.77) is proved.

4.22 Lemma. *The objective* (4.77a) *is a strictly pseudo-concave function in x.*

Proof. The first step of the proof is to show that the numerator of (4.77a) is an S-shaped function. Upon defining for convenience $h = \alpha/U$, it is shown that the second derivative of the numerator of (4.77a) has a unique zero given by

$$\begin{aligned} \mathrm{E}(h)\,(n-1) + \mathrm{E}\big(h^2\big)\,(n-1)x_0^n &= \mathrm{E}\big(h^2\big)\,n x_0^n \\ \left(\frac{(n-1)\mathrm{E}(h)}{\mathrm{E}(h^2)}\right)^{1/n} &= x_0 \\ \mathrm{E}\big(h^2\big)\,n x_0^n &\Leftrightarrow x_0\ . \end{aligned} \tag{4.78}$$

Moreover, it is easy to verify that the second derivative of the numerator of (4.77a) is strictly positive for $x < x_0$ and strictly negative for $x > x_0$ and thus Definition 4.21 is fulfilled.

At this point, the pseudo-concavity of (4.77) is shown by considering separately the intervals $x > x_0$ and $x < x_0$, proving that (4.77a) is strictly pseudo-concave in both intervals. This, in turn, implies Lemma 4.22 since (4.77a) is also differentiable in x_0.

For $x > x_0$ (4.77a) is the ratio between a strictly concave function and an affine function, and thus it is pseudo-concave.

For $0 < x < x_0$ it is proven that (4.77a) is a strictly increasing function, which implies the pseudo-concavity property. Computing the first derivative of (4.77a) shows that it is positive if

$$\underbrace{\mathrm{E}\left(\frac{nhx^n}{1+hx^n} - \log_2(1+hx^n)\right)}_{g_1(x)} + \underbrace{\mathrm{E}\left(\frac{nUP_{\mathrm{c,nl}}x^{n-1}}{1+hx^n}\right)}_{g_2(x)} > 0\ . \tag{4.79}$$

It is first noticed that g_2 is always positive for $x > 0$. As for g_1, when defining $f(x) = \mathrm{E}(\log_2(1+hx^n))$ it is observed that it can be written as

$$g_1(x) = -f(x) + x\frac{df(x)}{dx}\ . \tag{4.80}$$

At this point recall that in the range $0 < x < x_0$ the numerator of (4.77a) is convex and so will be $f(x)$. As a result, $g_1(x) \geq 0$ directly follows from the first-order condition for the convexity of a function. ■

After showing the above results, it is finally possible to employ the same arguments as in Proposition 4.15 to derive the solution of problem (4.77). Specifically, the following result holds.

4.23 Proposition. *For any given U (4.77a) has a unique stationary point x_s and there exists a unique x_{min} such that (4.77c) is fulfilled with equality. Problem (4.77a) is feasible if and only if $x_{min} \leq P_{max}^{1/n}$ and in this case it admits the unique solution*

$$x^* = \max(x_{min}, \min(P_{max}^{1/n}, x_s)) \ . \tag{4.81}$$

Proof. By virtue of Lemma 4.22, it holds that (4.77a) has a unique stationary point x_s. Moreover, it is easy to see that for any $n \geq 1$ (4.77c) is strictly increasing in $x \geq 0$ and thus there will be a unique x_min such that

$$\mathrm{E}\left(U \log_2\left(1 + \frac{x_\mathrm{min}^n \alpha}{U}\right)\right) = R_\mathrm{min} \ . \tag{4.82}$$

Then, the rest of the proof follows by a similar argument as in Proposition 4.15. ■

Thus, based on (4.81), the optimal transmit power is obtained as:

$$P^* = \left(\max\left(x_\mathrm{min}, \min\left(P_\mathrm{max}^{1/n}, x_\mathrm{s}\right)\right)\right)^n \ . \tag{4.83}$$

Joint Optimization of U and P. In order to jointly optimize the pair (U, P), an important difference with respect to the case in which a linear power model is used needs to be highlighted. Specifically, as discussed in Section 4.2.5, using a linear power model leads to problem (4.70) for which the stationary point of the objective function z_s, and the point z_min that fulfilled the QoS constraint with equality did not depend on the number of users U. This allowed us to develop Algorithm 3 in which z_s and z_min has to be computed just once. Instead, due to the more involved mathematical structure of (4.83) the quantities x_s and x_min in (4.74) depend on the number of users U and thus, they need to be computed for each value of U. This leads to the resource allocation in Algorithm 4.

Algorithm 4: Global optimization of (U, P) with power model (4.58)

```
for $U = U_{\min} : U_{\max}$ do
    Compute $x_s$ as the unique stationary point of (4.77a);
    Compute $x_{\min}$ as the unique solution of (4.82);
    Compute $P^*$ as in (4.83);
    $\mathrm{EE}_U = \mathrm{EE}(U, P^*)$;
    if $\mathrm{EE}_U > \mathrm{EE}_{U-1}$ then
        $(\widetilde{U}, \widetilde{P}) = (U, P^*)$;
    end if
end for
```

4.24 Remark. It should be stressed that, even though Algorithm 4 involves computing x_s and $x_{\min}$ for all possible values of U, the resulting computational complexity is still affordable. Indeed, the maximum number of users $U_{\max}$ that can be served typically does not reach unmanageably large numbers and, more importantly, the optimal pair (U, P) does not depend on the instantaneous channel gains but only on the channel statistics. Thus, the optimal pair (U, P) needs to be updated following the slow-fading channel variations. Finally, the exhaustive search is only over U, whereas for each $U \in \{U_{\min}, \ldots, U_{\max}\}$ the corresponding optimal power can be easily computed by means of (4.83).

Next, the performances of the proposed algorithms are unveiled in mmWave large-scale MIMO systems. First, the performances of Algorithm 3 and Algorithm 4 are studied in Figure 4.13 and Figure 4.14 under various QoS constraints for the MU MIMO DL with SM [76]. Both algorithms are developed in order to globally optimize the total transmit power and the number of users (U, P). In order to solve the optimization program the upper bound of the ergodic achievable sum-rate developed in (4.56) is considered. For the simulations 5000 independent channel realizations are generated following the geometry-based channel model with $C = 10$ scattering clusters and a bandwidth $W = 1\,\mathrm{MHz}$. The number of AAs at the BS is chosen as $N_A = 2$ with $N_T = 8$ and $N_T = 64$ transmit antennas in order to observe the effect of the number of antennas on the system performance. It is assumed that all mobile users are equipped with a single antenna. Therefore, only digital combining is employed at the receiver side while hybrid beamforming is performed at the transmitter side. In Figure 4.13, the achieved EE is reported by jointly allocating the transmit power

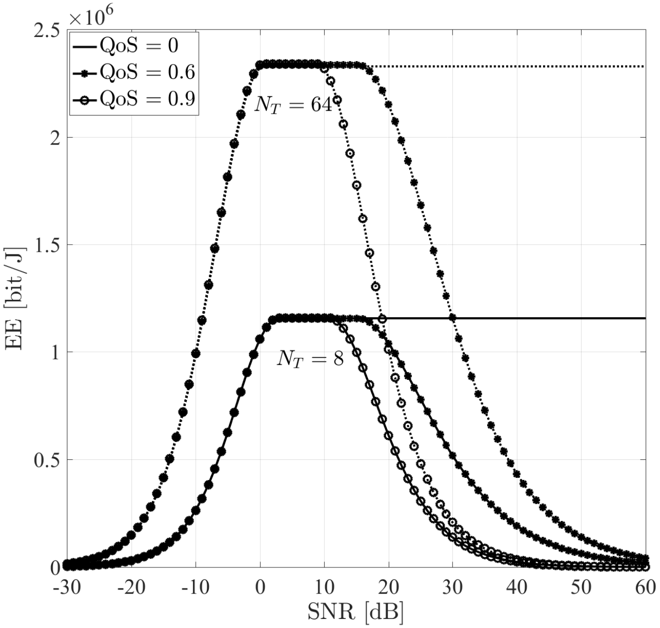


Figure 4.13: Energy efficiency maximization of the HBFSM scheme considering Algorithm 3 ($N_{\mathrm{A}} = 2$, $N_{\mathrm{R}} = 1$, $P_{\mathrm{c}} = 1$ Watt, $U_{\min} = 2$, $U_{\max} = 10$ with BPSK).

P and the number of active RF chains (users) U according to Algorithm 3 which is developed by considering a standard power consumption model. Figure 4.14 depicts the performance of the system according to Algorithm 4 considering the power consumption model that accounts also for the non-linearity of the power amplifier with $n = 2$. Moreover, various QoS requirements are enforced in order to serve the user with the minimum rate $R_{\min} = \mathrm{QoS} R_{\max}$ where $R_{\max}$ is the maximum achievable ergodic rate, which is achieved by setting $P = P_{\max}$ and $U = U_{\max}$.

Figure 4.13 and Figure 4.14 show that increasing the number of total transmit antennas N_{T} results in better performance for both algorithms. Since the wavelength of the signal is short enough to design and pack a large number of antennas in a small space it is possible to deploy large AAs in a confined space

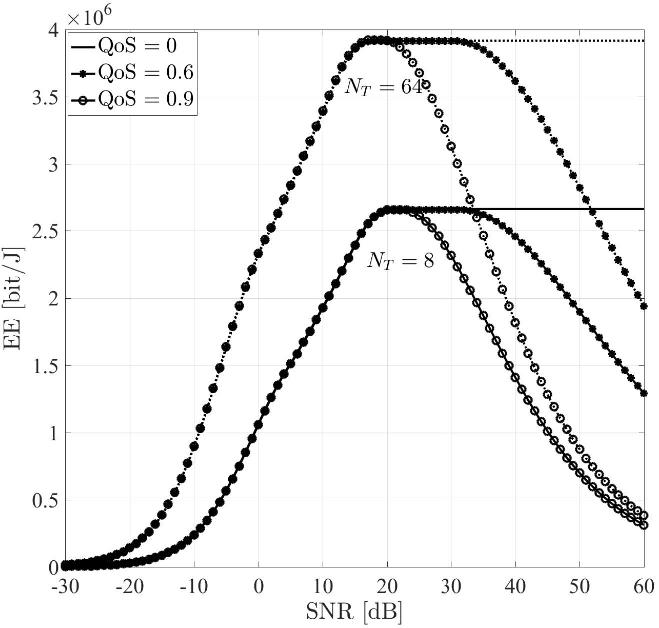


Figure 4.14: Energy efficiency maximization of the HBFSM scheme considering Algorithm 4 ($N_A = 2$, $N_R = 1$, $P_c = 1$ Watt, $U_{\min} = 2$, $U_{\max} = 10$ with BPSK).

at the BS to enjoy higher beamforming gains. Thus, the transmitted beams can be directed to the intended user more precisely, which results in less IUI and more robust communication. When there is no minimum rate requirement, EE increases up to the optimal point with (U^*, P^*) and remains constant when increasing the transmit power budget. In contrast, the numerical results verify that when the QoS constraint is enforced EE drops if the transmit power budget is further increased, due to the fact that the transmit power needs to be increased beyond the EE-maximizing power level in order to fulfill the QoS requirement. As the power budget increases $R_{\max}$ increases as well and eventually the optimum power that satisfies the rate requirement becomes the maximum power $P^* = P_{\max}$. This result also confirms that the EE is not monotonically increasing in U and P, admitting in general a maximizer $(U^* \neq U_{\max}, P^* \neq P_{\max})$.

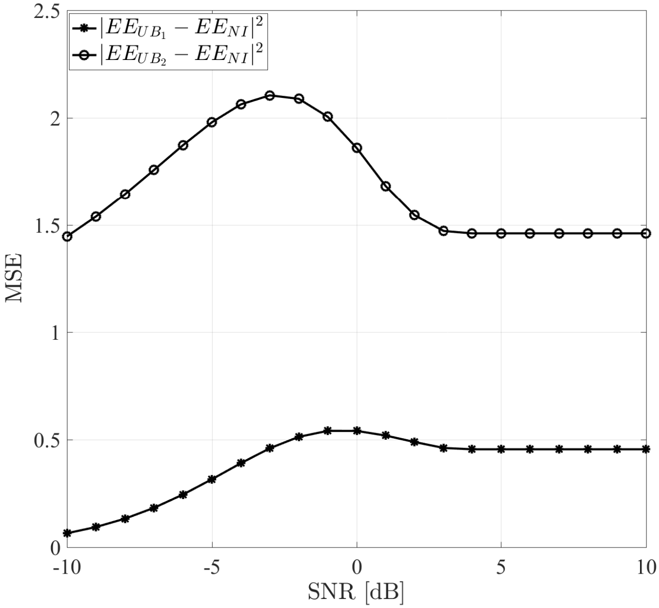


Figure 4.15: Mean-square error of the energy efficiency of the HBFSM scheme with ergodic rate upper bound.

Next, the accuracy of the proposed upper-bound of the ergodic achievable sum-rate is studied in terms of the EE value. Specifically, Figure 4.15 considers $N_\mathrm{A} = 2$, $N_\mathrm{T} = 8$, and shows the mean square error (MSE) between the EE_NI obtained using the true ergodic achievable rate value, computed by numerical integration, and the EE obtained using [76]:

- the proposed bound in (4.56), labeled as $\mathrm{EE}_{\mathrm{UB}_1}$,
- the bound proposed in [71], labeled as $\mathrm{EE}_{\mathrm{UB}_2}$, which is the only other bound available in the literature that is suitable for optimization purposes.

The results show that the newly proposed bound achieves a lower MSE, thus being more accurate than that from [71] and motivating its use for an energy-efficient system design.

Moreover, Figure 4.16 compares the performances of two schemes given in

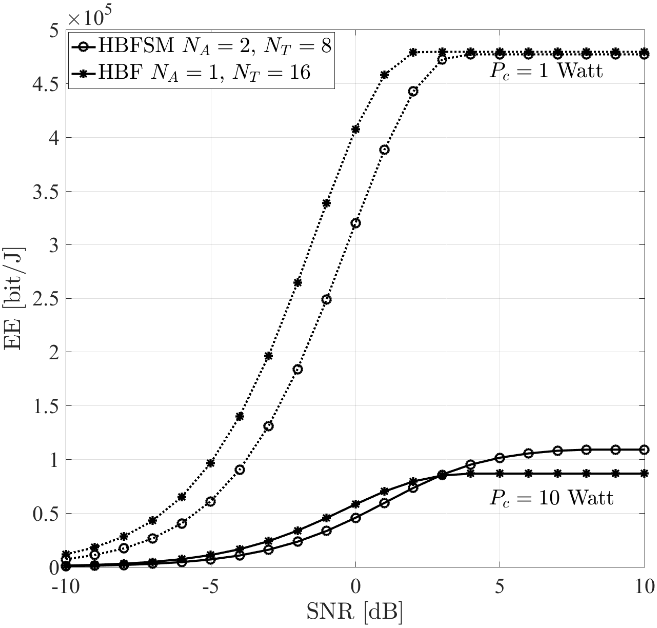


Figure 4.16: Energy efficiency comparison of the HBFSM and HBF schemes.

Section 4.2 and its counterpart without SM in Appendix C, respectively [76]. Both schemes employ hybrid beamforming that grants low power consumption compared to the fully digital systems. This is due to the fact that the number of RF chains is lower than the number of antennas. The numerical integration method is adopted in order to compare the HBF with and without SM introduced in Section 4.2 and Appendix C, respectively. The EE performance of the two schemes is compared in Figure 4.16 where the rate of a single user is computed by numerical integration over 5000 channel realizations. Throughout the simulations a BS with 16 transmit antennas is considered. As for the HBFSM case, the available transmit antennas are divided into two groups to form $N_{\mathrm{A}} = 2$ AAs in order to employ SM. Therefore, only $N_{\mathrm{T}} = 8$ transmit antennas are used to perform beamforming. On the contrary, all $N_{\mathrm{T}} = 16$ antennas are active in the HBF system to transmit M-ary modulated symbols. Although the achievable rate is only $\log_2 M$ in the HBF case the fact that the number of active antennas to perform beamforming is doubled results in better performance in the low-SNR range due to the beamforming gain when the static hardware power dissipation is $P_\mathrm{c} = 1\,\mathrm{Watt}$, as depicted in Figure 4.16. When the static hardware power dissipation is low enough the optimum number of users and the transmit power pair (U^*, P^*) is lower than the values that achieve the maximum rate. Finally, the EE of the two schemes converge to the global optimum with increased transmit power budget. In addition, the performance gap between HBF with and without SM schemes becomes smaller when the static hardware power dissipation is increased to $P_\mathrm{c} = 10\,\mathrm{Watt}$. The system model detailed in Section 4.2 features the SM scheme that allows transmitting spatially-modulated bits mapped onto the AA index. The spatially-modulated bits are transmitted inherently with the knowledge of the activated AA to transmit M-ary modulation symbol to the targeted user. When the user detects the AA index of the transmission the spatially-modulated bits are also detected. As a result, the radiated power is consumed only to transmit the M-ary modulated bits while the spatially-modulated bits are transmitted as well. Compared to HBF systems HBFSM offers a larger SE with the same transmit power consumption. Since the optimum number of users and the transmit power pair (U^*, P^*) is high enough to achieve the maximum rate $\log_2 N_\mathrm{A} + \log_2 M$ the HBFSM scheme outperforms the HBF scheme as observed in Figure 4.16. These results show, a larger static hardware power dissipation P_c results in a larger optimal transmit power P^*. Therefore, the system operates in the SNR region that results in the maximum achievable rate.

4.3 Summary

This chapter addressed the problem of designing a spectrum- and energy-efficient transceiver that operates in mmWave bands. First, an UT architecture was designed that employs analog beamforming together with SM. The system comprises two stages: a switching stage that intelligently allocates the UT antennas on the phase shifter groups to minimize the spatial correlation and a phase shifting stage that maximizes the beamforming gain to combat the path-loss. A computationally efficient optimization algorithm was developed to configure the system. The flexibility of the architecture enables optimization of the hybrid transceiver at any SNR region. Only one group of phase shifters is activated to attain a high beamforming gain at the low-SNR region, while the number of groups increases at the high-SNR region. This also increases the spatial rate since SM bits are mapped onto the AA indices. Moreover, the number of phase shifters per group decreases at a high-SNR region as optimizing the EE implies reducing the total number of phase shifters. The performance of the developed system model has been evaluated under a stochastic and a deterministic channel model. Extensive simulation results have shown that the considered stochastic channel model is accurate when the number of clusters is chosen to emulate the poor scattering nature of mmWave bands.

Furthermore, a transceiver design was developed that employs hybrid beamforming together with SM. The analog beamformer was designed to focus the transmitted beam only into the direction of the intended UT so that the transmit power per user is focused only for the communication with that user. The analog beamformer is chosen to transmit the conventionally modulated signal based on the spatially-modulated bits. Moreover, the digital precoder was designed to eliminate the IUI by exploiting ZF technique. Each user employs digital combining that matches with the activated AA and nulls out the remaining directions which enables to detect the SM bits. The system model was developed to calculate hybrid beamformers and the digital combiners at the BS. Only the digital combiners of each AA–user pair are fed back to the related UTs during the DL training phase, thus a low complexity user architecture is sufficient to achieve higher performance. Moreover, for an analog beamformer the effect of the codebook selection on the performance has been studied and corroborated that increasing the codebook size enables to achieve optimum performance. Then, an approximation for the achievable rate considering a finite input alphabet has been derived which is then utilized to optimize the EE of the developed system model. An optimization framework was developed for the EE of the system to jointly optimize the total transmit power and the number of served users. It has been corroborated with extensive simulation results that the developed optimization algorithm achieves the global optimum. Moreover, the proposed transceiver architecture provides a more energy-efficient system model compared to the HBF systems without SM.

Chapter 5

Conclusion and Outlook

5.1 Conclusion

This thesis addressed the problem of designing a spectrum- and energy-efficient system by exploiting IM techniques. The number of devices in wireless communication networks is ever-growing. Recent developments in the technology result in a heterogeneous network architecture that comprises not only mobile terminals but also IoT, autonomous cars, etc. While the network gets perpetually more crowded, the concern on the power consumption has risen significantly as well. As discussed throughout this thesis, IM techniques enable activating already existing features of the system as a resource that can convey information. This new source of information is exploited without additional transmit power dissipation.

By realizing the possible virtues, the IM technique was first exploited for the OFDM system. An MU wireless communication network that employs a OFDM-IM system was designed in which the IM technique was utilized on the subcarriers of the multicarrier system. In this respect, only the predefined number of subcarriers is activated to carry the conventionally modulated symbols where the subcarrier activation pattern conveys the part of information without power consumption. A similar procedure was employed in a GFDM system to exploit the advantages brought by IM techniques. In this case, the subcarriers that belong to a subsymbol are utilized to carry the bilateral information.

Since an MU network was considered where multiple UTs communicate with the BS at the same time and in the same frequency resources, mitigating IUI has the utmost importance. Therefore, linear MMSE precoders and combiners were designed to already pre-process the signal at the BS side to enable a low complexity UT design. The detection of the subcarrier activation pattern is crucial to be able to further decode the conventionally modulated symbols carried by the active subcarriers. Thus, a low complexity decoder was designed to detect the subcarrier activation pattern by sensing the received power, then proceeding to decode the conventionally modulated symbols on the estimated subcarriers. The performance of the developed OFDM-IM and GFDM-IM system models was corroborated by extensive simulations and established that exploiting IM techniques improves the resulting system performance. Moreover, the developed OFDM-IM and GFDM-IM system models provide flexibility that can be adjusted to the network requirements. The number of groups, the number of available subcarriers, and the number of active subcarriers can be configured in order to improve the SE, EE, PAPR, and error performance.

Next, an energy-efficient UT architecture that operates in the mmWave channel environment was developed. As it was elaborated, large-scale MIMO systems are one of the technologies that will be deployed in future networks. Especially for the systems that exploit the empty spectrum in high frequencies, terminals with many antennas will be utilized to cope with the high path-loss of mmWave channels. Since the UTs are portable devices that have space limitations and battery life constraints, EE is one of the fundamental requirements in the system design. Thus, a UT was developed that comprises a low number of bulky and high-power-consuming hardware elements, e.g., RF chain, DAC, and PA, while utilizing more simple and low-energy-consuming elements such as antennas, switches, and phase shifters. An algorithm was developed that configures the hardware architecture considering the channel statistics to maximize the EE subject to the SE of the UT. This has led to devising a transmission scheme that exploits the antennas that have the least correlated channels to form the AAs. Moreover, spatially modulated bits are mapped onto the indices of AAs that are utilized to transmit conventionally modulated bits by employing analog beamforming. Although the UT is equipped with a single RF chain that enables transmitting a single data stream, more information bits are transmitted without consuming transmit power by the SM scheme to satisfy the SE requirement as well. The performance of the developed system model was compared with the performance of a UT that exploits SM without analog beamforming. The results have shown that utilizing analog beamforming improves the system performance and reinforces the system against the disadvantages of the mmWave channels.

As it has been corroborated in the literature, fully digital signal processing achieves optimum performance. However, scaling up the number of antennas to combat the poor scattering environment of the mmWave bands prevents from employing a large number of hardware elements due to the power consumption constraints. Although BSs have more resources at their disposal, measures should be taken to keep the power consumption in the required level to satisfy the EE demands. Hence, a BS architecture that exploits hybrid beamforming techniques was developed. The designed system model allows transmitting multiple streams at the same time and in the same frequency resources of the subscribers of the network. The designed BS selects the AA to transmit the conventionally modulated symbol according to the SM part of the information bits. The design of a low complexity system model with the hybrid beamformer at the BS and the digital combiner at the UT has not been straightforward to tackle with the IUI and the detection of the SM bits. Hence, the digital signal processing was designed by taking into account the channel, interference of the users, and SM scheme by exploiting ZF methods. Moreover, analog beamformers have been designed to enable transmitting multiple streams from an AA and directing the beam to the intended user which also helps to reduce the interference. The performance improvement of the developed system model was corroborated by numerically comparing it with a system that employs only digital signal processing. Next, an optimization framework for EE maximization for the proposed system model was developed. Algorithms were designed to optimize the transmit power and the number of active RF chains (users), subject to maximum power constraints, maximum and minimum number of UTs to serve, and minimum achievable rate guarantee. The optimization was carried out considering both the standard linear power consumption model, as well as the case in which the transmit amplifier operates in the non-linear region. Both algorithms enjoy global optimality and require an affordable computational complexity. Moreover, it was shown that superior energy-efficient performance can be obtained by employing SM with hybrid beamforming.

5.2 Outlook

This thesis developed various system models to meet the EE and SE demands of future wireless networks. Each and every considered scenario brings its own challenges to the system design. Therefore, each system model was handled separately to address the challenges and to develop a low-complexity system model.

First, multicarrier systems that exploit IM technique have been designed considering low frequencies. As there is a large available spectrum in mmWave bands, there has been research on the hybrid beamforming techniques for multicarrier systems. Hence, the developed OFDM-IM and GFDM-IM systems can be further studied to combine them with hybrid beamforming techniques and to operate in mmWave bands. Since the antennas that are located close to each other suffer more from spatial correlation in mmWave bands compared to the lower frequency bands, the number of subcarriers that will be activated during the data transmission can be optimized considering the channel statistics.

In the design of the low-power-consuming UT that operates at high frequencies, a fully digital BS has been assumed. As it has been elaborated, power consumption at the BS should be taken into account as a design parameter since the number of UTs to be served is ever-increasing. Hence, the design of a transceiver for the UL transmission phase in which the BS employs hybrid beamforming to process the received signal is an interesting research direction.

Furthermore, the developed transceiver architecture for an MU network that operates in mmWave bands has been shown to improve the performance in comparison to existing works. In this system design, AAs were grouped in a linear manner without considering the channel quality. When the communication in mmWave bands is considered, it is highly possible that the antennas which are located in the close proximity result in having spatially correlated channels. As done for the UL transmission phase, a grouping algorithm can be developed to form the AAs with the antennas that have the least correlated channels among each other. It can be already foreseen that this will be a challenging task since each UT's channel is diversified. For the developed system model, the UTs were assumed to have fully digital architectures. In order to reduce the power consumption of the hardware circuitry and enjoy the array gain at the receiver side as well, UTs with hybrid architecture can be designed.

In addition, the design of a generic system model with all aforementioned features is indeed an attractive research topic. A hybrid MU transceiver that can operate in wideband mmWave channels with high performance and which exploits IM techniques both on the indices of subcarriers and AAs would be useful in various scenarios. The resulting system model could be configured according to the system requirements possibly in cooperation with SDN technology.

Appendix A

Quantization Error Derivations

In the sequel, the metric ball volume coefficient v in (4.46) is computed for the developed system parameters. Let $\mathcal{F}_{\mathrm{A}} \in G(N_{\mathrm{T}}, a)$ and $\mathcal{F}_{\mathrm{B}} \in G(N_{\mathrm{T}}, b)$ be the Grassmann manifold which lies on N_{T}-dimensional Euclidean space on a- and b-dimensional planes, respectively. Then the volume formula for a metric ball in the Grassmann manifold is [93]

$$v = \begin{cases} \frac{1}{\Gamma(a(N_{\mathrm{T}}-b)+1)} \prod_{i=1}^{N_{\mathrm{T}}-b} \frac{\Gamma(N_{\mathrm{T}}-i+1)}{\Gamma(N_{\mathrm{T}}-a-i+1)} & \text{if} \quad N_{\mathrm{T}} \leq a+b, \\ \frac{1}{\Gamma(a(N_{\mathrm{T}}-b)+1)} \prod_{i=1}^{a} \frac{\Gamma(N_{\mathrm{T}}-i+1)}{\Gamma(b-i+1)} & \text{if} \quad N_{\mathrm{T}} \geq a+b. \end{cases} \tag{A.1}$$

For the considered scenario, the codebooks $\mathcal{F}_{\mathrm{A}}$ and $\mathcal{F}_{\mathrm{B}}$ lie on 1-dimensional plane since the codewords $\mathbf{f}$ have the dimension $N_{\mathrm{T}} \times 1$. This degrades the problem in the following form

$$v = \begin{cases} \frac{1}{\Gamma((N_{\mathrm{T}}-1)+1)} \prod_{i=1}^{N_{\mathrm{T}}-1} \frac{\Gamma(N_{\mathrm{T}}-i+1)}{\Gamma(N_{\mathrm{T}}-1-i+1)} & \text{if} \quad N_{\mathrm{T}} \leq 2, \\ \frac{1}{\Gamma((N_{\mathrm{T}}-1)+1)} \prod_{i=1}^{1} \frac{\Gamma(N_{\mathrm{T}}-i+1)}{\Gamma(1-i+1)} & \text{if} \quad N_{\mathrm{T}} \geq 2. \end{cases} \tag{A.2}$$

Since the Γ function is defined as $\Gamma(N) = (N-1)!$ the two cases become the following after the manipulations

$$v = \begin{cases} \frac{1}{\Gamma(N_{\mathrm{T}})} \prod_{i=1}^{N_{\mathrm{T}}-1} (N_{\mathrm{T}} - i) & \text{if} \quad N_{\mathrm{T}} \leq 2, \\ 1 & \text{if} \quad N_{\mathrm{T}} \geq 2. \end{cases} \tag{A.3}$$

The derivation of the metric ball volume can be further manipulated since the first case of (A.3) has only $N_{\mathrm{T}} = 1$ different form the second case. In that case, the lower limit of the product operator is larger than the upper limit. There is no integer i that satisfies the condition $1 \leq i \leq N_{\mathrm{T}} - 1 = 0$, thus this is an empty product. In order to make sure that associativity laws are respected, empty products are defined to be equal to 1. Hence the metric ball volume is $v = 1$ for all values of $N_{\mathrm{T}} > 0$.

Appendix B

On the Achievable Rate of Gaussian Mixtures

B.1 The Conditional Density Function

By exploiting information theoretical methods, the conditional density function of the random variable in (4.17) can be derived as follows

$$
\begin{aligned}
& h\Big(y_k \Big| y_{k-1}, \dots, y_1, \mathbf{t}_{\mathrm{j}}^{(1:k)}\Big) \\
& \overset{(a)}{=} h\Big(y_1, \dots, y_k \Big| \mathbf{t}_{\mathrm{j}}^{(1:k)}\Big) - h\Big(y_1, \dots, y_{k-1} \Big| \mathbf{t}_{\mathrm{j}}^{(1:k)}\Big) \\
& \overset{(b)}{=} \log_2 \pi \mathrm{e} \frac{\Big| \alpha P \mathbf{t}_{\mathrm{j}}^{(1:k)} \mathbf{t}_{\mathrm{j}}^{(1:k)^H} + \mathbf{R}_{\mathrm{nn}}^{(1:k,1:k)} \Big|}{\Big| \alpha P \mathbf{t}_{\mathrm{j}}^{(1:k-1)} \mathbf{t}_{\mathrm{j}}^{(1:k-1)^H} + \mathbf{R}_{\mathrm{nn}}^{(1:k-1,1:k-1)} \Big|} \\
& \overset{(c)}{=} \log_2 \pi \mathrm{e} \frac{\Big| \mathbf{R}_{\mathrm{nn}}^{(1:k,1:k)} \Big|}{\Big| \mathbf{R}_{\mathrm{nn}}^{(1:k-1,1:k-1)} \Big|} \frac{1 + \alpha P \mathbf{t}_{\mathrm{j}}^{(1:k)T} \mathbf{R}_{\mathrm{nn}}^{(1:k,1:k)^{-1}} \mathbf{t}_{\mathrm{j}}^{(1:k)}}{1 + \alpha P \mathbf{t}_{\mathrm{j}}^{(1:k-1)^T} \mathbf{R}_{\mathrm{nn}}^{(1:k-1,1:k-1)^{-1}} \mathbf{t}_{\mathrm{j}}^{(1:k-1)}} \\
& \overset{(d)}{=} \log_2 \pi \mathrm{e} \Big(\sigma_K^2 + P_{i,K} \Big)
\end{aligned}
$$

where

$$
\sigma_K^2 = \Big| \mathbf{R}_{\mathrm{nn}}^{(1:k,1:k)} \Big| \Big/ \Big| \mathbf{R}_{\mathrm{nn}}^{(1:k-1,1:k-1)} \Big|
$$

and

$$
P_{j,K} = \sigma_K^2 \alpha P \frac{\mathbf{t}_{\mathrm{j}}^{(1:k)^T} \mathbf{R}_{\mathrm{nn}}^{(1:k,1:k)^{-1}} \mathbf{t}_{\mathrm{j}}^{(1:k)} - \mathbf{t}_{\mathrm{j}}^{(1:k-1)^T} \mathbf{R}_{\mathrm{nn}}^{(1:k-1,1:k-1)^{-1}} \mathbf{t}_{\mathrm{j}}^{(1:k-1)}}{1 + \alpha P \mathbf{t}_{\mathrm{j}}^{(1:k-1)^T} \mathbf{R}_{\mathrm{nn}}^{(1:k-1,1:k-1)^{-1}} \mathbf{t}_{\mathrm{j}}^{(1:k-1)}}
$$

Therein, (a) follows from chain rule of differential entropy and (b) from the Gaussian distribution, (c) follows from the identity $|\mathbf{A} + \mathbf{t}\mathbf{t}^{\mathrm{H}}| = |\mathbf{A}|(1 + \mathbf{t}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{t})$ and (d) follows from noise and signal power separation. Thus,

$$
f\Big(y_k \Big| y_{k-1}, \dots, y_1, \mathbf{t}_{\mathrm{j}}^{(1:k)}\Big) = \mathcal{CN}\Big(0,\, \sigma_K^2 + P_{j,K}\Big).
$$

B.2 Tight Bounds on the Differential Entropy

Tight bounds on the differential entropy of Gaussian mixtures are derived in [96]. They are applicable when the Gaussian components of the mixtures differ in their mean values only. In order to achieve tight upper and lower bounds on the differential entropy of the Gaussian mixture y_i with $L = N_\mathrm{A} M$ equiprobable components in (4.29) and hence the corresponding achievable rate in (4.48) results of [96, Theorem 1] are adopted in this section.

B.1 Theorem. *Let y_i be a Gaussian Mixture random variable with mean values $x_1 < x_2 < \cdots < x_L$, a probability sequence $p_1 = p_2 = \cdots = p_L = 1/L$, and variance 1. Define*

$$\begin{aligned}
a_l &= \frac{1}{L}\exp\Big\{-\frac{x_l^2}{2}\Big\}, && && l = 1,\ldots,L,\\
\hat{b}_l &= \frac{a_l}{a_L}, && \hat{c}_l = x_l - x_L, && l = 1,\ldots,L-1,\\
\check{b}_l &= \frac{a_{l+1}}{a_1}, && \check{c}_l = x_1 - x_{l+1}, && l = 1,\ldots,L-1,
\end{aligned}$$

$$\begin{aligned}
\hat{\eta}(u) &= \ln\bigg(1 + \sum_{l=1}^{L-1} \hat{b}_l \exp\{\hat{c}_l u\}\bigg), && u \geq 0,\\
\check{\eta}(u) &= \ln\bigg(1 + \sum_{l=1}^{L-1} \check{b}_l \exp\{\check{c}_l u\}\bigg), && u \geq 0.
\end{aligned}$$

Let $\hat{u} > 0$ be the unique root of $\hat{\eta}(u) = \ln 2$ if $\sum_{l=1}^{L-1} \hat{b}_l > 1$ and $\hat{u} = 0$ otherwise. Similarly, let $\check{u} > 0$ be the unique root of $\check{\eta}(u) = \ln 2$ if $\sum_{l=1}^{L-1} \check{b}_l > 1$ and $\check{u} = 0$ otherwise. In addition, define $\hat{\theta}_{i_1,\ldots,i_{L-1}}$, $\check{\theta}_{i_1,\ldots,i_{L-1}}$ for $i_1,\ldots,i_{L-1} \geq 0$, and v_n for $n \geq 1$ as

$$\hat{\theta}_{i_1,\ldots,i_{L-1}} = \sum_{l=1}^{L} a_l \exp\bigg\{\frac{1}{2}\bigg(x_l + \sum_{k=1}^{L-1} i_k \hat{c}_k\bigg)^2\bigg\} Q\bigg(\hat{u} - x_l - \sum_{k=1}^{L-1} i_k \hat{c}_k\bigg),$$

$$\check{\theta}_{i_1,\ldots,i_{L-1}} = \sum_{l=1}^{L} a_l \exp\bigg\{\frac{1}{2}\bigg(x_l - \sum_{k=1}^{L-1} i_k \check{c}_k\bigg)^2\bigg\} Q\bigg(\check{u} + x_l - \sum_{k=1}^{L-1} i_k \check{c}_k\bigg),$$

$$v_n = \sum_{\substack{i_1,\ldots,i_{L-1}\geq 0\\ i_1+\cdots+i_{L-1}=n}} \frac{(n-1)!}{\prod_{l=1}^{L-1} i_l!}\bigg(\hat{\theta}_{i_1,\ldots,i_{L-1}} \prod_{l=1}^{L-1} \hat{b}_l^{i_l} + \check{\theta}_{i_1,\ldots,i_{L-1}} \prod_{l=1}^{L-1} \check{b}_l^{i_l}\bigg)$$

Then for any $N, N' \geq 1$ *it holds*

$$(\gamma + \alpha_{N,N'}) \log_2(e) \leq h(y_i) \leq (\gamma + \beta_{N,N'}) \log_2(e), \qquad \text{(B.1)}$$

where

$$\gamma = \ln \frac{\sqrt{2\pi e}}{a_L} + \frac{1}{2L} \sum_{l=1}^{L} x_l^2 - \ln \frac{\hat{b}_1}{L} \sum_{l=1}^{L} Q(x_l) - \frac{x_L}{L} \sum_{l=1}^{L} x_l$$
$$+ (x_L - x_1) \sum_{l=1}^{L} \left(\frac{x_l}{L} Q(x_l) - \frac{a_l}{\sqrt{2\pi}} \right),$$

$$\alpha_{N,N'} = \sum_{n=1}^{2N-1} (-1)^n v_n - \sum_{n=0}^{N'-1} \big(\hat{\eta}(\hat{u}_n)\, \hat{\xi}_n + \check{\eta}(\check{u}_n)\, \check{\xi}_n \big),$$

$$\beta_{N,N'} = \sum_{n=1}^{2N} (-1)^n v_n - \sum_{n=0}^{N'-1} \big(\hat{\eta}(\hat{u}_{n+1})\, \hat{\xi}_n + \check{\eta}(\check{u}_{n+1})\, \check{\xi}_n \big),$$

and with $\hat{u}_n = n\hat{u}/N'$, $\check{u}_n = n\check{u}/N'$ *for all* $n \in \{0, 1, \ldots, N'\}$ *further*

$$\hat{\xi}_n = \frac{1}{L} \sum_{l=1}^{L} \big(Q(\hat{u}_n - x_l) - Q(\hat{u}_{n+1} - x_l) \big),$$

$$\check{\xi}_n = \frac{1}{L} \sum_{l=1}^{L} \big(Q(\check{u}_n + x_l) - Q(\check{u}_{n+1} + x_l) \big),$$

The gaps $h(y_i) - (\gamma + \alpha_{N,N'}) \log_2(e)$ *and* $(\gamma + \beta_{N,N'}) \log_2(e) - h(y_i)$ *in* (B.1) *are bounded from above by* $\varepsilon_{N,N'}$ *defined by*

$$\varepsilon_{N,N'} = v_{2N-1} + \sum_{l=0}^{N'-1} \big(\hat{\eta}(\hat{u}_n) - \hat{\eta}(\hat{u}_{n+1}) \big) \hat{\xi}_n + \sum_{l=0}^{N'-1} \big(\check{\eta}(\check{u}_n) - \check{\eta}(\check{u}_{n+1}) \big) \check{\xi}_n$$

The term $\varepsilon_{N,N'}$ *vanishes at least as fast as* $O(1/N) + O(1/N')$.

Proof. See [96, Theorem 1]. ■

B.2 Remark. Let y_i be a Gaussian Mixture random variable with the same mean values and probability sequence as in Theorem B.1 but with variance σ^2. Defining $\tilde{y}_i = y_i/\sigma$ leads to a Gaussian Mixture random variable with mean values $x_1/\sigma < x_2/\sigma < \cdots < x_L/\sigma$ the same probability sequence as that of y_i and variance 1. Then, one can use Theorem B.1 to bound the differential entropy $h(y_i) = h(\tilde{y}_i) + \log_2(\sigma)$.

B.3 A Bound on the Achievable Rate

Let X be a real-valued Gaussian mixture with the distribution function

$$f_X(x) = \frac{1}{L}\sum_{l=1}^{L} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(x-\mu_l)^2}{2\sigma^2} \right\},$$

where μ_l and σ^2 are the mean and variance of the l-th Gaussian random variable X_l of the mixture. The mean of X can be computed as

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \frac{1}{L}\sum_{l=1}^{L}\int_{-\infty}^{\infty} x f_{X_l}(x)\,dx = \frac{1}{L}\sum_{l=1}^{L}\mu_l$$

Further, if the second moment $\mathrm{E}\big(X^2\big)$ is known as well, then the variance of X can be computed by $\mathrm{var}(X) = \mathrm{E}\big(X^2\big) - \mathrm{E}(X)^2$. Since

$$\mathrm{E}\big(X^2\big) = \frac{1}{L}\sum_{l=1}^{L}\int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(x-\mu_l)^2}{2\sigma^2} \right\} dx = \frac{1}{L}\sum_{l=1}^{L}\left(\sigma^2 + \mu_l^2\right)$$

the variance of the Gaussian mixture X is given by

$$\mathrm{var}(X) = \sigma^2 + \frac{1}{L}\sum_{l=1}^{L}\mu_l^2 - \frac{1}{L^2}\left(\sum_{l=1}^{L}\mu_l\right)^2.$$

Now, let $Z = X + jY$ be a complex random variable of two independent Gaussian mixtures X and Y with the same number L of equiprobable Gaussian distributions each with expected values $\mu_{X,l}$ and $\mu_{Y,l}$ as well as an identical variance σ^2. Then, the mean of Z can be written as

$$\mathrm{E}(Z) = \mathrm{E}(X) + j\mathrm{E}(Y) = \frac{1}{L}\sum_{l=1}^{L}\left(\mu_{X,l} + j\mu_{Y,l}\right) = \frac{1}{L}\sum_{l=1}^{L}\mu_l$$

and the variance of Z with $\mathrm{var}(Z) = \mathrm{var}(X) + \mathrm{var}(Y)$ further

$$\mathrm{var}(Z) = 2\sigma^2 + \frac{1}{L}\sum_{l=1}^{L}|\mu_l|^2 - \frac{1}{L^2}\left(\sum_{l=1}^{L}\Re\{\mu_l\}\right)^2 - \frac{1}{L^2}\left(\sum_{l=1}^{L}\Im\{\mu_l\}\right)^2 \tag{B.2}$$

Based on (B.2) and the fact that the differential entropy of any random variable with finite variance is less than or equal to the differential entropy of a Gaussian random variable with the same variance [97], the variance of the random variable in (4.29) is given as

$$\begin{aligned}\mathrm{var}(Y_i) = \sigma^2 &+ \frac{1}{MN_{\mathrm{A}}}\sum_{m=1}^{M}\sum_{j=1}^{N_{\mathrm{A}}}\left|\sqrt{P_i}\mathbf{w}_{j,i}^{\mathrm{H}}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m\right|^2 \\ &- \frac{1}{M^2N_{\mathrm{A}}^2}\left(\sum_{m=1}^{M}\sum_{j=1}^{N_{\mathrm{A}}}\Re\{\sqrt{P_i}\mathbf{w}_{j,i}^{\mathrm{H}}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m\}\right)^2 \\ &- \frac{1}{M^2N_{\mathrm{A}}^2}\left(\sum_{m=1}^{M}\sum_{j=1}^{N_{\mathrm{A}}}\Im\{\sqrt{P_i}\mathbf{w}_{j,i}^{\mathrm{H}}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m\}\right)^2 .\end{aligned} \tag{B.3}$$

Hence, the achievable rate $\tilde{R}_i$ of user i in (4.52) is upper bounded by applying the variance (B.3) to $\log_2(\pi\mathrm{e}\,\mathrm{var}(Y_i))$ and interchanging $P_i = P/U$, i.e.

$$\begin{aligned}\bar{R}_i = \log_2 \pi\mathrm{e}\Bigg[\sigma^2 &+ \frac{P}{UMN_A}\sum_{m=1}^{M}\sum_{j=1}^{N_A}\left|\mathbf{w}_{j,i}^{\mathrm{H}}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m\right|^2 \\ &- \frac{P}{UM^2N_A^2}\left(\sum_{m=1}^{M}\sum_{j=1}^{N_A}\Re\{\mathbf{w}_{j,i}^{\mathrm{H}}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m\}\right)^2 \\ &- \frac{P}{UM^2N_A^2}\left(\sum_{m=1}^{M}\sum_{j=1}^{N_A}\Im\{\mathbf{w}_{j,i}^{\mathrm{H}}\mathbf{H}_{j,i}\mathbf{f}_{j,i}x_m\}\right)^2\Bigg] \\ &- \log_2(\pi\mathrm{e}\sigma^2).\end{aligned}$$

Appendix C

Multiuser MIMO Downlink without Spatial Modulation

This section describes a similar system model as that described in Section 4.2 , with the only difference that no SM is employed, but instead all the available transmit antennas are used in order to transmit data to U users in parallel. This specific system can be thought as one of the configurations of HBFSM where instead of transmitting the M-ary modulated symbol from different antenna arrays according to the SM bits, all the antenna arrays are activated during the transmission, therefore no SM bits are transmitted. A transmitter is considered that deploys hybrid beamforming to communicate with UTs that deploy digital combining. In this case, the received signal of the i-th user is

$$\mathbf{r}_i = \sqrt{P}\sum_{j=1}^{U}\mathbf{H}_i\mathbf{f}_j\mathbf{p}_j\mathbf{x} + \mathbf{n}_i, \tag{C.1}$$

where P is the total power transmitted by the BS, $\mathbf{H}_i \in \mathbb{C}^{N_\mathrm{R}\times N_\mathrm{A}N_\mathrm{T}}$ is the channel matrix between BS and the i-th user. Furthermore, $\mathbf{f}_j \in \mathbb{C}^{N_\mathrm{A}N_\mathrm{T}\times 1}$ and $\mathbf{p}_j \in \mathbb{C}^{1\times U}$ are the analog beamformer and the digital precoder for the i-th user. $\mathbf{x} \in \mathbb{C}^{U\times 1}$ is the data vector containing the M-ary modulated symbols of all the users, and finally $\mathbf{n}_i \in \mathbb{C}^{N_\mathrm{R}\times 1}$ is the noise vector with i.i.d. circularly symmetric complex Gaussian elements $\mathcal{CN}(0, \sigma^2)$. Similar to the Section 4.2.1, the resulting post-processed signal at each individual user is

$$\begin{aligned} y_i &= \mathbf{w}_i^\mathrm{H}\mathbf{r}_i \\ &= \sqrt{P_i}\mathbf{w}_i^\mathrm{H}\mathbf{H}_i\mathbf{f}_i x_i + \mathbf{w}_i^\mathrm{H}\mathbf{n}_i. \end{aligned} \tag{C.2}$$

In order to derive the achievable rate of the system, the mutual information between transmitted and received symbols $I(y_i; x_i) = h(y_i) - h(z_i)$ is attained where x_i in this case is the M-ary modulated symbol, z_i is the noise term after receive combining. Since the modulated symbol x_i belongs to a finite input constellation, the resulting signal y_i follows a Gaussian mixture distribution

$$f_{Y_i}(y_i) = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{\pi}\exp\left\{-\frac{|y_i - \sqrt{P_i}\mathbf{w}_i^{\mathrm{H}}\mathbf{H}_i\mathbf{f}_i x_m|^2}{\sigma^2}\right\}, \tag{C.3}$$

where x_m denotes the m-th symbol from the M-ary constellation diagram, and σ^2 is the received noise power.

Following Proposition 4.13, the achievable rate R_i of user i assuming discrete-symbol input, is upper-bounded as

$$\begin{aligned}\tilde{R}_i \le \log_2 \pi\mathrm{e}\Bigg[\sigma^2 + \frac{P}{UM}\sum_{m=1}^{M}|\mathbf{w}_i^{\mathrm{H}}\mathbf{H}_i\mathbf{f}_i x_m|^2 - \frac{P}{UM^2}\left(\sum_{m=1}^{M}\Re\{\mathbf{w}_i^{\mathrm{H}}\mathbf{H}_i\mathbf{f}_i x_m\}\right)^2 \\ - \frac{P}{UM^2}\left(\sum_{m=1}^{M}\Im\{\mathbf{w}_i^{\mathrm{H}}\mathbf{H}_i\mathbf{f}_i x_m\}\right)^2\Bigg] - \log_2 \pi\mathrm{e}\sigma^2 = \bar{R}_i\end{aligned} \tag{C.4}$$

Furthermore, the right-hand-side of Equation (C.4) can be expressed as

$$\begin{aligned}\bar{R}_i &= \log_2\left(1 + \frac{P}{U}a_i\right), \\ a_i &= \sum_{m=1}^{M}\frac{|b_{m,i}|^2}{M} - \left(\sum_{m=1}^{M}\frac{\Re\{b_{m,i}\}}{M}\right)^2 - \left(\sum_{m=1}^{M}\frac{\Im\{b_{m,i}\}}{M}\right)^2, \\ b_{m,i} &= \frac{\mathbf{w}_i^{\mathrm{H}}\mathbf{H}_i\mathbf{f}_i x_m}{\sigma}\end{aligned} \tag{C.5}$$

Finally, the ergodic achievable sum-rate is

$$R_s = \mathrm{E}\left(\sum_{i=1}^{U}\bar{R}_i\right) = UR, \tag{C.6}$$

similar to the Section 4.2.1, all $\{a_i\}_{i=1}^{U}$ are identically distributed, motivated by identical distribution of the fading channels experienced by the U users. Therefore, the ergodic achievable sum-rate can be expressed as $R_s = UR$, where R is the ergodic achievable rate of any user.

Bibliography

[1] International Telecommunication Union (ITU). IMT traffic estimates for the years 2020 to 2030. Report ITU-R M.2370-0, July 2015.

[2] 3rd Generation Partnership Project (3GPP). Procedures for the 5G system. Technical Report 23.502, Version 15.2.0, 3GPP, June 2018.

[3] NGMN 5G White Paper – A deliverable by the NGMN Alliance. https://www.ngmn.org/5g-white-paper/5g-white-paper.html, February 2015.

[4] Thomas L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, November 2010.

[5] Yong Niu, Yong Li, Depeng Jin, Li Su, and Athanasios V. Vasilakos. A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges. *Wireless Networks*, 21(8):2657–2676, November 2015.

[6] Robin Gerzaguet, Nikolaos Bartzoudis, Leonardo Gomes Baltar, Vincent Berg, Jean-Baptiste Doré, Dimitri Kténas, Oriol Font-Bach, Xavier Mestre, Miquel Payaró, Michael Färber, and Kilian Roth. The 5G candidate waveform race: A comparison of complexity and performance. *EURASIP Journal on Wireless Communications and Networking*, 2017(1):13, January 2017.

[7] TERRANOVA Consortium. Wireless terahertz system applications for networks beyond 5G. White Paper. https://ict-terranova.eu/white-papers/, March 2019.

[8] Bin Li, Zesong Fei, and Yan Zhang. UAV communications for 5G and beyond: Recent advances and future trends. *IEEE Internet of Things Journal*, 6(2):2241–2263, April 2019.

[9] Diego Kreutz, Fernando M. V. Ramos, Paulo Esteves Veríssimo, Christian Esteve Rothenberg, Siamak Azodolmolky, and Steve Uhlig. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1):14–76, January 2015.

[10] Rashid Mijumbi, Joan Serrat, Juan-Luis Gorricho, Niels Bouten, Filip De Turck, and Raouf Boutaba. Network function virtualization: State-of-the-art and research challenges. *IEEE Communications Surveys & Tutorials*, 18(1):236–262, January 2016.

[11] 3rd Generation Partnership Project (3GPP). Evolved universal terrestrial radio access (E-UTRA); radio resource control (RRC); protocol specification. Technical Report 36.331, Version 14.2.2, 3GPP, April 2017.

[12] Rami Abu-alhiga and Harald Haas. Subcarrier-index modulation OFDM. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2009.

[13] Dobroslav A. Tsonev, Sinan Sinanović, and Harald Haas. Enhanced subcarrier index modulation (SIM) OFDM. In *IEEE Global Communications Conference (GLOBECOM), Multimedia Communications Workshop – Enabling Green Wireless Multimedia Communications (MMCOM)*, December 2011.

[14] Ertuğrul Başar, Ümit Aygölü, Erdal Panayırcı, and H. Vincent Poor. Orthogonal frequency division multiplexing with index modulation. *IEEE Transactions on Signal Processing*, 61(22):5536–5549, November 2013.

[15] Yue Xiao, Shunshun Wang, Lilin Dan, Xia Lei, Ping Yang, and Wei Xiang. OFDM with interleaved subcarrier-index modulation. *IEEE Communications Letters*, 18(8):1447–1450, August 2014.

[16] Miaowen Wen, Xiang Cheng, Meng Ma, Bingli Jiao, and H. Vincent Poor. On the achievable rate of OFDM with index modulation. *IEEE Transactions on Signal Processing*, 64(8):1919–1932, April 2016.

[17] Youngwook Ko. A tight upper bound on bit error rate of joint OFDM and multi-carrier index keying. *IEEE Communications Letters*, 18(10):1763–1766, October 2014.

[18] Thien Van Luong and Youngwook Ko. The BER analysis of MRC-aided greedy detection for OFDM-IM in presence of uncertain CSI. *IEEE Wireless Communications Letters*, 7(4):566–569, August 2018.

[19] Abdulrahman Ikram Siddiq. Low complexity OFDM-IM detector by encoding all possible subcarrier activation patterns. *IEEE Communications Letters*, 20(3):446–449, March 2016.

[20] Ertuğrul Başar. Multiple-input multiple-output OFDM with index modulation. *IEEE Signal Processing Letters*, 22(12):2259–2263, December 2015.

[21] Tanumay Datta, Harsha S. Eshwaraiah, and Ananthanarayanan Chockalingam. Generalized space-and-frequency index modulation. *IEEE Transactions on Vehicular Technology*, 65(7):4911–4924, July 2016.

[22] Chakrapani Bathula, Lakshmi N. Theagarajan, and Ananthanarayanan Chockalingam. Generalized space-frequency index modulation: Low-complexity encoding and detection. In *IEEE Global Communications Conference (GLOBECOM), Technical Workshop on Emerging Technologies for 5G Wireless Cellular Networks (TW-14)*, December 2015.

[23] Huiying Zhu, Wenjin Wang, Qing Huang, and Xiqi Gao. Subcarrier index modulation OFDM for multiuser MIMO systems with iterative detection. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2016.

[24] Huiying Zhu, Wenjin Wang, Qing Huang, and Xiqi Gao. Uplink transceiver for subcarrier index modulation OFDM in massive MIMO systems with imperfect channel state information. In *International Conference on Wireless Communications and Signal Processing (WCSP)*, October 2016.

[25] Ertuğrul Başar. Index modulation techniques for 5G wireless networks. *IEEE Communications Magazine*, 54(7):168–175, July 2016.

[26] Ertuğrul Başar, Miaowen Wen, Marco Di Renzo, Raed Mesleh, Yue Xiao, and Harald Haas. Index modulation techniques for next-generation wireless networks. *IEEE Access*, 5:16693–16746, August 2017.

[27] Gerhard P. Fettweis, Marco Krondorf, and Steffen Bittner. GFDM – generalized frequency division multiplexing. In *IEEE Vehicular Technology Conference (VTC)*, April 2009.

[28] Nicola Michailow, Ivan Gaspar, Stefan Krone, Michael Lentmaier, and Gerhard P. Fettweis. Generalized frequency division multiplexing: Analysis of an alternative multi-carrier technique for next generation cellular systems. In *International Symposium on Wireless Communication Systems (ISWCS)*, August 2012.

[29] Nicola Michailow, Stefan Krone, Michael Lentmaier, and Gerhard P. Fettweis. Bit error rate performance of generalized frequency division multiplexing. In *IEEE Vehicular Technology Conference (VTC)*, September 2012.

[30] Nicola Michailow, Maximilian Matthé, Ivan S. Gaspar, Ainoa Navarro Caldevilla, Luciano L. Mendes, Andreas Festag, and Gerhard P. Fettweis. Generalized frequency division multiplexing for 5th generation cellular networks. *IEEE Transactions on Communications*, 62(9):3045–3061, September 2014.

[31] Ersin Öztürk, Ertuğrul Başar, and Hakan Ali Çırpan. Generalized frequency division multiplexing with index modulation. In *IEEE Global Communications Conference (GLOBECOM)*, December 2016.

[32] Yuan Niu and Jianping Zheng. Message passing algorithm for GFDM-IM detection. In *IEEE International Conference on Digital Signal Processing (DSP)*, November 2018.

[33] Ersin Öztürk, Ertuğrul Başar, and Hakan Ali Çırpan. Generalized frequency division multiplexing with flexible index modulation numerology. *IEEE Signal Processing Letters*, 25(10):1480–1484, October 2018.

[34] Ersin Öztürk, Ertuğrul Başar, and Hakan Ali Çırpan. Generalized frequency division multiplexing with space and frequency index modulation. In *IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, June 2017.

[35] Ersin Öztürk, Ertuğrul Başar, and Hakan Ali Çırpan. Spatial modulation GFDM: A low complexity MIMO-GFDM system for 5G wireless networks. In *IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, June 2016.

[36] Ersin Öztürk, Ertuğrul Başar, and Hakan Ali Çırpan. Generalized frequency division multiplexing with flexible index modulation. *IEEE Access*, 5:24727–24746, October 2017.

[37] Merve Yüzgeçcioğlu and Eduard A. Jorswieck. Uplink and downlink transceiver design for OFDM with index modulation in multi-user networks. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, October 2017.

[38] Merve Yüzgeçcioğlu and Eduard A. Jorswieck. Transceiver design for GFDM with index modulation in multi-user networks. In *International ITG Workshop on Smart Antennas (WSA)*, March 2018.

[39] Theodore S. Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N. Wong, Jocelyn K. Schulz, Mathew Samimi, and Felix Gutierrez Jr. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE Access*, 1:335–349, May 2013.

[40] Theodore S. Rappaport, Robert W. Heath Jr., Robert C. Daniels, and James N. Murdock. *Millimeter wave wireless communications*. Prentice Hall, September 2014.

[41] Zhouyue Pi and Farooq Khan. An introduction to millimeter-wave mobile broadband systems. *IEEE Communications Magazine*, 49(6):101–107, June 2011.

[42] Sundeep Rangan, Theodore S. Rappaport, and Elza Erkip. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proceedings of the IEEE*, 102(3):366–385, March 2014.

[43] Thomas L. Marzetta. Massive MIMO: An introduction. *Bell Labs Journal*, 20:11–22, March 2015.

[44] Fredrik Rusek, Daniel Persson, Buon Kiong Lau, Erik G. Larsson, Thomas L. Marzetta, Ove Edfors, and Fredrik Tufvesson. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1):40–60, January 2013.

[45] Thomas L. Marzetta, Erik G. Larsson, and Hong Yang. *Fundamentals of massive MIMO*. Cambridge University Press, November 2016.

[46] Erik G. Larsson, Ove Edfors, Fredrik Tufvesson, and Thomas L. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, February 2014.

[47] Luca Sanguinetti, Emil Björnson, Mérouane Debbah, and Aris L. Moustakas. Optimal linear precoding in multi-user MIMO systems: A large system analysis. In *IEEE Global Communications Conference (GLOBECOM)*, December 2014.

[48] Quentin H. Spencer, A. Lee Swindlehurst, and Martin Haardt. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *IEEE Transactions on Signal Processing*, 52(2):461–471, February 2004.

[49] Omar El Ayach, Sridhar Rajagopal, Shadi Abu-Surra, Zhouyue Pi, and Robert W. Heath Jr. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, March 2014.

[50] Zihuan Wang, Qian Liu, Ming Li, and Wolfgang Kellerer. Energy efficient analog beamformer design for mmWave multicast transmission. *IEEE Transactions on Green Communications and Networking*, 3(2):552–564, June 2019.

[51] Wonil Roh, Ji-Yun Seol, JeongHo Park, Byunghwan Lee, Jaekon Lee, Yungsoo Kim, Jaeweon Cho, Kyungwhoon Cheun, and Farshid Aryanfar. Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results. *IEEE Communications Magazine*, 52(2):106–113, February 2014.

[52] Irfan Ahmed, Hedi Khammari, Adnan Shahid, Ahmed Musa, Kwang Soon Kim, Eli De Poorter, and Ingrid Moerman. A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives. *IEEE Communications Surveys & Tutorials*, 20(4):3060–3097, June 2018.

[53] Foad Sohrabi and Wei Yu. Hybrid digital and analog beamforming design for large-scale MIMO systems. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015.

[54] Sohail Payami, Mir Ghoraishi, and Mehrdad Dianati. Hybrid beamforming for large antenna arrays with phase shifter selection. *IEEE Transactions on Wireless Communications*, 15(11):7258–7271, November 2016.

[55] Ang Li and Christos Masouros. Hybrid precoding and combining design for millimeter-wave multi-user MIMO based on SVD. In *IEEE International Conference on Communications (ICC)*, May 2017.

[56] Shiwen He, Jiaheng Wang, Yongming Huang, Björn E. Ottersten, and Wei Hong. Codebook-based hybrid precoding for millimeter wave multiuser systems. *IEEE Transactions on Signal Processing*, 65(20):5289–5304, October 2017.

[57] Ahmed Alkhateeb, Geert Leus, and Robert W. Heath Jr. Limited feedback hybrid precoding for multi-user millimeter wave systems. *IEEE Transactions on Wireless Communications*, 14(11):6481–6494, November 2015.

[58] Xiumei Yang, Mengying Zhang, Huaxia Chen, Ming-Tuo Zhou, and Yang Yang. Low-complexity hybrid precoding for energy-efficient mmWave transmission. In *International Conference on Wireless Communications and Signal Processing (WCSP)*, October 2017.

[59] R. López-Valcarce, N. González-Prelcic, C. Rusu, and Robert W. Heath Jr. Hybrid precoders and combiners for mmWave MIMO systems with per-antenna power constraints. In *IEEE Global Communications Conference (GLOBECOM)*, December 2016.

[60] Xiaoyong Wu, Danpu Liu, and Fangfang Yin. Hybrid beamforming for multi-user massive MIMO systems. *IEEE Transactions on Communications*, 66(9):3879–3891, September 2018.

[61] Duy H. N. Nguyen, Long Bao Le, Tho Le-Ngoc, and Robert W. Heath Jr. Hybrid MMSE precoding and combining designs for mmWave multiuser systems. *IEEE Access*, 5:19167–19181, September 2017.

[62] Roi Méndez-Rial, Cristian Rusu, Ahmed Alkhateeb, Nuria González-Prelcic, and Robert W. Heath Jr. Channel estimation and hybrid combining for mmWave: Phase shifters or switches? In *Information Theory and Applications Workshop (ITA)*, February 2015.

[63] Xiaohu Ge, Yang Sun, Hamid Gharavi, and John Thompson. Joint optimization of computation and communication power in multi-user massive MIMO systems. *IEEE Transactions on Wireless Communications*, 17(6):4051–4063, June 2018.

[64] Xinyu Gao, Linglong Dai, Shuangfeng Han, Chih-Lin I, and Robert W. Heath Jr. Energy-efficient hybrid analog and digital precoding for

mmWave MIMO systems with large antenna arrays. *IEEE Journal on Selected Areas in Communications*, 34(4):998–1009, April 2016.

[65] Raed Y. Mesleh, Harald Haas, Sinan Sinanović, Chang W. Ahn, and Sangboh Yun. Spatial modulation. *IEEE Transactions on Vehicular Technology*, 57(4):2228–2241, July 2008.

[66] Marco Di Renzo, Harald Haas, Ali Ghrayeb, Shinya Sugiura, and Lajos Hanzo. Spatial modulation for generalized MIMO: Challenges, opportunities and implementation. *Proceedings of the IEEE*, 102(1):56–103, January 2014.

[67] Ming-Chun Lee, Wei-Ho Chung, and Ta-Sung Lee. Generalized precoder design formulation and iterative algorithm for spatial modulation in MIMO systems with CSIT. *IEEE Transactions on Communications*, 63(4):1230–1244, April 2015.

[68] Naoki Ishikawa, Rakshith Rajashekar, Shinya Sugiura, and Lajos Hanzo. Generalized-spatial-modulation-based reduced-RF-chain millimeter-wave communications. *IEEE Transactions on Vehicular Technology*, 66(1):879–883, January 2017.

[69] Sandeep Narayanan, Marium J. Chaudhry, Athanasios Stavridis, Marco Di Renzo, Fabio Graziosi, and Harald Haas. Multi-user spatial modulation MIMO. In *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2014.

[70] Ming-Chun Lee and Wei-Ho Chun. Transmitter design for analog beamforming aided spatial modulation in millimeter wave MIMO systems. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2016.

[71] Yaping Cui, Xuming Fang, and Li Yan. Hybrid spatial modulation beamforming for mmWave railway communication systems. *IEEE Transactions on Vehicular Technology*, 65(12):9597–9606, December 2016.

[72] Merve Yüzgeçcioğlu and Eduard A. Jorswieck. Hybrid beamforming with spatial modulation in multi-user massive MIMO mmWave networks. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, October 2017.

[73] Ahmed Raafat, Merve Yüzgeçcioğlu, M. Zahid Aslam, Adrian Agustin, Josep Vidal, Eduard A. Jorswieck, and Yoann Corre. Energy efficient transmit-receive spatial modulation for uplink-downlink large-scale MIMO systems. In *IEEE Global Communications Conference (GLOBECOM)*, December 2018.

[74] Ahmed Raafat, Merve Sefunç, Adrian Agustin, Josep Vidal, Eduard A. Jorswieck, and Yoann Corre. Energy efficient transmit-receive hybrid spatial modulation for large-scale MIMO systems. *IEEE Transactions on Communications*, December 2019.

[75] Merve Yüzgeçcioğlu, Alessio Zappone, and Eduard A. Jorswieck. Energy efficiency in hybrid beamforming large-scale mmWave multiuser MIMO with spatial modulation. In *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018.

[76] Merve Sefunç, Alessio Zappone, and Eduard A. Jorswieck. Energy efficiency of mmWave MIMO systems with spatial modulation and hybrid beamforming. *IEEE Transactions on Green Communications and Networking*, October 2019.

[77] Foad Sohrabi and Wei Yu. Hybrid analog and digital beamforming for OFDM-based large-scale MIMO systems. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2016.

[78] Foad Sohrabi and Wei Yu. Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays. *IEEE Journal on Selected Areas in Communications*, 35(7):1432–1443, July 2017.

[79] Andreas F. Molisch, Vishnu V. Ratnam, Shengqian Han, Zheda Li, Sinh L. H. Nguyen, Linsheng Li, and Katsuyuki Haneda. Hybrid beamforming for massive MIMO: A survey. *IEEE Communications Magazine*, 55(9):134–141, September 2017.

[80] Marco Di Renzo, Harold Haas, and Peter M. Grant. Spatial modulation for multiple-antenna wireless systems: A survey. *IEEE Communications Magazine*, 49(12):182–191, December 2011.

[81] Mustafa Riza Akdeniz, Yuanpeng Liu, Mathew K. Samimi, Shu Sun, Sundeep Rangan, Theodore S. Rappaport, and Elza Erkip. Millimeter wave

channel modeling and cellular capacity evaluation. *IEEE Journal on Selected Areas in Communications*, 32(6):1164–1179, June 2014.

[82] Adel A. M. Saleh and Reinaldo A. Valenzuela. A statistical model for indoor multipath propogation. *IEEE Journal on Selected Areas in Communications*, 5(2):128–137, February 1987.

[83] Nafiseh Shariati, Emil Björnson, Mats Bengtsson, and Mérouane Debbah. Low-complexity channel estimation in large-scale MIMO using polynomial expansion. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2013.

[84] Ahmed Alkhateeb, Omar El Ayach, Geert Leus, and Robert W. Heath Jr. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831–846, October 2014.

[85] Nati Dinur and Dov Wulich. Peak-to-average power ratio in high-order OFDM. *IEEE Transactions on Communications*, 49(6):1063–1072, June 2001.

[86] Alok Aggarwal and Teresa H. Meng. Minimizing the peak-to-average power ratio of OFDM signals using convex optimization. *IEEE Transactions on Signal Processing*, 54(8):3099–3110, August 2006.

[87] Yoann Corre and Yves Lostanlen. Three-dimensional urban EM wave propagation model for radio network planning and optimization over large areas. *IEEE Transactions on Vehicular Technology*, 58(7):3112–3123, September 2009.

[88] Sherif Shakib, Hyun-Chul Park, Jeremy Dunworth, Vladimir Aparin, and Kamran Entesari. A highly efficient and linear power amplifier for 28-GHz 5G phased array radios in 28-nm CMOS. *IEEE Journal of Solid-State Circuit*, 51(12):3020–3036, December 2016.

[89] Quan Zhou and Huaiyu Dai. Joint antenna selection and link adaptation for MIMO systems. *IEEE Transactions on Vehicular Technology*, 55(1):243–255, January 2006.

[90] Qi Zhang, Shi Jin, Matthew McKay, David Morales-Jimenez, and Hongbo Zhu. Power allocation schemes for multicell massive MIMO systems. *IEEE*

Transactions on Wireless Communications, 14(11):5941–5955, November 2015.

[91] Omar El Ayach, Robert W. Heath Jr., Shadi Abu-Surra, Sridhar Rajagopal, and Zhouyue Pi. The capacity optimality of beamsteering in large millimeter wave MIMO systems. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2012.

[92] Jung-Chieh Chen. Hybrid beamforming with discrete phase shifters for millimeter-wave massive MIMO systems. *IEEE Transactions on Vehicular Technology*, 66(8):7604–7608, August 2017.

[93] Wei Dai, Youjian Liu, and Brian Rider. Quantization bounds on Grassmann manifolds and applications to MIMO communications. *IEEE Transactions on Information Theory*, 54(3):1108–1123, March 2008.

[94] Rajesh T. Krishnamachari and Mahesh K. Varanasi. On the geometry and quantization of manifolds of positive semi-definite matrices. *IEEE Transactions on Signal Processing*, 61(18):4587–4599, September 2013.

[95] Pan Cao, Alessio Zappone, and Eduard A. Jorswieck. Grouping-based interference alignment with IA-cell assignment in multi-cell MIMO MAC under limited feedback. *IEEE Transactions on Signal Processing*, 64(5):1336–1351, March 2016.

[96] Kamyar Moshksar and Amir K. Khandani. Arbitrarily tight bounds on differential entropy of Gaussian mixtures. *IEEE Transactions on Information Theory*, 62(6):3340–3354, June 2016.

[97] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, July 2006.

[98] Daniel Persson, Thomas Eriksson, and Erik G. Larsson. Amplifier-aware multiple-input multiple-output power allocation. *IEEE Communications Letters*, 17(6):1112–1115, June 2013.

[99] M. M. Aftab Hossain, Konstantinos Koufos, and Riku Jäntti. Minimum-energy power and rate control for fair scheduling in the cellular downlink under flow level delay constraint. *IEEE Transactions on Wireless Communications*, 12(7):3253–3263, July 2013.

[100] Qimei Cui, Tianpeng Yuan, and Wei Ni. Energy-efficient two-way relaying under non-ideal power amplifiers. *IEEE Transactions on Vehicular Technology*, 66(2):1257–1270, February 2017.

[101] Alessio Zappone and Eduard A. Jorswieck. Energy efficiency in wireless networks via fractional programming theory. *Foundations and Trends® in Communications and Information Theory*, 11(3-4):185–396, June 2015.

[102] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, March 2004.