Hybrid Precoding Algorithms for Millimeter-Wave Massive MIMO Systems

A Thesis Submitted to The University of Kent

For The Degree of Doctor of Philosophy

In Electronic Engineering

By

Osama Alluhaibi

March, 2018

Supervisor

Professor Jiangzhou Wang

Dedication

To the heroes in the Iraqi army and the popular mobilization forces who have sacrificed themselves to save our lives from terrorism.

Acknowledgements

First of all, I thank Allah for all his blessings, and for giving me strength and ability to complete this work. I would like to express my deepest appreciation to my supervisor, Professor Jiangzhou Wang, for his constant inspiration, support, and guidance during my Ph.D program at the University of Kent. I have benefited tremendously from his enthusiasm and constructive criticism. I am indebted to his for carefully reading my manuscripts and patiently working with me on my writing. I would not have been able to complete this thesis without the benefit of his vast knowledge, experience, and valuable comments. I would also like to thank my co-supervisor, Dr. Huiling Zhu, for her help advice, insightful questions and constructive feedback on my research. Through our conversations, her deep and broad knowledge of life has been a great source of inspiration to me. I also would like to thank Dr. Qasim Z. Ahmed from university of Huddersfield, for his suggestions, insightful questions, constructive feedback and kind help during this research work. I would like to thank all my past and current colleagues: Hassan Osman, Ali J. Mahbas, Assad Daghal, Manish Nair, Wael Ali, Ramiz Raoof, Yan Kai, Hind Albasry, R. Husbands, A. Khadka and Estela, for the wonderful time at the Communication Research Group. I have enjoyed many discussions with them on various technical and non-technical topics. I gratefully acknowledge the Iraqi government, Ministry of Higher Education and Scientist Research and Kirkuk University for their financial support. Without it, I would not have been able to study in United Kingdom and complete my Ph.D program. Finally, my deepest gratitude and love belongs to all of my family members, especially my Parents, for their tremendous understanding, support, care, encouragement, and love during this long-term study. Without them, this thesis would not have been completed.

List of Publications and Submissions

The research carried out during this PhD resulted in the following publications:

A. Journal Papers

1. Osama Alluhaibi, Qasim Zeeshan Ahmed, Cunhua Pan, Junyuan Wang, Huiling Zhu, and Jiangzhou Wang, "Hybrid Precoding Algorithms for Millimetre Wave Large Antenna Array Systems",
(Accepted with major revision in IEEE Transactions on Wireless Communications, 2017.)

2. Osama Alluhaibi, Junyuan Wang, Cunhua Pan, Huiling Zhu, and Jiangzhou Wang, "Impact of Overlapped AoAs on the Achievable Uplink Rate of Hybrid Beamforming for Massive MIMO mm-Wave Systems",
(Under revision to be submitted in IEEE Transactions on Vehicular Technology Correspondence, 2018.)

B. Letter Papers

1. Osama Alluhaibi, Qasim Zeeshan Ahmed, and Jiangzhou Wang, "Bisection and Power Iteration Algorithms Design Energy-Efficient Hybrid Precoding for mm-Wave Systems.",
(Under revision to be submitted in IEEE Wireless Communication Letter, 2018.)

C. Conference Papers

1. Osama Alluhaibi and Q. Z. Ahmed and J. Wang and H. Zhu "Hybrid digital-to-analog precoding design for mm-wave systems" ,
(Accepted in IEEE International Conference on Communications (ICC) 2017,

no. 7996401, pp. 1-6, May 2017)

2. Osama Alluhaibi, Qasim Ahmed, Cunhua Pan, and H. Zhu "Hybrid Digital-to-Analog Beamforming Approaches to Maximise the Capacity of mm-Wave Systems",
(Accepted in IEEE Vehicular Technology Conference (VTC) 2017 spring, no. 8108385, pp. 1-5, June 2017)

3. Osama Alluhaibi, Qasim Zeeshan Ahmed, Cunhua Pan, and H. Zhu "Capacity Maximisation for Hybrid Digital -to-Analog Beamforming mm-Wave Systems",
(Accepted in IEEE Globecom 2016, no. 7841649, pp. 1-6, Dec. 2016)

4. Osama Alluhaibi, and Qasim Zeeshan Ahmed, "Multi-User Hybrid Precoding and Decoding Design for mm-Wave Large Antenna Systems",
(Accepted in IEEE Vehicular Technology Conference (VTC) 2018 spring)

D. Not Included Papers

The following publication by the author is not included in the thesis because it is not fit within the main scope of this thesis:

1. Estela Carmona Cejudo and H. Zhu and Osama Alluhaibi "On the Power Allocation and Constellation Selection in Downlink NOMA" ,
(Accepted in IEEE Vehicular Technology Conference (VTC) fall, Sep. 2017))

Abstract

The large available spectrum efficiency and wider bandwidth at millimeter wave (mm-Wave) frequencies can enable the gigabit-per-second data rates needed for next generation wireless systems. To compensate for the high propagation loss at mm-Wave bands, multiple-input multiple-output (MIMO) with a large number of antennas are usually employed to enable beamforming. Therefore, a combination of the large number of antennas which can be called massive MIMO technology with mm-Wave bands are considered as one solution for substantially increasing the data rate for future wireless communication systems. The theoretical benefits of large antenna array systems are based on the fact that the number of radio frequency (RF) chains is equivalent to the number of antennas in the conventional wireless communication system which is known as the fully digital system. Nevertheless, implementing a large number of RF chains can be problematic since it increases the system cost, power consumption, high complexity and lowers power efficiency. The overall objective of this thesis is to provide simple and effective hybrid D/A precoding and combing for mm-Wave large antenna array systems.

Firstly, hybrid D/A precoding with a small number of RF chains is being considered for mm-Wave large antenna array systems. Currently, two types of antenna structures, fully-connected antenna array and partially-connected antenna array structures are adopted in the literature. Considering that each antenna array structure has its own practical advantage, in this thesis, by addressing both structures hybrid D/A precoding algorithms are proposed with target of maximizing the system's spectral efficiency with low computational complexity. For a fully-connected antenna array structure, the precoding design is formulated as an optimization problem to minimize the Euclidean distance between the hybrid D/A and the fully digital system. For a partially-connected antenna array structures, the hybrid D/A precoding is formulated as a joint D/A optimization to maximize the spectral efficiency of the system. This work further develops hybrid D/A precoding designs for mm-Wave multi-user systems based on maximizing the sum rate of the system directly. It will be shown that the proposed algorithms outperform the existing hybrid D/A precoding algorithms for the two types of structures, in terms of the spectral efficiency.

Secondly, energy efficient and low complexity hybrid D/A system for mm-Wave large antenna array systems is proposed to reduce the power consumption at the system. The energy efficiency criteria is formulated as fractional programming maximization problem. The target is to find the optimal number of RF chains as the RF chains consume a high energy at the system. Therefore, the effective optimal number of RF chains of the system is found by proposing a simple search algorithm. Then, two methods are proposed for designing low complexity analog and digital precoders and combiners. The presented solutions for the hybrid D/A system are shown to be effective, as these approaches can achieve high energy efficiency, and low computational complexity as compared to the existing algorithms in hybrid D/A paradigms. Ultimately, the proposed D/A precoders and combiners based on the fully-connected antenna array attain an asymptotically optimal achievable spectral efficiency to that of the fully digital system.

Thirdly, uplink multi-user hybrid D/A precoding and combining design for mm-Wave large antenna array systems is investigated. The intrinsic focus of this work is to reduce the interference of the system in the analog and digital precoders and combiners. Considering the possibility that uplink transmissions from different users can go through the paths sharing the same physical scatters, some transmission paths of different users may have overlapped angle of arrivals (AoAs) at the base station. Under this circumstance, the correlation between the channel vectors also increases highly, which affects the achievable uplink rate severely. The underlying concentrate of this work is to reduce the interference caused by users sharing the same scatterers during simultaneous uplink transmission. Therefore, in this thesis, by taking account the channel correlation between users sharing the overlapped AoAs, the genuine focus is on substantially maximizing the desired signal of a piratical user while reducing the system interference. Furthermore, the channel estimation is investigated by designing a two-step procedure. The strongest power point in each scattering point is detected and the accuracy is improved by employing an angler domain scheme. Extensive simulations demonstrate that the achievable uplink rate of our proposed algorithms surpasses the achievable uplink rate of the existing algorithms in hybrid D/A paradigms in the practical scenarios.

Contents

List of Figures

List of Tables

Abbreviations

2D . . . . Two dimensional

3D . . . . Three dimensional

4G . . . . Forth Generation

5G . . . . Fifth Generation

A . . . . Azmiuth

ADC . . . . Analog-to-Digital converter

AF . . . . Array factor

AoA . . . . Angle of Arrival

AoD . . . . Angle of Departure

APSO . . . . Alternating minimization partial swarm optimization

ASO . . . . Alternating minimization Stiefel manifold

BB . . . . Barzilai-Borwein

BER . . . . Bit-Error-Rate

BF . . . . Beamforming

BI . . . . Bisection algorithm

BS . . . . Base station

CMO .... Complex circle manifold optimization

CSI .... Channel state information

D-A .... Digital-to-analog precoding

DAC .... Digital-to-Analog conversion

EE .... Energy Efficiency

EVD .... Eigenvalue decomposition

FDD .... Frequency division duplex

GHP .... Greedy hybrid precoding algorithm

GS .... Gram-Schmidt algorithm

HBT .... Hierarchical beamforming training algorithm

HBT .... Independent and identically distributed

ICT .... Information and communication technology

LS .... Least square solution

LTE .... Long term evolution

MIMO ...... Multi-input multi-output

MISO .... Multiple-input single-output

MMSE ...... Minimum mean-square error

mm-Wave ... Millimeter Wave

MO-ALtMin Alternating minimization based manifold

MRC .... Maximum ratio combining

MSE .... Mean Squared Error

OMP .... Orthogonal matching pursuit

PE-ALtMin . Alternating minimization procrustes solution

PI .... Power Iteration algorithm

PSO .... Partial swarm optimization

PSs .... Phase shifters

RF .... Radio frequency

ROSt .... Riemannian optimisation method based on Stiefel manifold

Rx .... Receive antenna

SMW .... Sherman-Morrison-Woodbury

SINR .... Signal to Interference plus Noise Ratio

SNR .... Signal-to-noise ratio

SPSO .... Partial swarm optimization

SVD .... Singular value decomposition

TDD .... Time division duplex

Tx .... Transmit antenna

UPA .... Uniform planar antenna

WSMSE .... Weighted sum mean square error

Z .... Zenith

ZF .... Zero-forcing

Mathematical Notation

$\boldsymbol{X}$ .... A bold uppercase lette denote as matrices.

$\boldsymbol{x}$ .... A bold uppercase lette denote as vectors.

$\boldsymbol{X}^*$ .... Conjugate of matrix $\boldsymbol{X}$

$\nabla f(\boldsymbol{X})$ .... Denotes the classical gradient of f seen as a function in an Euclidean space.

$\mathbf{det}(\boldsymbol{X})$ .... Determinant of matrix $\boldsymbol{X}$.

$\mathbf{blkdiag}(\cdot)$ ... Diagonal matrix.

$\mathbb{E}(\cdot)$ .... Expectation operator.

$||\cdot||_F$ .... Frobenius norm of a matrix.

$\mathbf{arg}(\boldsymbol{X})$ .... Generates a matrix containing the phases of the entries.

$\odot$ .... Hadamard product

$\boldsymbol{X}^H$ .... Hermitian of matrix $\boldsymbol{X}$

$\boldsymbol{I}_N$ .... Identity matrix.

$\boldsymbol{X}^{-1}$ .... Inverse of matrix $\boldsymbol{X}$

$\mathbf{vec}(\cdot)$ .... Maps the $P \times X$ matrix to a $PX \times 1$ vector.

$||\cdot||_1$ .... Norm-1 of a vector.

$||\cdot||_2$ .... Norm-2 of a vector.

$\mathbb{R}\{\cdot\}$ .... Real part of the input value.

$\mathbf{Tr}(\boldsymbol{X})$ .... Trace of matrix $\boldsymbol{X}$.

$\boldsymbol{X}^T$ .... Transpose of matrix $\boldsymbol{X}$

Chapter 1

Introduction

Contents

According to the latest Cisco Visual Networking Index [1], global mobile data traffic will increase sevenfold between 2016 and 2021. The high penetration rate of mobile devices and new applications, such as virtual reality, autonomous vehicles, and telemedicine, require high data rates and large capacity in cellular communication systems. To accommodate such a large amount in mobile traffic, the fifth-generation (5G) of the mobile communication system has been investigated by other researcher. One of the promising solutions for substantially improving the spectrum utilization efficiency in 5G is to deploy a large number of antennas which is practically called massive multi-input multi-output (MIMO) systems [2–8] with millimetre-wave (mm-Wave) bands (frequency ranges from 30 GHz to 300 GHz) that has a large available bandwidth [9–15]. The extreme short wavelengths (i.e., less than 10 mm) of mm-Wave signals allow a large number of antennas to be packed in a small physical space. Despite the benefits of combing the large number of antennas (massive MIMO) system with the mm-Wave communication system, the implementation of such systems is expensive, also requires high computational complexity and high power consumption. Notoriously, this thesis aims to investigate some of the key challenges of precoding and combining in mm-Wave large antenna array systems and propose low cost and simple solutions towards the implementation of such systems.

1.1 Scope and Objectives

In general, the hardware architecture of beamformers mm-Wave large antenna array systems can be divided into three categories as fully digital, analog and hybrid digital-to-analog (D/A) precoding and combining systems. In conventional fully digital wireless systems, the number of antennas is equal to the number of radio frequency (RF) chains, so that the maximum performance gain is achieved by using the fully digital beamformer scheme. However, if the fully digital system is adopted in the mm-Wave large antenna array systems, the associated large number of RF chains will consume extremely high power and the hardware complexity will increase dramatically [10,14]. A simple approach is to employ a pure analog precoder system which uses only one RF chain to support a large number of antennas. The analog precoder controls the phase of each antenna and forms beams through low-cost phase shifters [16]. However, the achievable spectral efficiency of the analog precoding and combining system is much lower than that of the fully digital system [10]. This is because, the spatial multiplexing gains are not achievable in the analog precoding and combining system [9, 10, 17–23]. Therefore, a hybrid D/A precoding and combining system which comprises a small number of RF chains and a large number of antennas is proposed to balance the achievable spectral efficiency, power consumption, and the implementation complexity [10, 17–23].

The architecture of the hybrid D/A system is designed based on a small number of RF chains which is known as the digital part. These RF chains are connected to the large antenna array through a network of phase shifters which is called the analog part. Thereupon, the design of hybrid D/A precoding and combining is formulated as an optimization problem with a constant modulus constraint (i.e., all the non-zero elements in the analog precoder and combiner have the same amplitudes) for all the phase shifters. This constant modulus is known as a non-convex constraint which makes the optimization problem difficult to solve. As it will be discussed in the next chapters, to simplify this constraint, researchers in academia and industry proposed to impose extra constraints on the system to facilitate the design of analog precoders and combiners. Due to the complex nature of the problem, such kind of approaches may result in system capacity and spectral efficiency loss. Furthermore, the implementation complexity of algorithms requires to be considered, especially when the number of antennas is large. Since hybrid D/A mm-Wave large antenna array systems demand a large number of phase shifters, the power consumption of the system remains high. Consequently, the energy efficiency of hybrid D/A mm-Wave large antenna array systems needs to be optimized.

To extend to the uplink multi-user case in order to improve the maximum achievable rate of mm-Wave large antenna array systems, various hybrid D/A precoding and combining schemes have been proposed [24–27]. Furthermore, all the existing algorithms in [24–27] have assumed that only the digital part based on zero-forcing (ZF) or the minimum mean-square error (MMSE) can easily cancel all the interference in the hybrid D/A mm-Wave large antenna array system. Due to the digital domain restriction to fewer RF chains, such approaches [24–27] cannot cancel all the system interference which severely limits the achievable uplink rate in the current hybrid D/A mm-Wave large antenna array systems. Hence, it is of paramount importance that the D/A precoders and combiners should be designed carefully to mitigate all the system interference in both the analog and digital precoders and combiners.

Motivated by the aforementioned discussions and challenges, this thesis investigates hybrid D/A precoders and combiners problems in mm-Wave large antenna array systems. The main objectives of this research are:

1. Designing low complexity algorithms and asymptotically optimal hybrid D/A precoding and combining for the single user and multi-user mm-Wave large antenna array system.

2. Constructing an energy-efficient hybrid D/A mm-Wave large antenna array system. The effect of the resolution of the phase shifters on the energy efficiency is also investigated in this process.

3. Designing an optimal analog and digital precoders and combiners for uplink multi-

user hybrid D/A mm-Wave large antenna systems that help in maximizing the effective gain for a particular user and concurrently reducing the system interference in both the analog and digital precoders and combiners.

4. Investigating the estimation channel by designing a two-step multi-user scheme for hybrid D/A mm-Wave large antenna array systems. The aim of this scheme is to attain lower error estimated channels as compared to the recent existing methods.

1.2 Overview of Contributions

Implementing mm-Wave large antenna array systems in practice demands addressing the key challenges discussed in Section 1.1. The problems tackled in this thesis lie on the intersection of these challenges. The primary contributions of this thesis are summarized as follows:

1. We propose efficient low complexity algorithms for hybrid D/A precoding and combining mm-Wave large antenna array systems. Two types of hybrid D/A architectures namely, fully-connected antenna array and partially-connected antenna array structures, have been adopted in the literature. Considering the fact that each structure has its own practical advantages, in this thesis, novel hybrid D/A precoding and combining algorithms that can achieve high system capacity with low computational complexity are proposed for both structures. For the fully-connected antenna array structure, the precoding and combining design is formulated as an optimization problem to minimize the Euclidean distance between the hybrid D/A and fully digital system. For the partially-connected antenna array structure, hybrid D/A precoding and combining is formulated as a joint D/A optimization problem to maximize the spectral efficiency of the system.

   Publications related to the contribution number.1:

   A. Journal Papers

   1. Osama Alluhaibi, Qasim Zeeshan Ahmed, Cunhua Pan, Junyuan Wang, Huiling Zhu, and Jiangzhou Wang, "Hybrid Precoding Algorithms for Millimetre Wave Large Antenna Array Systems", **(Accepted in IEEE Transactions on Wireless**

Communications with major revision, 2017.)

B. Conference Papers

1. Osama Alluhaibi and Q. Z. Ahmed and J. Wang and H. Zhu ”Hybrid digital-to-analog precoding design for mm-wave systems” ,
(Accepted in IEEE International Conference on Communications (ICC) 2017, no.7996401, pp.1-6, May 2017)

2. Osama Alluhaibi, Qasim Ahmed, Cunhua Pan, ”Hybrid Digital-to-Analog Beamforming Approaches to Maximise the Capacity of mm-Wave Systems”,
(Accepted in IEEE Vehicular Technology Conference (VTC) 2017 spring, no. 8108385, pp.1-5, June 2017)

3. Osama Alluhaibi, Qasim Zeeshan Ahmed, Cunhua Pan, et al, ”Capacity Maximisation for Hybrid Digital -to-Analog Beamforming mm-Wave Systems”,
(Accepted in IEEE Globecom 2016, no.7841649, pp.1-6, Dec. 2016)

2. One of the main points of proposing the hybrid D/A precoding and combining for the mm-Wave large antenna array system is to reduce the power consumption of the conventional fully digital large antenna array system. However, the RF components in the hybrid D/A system consume up to 70% of the total transceiver power. Therefore, in this thesis an energy efficient hybrid D/A precoding and combining mm-Wave large antenna array system is designed through optimizing the number of RF chains and the number of data streams while optimizing the D/A precoder and combiner matrices. To achieve the target of this optimization problem, we propose simple methods for initially, finding the optimal number of RF chains and then, designing the low complexity D/A precoder and combiner matrices based on fully-connected antenna array structure. Further, we analyse the proposed hybrid precoding and combining strategy under a finite number of quantized bits in the analog precoding and combining to show the impacts of limited number of bits on the energy efficiency and spectral efficiency.

Publications related to the contribution number.2:

B. Letter Paper

1. Osama Alluhaibi, Qasim Zeeshan Ahmed, and Jiangzhou Wang, ”Bisection and

Power Iteration Algorithms Design Energy-Efficient Hybrid Precoding for mm-Wave Systems.”,

(Under revision to be submitted in IEEE Wireless Communication Letter, 2018.)

3. We propose novel precoders and combiners for uplink multi-user hybrid D/A beamforming mm-Wave large antenna array systems, where the number of RF chains is much less than the number of antennas equipped at each transceiver. It has been demonstrated that the rough surfaces of buildings and small building rifts can cause diffused scattering in mm-Wave channels (short distance propagation), which makes the propagating wave attenuate significantly as the wavelength shrinks. In this environment, some propagation uplink paths of multiple users are transmitted via physical common scatters regime, some of these paths may arrive with overlapped angle of arrivals (AoAs) at the base station. In this scenario, the correlation between the channel vectors also increases highly which affects the achievable uplink rate severely. This limits the achievable uplink rate in the current hybrid D/A mm-Wave systems. Therefore, it is of paramount importance that the hybrid D/A precoders and combiners should be designed precisely to mitigate the interference in both analog and digital parts.

   Publications related to the contribution number.3:

   A. Journal Papers

   1. Osama Alluhaibi, Junyuan Wang, Cunhua Pan, Huiling Zhu, and Jiangzhou Wang, ”Impact of Overlapped AoAs on the Achievable Uplink Rate of Hybrid Beamforming for Massive MIMO mm-Wave Systems”,

   (Under revision to be submitted in IEEE Transactions on Vehicular Technology Correspondence, 2018.)

   B. Conference Papers

   1. Osama Alluhaibi, and Qasim Zeeshan Ahmed, ”Multi-User Hybrid Precoding and Decoding Design for mm-Wave Large Antenna Systems”,

   (Accepted in IEEE Vehicular Technology Conference (VTC) 2018 spring)

4. We propose an efficient channel estimation algorithm for mm-Wave large antenna array systems with hybrid D/A architectures. Leveraging the mm-Wave channel characteristics, we develop a two-steps scheme that is based on firstly detecting the strongest power point among multi-power points in the scattering scenarios. Secondly, the accuracy of the estimation process is considered by proposing a search method based on searching for the strongest points for all the antenna array. These two-steps allow our procedure to avoid the drawback of lacking the accuracy of estimating the mm-Wave channels.

   Publications related to the contribution number. 4:

   A. Journal Papers

   1. Osama Alluhaibi, Junyuan Wang, Cunhua Pan, Huiling Zhu, and Jiangzhou Wang, "Impact of Overlapped AoAs on the Achievable Uplink Rate of Hybrid Beamforming for Massive MIMO mm-Wave Systems",

   (Under revision to be submitted in IEEE Transactions on Vehicular Technology Correspondence, 2018.)

1.3 Overview of the Thesis

This thesis is organized as following: Chapter 2 discusses the background in the context of hybrid D/A precoding and combining mm-Wave large antenna array systems. In Chapter 3, we propose D/A precoders and combiners for the single user and multi-user hybrid D/A mm-Wave large antenna array system. Then in Chapter 4, we design an energy efficient hybrid D/A precoding and combining for mm-Wave large antenna array systems. Ultimately, this is then extended in Chapter 5 for designing hybrid D/A precoders and combiners for uplink multi-user mm-Wave large antenna array systems. Finally, the conclusions and future works are presented in Chapter 6.

Chapter 2

Overview on Beamforming in Millimetre-Wave Systems

Contents

2.1 Overview

The foreseeable future of wireless communication systems are demanded to provide high data rates and increasing the capacity of the system to assist large number of users for using different range of applications. In general, the capacity of a user communication system relies on the bandwidth of the system and signal-to-noise ratio (SNR). In cellular communication system, transmitting a high power signal over a large bandwidth can attain the target of achieving a high data rate in the system. Nevertheless, such a system is not possible so far as the system consumes high power, and has high inter-cell and intra-cell interference which affects the performance of the system directly. Hence, it is of paramount importance that the system efficiently utilizes the large spectrum and provides a reasonable SNR level at the receiver side. This contributes to attaining the target of high data rates in the next generation wireless communication system. The presence of a large-scale antenna array of tens or even hundreds of elements has been recognized as an essential component which is also called massive MIMO system to meet the explosive data rates and capacity increase requirement [2–8]. Large antenna array system exploits the multipath structure of the propagation channel which leads to efficient use of the basic resources such as time, spectrum, code and space [2, 3, 7]. Developing multiple antenna array technology opens possibilities for a large variety of features for example diversity gain, array gain, and spatial multiplexing gain. Another important achievement of utilizing large antenna array systems is to attain a high spectral efficiency by exploiting the spatial multiplexing gain and transmitting multiple data streams over the same time and frequency slot by spatial filtering. This is provided by proposing precoding and combining techniques at the transmitter and receiver, respectively [9, 28].

One of the substantial solutions to significantly improve the received SNR is to focus the transmitted signal power towards the desired user by exploiting directional beamforming solutions. Moreover, the total number of transmitted data streams by using the wireless communication system is limited by the total number of antennas that are implemented at the base station and the user. Essentially, equipping the base station with a large number of antennas employs the use of antenna arrays in the range of a few hundred antennas simultaneously serving many users in the same frequency resource and also attaining high data rates. Furthermore, large antenna array can effectively generate a very narrow beams by implementing beamforming techniques [5] as will be shown in Chapter 3. In this case, a high level of the transmitted power can be intended towards the desired user which improves the received SNR concurrently low interference among the all users in the system is achieved as well.

The prior advantages of equipping the base station with a large number of antennas, which is also known as massive MIMO systems, has stimulated many researchers in industry and academia to contemplate this technology as one of the primary candidates for next mobile wireless communication systems e.g., 5G. An overview of such systems and their advantages, challenges and applications are presented in the following sections.

2.2 MIMO with Large Antenna Array Systems

A large antenna array system is known as a massive MIMO system, is a new research field, where BS are equipped with a very large number of antennas as compared to previously considered systems [2–8]. Without any doubt, the main objective of utilizing such system is to scale up the benefit of the conventional low number of antennas which is known as MIMO but on a much larger scale. In the future wireless mobile communication system, large antenna array will be energy-efficient, secure, robust, and efficient in spectrum use. Fortunately, this can be achieved by deploying a significant large number of antennas per area, since spectral resources are scarce [29]. Over the last couple of years, the large antenna array system has gone from being a theoretical concept to becoming one of the most promising ingredients of the emerging future mobile wireless communication technology [30]. This is because, it provides a way to improve the area spectral efficiency (bit/s/Hz/area) under realistic conditions, by upgrading existing base stations. For large antenna array system, to be capable of providing its full advantages, accurate and instantaneous, channel state information (CSI) is demanded at the base station. Substantially, the large antenna array system was originally envisioned for time division duplex (TDD) operation [31], but can potentially be applied also in frequency division duplex (FDD) operation [32]. According to this, the estimates of the channel during uplink responses need to be available at the base station in order to allow the base station to apply the same channel parameters during the downlink transmission.

One of the challenges of deploying large number of antennas at the system is the physical aperture size of the antenna array. Due to antenna size constraints of being too large, deploying large number of antennas cannot be possible at the system with the conventional linear arrays. Thus there is a growing interest in a planar array which is known also as two dimensional (2D) arrays [33]. The definition of the linear arrays is based on the antenna elements are arranged in a singular straight line, while, the 2D ar-

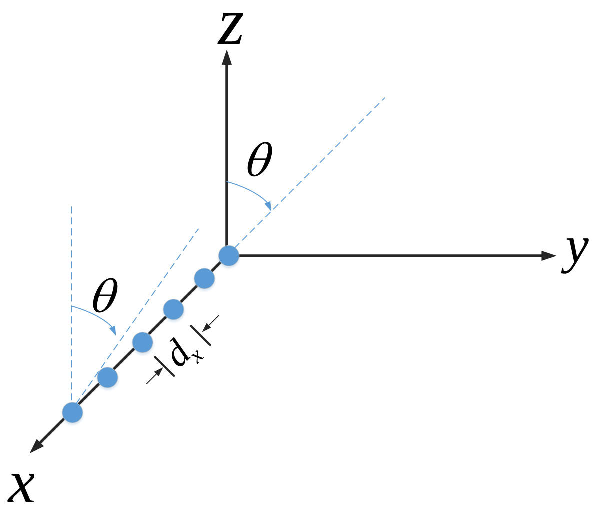


Figure 2.1: The linear array geometry of N elements along the x-axis.

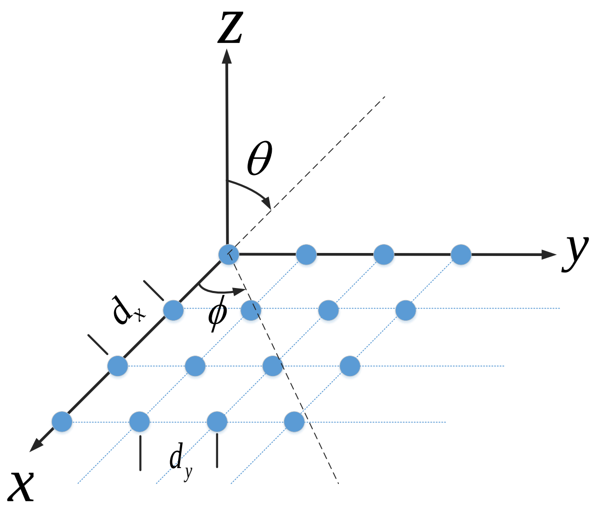


Figure 2.2: The 2D planar array geometry of $N_x \times N_y$ elements along the x-axis and y-axis.

rays arrange the antenna elements in a geometric grid. Fig. 2.1 and Fig. 2.2 illustrate the antenna elements structures in both the linear arrays and the planar arrays, respectively. Undoubtedly, the idea of equipping large number of antennas at the system by using the linear arrays architecture is critical as the physical aperture size is too large in terms of the length. Therefore, 2D arrays with relatively acceptable in size is proposed to make the large antenna system physically possible at the system. Furthermore, the anticipated results of utilizing 2D arrays structure for large antenna array systems allows the transmitted signal power to be formed in 3D direction, as the base station can spatially steer the power in the horizontal and vertical directions.

2.3 Millimetre Wave (mm-Wave) Systems

The radio frequency is divided into categories of frequency and wavelength bands ranging from 30 Hz to 3000 GHz. Most of current communication systems operate below 6 GHz band [9, 34]. While, the range between 30-300 GHz is referred to the mm-Wave bands [35, 36]. The mm-Wave transmission promises the distinct advantage of exploiting the huge amounts of unused (and possibly unlicensed) spectrum in those bands - around 200 times more than conventional cellular systems [35, 36]. Furthermore, such a system have demonstrated the ability to achieve gigabit-per-second data rates at a distance up to one kilometre in an urban mobile environment. Hence, it is well-known that the wavelength of a frequency is inversely proportional to the frequency value which implies that the wavelength of the mm-Wave bands is shorter than the current utilized bands (i.e., less than 6 GHz). Due to the nature of electromagnetic waves at the short wavelength, deployment of mm-Wave remains a largely unexplored frontier in mobile communication, and many challenges are required to be addressed. Beginning with signal propagation and channel properties, it is proved that the mm-Wave signals suffer unwieldy from huge path loss, penetration loss, rain effect, and atmospheric absorption. On the other hand, these short wavelength (i.e., less than 10 mm) allow us to pack a large number of antennas, resulting in strong diversity gains which can overcome the high path loss and the environment effects. Due to this reason, mm-Wave bands can be an excellent candidate for future wireless communication system to have a large number of antennas in a small physical 2D aperture size.

Unlike the current wireless communication system such as e long term evolution (LTE) [37], mm-Wave bands provide a short transmission distance which however support a low mobility environments. Further, the mm-Wave channel measurements demonstrated that there is a small number of multipath as compared to the exiting rich multipath channels [11, 38–40]. To eliminate the impact of this issue, directional antennas techniques at the transmitter and receiver are demanded to be employed to cooperate in generating a perfect beam alignment between the transmitter and receiver. Therefore, for extracting the full benefit of utilizing mm-Wave systems, directional antennas with beamsteering capability are required for the system. As mentioned earlier that the mm-Wave system allows to equip a large number of antennas which generates a narrow steerable beamforming. In this case, beamforming algorithms for mm-Wave large antenna array systems are required to be proposed which can help in successfully creating the directional transmission. Therefore, an overview of beamforming techniques for mm-Wave large antenna array system will be discussed in the next section.

2.4 Beamforming in mm-Wave Large Antenna Array Systems

Beamforming is a technique referring to the antenna's capability to focus energy in a specific direction in the space. Furthermore, this technique allows the antenna's energy to be focused towards a desired user, whilst attenuating signals or producing nulls in undesired directions. The principle concept of generating beamforing is based on manipulating the amplitude and phase of the transmitted signals from each antenna for the purpose of performing beamforming such that the transmitted signals from a desired direction are added constructively or de-constructively. It is observed that it is common in the literature and also in this thesis that the term beamforming is utilized to denote a technique at both the transmitter and receiver to improve the received SNR. The precoding and combining terms are always applied when the spatial filter is designed to attain spatial multiplexing at the transmitter and the receiver, respectively.

To utilize the conventional idea of using multiple antennas at the system, each antenna elements is connected to the base station processor. In this system, a digital-to-analog conversion (DAC) or an analog-to-digital converter (ADC), signal mixing, amplifier and filter components are required either per antenna or per a group of antennas [2, 34]. The term that covers the combination of the required components to connect each antenna elements or group of antennas in the conventional multiple antennas to the processor are called radio frequency (RF) chains. Typically, a system that requires to connect one RF chain with each antenna in the system is called a fully digital beamforming system [10,41]. Simultaneously, at the system that group of antenna elements are connected to one RF chain, the system known as an analog beamforming system [17, 20].

In the fully digital beamforming, the signals from each RF chain are sampled and stored, then the beam is formed by summing of the appropriate samples. The RF information at each element is converted into two streams of binary complex baseband signals, representing the in-phase and quadrature amplitude or 90 phase shifter components. However, this conventional fully digital precoding requires one RF chain for each antenna, thus in large antenna array systems such system induces huge hardware cost and power consumption which makes fully digital precoding difficult to realize in practice [10, 17–23]. In fact, this is particularly crucial at mm-Wave systems, where a large number of antennas are demanded [5, 10, 17–23, 28].

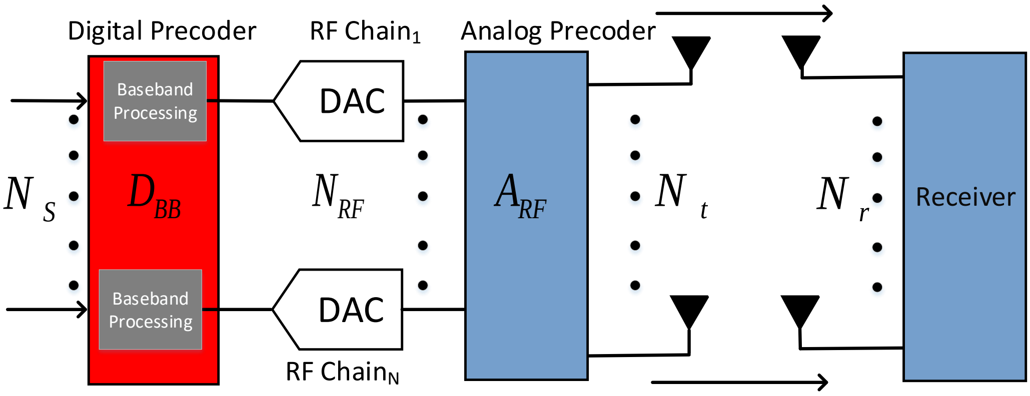
To handle this challenge, an initially proposed solution was to use only the analog domain, which is designed based on controlling the phase of each antenna and forming beams through low-cost phase shifters [16] resulting in proposing analog beamforming architectures. The analog beamforming is one of the simplest approaches for applying large antenna array in mm-Wave systems. It can be utilized at both the transmitter and receiver. It is a defacto solution supported in IEEE 802.11ad [42]. The configuration of the analog system is often implemented by exploiting a network of digitally controlled phase shifters. The phase shifter weights are adaptively adjusted using digital signal processing utilizing a specific strategy to steer the beam and meet a given objective, for example to maximize received signal power. Nevertheless, this solution are mainly limited to a single data stream transmission. Since the hardware constraints are applied on the analog precoder such as the availability of only quantized phase shifters, the extension for a multi-stream processing is very difficult to design. In this case, the achievable spectral efficiency of the analog precoding system that is achieved is much lower than that of the fully digital precoding system [10]. Therefore, a hybrid D-A beamforming system which comprises a small number of RF chains and large number of antennas have been proposed to balance the achievable spectral efficiency, power consumption, and the implementation complexity [10, 17–23].

In the following section more details regarding the hybrid D-A beamforming challenges, feasibilities, and different type of structures are provided.

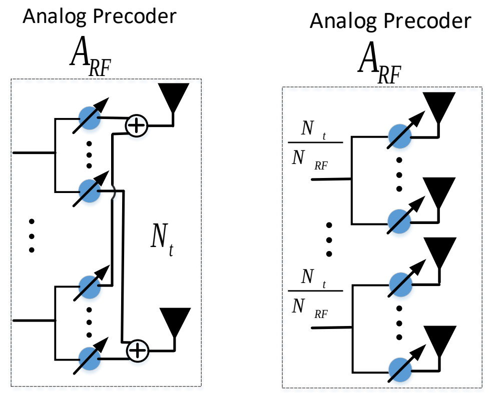
2.5 Hybrid Precoding and Combining for mm-Wave Large Antenna Array Systems

Large antenna wireless systems have been proposed as a leading candidate for 5G wireless access [10, 17–23]. Along with mm-Wave technologies, that were recently recognized as essential for providing more spectrum [35, 36], it offers higher data rates and capacities than traditional wireless communication systems. The utilization of large antenna arrays at both the transmitter and receiver holds the potential for higher array gain than the previous existing systems [2,34,37]. To exploit this gain, precoding and combining techniques are utilized in mm-Wave large antenna array systems. Traditionally, these techniques are implemented in the baseband (e.g., digital domain), these methods demand a dedicated RF chain hardware per antenna. However, in the case that a massive number of antennas is taking into account, this results in a huge computational complexity and cost, as the RF components are expensive and demand high power consumption. Note, it is feasible to propose economical hardware that will utilize the potential acquisition from a large number of cheap antenna elements exploiting a small number of RF chains. To target this goal, different hybrid D/A schemes have been proposed [10, 17–23].

In hybrid D/A precoding and combining, the function of the precoder is split between the digital and analog domains as illustrated in Fig. 2.3: (i) a low dimensional digital precoder operates on the transmitted signal at the baseband; (ii) an analog precoder then maps the small number of digital outputs to a large number of antennas; the same structure can be performed at the transmitter side and the receiver side as well. Common analog architectures are based on analog phase shifters networks. The specific schemes can vary according to the power and area budget. Two main families of hybrid D-A architectures are proposed, which are introduced as fully-connected antenna array and partially-connected antenna array structures. Fig. 2.3(b) and Fig. 2.3(c) illustrate the



(a) Block diagram of hybrid D-A precoding based mm-Wave large antenna system



(b) Fully- connected antenna array.

(c) Partially-connected antenna array

Figure 2.3: A diagram of hybrid D-A precoding based mm-Wave systems when RF chain is **(b)** connected to all N_t transmit antennas, or **(c)** connected to a partially-connected antenna array with $\frac{N_t}{N_{RF}}$ Tx antennas.

structure of the hybrid D/A fully- and partially-connected antenna array systems. Fully-connected antenna array networks offer a mapping from each antenna to each RF chain and allow to maximize the precoding and combining gain as shown in Fig. 2.3(b). While in the partially-connected antenna array scheme, a reduced number of analog components is utilized and clarified in Fig. 2.3(c). This degrades the spectral efficiency but provides lower power consumption and lower hardware complexity. In the data phase of each coherence interval of the large antenna communication process, the transmitter sends multiple data streams to the receiver over the constant and known channel, using precoding techniques. At the receiver side, a combiner is used to estimate the data vector from the received signal at the antennas. The precoder and combiner are chosen to optimize some desired performance measure, such as estimation error or spectral efficiency of the system. Unlike the fully digital system case, when considering a hybrid precoding, the precoder and combiner matrices cannot have arbitrary entries, but are constrained according to the specific hardware choice. For example, when using a phase shifter network at the analog side, only unimodular matrices for the analog precoder are considered. The goal then is to optimize the performance measure over all pairs of digital and analog precoder matrices, which yields a non-convex difficult optimization problem. The useful of applying the digital precoder in the hybrid D-A system is to correct for lack of precision in the analog, for example to cancel residual multi-stream interference. This allows hybrid precoding to approach the performance of the unconstrained solutions which is known as the fully digital system.

In general, the hybrid D/A precoder based on partially-connected antenna array structure is more suitable for practical applications. In theoretical works, however, the hybrid D/A precoder based on fully-connected antenna array phase shifter network model is frequently utilized. It is noted that the hybrid precoding design problems have been investigated based on various criteria in different scenarios with different structures i.e., fully- and partially-connected antenna array system. For example, maximization of spectral efficiency can be considered in single user and multiuser, with perfect and imperfect channel state information, with joint and separate design the digital and analog precoders and combiners.

In the following Chapters 3, 4 and 5, the existing solutions for hybrid D/A precoding systems design problems have been examined, and the proposed solutions for designing hybrid D/A precoding systems are presented.

Chapter 3

Designing Hybrid Precoding for Millimetre Wave Large Antenna Array Systems

Contents

3.1 Overview

Hybrid digital-to-analog (D/A) precoding with a limited number of RF chains has become popular for mm-Wave systems. Two types of architectures namely, fully and partially-connected antenna array architectures, have been studied in the literature. In this chapter, novel hybrid D/A precoding algorithms that can achieve high system spectral efficiency with low computational complexity are proposed for both these structures. For the fully-connected antenna array, the precoding design is formulated as an optimization problem to minimize the Euclidean distance between the hybrid D/A and fully digital system. An alternative minimization algorithm is then proposed based on Stiefel manifold (ASO). For the partially-antenna array, a joint D/A precoding optimization problem is formulated to maximize the system spectral efficiency. An optimal algorithm is proposed based on the principle of complex circle manifold optimization (CMO). Both ASO and CMO achieve high spectral efficiency for these structures, although their computational complexities are high. To reduce the complexity of the proposed ASO and CMO algorithms, particle swarm optimization (PSO) is designed for both these structures. The alternative minimization based on PSO (APSO) is tailored for the fully-connected antenna array, while PSO in the partially-connected antenna array (SPSO) is designed to jointly optimize D/A precoder.

This chapter further develops hybrid D/A precoding designs for mm-Wave multi-user systems based on maximizing the sum rate of the system directly. Simulation results show that our proposed algorithms outperform the existing hybrid D/A precoding algorithms for the two types of structures, in terms of spectral efficiency. Finally, simulation results show that our proposed APSO and SPSO algorithms have lower complexity than the existing algorithms.

3.2 Introduction

The high prevalence of smartphones has generated a continuing demand for increasing data rate in wireless systems. One of the promising solutions to handle this challenge is to deploy multiple antenna systems [3, 8, 44, 45] and mm-Wave systems with large available bandwidth [9–14]. The extreme short wavelengths (i.e., less than 10 mm) of mm-Wave signals allow a large number of antennas to be packed in a small physical space [10,11]. As a result, it is possible to achieve a high array gain which compensates for severe path loss and atmospheric absorption experienced by mm-Wave signal propagation [10–14, 34]. In conventional wireless systems, the number of antennas is equal to the number of RF chains, so that the maximum performance gain is achieved by using the fully digital precoder scheme. However, if the fully digital precoder scheme is adopted in the mm-Wave systems, the associated large number of RF chains will consume extremely high power and the hardware complexity will increase dramatically [10, 14]. A simple approach is to employ a pure analog precoder system which uses only one RF chain to support a large number of antennas. The analog precoder controls the phase of each antenna and forms beams through low-cost phase shifters [16]. However, the achievable spectral efficiency of the analog precoding system is much lower than that of the fully digital precoding system [10]. Therefore, a hybrid D/A precoding system which comprises of a small number of RF chains and large number of antennas have been proposed to balance the achievable spectral efficiency, power consumption, and the implementation complexity [10, 17–22, 34, 41].

The principle of the hybrid D/A precoding system is to first process the signals in a low-dimensional digital precoder and then push the digital output signals through high dimensional analog precoders [10,17–22]. There are two types of structures for the hybrid D/A precoding. The first structure is a fully-connected antenna array structure, where each RF chain is connected to all the transmit antennas [10, 14]. The fully-connected antenna array structure achieves full analog bemforming (BF) gain, even though the number of RF chains is smaller than the number of antennas. Nevertheless, the computational complexity for the analog precoder is high, since the number of phase shifters to be optimized scales linearly with both the number of RF chains and the number of antennas [20,21]. The second type is a partially-connected antenna array structure, which can reduce the computational complexity at the cost of BF gains [17, 18, 20–22]. Under this structure, each RF chain is connected to a specific partially-connected antenna array, so that the operation cost and computational complexity are much lower than that of the fully-connected antenna array at the expense of lower spectral efficiency.

In the design of fully-connected antenna array structure, the performance of the hybrid D/A precoding mm-Wave system is generally evaluated by comparing the system spectral efficiency with that of the fully digital system. The design of the hybrid D/A precoding is formulated as an optimization problem with a constant modulus constraint (i.e., all the non-zero elements in the analog precoder have the same amplitudes) for all the phase shifters, which is a non-convex constraint making the optimization problem difficult to solve. To simplify this constraint, researchers proposed to impose extra constraints on the system to facilitate the design of analog precoders [10, 15]. However, such kind of approaches may result in system spectral efficiency loss. Furthermore, implementation complexity of the algorithms needs to be considered, especially when the number of antennas is large.

Sparse precoding based on an orthogonal matching pursuit (OMP) algorithm is first proposed in [10] and widely adopted in the recent papers [15]. The sparse-scattering property of mm-Wave channels is fully exploited by OMP, which formulates the design of hybrid D/A precoders as a sparsity constrained matrix reconstruction problem. In OMP, the columns of the analog precoding matrix are selected from some candidate vectors, such as array response vector of the channel [12,15] or a pre-defined dictionary. However, when designing the analog precoder, the optimal precoder for the corresponding fully digital system is derived from a singular value decomposition (SVD) of the channel. In general, the dimension of the analog precoder is much higher than that of the digital precoder. Hence, the complexity of the OMP algorithm is dominated by the analog part, which is quite high. A near-optimal hybrid precoder can be found via an optimization that resembles the problem of sparse signal recovery. However, to perfectly achieve the performance of fully digital precoding this method requires the number of RF chains to be twice that of the number of data streams. For the case when the number of RF chains is equal to the number of data streams, OMP suffers from performance loss, as the Euclidean distance between the optimal solution and OMP is not tight any more. To solve this issue, greedy hybrid precoding (GHP) is proposed in [41]. The aim of GHP is to reduce the complexity of OMP and avoid any constraints on the number of RF chains. That can be achieved in GHP by replacing the correlation operations over a dictionary with the element-wise normalization of the first singular vector of the residual. In this case, GHP avoids any additional constraints on the number of RF chains. This method attains higher spectral efficiency than OMP and its achievable spectral efficiency is closer to the fully digital system. Unfortunately, the residual is needed for designing the analog precoder in GHP which is obtained by using SVD operation two times. This operation incurs high computational complexity, especially when a large number of antennas are equipped in the mm-Wave system.

In [19], a popular algorithm, named manifold optimization based on alternating minimization (MO-AltMin) algorithm, is introduced to achieve a spectral efficiency that is tightly close to the fully digital system. To satisfy the constant modulus constraint directly, the analog precoding matrix is relaxed into a circle complex manifold optimization based on the conjugate gradient algorithm. This algorithm does not need any pre-defined candidate set for analog precoders (e.g., OMP). Furthermore, MO-AltMin minimizes the Euclidean distance very tightly rather than minimizing an upper bound or considering sparsity optimization, which makes MO-AltMin attain higher spectral efficiency than OMP and GHP. Like GHP, MO-AltMin also enjoys the flexibility of having any number of RF chains in the system. Although the MO-AltMin algorithm approaches the performance of the fully digital system, its associated high complexity hampers its practical application. The complexity of MO-AltMin is high because the analog precoders in each iteration are updated via three stages. First, a conjugate gradient descent search is required, where the Armijo backtracking step-size is performed. Second, one needs to search repeatedly to find the local optimum with zero gradient of the cost function. Third, the Kronecker product of large-sized matrices is involved in the MO-AltMin algorithm. [19] has developed another algorithm, called PE-AltMin by using the procrustes solution based on SVD, where the orthogonal property is imposed to design the digital precoder, and phase extraction is used to design the analog precoder. However, in the case of the number of RF chains being equal to the number of data streams, PE-AltMin achieves higher spectral efficiency than OMP, and even close to the spectral efficiency of MO-AltMin. Although PE-AltMin quantifies the gap between the proposed analog precoder solution and the optimal solution, the PE-AltMin performs SVD operation two times to design the analog precoder, which yields a high complexity when a large number of antennas are utilized.

Although the fully-connected antenna array system can achieve good performance, the computational complexity of designing the hybrid D/A precoding is high. Therefore, another alternative structure named partially-connected antenna array structure attracts great interest due to its lower computational complexity when designing the hybrid D/A precodings [17, 18, 20–22]. In the current state of the art of the hybrid D/A precoding for the partially-connected antenna array system, the digital precoder is generally fixed as an identity matrix while the analog precoder is exactly the normalized Hermitian of the channel matrix [17,18,22]. Since the digital and analog precoders are not jointly designed, the methods in [17,18,22] led to sub-optimal solutions. Recent investigations on joint D/A precoders have been proposed using an iterative algorithm based on SVD [20, 21], which optimizes every RF chain successively. The complexity of the SVD-based hybrid D/A algorithm is still high because of matrix inversion [46] and the solution of the SVD-based hybrid D/A algorithm cannot be directly applied to the system as it does not satisfy the constant modulus constraint [20, 21].

For the downlink of K-user where single antenna is allocated at each user, it is introduced in [13] that hybrid precoder with K RF chains at the base station can achieve a reasonable sum rate as compared to the sum rate of fully digital ZF systems. The analog precoder involves the idea of realizing a scaled version of an arbitrary complex vector as the addition of two phase shift vectors. In this design, this approach fails to explore the sparsity structure of mm-Wave channels for complexity reduction which results in the implementation of such system may not be hardware efficient as it doubles the demanded phase shifters. Moreover, the digital precoder based on ZF method with effective analog channel is proposed. Since most of recent works [10, 17–22, 34, 41] showed that mm-Wave systems work at low and medium SNR levels, and it is known that ZF performs poorly at low SNR regime [27], thus ZF digital precoder in [13] scarifies performance at mm-Wave system. Another limitation is that the designed of the hybrid precoding in [13] is based on the fact that the number of RF chains is twice than the number of data streams in the multi-user system. Therefore, in the case that the number of RF chains equals the number of data stream there is a gap between the sum rate achieved with this particular hybrid precoding design and the fully digital system.

As introduced above, the existing hybrid D/A precoding algorithms either have extremely high complexity, or significantly sacrifice the system spectral efficiency. Therefore, this chapter aims to develop novel hybrid D/A precoding algorithms for large-scale antenna mm-Wave systems by taking into account both system spectral efficiency and the computational complexity for fully- and partially-connected antenna array structures. For the fully-connected antenna array system, the spectral efficiency of the fully digital system is used as the performance benchmark for the design of hybrid D/A precoding. Then, the precoding design is formulated as an optimization problem to minimize the Euclidean distance between the hybrid D/A system spectral efficiency and fully digital system spectral efficiency. Due to the non-convexity of this optimization problem, the analog and digital precoders are firstly designed separately, and then optimized together by employing an alternative optimization technique. We also consider the partially-connected antenna array structure where the digital and analog precoders based algorithms can be jointly designed. Specifically, the contributions of this chapter are summarized as follows:

1. An alternating minimization based on a Stiefel manifold algorithm (ASO) is proposed for the analog precoder in the fully-connected antenna array structure. In

order to guarantee the constant modulus constraint, a convex Euclidean distance problem is proposed. Compared to the current algorithms (e.g., OMP and GHP), the ASO algorithm does not require any pre-determined candidate vectors for forming the analog precoding matrix. The proposed ASO requires much lower computational complexity than MO-AltMin. Firstly, the SVD operation is required in the curvilinear search for the MO-AltMin, which is not necessary in ASO. Secondly, when determining the step size in MO-AltMin, higher complexity is required due to the additional constraints in MO-AltMin problem. In contrast to MO-AltMin, Barzilai-Borwein (BB) step-size in our proposed ASO does not require any additional constraints and also provides fast convergence speed. Simulation results show that the spectral efficiency of the proposed ASO algorithm surpasses OMP, GHP, and PE-AltMin, and coincides with MO-AltMin which is very close to that of the fully digital system. In contrast to OMP and PE-AltMin, ASO does not place any constraints on the number of RF chains.

2. Although the ASO algorithm has good performance and its complexity is lower than the popular MO-AltMin algorithm, the complexity is still higher than some of the existing algorithms (e.g., GHP and PE-AltMin). This is because for ASO, three operations are needed in each iteration. Therefore, for systems with a large number of antennas, another analog precoding algorithm is proposed by applying particle swarm optimization into alternating minimization which is abbreviated as APSO. This algorithm is an evolutionary approach, which refines the estimates through a group of agents searching for the solution space. In each iteration, APSO only requires a cost-function and does not require any differentiation, matrix inversion or pre-defined candidates (e.g. OMP). Consequently, APSO is very flexible and feasible in terms of the number of RF chains but also has good spectral efficiency performance and lower computational complexity. The reason is that APSO depends only on the number of agents which are used to find the optimal solution. In fact, the number of agents is directly proportional to the computational complexity and also to the spectral efficiency of the system. With a smaller number of agents, APSO has much lower complexity than all the existing algorithms. Interestingly, by

increasing the number of agents the spectral efficiency performance of the system increases until approaches those achieved by GHP and PE-AltMin, and it has lower complexity than the latter algorithms.

3. For the purpose of implementation for the partially-connected antenna array, the key issue to be addressed in the designing of the joint hybrid D/A precoding is to satisfy the constant modulus constraint without scarifying much of the system spectral efficiency. The first algorithm proposed in this chapter is based on the principle of complex circle manifold optimization (CMO) which directly satisfies the constant modulus constraint by converting the precoding matrix as a vector and laying the elements of the vector into a complex circle. CMO achieves higher spectral efficiency than the existing algorithms. In order to find the optimal solution in CMO, Riemannian gradient and retraction are preformed. Retraction causes the computational complexity of the algorithm to be high.

4. For the partially-connected antenna array structure, particle swarm optimization (SPSO) algorithm is proposed. Due to the flexibility advantages associated with PSO, SPSO can be implemented in the practical system. Simulation results demonstrate that the proposed CMO and SPSO algorithms outperform the existing algorithms in [16–18,22]. Numerical results show that SPSO has much lower complexity than the discussed hybrid D/A algorithms.

5. The designed of the single-user system for both structures is extended to a multi-user system where the hybrid precoder architecture has multiple RF chains for multi-user data streams detection. By generalizing the proposed algorithms i.e., ASO, APSO, CMO, and SPSO for the single-user system, the analog precoder matrix is designed for maximizing the sum rate of the multi-user system. The digital counterpart follows the conventional digital MMSE design based on the effective downlink channels to support spatial multiplexing by decoupling data streams from diverse desired users. The main idea for the multi-user system is to maximize the sum rate directly. Consequently, this designs can bypass the steps of (i) deriving the fully digital precoder via SVD; (ii) avoiding the used of the alternating minimization

method. This helps in limiting the complexity of the optimization problem when designing multi-user hybrid precoding mm-Wave systems.

3.3 System Model

Fig. 3.1(a) shows a downlink mm-Wave system with a large antenna array for hybrid D/A precoding. The base station (BS) sends N_s data streams via N_t transmit antennas to the receiver which consists of N_r receiver antennas. It is assumed that there are N_{RF}^t and N_{RF}^r RF chains at the BS and the receiver, respectively. The digital precoder is denoted as $\boldsymbol{D}_{BB}$ and the analog precoder is denoted as $\boldsymbol{A}_{RF}$. The analog precoder is designed under the condition that all the non-zero elements of $\boldsymbol{A}_{RF}$ have the same amplitude, but different phase shifts, while the digital precoder satisfies the total transmit power constraint by normalizing $\|\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F^2 = N_s$. The dimensions of $\boldsymbol{D}_{BB}$ and $\boldsymbol{A}_{RF}$ are $N_{RF}^t \times N_s$, and $N_t \times N_{RF}^t$, respectively. Note that, both $\boldsymbol{D}_{BB}$ and $\boldsymbol{A}_{RF}$ may not be square matrices as $N_s \leq N_{RF}^t$ and $N_t >> N_{RF}^t$.

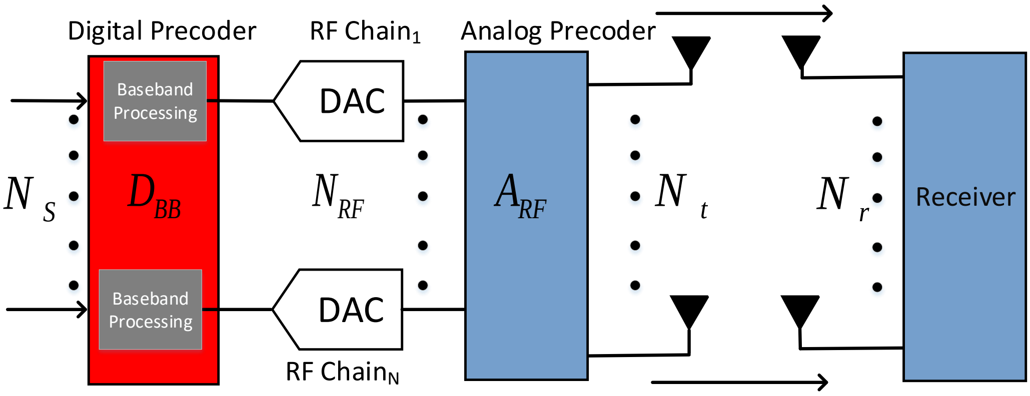
3.3.1 Channel Model

The mm-Wave channel is using 3D Saleh-Valenzuela geometry-statistical channel model an adopted in [20, 47] which is represented as

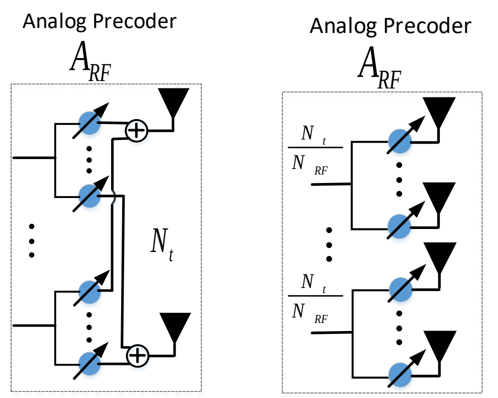
$$\boldsymbol{H} = \sqrt{\frac{N_t N_r P_{n_t}}{L}} \sum_{l=1}^{L} \left(h_{n_r,n_t,l}^{3D}\, \boldsymbol{sv}_r(\theta_l^r, \phi_l^r)\, \boldsymbol{sv}_t^H(\theta_l^t, \phi_l^t) \right), \tag{3.1}$$

where $\boldsymbol{H} \in \mathbb{C}^{N_r \times N_t}$, and L is the number of multipaths [20]. P_{n_t} is the power of the n_t-th Tx antenna and is calculated by assuming a single slope exponential power delay profile by [47]. The superscript (and subscript) t (and r) represents transmitter (and receiver) respectively. The 3D channel gain h_{n_r,n_t}^{3D} from the n_t-th transmit (Tx) antenna element to the n_r-th receive antenna (Rx) experiencing the l-th multipath is given by

$$\begin{aligned} h_{n_r,n_t,l}^{3D} &= \begin{bmatrix} F_{Rx,n_r,\mathcal{Z}}(\theta_{l,n_r}^r, \phi_{l,n_r}^r) \\ F_{Rx,n_r,\mathcal{A}}(\theta_{l,n_r}^r, \phi_{l,n_r}^r) \end{bmatrix}^T \begin{bmatrix} e^{j\phi_l^{\mathcal{ZZ}}} & \sqrt{\kappa_{n_t}^{-1}} e^{j\phi_l^{\mathcal{ZA}}} \\ \sqrt{\kappa_{n_t}^{-1}} e^{j\phi_l^{\mathcal{AZ}}} & e^{j\phi_l^{\mathcal{AA}}} \end{bmatrix} \begin{bmatrix} F_{Tx,n_t,\mathcal{Z}}(\theta_{l,n_t}^t, \phi_{l,n_t}^t) \\ F_{Tx,n_t,\mathcal{A}}(\theta_{l,n_t}^t, \phi_{l,n_t}^t) \end{bmatrix} \\ &\times\ e^{(j2\pi\lambda^{-1}(\bar{\boldsymbol{\varphi}}_{l,n_r}\bar{\boldsymbol{r}}_{Rx}))} e^{(j2\pi\lambda^{-1}(\bar{\boldsymbol{\phi}}_{l,n_t}\bar{\boldsymbol{r}}_{Tx}))}, \end{aligned} \tag{3.2}$$



(a) Block diagram of hybrid D-A precoding based mm-Wave large antenna system



(b) Fully- connected antenna array.

(c) Partially-connected antenna array

Figure 3.1: A diagram of hybrid D/A precoding based mm-Wave systems when RF chain is **(b)** connected to all N_t transmit antennas, or **(c)** connected to a partially-connected antenna array with $\frac{N_t}{N_{RF}}$ Tx antennas.

where λ is the wavelength, and $\theta_l^t(\phi_l^t)$ and $\theta_l^r(\phi_l^r)$ represent the zenith (azimuth) angles of departure and arrival (AoD) and (AoA), respectively. $F_{Rx,n_r,\mathcal{Z}}$ and $F_{Rx,n_r,\mathcal{A}}$ are the Rx beam pattern for the zenith (Z) and azimuth (A) polarization. $F_{Tx,n_t,\mathcal{Z}}$ and $F_{Tx,n_t,\mathcal{A}}$, are the Tx beam pattern for the n_t-th Tx antenna. $\phi_l^{\mathcal{Z}\mathcal{Z}}, \phi_l^{\mathcal{Z}\mathcal{A}}, \phi_l^{\mathcal{A}\mathcal{Z}}$, and $\phi_l^{\mathcal{A}\mathcal{A}}$ are the initial random phases for zenith $(\mathcal{Z}\mathcal{Z})$, cross $(\mathcal{Z}\mathcal{A}, \mathcal{A}\mathcal{Z})$, and azimuth $(\mathcal{A}\mathcal{A})$ polarization for the l-th multipath. κ_{n_t} is the intra-cluster Rician K-factor associated with the n_t-th Tx antenna [47]. The receive antenna array phase offset of the l-th multipath is represented as $e^{(j2\pi\lambda^{-1}(\bar{\varphi}_{l,n_r}\cdot\bar{\boldsymbol{r}}_{Rx}))}$ where $\bar{\varphi}_{l,n_r}$ is the arrival spherical unit vector and $\bar{\boldsymbol{r}}_{Rx}$ is the position vector of the Rx antenna of the user located at the origin which is yielded by

$$\begin{aligned}\bar{\boldsymbol{\varphi}}_{l,n_r} &= [\cos\vartheta_{l,n_r}\cos\varphi_{l,n_r}, \cos\vartheta_{l,n_r}\sin\varphi_{l,n_r}, \sin\vartheta_{l,n_r}]^T,\\ \bar{\boldsymbol{r}}_{Rx} &= [x_{Rx} = N_{r_x}u_x, y_{Rx} = h_U + N_{r_y}u_y, z_{Rx} = 0]^T,\end{aligned} \tag{3.3}$$

where ϑ_{l,n_r} and φ_{l,n_r} are the zenith and azimuth AoA. h_U is the height of the user from the ground. $e^{(j2\pi\lambda^{-1}(\bar{\boldsymbol{\phi}}_{l,n_t}\cdot\bar{\boldsymbol{r}}_{Tx}))}$ is the array phase offset for the n_t-th Tx antenna and l-th multipath. $\bar{\boldsymbol{\phi}}_{l,n_t}$ is the departure spherical unit vector of the l-th multipath and n_t-th Tx antenna. $\bar{\boldsymbol{r}}_{Tx}$ is the position vector of the Tx antenna element relative to the user. $\bar{\boldsymbol{\phi}}_{l,n_t}$ and $\bar{\boldsymbol{r}}_{Tx}$ are introduced as

$$\begin{aligned}\bar{\boldsymbol{\phi}}_{l,n_t} &= [\cos\theta_{l,n_t}\cos\phi_{l,n_t}, \cos\theta_{l,n_t}\sin\phi_{l,n_t}, \sin\theta_{l,n_t}]^T,\\ \bar{\boldsymbol{r}}_{Tx} &= [x_{Tx} = N_{t_x}u_x, y_{Tx} = h_B + N_{t_y}u_y, z_{Tx} = 0]^T,\end{aligned} \tag{3.4}$$

where θ_{l,n_t} and ϕ_{l,n_t} are the zenith and azimuth AoD, and h_B denotes the height of the BS.

As the uniform planar antenna (UPA) structure [47] is a preferred choice for 3D channel, $N_t = N_{t_x} \times N_{t_y}$, and $N_r = N_{r_x} \times N_{r_y}$, where (N_{t_x}, N_{r_x}) and (N_{t_y}, N_{r_y}) represent the numbers of antennas in the x-axis and y-axis. The array response $\boldsymbol{sv}_b$ in (3.1), where $b \in \{r\,,\,t\}$ represents arrival at the receiver (r) or departure at the transmitter (t), is given by [10, 47]

$$\boldsymbol{sv}_b(\theta_l^b, \phi_l^b) = \text{vec}[\boldsymbol{sv}_{b_x}(\Psi)\boldsymbol{sv}_{b_y}^H(\Phi)], \quad b \in [r\,,\,t], \tag{3.5}$$

where the directions of the array responses $\boldsymbol{sv}_{b_x}$ and $\boldsymbol{sv}_{b_y}$ are represented by $\boldsymbol{sv}_{b_q}$ and devised as

$$\boldsymbol{sv}_{b_q}(\Omega) = \frac{1}{\sqrt{N_{b_q}}}\left[1, e^{j\Omega}, \cdots, e^{j(N_{b_q}-1)\Omega}\right]^T, \quad q \in [x\,,\,y], \tag{3.6}$$

where $\Omega \in \{\Psi, \Phi\}$ and the values of Ψ and Φ are given by

$$\Psi \;=\; -2\pi\lambda^{-1}u_x\sin(\theta_l^b)\cos(\phi_l^b), \quad \Phi = -2\pi\lambda^{-1}u_y\sin(\theta_l^b)\sin(\phi_l^b), \quad b \in [r\,,\,t], \tag{3.7}$$

where u_x and u_y are the inter-element distances in the x-axis and y-axis, respectively.

3.3.2 Received Signal of Hybrid D/A BF System

The received signal for all N_{RF} receive antennas $\boldsymbol{y} = [y_1, y_2, \cdots, y_{N_{RF}}]^T$ at the user is expressed as

$$\boldsymbol{y} = \sqrt{P_{\max}}\boldsymbol{W}_{BB}^{H}\boldsymbol{W}_{RF}^{H}\boldsymbol{H}\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\boldsymbol{s} + \boldsymbol{W}_{BB}^{H}\boldsymbol{W}_{RF}^{H}\boldsymbol{n}, \tag{3.8}$$

where $P_{\max}$ represents the average received power. A linear precoder is applied to the symbol vector $\boldsymbol{s} \in \mathbb{C}^{N_s}$ with $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^H] = \frac{1}{N_s}\boldsymbol{I}_{N_s}$. $\boldsymbol{n} \in \mathbb{C}^{N_r}$ is the zero-mean complex Gaussian noise vector with independent and identically distributed (i.i.d) elements with $\mathbb{E}[\boldsymbol{n}\boldsymbol{n}^H] = \sigma^2\boldsymbol{I}_{N_r}$. The analog combiner is denoted as $\boldsymbol{W}_{RF}$ and the digital combiner is represented as $\boldsymbol{W}_{BB}$. The dimension of $\boldsymbol{W}_{RF}$ and $\boldsymbol{W}_{BB}$ are $N_r \times N_{RF}^r$ and $N_{RF}^r \times N_s$. The analog combiner is realized by the phase shifters which can only adjust the phase of the signal without changing its amplitude, therefore, all the non-zero elements of $\boldsymbol{W}_{RF}$ have the same amplitude. The objective is to design $\boldsymbol{A}_{RF}$, $\boldsymbol{D}_{BB}$, $\boldsymbol{W}_{RF}$ and $\boldsymbol{W}_{BB}$ to maximize the spectral efficiency C that is given by

$$C = \log_2 \left| \boldsymbol{I}_{N_s} + \frac{P_{\max}}{N_s\sigma^2}\boldsymbol{R}^{-1}\boldsymbol{W}_{BB}^{H}\boldsymbol{W}_{RF}^{H}\boldsymbol{H}\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\boldsymbol{D}_{BB}^{H}\boldsymbol{A}_{RF}^{H}\boldsymbol{H}^{H}\boldsymbol{W}_{RF}\boldsymbol{W}_{BB} \right|, \tag{3.9}$$

where $\boldsymbol{R}^{-1} = \boldsymbol{W}_{BB}^{H}\boldsymbol{W}_{RF}^{H}\boldsymbol{W}_{RF}\boldsymbol{W}_{BB}$ is the processed noise covariance matrix after combining. To simplify the unknown parameters $\boldsymbol{A}_{RF}, \boldsymbol{D}_{BB}, \boldsymbol{W}_{RF}$ and $\boldsymbol{W}_{BB}$, it has been confirmed that the designed of precoders and combiners can be decoupled into two sub-problems [10, 19, 41], i.e., to design the combiner of hybrid D/A systems a similar mathematical procedure can be used except the extra transmitter power constraint at the transmitter. Therefore, the structure of this chapter is mainly focused on the precoder side and the proposed algorithms can be applied for the combiner. Note here as in the previous proposed hybrid precoding [10, 13, 19, 41, 48], perfect channel state information (CSI) of $\boldsymbol{H}$ will be assumed in this chapter.

Due to the fact that the imperfection of the channel state information (CSI) at each user has a trivial impact on the system performance, we studied this impact on the proposed algorithms. The impact of the imperfect CSI on the hybrid D/A precoding mm-Wave system can be evaluated thus the estimated channel matrix $\widetilde{\boldsymbol{H}}$ can be given as

$$\widetilde{\boldsymbol{H}} = \eta\boldsymbol{H} + \sqrt{1-\eta^2}\boldsymbol{\Delta}_{err}, \tag{3.10}$$

where $\boldsymbol{H}$ is the accurate channel matrix as represented in (1), $\eta \in [0,1]$ accounts for the CSI accuracy. Moreover, the estimation variance parameter η is utilized to reflect the accuracy of CSI the channel estimation, e.g., $\eta = 1$ represents the perfect CSI, whereas

$\eta = 0$ corresponds to the case that the CSI is completely uncorrelated with the actual channel response. $\boldsymbol{\Delta}_{err}$ is a channel estimation error matrix, and each element is generated as an independent identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance $\mathcal{CN}(0, 1)$.

***Remark* 1.** *Although assuming complete channel information knowledge at the transmitter, the CSI can be estimated by using compressive sensing method by exploiting the sparsity properties of the mm-Wave channel, where only limited channel paths contributes the channel. Hence, only a few number of channel gains and angles need to be fed back to the base station. As a result, it is feasible to acquire the full CSI with high accuracy [48].*

3.4 Proposed Fully-Connected Antenna Array Hybrid D/A Precoding

When designing the hybrid D/A precoding for the fully-connected antenna array system, each RF chain is connected to all transmit antennas, as shown in Fig. 3.1(b). The optimal solution of $\boldsymbol{A}_{RF}$ and $\boldsymbol{D}_{BB}$ can be derived by solving the following optimization problem

$$\underset{\boldsymbol{A}_{RF}, \boldsymbol{D}_{BB}}{\text{minimize}} \quad \|\boldsymbol{F}_{opt} - \boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F^2, \tag{3.11a}$$

$$\text{subject to} \quad |(\boldsymbol{A}_{RF})_{i,j}| = 1, \forall i, j, \tag{3.11b}$$

$$\|\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F^2 = N_s, \tag{3.11c}$$

where $\boldsymbol{F}_{opt}$ is the optimal fully digital precoder obtained through the SVD operation of the channel matrix. Constraint (3.11a) directly leads to maximum spectral efficiency of the hybrid D/A precoding system. Therefore, the spectral efficiency of hybrid D/A fully-connected antenna array is compared with the fully digital system [10, 13]. (3.11b) is the constant modulus constraint, restricting all the non-zero elements to the same amplitude, while (3.11c) represents the normalized transmit power constraint. The main challenge in solving the above problem is that it is a non-convex optimization problem since the constraint in (3.11b) is non-convex [10, 17–22].

Minimizing over both variables simultaneously is intractable. However, minimization with respect to one variable while keeping the other variables fixed also known as the alternating minimization approach [19, 49], leads to straightforward and analytical solutions. Therefore, instead of solving a single minimization problem over two variables, the alternating minimization approach is adopted. A solution for each of the alternating minimization steps can be developed (i) for the reduced-complexity structure, (ii) for the fully-connected antenna array structure [10, 19, 41], the non-convex problem can be expanded into a series of convex sub-problems. Specifically, in this case the optimization problem is divided into two sub-optimization-problems: (3.11a) without constraints for solving the digital precoder $\boldsymbol{D}_{BB}$, and (3.11a) subject to constraint (3.11b) for obtaining the analog precoder $\boldsymbol{A}_{RF}$. Correspondingly, the alternative optimization procedure will involve three phases. In Phase-I, the analog precoder $\boldsymbol{A}_{RF}$ is initialized randomly, the target is to optimize the digital precoder $\boldsymbol{D}_{BB}$, while in Phase-II, the digital precoder $\boldsymbol{D}_{BB}$ that is optimized from Phase-I, is given and analog precoder $\boldsymbol{A}_{RF}$ is optimized. In Phase-III, with $\boldsymbol{D}_{BB}$ and $\boldsymbol{A}_{RF}$ obtained in Phase-I and II, constraint (3.11c) is enforced to modify these two precoders, so that the constraints (3.11b) and (3.11c) are satisfied. The design for these three phases is presented as follows:

3.4.1 Phase-I: Digital Precoder $\boldsymbol{D}_{BB}$

In Phase-I, the goal is to design the digital precoder $\boldsymbol{D}_{BB}$ with preliminary analog precoder $\boldsymbol{A}_{RF}$. Notice that only the normalized power constraint (3.11c) affects the design of the digital precoder $\boldsymbol{D}_{BB}$. We first ignore this constraint and Problem (3.11a) reduces to

$$\underset{\boldsymbol{D}_{BB}}{\text{minimize}} \quad \|\boldsymbol{F}_{opt} - \boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F, \tag{3.12}$$

where the optimal $\boldsymbol{D}_{BB}$ is achieved through least squares, resulting in

$$\boldsymbol{D}_{BB} = (\boldsymbol{A}_{RF}^H \boldsymbol{A}_{RF})^{-1} \boldsymbol{A}_{RF}^H \boldsymbol{F}_{opt}. \tag{3.13}$$

After obtaining the digital precoder, the design of the analog precoder is discussed in Phase-II and the constraint (3.11c) will be recalled in Phase-III.

3.4.2 Phase-II: Analog Precoder A_{RF}

Given the digital precoder $\boldsymbol{D}_{BB}$, the analog precoder is optimized as follows:

$$\underset{\boldsymbol{A}_{RF}}{\text{minimize}} \quad \|\boldsymbol{F}_{opt} - \boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F^2 \ \text{ subject to } \quad |(\boldsymbol{A}_{RF})_{i,j}| = 1, \forall i, j. \tag{3.14}$$

Unfortunately, the constraint in Problem (3.14) is a non-convex constraint, as mentioned previously this constraint (3.11b) can incur a high computational complexity to the system to solve since it cannot be relaxed. A novel efficient method to deal with this constraint is to convert this constraint into an additional unitary condition such that $\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_{N_{RF}}$, resulting in the reformulations of Problem (3.14) as

$$\underset{\boldsymbol{X}}{\text{minimize}} \quad \|\boldsymbol{F}_{opt} - \boldsymbol{X}\boldsymbol{D}_{BB}\|_F^2 \ \text{ subject to } \quad \boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_{N_{RF}}, \tag{3.15}$$

where $\boldsymbol{X}$ is the optimization parameter that needs to be optimized. The constraint set $\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_{N_{RF}}$ is a submanifold of $\mathbb{C}^{N_t \times N_{RF}^t}$ called the Stiefel manifold. The main purpose of proposing the unitary condition is, in (3.14) the constraint is based on non-convex constant modulus which is hard to solve directly. In this case, we develop new solutions for the hybrid precoder design problem. Before proposing method to approach the problem in (3.14), we show several properties of the optimum unconstrained precoder ($\boldsymbol{F}_{opt}$). Initially, it can be noticed that $\boldsymbol{F}_{opt}$ is a semi-unitary matrix, as it corresponds to the first columns $\boldsymbol{N}_s$ of the right SVD of $\boldsymbol{H}$. Since $\boldsymbol{F}_{opt}^H\boldsymbol{F}_{opt} = \boldsymbol{I}$, its factorization should also have a semi-unitary structure, $(A_{RF}D_{BB}) * (A_{RF}D_{BB}) \approx \boldsymbol{I}$ [10, 19]. This observation justifies the relaxation unitary constraints that we will contemplate when solving (3.14). This can help us to propose problem (3.15) that involves matrix $\boldsymbol{X}$ based on the unitary constraint. Now, proposing (3.15) is found to simplify the analog preocder design, then this design is applied at the Frobenius norm (3.14) to reduce the gap between the unconstrained precoder $\boldsymbol{F}_{opt}$ and the hybrid precoders. Therefore, (3.15) used as a criteria to simplify the non-convex constraint and make sure that $\boldsymbol{X}$ is an optimal solution for an initial stage, then extracting the phases of $\boldsymbol{X}$ to design $\boldsymbol{A}_{RF}$ which is used in (3.14) to find the performance of the hybrid precoding as compared to the unconstrained fully digital precoder.

In the upcoming subsections, optimal $\boldsymbol{X}$ is determined by the following two proposed algorithms. After obtaining $\boldsymbol{X}$, an Euclidean distance problem is intended and formulated between the analog precoder $\boldsymbol{A}_{RF}$ and the known optimal $\boldsymbol{X}$ in order to help in designing the analog precoder $\boldsymbol{A}_{RF}$ which satisfies the constant modulus constraint in (3.11b).

3.4.2.1 Algorithm 1 Stiefel Manifold optimization based algorithm

In the first algorithm, a Stiefel manifold optimization based geometry method is proposed which satisfies the unitary condition i.e., $\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_{N_{RF}}$, and determines a near optimal $\boldsymbol{X} \in \mathbb{C}^{N_t \times N_{RF}^t}$ solution to (3.15). This algorithm operates by finding the gradient of (3.15) in the tangent plane of the manifold at the point $\boldsymbol{X}^\alpha$ as shown in Fig. 3.2, where α denotes the number of iterations. A curve is then found on this manifold along the projected negative gradient. Curvilinear search is proposed along this curve to evaluate the next point for iteration which is $\boldsymbol{X}^{\alpha+1}$. The reason for selecting curvilinear search over retraction is its lower complexity and its convergence to a stationary point. The Stiefel manifold optimization consists of five steps; 1) manifold structuring; 2) tangent space refinement; 3) Riemannian gradient; 4) curvilinear search; and 5) step-size calculation. Let us discuss these steps in detail.

Step-1 *Manifold Structuring:* The standard notation for representing a Stiefel manifold is $\mathcal{V}_{N_{RF}^t}^{N_t}$. In the manifold theory, the embedded submanifold is a smooth subset of an ambient space $\mathbb{C}^{N_t \times N_{RF}^t}$, and the Stiefel manifold will be viewed as an embedded sub-manifold.

Step-2 *Tangent Space Refinement of a Manifold Structure:* The second step is to comprehend the tangent space of the Stiefel manifold $\mathcal{V}_{N_{RF}^t}^{N_t}$ at a current point $\boldsymbol{X}$. Each point $\boldsymbol{X}$ in an embedded sub-manifold having a tangent space which is represented as $\mathcal{T}_{\boldsymbol{X}}\mathcal{V}_{N_{RF}^t}^{N_t}$. The tangent space consists of a tangent vector $\boldsymbol{\xi}$ of the curves, thus, the tangent space at point $\boldsymbol{X}$ can be written as

$$\mathcal{T}_{\boldsymbol{X}}\mathcal{V}_{N_{RF}^t}^{N_t} = \{\boldsymbol{Z} \in \mathbb{C}^{N_t \times N_{RF}^t} : \mathbb{R}\{\boldsymbol{Z} \odot \boldsymbol{X}^H\} = \mathbf{0}_{N_t \times N_{RF}^t}\}. \tag{3.16}$$

where $\boldsymbol{Z} \in \mathcal{T}_{\boldsymbol{X}}\mathcal{V}_{N_{RF}^t}^{N_t}$, is an element of $\mathbb{C}^{N_t \times N_{RF}^t}$, satisfying (15). In a manifold, each tangent space is given with an inner product. In this algorithm, Riemannian metric is used as the inner product which varies smoothly from point to point and is given by

$$\langle \boldsymbol{Z}_1, \boldsymbol{Z}_2 \rangle = \mathrm{Tr}(\boldsymbol{Z}_1^H \boldsymbol{Z}_2). \tag{3.17}$$

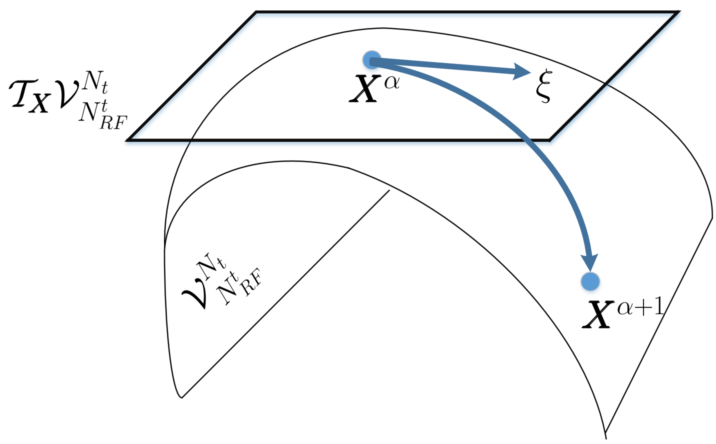


Figure 3.2: Block diagram of an illustration of tangent space and normal space of Stiefel manifold.

Step-3 *The Riemannian Gradient:* The computation of the gradient descent direction is based on the Riemannian gradient structure of the Stiefel manifold. Generally, the Riemannian gradient of f at point $\boldsymbol{X}$ is simply the linear projection of the ordinary gradient of f onto the tangent space. Our Stiefel manifold is a Riemannian submanifold of the Euclidean space. Therefore, one of the tangent vectors is associated with a negative Riemannian gradient which is expressed as the direction of the greatest decrease in the function. Intuitively speaking, this algorithm consists of choosing the search direction as the negative Riemannian gradient in the tangent space at the current iterate, computing the line search, move in this direction and then get back to the manifold using the retraction. In fact, the Riemannian gradient at $\boldsymbol{X}$ is defined as a tangent vector, grad$f(\boldsymbol{X})$, obtained by projecting the Euclidean gradient, $\bigtriangledown f(\boldsymbol{X}) \in \mathbb{C}^{N_t \times N_{RF}^t}$, onto the tangent space $\mathcal{T}_{\boldsymbol{X}}\mathcal{V}_{N_{RF}^t}^{N_t}$ to be

$$\text{grad} f|_{\mathcal{V}_{N_{RF}^t}^{N_t}}(\boldsymbol{X}) = \mathbf{Proj}_{\boldsymbol{X}} \bigtriangledown f(\boldsymbol{X}) = (\boldsymbol{I} - \boldsymbol{X}\boldsymbol{X}^H) \bigtriangledown f(\boldsymbol{X}), \tag{3.18}$$

where $\bigtriangledown f(\boldsymbol{X})$, named the Euclidean gradient of (3.15), is derived as

$$\begin{aligned}\bigtriangledown f(\boldsymbol{X}) &= \frac{\partial}{\partial \boldsymbol{X}} \quad \text{Tr}\left\{\|\boldsymbol{F}_{\text{opt}} - \boldsymbol{X}\boldsymbol{D}_{\text{BB}}\|_{\text{F}}^2\right\}, \\ &= \frac{\partial}{\partial \boldsymbol{X}}\text{Tr}[\boldsymbol{F}_{\text{opt}}\boldsymbol{F}_{\text{opt}}^{\text{H}} - \boldsymbol{X}\boldsymbol{D}_{\text{BB}}\boldsymbol{F}_{\text{opt}}^{\text{H}} - \boldsymbol{F}_{\text{opt}}\boldsymbol{D}_{\text{BB}}^{\text{H}}\boldsymbol{X}^{\text{H}} - \boldsymbol{A}_{\text{RF}}\boldsymbol{D}_{\text{BB}}\boldsymbol{D}_{\text{BB}}^{\text{H}}\boldsymbol{X}^{\text{H}}], \\ &= -\boldsymbol{F}_{opt}\boldsymbol{D}_{BB}^{H} + \boldsymbol{X}\boldsymbol{D}_{BB}\boldsymbol{D}_{BB}^{H}.\end{aligned} \tag{3.19}$$

Step-4 *The Curvilinear Search:* The ideal method to find the next points on the manifold is called retraction [50]. It is worth mentioning that calculating retraction which is mathematically equivalent to exponential mapping in Stiefel manifold optimization requires SVD, which involves high computational complexity. Therefore, to find the next point $\boldsymbol{X}^{\alpha+1}(\tau)$, Cayley transformation is utilized such that

$$\boldsymbol{X}^{\alpha+1}(\tau) \;=\; (\boldsymbol{I}+\frac{\tau}{2}\boldsymbol{W})^{-1}(\boldsymbol{I}-\frac{\tau}{2}\boldsymbol{W})\boldsymbol{X}^{\alpha}, \tag{3.20}$$

where $\boldsymbol{W}=\boldsymbol{X}\bigtriangledown f^{H}(\boldsymbol{X})-\bigtriangledown f(\boldsymbol{X})\boldsymbol{X}^{H}$ and τ is the step-size. For brevity, the main parts of the minimization algorithm by applying curvilinear search $\boldsymbol{X}(\tau)$ are: 1) provide an initial point $\boldsymbol{X}^{\alpha}$; 2) for $\alpha=1,\cdots$, generate new point $\boldsymbol{X}^{\alpha+1}$ from $\boldsymbol{X}^{\alpha}$ via a curvilinear search along $\boldsymbol{X}^{\alpha+1}(\tau)=(\boldsymbol{I}+\frac{\tau}{2}\boldsymbol{W})^{-1}(\boldsymbol{I}-\frac{\tau}{2}\boldsymbol{W})\boldsymbol{X}^{\alpha}$. As invert $(\boldsymbol{I}+\frac{\tau}{2}\boldsymbol{W})^{-1}\in\mathbb{C}^{N_t\times N_t}$ is quite complex due to a large number of antennas, by applying Sherman-Morrison-Woodbury (SMW) formula [50], (3.20) is expressed as

$$\boldsymbol{X}^{\alpha+1}(\tau)=\boldsymbol{X}^{\alpha}-\tau^{\alpha}\boldsymbol{K}(\boldsymbol{I}+\frac{\tau}{2}\boldsymbol{P}^{H}\boldsymbol{K})^{-1}\boldsymbol{P}^{H}\boldsymbol{X}^{\alpha}, \tag{3.21}$$

where $\boldsymbol{K}=[\bigtriangledown f(\boldsymbol{X}),\boldsymbol{U}]$ and $\boldsymbol{P}=[\boldsymbol{U},-\bigtriangledown f(\boldsymbol{X})]$. The details are presented in Appendix A. From (3.21), it is observed that $(\boldsymbol{I}+\boldsymbol{P}^{H}\boldsymbol{K})$ needs to be inverted. However, the dimension of $(\boldsymbol{I}+\boldsymbol{P}^{H}\boldsymbol{K})$ is only $2N_{RF}^{t}\times 2N_{RF}^{t}$ where $N_{RF}^{t} << N_t$, therefore, resulting in lower computational complexity.

Step-5 *Calculating Step-Size:* Since the procedure of finding the suitable step size in (3.21) introduces additional constraints, therefore, the Barzilai-Borwein (BB) step-size is employed and applied to (3.21) to speed up the gradient computation. BB chooses a simple step-size as

$$\tau^{\alpha} \;=\; \frac{\mathrm{Tr}\Big((\triangle\boldsymbol{X})^{T}\triangle\boldsymbol{X}\Big)}{|\mathrm{Tr}\Big((\triangle\boldsymbol{X})^{T}\bigtriangledown(\triangle\boldsymbol{X})\Big)|}, \tag{3.22}$$

where $\triangle\boldsymbol{X}=\boldsymbol{X}^{\alpha}-\boldsymbol{X}^{\alpha-1}$ and $\bigtriangledown(\triangle\boldsymbol{X})=\bigtriangledown f(\boldsymbol{X}^{\alpha})-\bigtriangledown f(\boldsymbol{X}^{\alpha-1})$ [51]. In order to ensure the global convergence of our BB step-size, a globalization technique is proposed to guarantee global convergence by regulating the step-sizes. Next points of $\boldsymbol{X}^{\alpha+1}(\tau^{\alpha+1})$ are iteratively generated where $\tau^{\alpha+1}=\tau^{\alpha}\delta^{\varrho}$ and ϱ is the smallest integer satisfying

$$f(\boldsymbol{X}^{\alpha}(\tau))\leq\Upsilon^{\alpha}+\iota_1\tau^{\alpha}\bigtriangledown f(\boldsymbol{X}^{\alpha}(0)), \tag{3.23}$$

where each reference $\Upsilon^{\alpha+1}$ is possessed to be the convex combination of Υ^{α} and $f(\boldsymbol{X}^{\alpha+1})$ as $\Upsilon^{\alpha+1} = \frac{(\eta\varepsilon^{\alpha}\Upsilon^{\alpha}+f(\boldsymbol{X}^{\alpha+1}))}{\varepsilon^{\alpha+1}}$, where $\varepsilon^{\alpha+1} = \eta\varepsilon^{\alpha}+1$, $\varepsilon^{0} = 1$ and set $\eta_1 \in (0,1)$ and $\iota_1 \in (0,1)$.

Note that, in Algorithm 1 to build ASO, it is necessary to compare vectors which are produced in different tangent spaces. Essentially on manifolds, every tangent vector belongs to a tangent space specific to its root point on manifolds. Typically, these vectors from different tangent spaces cannot be compared. An efficient way is to apply a vector transport along with the chosen curvilinear search. Ideally, this happens when the gradient at the current iteration is combined with the search direction at the previous iteration. Therefore, the following simple procedure is used to transport a tangent vector from $\boldsymbol{X}^{\alpha}$ to $\boldsymbol{X}^{\alpha+1}$ which is represented as

$$\mathrm{Transp}_{\boldsymbol{X}^{\alpha}\to\boldsymbol{X}^{\alpha+1}} = (\boldsymbol{I} - \boldsymbol{X}^{\alpha+1}(\boldsymbol{X}^{\alpha+1})^{H}) \bigtriangledown f(\boldsymbol{X}). \tag{3.24}$$

Algorithm 1 in the top of the next page summaries ASO procedure to design $\boldsymbol{X}$. Importantly, the currently obtained $\boldsymbol{X}$ is just the solution to Problem (3.15) without considering the constant modulus constraint (3.11b). To consider this constraint (3.11b), we minimize the Euclidean distance between the known $\boldsymbol{X}$ and the analog precoder which is formulated as follows

$$\text{minimize} \quad \|\boldsymbol{A}_{RF} - \boldsymbol{X}\|_F^2 \quad \text{subject to} \quad |(\boldsymbol{A}_{RF})_{i,j}| = 1, \forall i,j. \tag{3.25}$$

Note that

$$\|\boldsymbol{A}_{RF} - \boldsymbol{X}\|_F^2 = 2N_{RF}^{t} - 2\mathrm{Tr}\{\mathbb{R}\{\boldsymbol{A}_{RF}\boldsymbol{X}\}\}. \tag{3.26}$$

Ideally, as all the elements in $\boldsymbol{A}_{RF}$ have to satisfy a constant modulus i.e., $\frac{1}{\sqrt{N_t}}$, the phases of all the elements in $\boldsymbol{A}_{RF}$ are the same as the elements of the corresponding orthogonal $\boldsymbol{X}$. This point is proven by utilizing the properties of complex numbers. As matrix $\boldsymbol{X}$ has N_{RF}^{t} orthogonal columns, and matrix $\boldsymbol{A}_{RF}$ needs constant modulus, i.e., $\frac{1}{\sqrt{N_t}}$, the minimum value can be achieved between $\boldsymbol{X}$ and $\boldsymbol{A}_{RF}$ when a matrix $\boldsymbol{A}_{RF}$ has the same phase components as $\boldsymbol{X}$, i.e.,

$$\boldsymbol{A}_{RF} = \frac{1}{\sqrt{N_t}} e^{j\,\arg(\boldsymbol{X})}. \tag{3.27}$$

***Remark* 2.** *Complexity Advantages over MO-AltMin algorithm [19]: Note that superior steps are adopted in Algorithm 1 ASO, which has a lower computational complexity. In*

Algorithm 1 Analog Precoding Based on Stiefel manifold (ASO).

Initialize $\varphi_{min} = 0$, $\varphi_{max} = \frac{C}{P_{\text{total}}}$ and small threshold value $\varpi > 0$;

while Calculate, $\mathbf{d}_0 = -\text{grad} f(\boldsymbol{X}^0)$; **do**

  Choose BB step size, $f(\mathrm{X}^\alpha(\tau)) \leq \Upsilon^\alpha + \iota_1 \tau f'(Y^\alpha(0))\,\mathbf{d}_0$;

  $\tau \leftarrow \delta\tau$;

  Update the next point $\boldsymbol{X}^{\alpha+1} \leftarrow \boldsymbol{X}(\tau)$, by using (3.21), $\varepsilon^{\alpha+1} \leftarrow \eta\varepsilon^\alpha + 1$ and $\Upsilon^{\alpha+1} \leftarrow \frac{(\eta\varepsilon^\alpha \Upsilon^\alpha + f(\boldsymbol{X}^{\alpha+1}))}{\varepsilon^{\alpha+1}}$;

  Calculate the gradient of the new point $\text{grad} f(\boldsymbol{X}^{\alpha+1})$; by utilizing (3.18) and (3.19);

  Transport to the new tangent space, $\text{Transp}_{\boldsymbol{X}^\alpha \to \boldsymbol{X}^{\alpha+1}}$;

  Set $\tau \leftarrow \max(\min(\tau^{\alpha+1,1}, \tau_K), \tau_k)$;

  $\alpha \leftarrow \alpha + 1$;

EndWhile, a stopping criterion triggers

contrast to the MO-AlitMin method where the SVD operation is needed in curvilinear search, this operation can be avoided in ASO by using Cayley transform. Different from MO-AltMin that applies Armijo backtracking step-size with additional constraints, BB step-size is utilized in ASO which does not require any additional constraint as shown in (3.22) and also provides fast convergence.

It is seen from (3.27) that the analog precoder phases of $\boldsymbol{A}_{RF}$ are extracted from the phases of the precoder $\boldsymbol{X}$, and it satisfies the constant modulus constraint in (3.11b). When the number of antennas is very large, the computational complexity of Algorithm 1 is quite high, and it is unwieldy to deal with the constant modulus constraint as will be analyzed and demonstrated in Sections V and VI. Therefore, a low-complexity algorithm to design $\boldsymbol{A}_{RF}$ is proposed in the following.

3.4.2.2 Algorithm 2 Particle Swarm optimization based algorithm

In the second algorithm, we still use the alternating minimization method to design $\boldsymbol{A}_{RF}$, but based on particle swarm optimization (APSO). APSO is a heuristic algorithm which is popular for solving optimization problems due to the fact that it is simple to implement, quickly converges to the desired solution and guarantees the near global solution. It is robust against local minima which makes it appealing for real-time applications [52]. In the framework of APSO, a collection of individuals, known as agents, move within the search space of the optimization problem. To seek better positions within the search space, the velocity of search agents are iteratively updated according to rules generated by social behavior of fish schooling or bird flocking. Furthermore, in each iteration of APSO, velocity of each agent is adjusted towards the best location and toward the best agent. At the end, the position of an agent represents the solution to the optimization problem. To devise the optimization problem, exploiting the same novel way and the reason of reducing Problem (3.14) into Problem (3.15), an $N_t \times N_{RF}^t$ dimensional matrix $\boldsymbol{Q}$ is introduced and replaced with $\boldsymbol{A}_{RF}$ to be optimized, resulting in re-designing of Problem (3.14) as

$$\underset{\boldsymbol{Q}}{\text{minimize}} \quad \|\boldsymbol{F}_{opt} - \boldsymbol{Q}\boldsymbol{D}_{BB}\|_F^2 \ \text{ subject to } \quad \boldsymbol{Q}^H\boldsymbol{Q} = \boldsymbol{I}_{N_{RF}}. \tag{3.28}$$

The framework details of the algorithm are structured as follows. Initialize E agents with random positions $\boldsymbol{Q}_1(0), \boldsymbol{Q}_2(0), \cdots, \boldsymbol{Q}_E(0)$, such that $\boldsymbol{Q}_e(0)^H\boldsymbol{Q}_e(0) = \boldsymbol{I}_{N_{RF}}$, where $e = \{1, 2, \cdots, E\}$. The position of the agent which minimizes (3.28) is represented as $\boldsymbol{F}_{\text{best}}$ and is the global best agent. After that, the velocity of all the E agents $\boldsymbol{V}_1(0), \boldsymbol{V}_2(0), \cdots, \boldsymbol{V}_E(0)$ are randomly initialized. After that, the velocity $\boldsymbol{V}_e$ and position $\boldsymbol{Q}_e$ of the e-th agent are updated as

$$\boldsymbol{V}_e(\alpha+1) = \boldsymbol{V}_e(\alpha) + c_1\bar{\boldsymbol{W}}_1 \odot (\boldsymbol{Q}_{e_{\text{best}}}(\alpha) - \boldsymbol{Q}_{e_{\text{current}}}(\alpha)) + c_2\bar{\boldsymbol{W}}_2 \odot (\boldsymbol{F}_{\text{best}} - \boldsymbol{Q}_{e_{\text{current}}}(\alpha)) \tag{3.29}$$

$$\boldsymbol{Q}_e(\alpha+1) = \boldsymbol{Q}_e(\alpha) + \boldsymbol{V}_e(\alpha+1), \qquad \text{e} = 1, 2, \cdots, \text{E}, \tag{3.30}$$

where $\bar{\boldsymbol{W}}_1$ and $\bar{\boldsymbol{W}}_2$ are uniformly distributed random matrices where each element is selected between 0 to 1. c_1 and c_2 are positive acceleration coefficients. $\boldsymbol{Q}_{e_{\text{best}}}$ is the best and $\boldsymbol{Q}_{e_{\text{current}}}$ is the current of the e-th agent. In the first iteration $\boldsymbol{Q}_{e_{\text{best}}} = \boldsymbol{Q}_{e_{\text{current}}}(0)$ and then this step is repeated for each agent. After that, each agent keeps track of its own best position, which is in terms of achieving the minimum value to Problem (3.28). Once the position of the e-th agent is updated, its fitness $\boldsymbol{Q}_e(\alpha+1)$ is evaluated, where the fitness is the reciprocal value of the objective function. If the updated fitness of the agent is more than the previous best-fitness of the agent, then $\boldsymbol{Q}_{e_{\text{best}}}(\alpha) = \boldsymbol{Q}_{e_{\text{current}}}(\alpha)$. Likewise, the best-fitness of the e-th agent is set to its current-fitness. Finally, we compare the fitness of all the agents, and the agent which minimizes (3.28) is compared with previous $\boldsymbol{F}_{\text{best}}$. The e-th agent which minimizes (3.28) becomes the global best agent $\boldsymbol{F}_{\text{best}}$. Once the algorithm converges to a global minimum, then $\boldsymbol{F}_{\text{best}} = \boldsymbol{Q}_{\text{best}}$. The main conditions to guarantee the convergence of PSO $c_1 + c_2 < 3.8$ [53]. Algorithm 2 on the top of the next page, provides APSO architecture to design $\boldsymbol{Q}$.

The solution $\boldsymbol{Q}$ that is obtained from APSO cannot be directly applied to Problem (3.14) as an optimal analog precoder $\boldsymbol{A}_{RF}$, because it does not satisfy the constant modulus constraint. To find the feasible analog precoder which is perfectly close to the optimal solution, we first normalize the obtained matrix $\boldsymbol{Q}$ via $\frac{\boldsymbol{Q}_{\text{best}}}{\sqrt{\boldsymbol{Q}^H_{\text{best}}\boldsymbol{Q}_{\text{best}}}}$. Next, the Euclidean distance between $\boldsymbol{Q}$ and the optimal analog precoder $\boldsymbol{A}_{RF}$ is formulated as $||\boldsymbol{A}_{RF} - \boldsymbol{Q}||_F^2$ and is made to approach zero. By exploiting the same approach as (3.25), the feasible solution is evaluated as

$$\boldsymbol{A}_{RF} = \frac{1}{\sqrt{N_t}} e^{j \arg(\boldsymbol{Q})}. \tag{3.31}$$

3.4.3 Phase-III: Alternating Minimization

In Phase-I and Phase-II, digital and analog precoders are designed separately. However, the power constraint in (3.11c) is not considered in Phase-I and Phase-II. We deal with this issue in Phase-III. Specifically, we normalize $\boldsymbol{D}_{BB}$ that is obtained in Phase-I by $\frac{\sqrt{N_s}\boldsymbol{D}_{BB}}{||\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}||_F}$. This normalized factor will not affect the design of hybrid precoders, as the Euclidean distance between $\boldsymbol{F}_{opt}$ and the hybrid precoders is sufficiently small. To explain this point clearly, in [19] lemma 1 illustrates that when the Euclidean distance of problem (3.11a) sufficiently small at step 3, where step 3 ignored the power constraint as formulated in (3.12), the normalization step will also achieve a small distance to the optimal unconstrained fully digital precoder. If it is clear enough, all $(\boldsymbol{F}_{opt})_i$, $i = 1, \cdots, N_s$, have unit l_2 norm which yields $||\boldsymbol{F}_{opt}||_F^2 = N_s$, $||\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}||_F^2 = N_s$ in (10c) [10,41]. Therefore, the normalization step ensures that the transmit power constraint is exactly satisfied. The detailed procedure in Phase-III is summarized in Algorithm 3.

In the proposed alternating minimization algorithm, the Euclidean distance in (3.11a) is minimized via steps 3 and 4. Observe that each step of the alternating minimization

Algorithm 2 Analog Precoding Based on Particle Swarm optimization (APSO).

Initialize E weights with random values random values $\boldsymbol{Q}_1(0), \boldsymbol{Q}_2(0), \cdots, \boldsymbol{Q}_E(0)$, such that $\boldsymbol{Q}_E(0)^H \boldsymbol{Q}_E(0) = \boldsymbol{I}_{N_{RF}}$;
Initialize the velocity of each weight $\boldsymbol{V}_1(0), \boldsymbol{V}_2(0), \cdots, \boldsymbol{V}_E(0)$;
Evaluate the fitness of each weight using (3.28). Find the global best weight, $\boldsymbol{F}_{\text{best}}$ among all and its fitness;
for $\alpha = 1, 2, \cdots;$ **do**
    for $e = 1, 2, \cdots, E;$ **do**
        Evaluate the velocity of the weight e using (3.29) and update agent e by utilizing (3.30);
        Evaluate the fitness of the agent e, using (3.28);
        if (weight-current-fitness < weight-current-fitness);
        (weight-best-fitness = weight-current-fitness) ;
        $\boldsymbol{Q}_{e_{\text{best}}} = \boldsymbol{Q}_{e_{\text{current}}}$;
    end for
    Find the global best-weight, $\boldsymbol{Q}_{\text{best}}$.
end for

process in Algorithm 3 treats the steps exactly. Thus, it can guarantee that no increase in the Frobenius norm can be attained and thus the error term is guaranteed to decrease monotonically with each iteration [10, 19]. Moreover, the objective function based on the Frobenius norm is also non-negative. Consequently, these two properties together guarantee that the proposed algorithms based on ASO and APSO can converge to a feasible solution.

Note that, the two algorithms, the optimal Stiefel manifold-based analog precoding ASO and the simplified PSO-based analogy precoding APSO, are proposed for analog precoder design. These algorithms are implemented for the fully-connected antenna array system. Let us now propose hybrid D/A precoder for partially-connected antenna array systems in the next section.

Algorithm 3 Analog Precoding Based on Particle Swarm optimization (APSO).

1: $\boldsymbol{F}_{opt}$;

2: Initialize $\boldsymbol{A}_{RF}$ with random phases.

3: **Iterations**, $\alpha = 1, 2, \cdots$;

4: Update $\boldsymbol{D}_{BB}$ by solving (3.13) where the solution is given $\boldsymbol{D}_{BB} = (\boldsymbol{A}_{RF}^H \boldsymbol{A}_{RF})^{-1} \boldsymbol{A}_{RF}^H \boldsymbol{F}_{opt}$.;

5: Fix $\boldsymbol{D}_{BB}$ and then optimize $\boldsymbol{A}_{RF}$ in (3.14) by utilizing either Algorithm 1, $\boldsymbol{A}_{RF} = \frac{1}{\sqrt{N_t}} e^{j \arg(\boldsymbol{X})}$; Or Algorithm 2, $\boldsymbol{A}_{RF} = \frac{1}{\sqrt{N_t}} e^{j \arg(\boldsymbol{Q})}$;

6: Normalize the digital part $\boldsymbol{D}_{BB}$ by factor $\frac{\sqrt{N_s}\boldsymbol{D}_{BB}}{||\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}||_F}$.

3.5 Proposed Partially-connected Antenna Array Hybrid D/A Precoding

The block diagram of partially-connected antenna array hybrid D/A precoding mm-Wave system is shown in Fig. 1(c). N_s data streams are passed through the N_{RF}^t RF chains and fed to $M = N_t/N_{RF}^t$ transmit antennas to perform analog precoding. The analog precoder is represented as $\boldsymbol{A}_{RF} = \text{blkdiag}[\bar{\boldsymbol{a}}_1, \bar{\boldsymbol{a}}_2, \cdots, \bar{\boldsymbol{a}}_{N_{RF}^t}] = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{N_{RF}^t}]$, where each column in the analog precoder is $\boldsymbol{a}_n = [\boldsymbol{0}_{1\times M(n-1)}, \bar{\boldsymbol{a}}_n^T, \boldsymbol{0}_{1\times M(N_{RF}^t - n)}]^T \in \mathbb{C}^{N_t \times 1}$, where $n = 1, 2, \cdots, N_{RF}^t$. The analog precoder vector is denoted as $\bar{\boldsymbol{a}}_n \in \mathbb{C}^{M\times 1}$, where all the elements of $\bar{\boldsymbol{a}}_n$ have the same amplitude but different phase shifts. In the partially-connected antenna array, $\boldsymbol{G}$ is defined as the joint hybrid precoding matrix, given by [54]

$$\boldsymbol{G} = \boldsymbol{A}_{RF}\boldsymbol{D}_{BB}. \tag{3.32}$$

In order to achieve the maximum capacity of the system, an appropriate $\boldsymbol{G}$ needs to be chosen from the following optimization problem [20, 21]

$$\boldsymbol{G}_{opt} = \underset{\boldsymbol{G}}{\text{argmax}} \quad \log_2 \det\left(\boldsymbol{I}_{N_r} + \frac{\boldsymbol{H}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{H}^H}{\sigma^2}\right), \tag{3.33}$$

where **C**1 and **C**2 are defined as:

C1: The Frobenius norm of $\boldsymbol{G}$ should satisfy $||\boldsymbol{G}||_F^2 = ||\text{blkdiag}[\bar{\boldsymbol{g}}_1, \cdots, \bar{\boldsymbol{g}}_{N_{RF}^t}]||_F^2 \leq N_{RF}^t$ to meet the total transmit power constraint. The joint hybrid precoding matrix is composed as $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_{N^t_{RF}}]$, where $\boldsymbol{g}_n = [\mathbf{0}_{1\times M(n-1)}, \bar{\boldsymbol{g}}_n^T, \mathbf{0}_{1\times M(N^t_{RF}-n)}]^T \in \mathbb{C}^{N_t\times 1}$ and $\bar{\boldsymbol{g}}_n = d_n \bar{\boldsymbol{a}}_n$.

C2: All the non-zero elements of $\boldsymbol{A}_{RF}$ have the same amplitude which is equal to 1.

The optimization problem in (3.33) is an $N_t \times N^t_{RF}$ matrix optimization problem, which is difficult to solve. Since $\boldsymbol{G}$ is a block diagonal matrix with N^t_{RF} column vectors, we can solve N^t_{RF} independent vector optimization problems instead of directly solving the original matrix optimization problem. As a result, the vector $\boldsymbol{g}_n$ can be optimized independently rather than the whole matrix $\boldsymbol{G}$. The constraints of the vector $\boldsymbol{g}_n$ can be obtained from constraints **C**1 and **C**2 as follows:

T1: The Frobenius norm of $\boldsymbol{g}_n$ should satisfy $||\boldsymbol{g}_n||_F^2 \leq 1$ to meet the transmit power constraint.

T2: All the non-zero elements of $\boldsymbol{g}_n$ have the same amplitude which is equal to 1.

Based on the above discussion, the optimization problem for vector $\boldsymbol{g}_n$ can be formulated as

$$\boldsymbol{g}_{n_{opt}} = \underset{\boldsymbol{g}_n}{\operatorname{argmax}} C(\bar{\boldsymbol{g}}_n) = \log_2\left(1 + \frac{\boldsymbol{g}_n^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{g}_n}{\sigma^2}\right), \tag{3.34}$$

In the following, two algorithms are proposed to solve Problem (3.34).

3.5.1 Complex Circle Manifold Optimization (CMO)

In this section, the optimization problem in (3.34) is solved by using a manifold optimization algorithm which has been widely used in geometry, topology and analysis [19, 55]. In fact, when a manifold reduces from a matrix to a vector, it becomes a complex circle manifold. Then, the precoder vector $\bar{\boldsymbol{g}}_n$ belongs to the complex circle manifold $\mathcal{V}_M^{cc} = \left\{\bar{\boldsymbol{g}}_n \in \mathbb{C}^{M\times 1} : \|\bar{g}_{n_1}\| = \cdots = \|\bar{g}_{n_M}\| = 1\right\}$ [56], which makes all the non-zero elements have the same amplitude satisfying (**T2**) i.e., the constant modulus constraint. Each $\bar{g}_m$ is a point on the complex plane $\mathbb{C}$ and is numerically represented in $\mathbb{R}^2$, where $m = 1, \cdots, M$. The detailed description of the complex circle manifold optimization (CMO) algorithm is summarized in Algorithm 4, which consists of three main steps: a tangent space, Riemannian gradient and retraction of complex circle manifold.

Let us first derive the tangent space. Each point in the complex circle manifold has a neighbourhood that is homomorphic to the Euclidean space. The notion of a tangent space $\mathcal{T}_{\bar{\boldsymbol{g}}_n}\mathcal{V}_M^{cc}$ of the manifold is a vector space which has the same dimension as the manifold that is composed of a tangent vector ξ of the curves. For example, for each point $\bar{\boldsymbol{g}}_n$ in the manifold $\mathcal{V}_M^{cc}$, there is a tangent space linked to it as

$$\mathcal{T}_{\bar{\boldsymbol{g}}_n}\mathcal{V}_M^{cc} = \{\boldsymbol{t} \in \mathbb{C}^{M\times 1} : \mathbb{R}\{\boldsymbol{t} \odot \bar{\boldsymbol{g}}_n^H\} = \boldsymbol{0}_{M\times 1}\}. \tag{3.35}$$

In the Euclidean space, one of the tangent vectors is associated with a negative Riemannian gradient which is expressed as the direction of the greatest increase of the function. The complex circle manifold is a submanifold, and the Riemannian gradient at $\bar{\boldsymbol{g}}_n$ is defined as a tangent vector $\mathrm{grad}f(\bar{\boldsymbol{g}}_n)$ obtained by projecting the Euclidean gradient $\bigtriangledown f(\bar{\boldsymbol{g}}_n) \in \mathbb{C}^{M\times 1}$ onto the tangent space $\mathcal{T}_{\bar{\boldsymbol{g}}_n}\mathcal{V}_M^{cc}$

$$\mathrm{grad}f(\bar{\boldsymbol{g}}_n) = \mathrm{Proj}_{(\bar{\boldsymbol{g}}_n)} \bigtriangledown f(\bar{\boldsymbol{g}}_n) = \bigtriangledown f(\bar{\boldsymbol{g}}_n) - \mathbb{R}\{\bigtriangledown f(\bar{\boldsymbol{g}}_n) \odot \bar{\boldsymbol{g}}_n^H\} \odot \bar{\boldsymbol{g}}_n, \tag{3.36}$$

where $\bigtriangledown f(\bar{\boldsymbol{g}}_n)$ can be obtained by using (3.34)

$$\bigtriangledown_{\bar{\boldsymbol{g}}} f(\bar{\boldsymbol{g}}_n) = \left(\frac{1}{\ln(2)\left(1 + \frac{\boldsymbol{g}_n^H \boldsymbol{H}^H \boldsymbol{H} \boldsymbol{g}_n}{\sigma^2}\right)} \frac{2(\boldsymbol{H}^H \boldsymbol{H} \boldsymbol{g}_n)}{\sigma^2} \right). \tag{3.37}$$

Ideal retraction is called the exponential mapping when it maps the tangent vector $\xi \in \mathcal{T}_{\bar{\boldsymbol{g}}_n}\mathcal{V}_M^{cc}$ to a point along a geodesic curve [55]. A geodesic on $\mathcal{V}_M^{cc}$ is defined as the shortest distance between two points in the manifold. In general, the exponential mapping $\mathrm{Exp}_{\bar{\boldsymbol{g}}_n}(\tau^\alpha \boldsymbol{p}^\alpha)$ of a tangent vector ξ in an embedded manifold $\mathcal{V}_M^{cc}$ is evaluated as

$$\mathrm{Exp}_{\bar{\boldsymbol{g}}_n} : \mathcal{T}_{\bar{\boldsymbol{g}}_n}\mathcal{V}_M^{cc} \to \mathcal{V}_M^{cc}, \mathrm{Exp}_{\bar{\boldsymbol{g}}_n}(\tau^\alpha \boldsymbol{p}^\alpha) = \exp\{-\varepsilon \mathbb{R}\{\bigtriangledown f(\bar{\boldsymbol{g}}_n)\bar{\boldsymbol{g}}_n^H - \bar{\boldsymbol{g}}_n \bigtriangledown f^H(\bar{\boldsymbol{g}}_n))\}\}\bar{\boldsymbol{g}}_n, \tag{3.38}$$

where $\exp(\cdot)$ is the matrix exponential and α is the iteration number. After finding the tangent space, Riemannian gradient and retraction of the complex circle manifold, we are ready to design $\bar{\boldsymbol{g}}_n$. In order to guarantee that the objective function is non-decreasing at each iteration, BB line step length of ε and Polak-Ribiere parameter are used [57]. On manifolds, every tangent vector belongs to a tangent space specific to its root point. Therefore, Polak-Ribiere step and the conjugate direction step can be applied to CMO. In addition, to combine them together, a mathematical tool that is capable of transporting a tangent vector $\boldsymbol{p}$ from $\bar{\boldsymbol{g}}_n^\alpha$ to $\bar{\boldsymbol{g}}_n^{\alpha+1}$ is introduced as follows

$$\mathrm{Transp}_{\bar{\boldsymbol{g}}_n^\alpha \to \bar{\boldsymbol{g}}_n^{\alpha+1}} : \boldsymbol{p} \to \boldsymbol{p} - \mathbb{R}\{\boldsymbol{p} \odot (\bar{\boldsymbol{g}}_n^{\alpha+1})^H\} \odot \bar{\boldsymbol{g}}_n^{\alpha+1}. \tag{3.39}$$

(3.39) is applied at step 7 of Algorithm 4 which is located in the top of this page. According to 4.3.1 in [56], the convergence of CMO can be guaranteed. The joint precoder $\bar{\boldsymbol{g}}_n$ that is obtained from CMO should be decomposed into $\bar{\boldsymbol{a}}_n$ and d_n, yielding

$$d_n = \frac{1}{M}, \quad \bar{\boldsymbol{a}}_n = \bar{\boldsymbol{g}}_n. \tag{3.40}$$

d_n is chosen to satisfy the power constraint $\mathbf{T}1$ and each RF chain has a unit power,

Algorithm 4 Complex Circle Manifold Optimization (CMO)

1: **for the n-th RF chain** $(n \leq N_{RF}^t)$ **do**

Require: $\boldsymbol{A}_{RF}$ with random phases.

2: $\boldsymbol{p}^0 = -\mathrm{grad} f(\bar{\boldsymbol{g}}_n^0)$

3: **Repeat**

4: $\varepsilon^\alpha = \mathrm{Linesearch}(\bar{\boldsymbol{g}}_n^\alpha, \boldsymbol{p}^\alpha, \bar{\boldsymbol{g}}_n^{\alpha+1})$ ▷ BB step size;

5: $\bar{\boldsymbol{g}}_n^{\alpha+1} = \mathrm{Exp}_{\bar{\boldsymbol{g}}_n^\alpha}(\tau^\alpha \boldsymbol{p}^\alpha)$ ▷ search for the next point by applying (3.38);

6: $\boldsymbol{\kappa}^{\alpha+1} = \mathrm{grad} f(\bar{\boldsymbol{g}}_n^{\alpha+1})$ ▷ Determine the gradient according to (3.37) and (3.36);

7: $(\boldsymbol{p}^\alpha)^+ = \mathrm{Transp}_{\bar{\boldsymbol{g}}_n^\alpha \to \bar{\boldsymbol{g}}_n^{\alpha+1}}$ ▷ find the vector transport of the Riemannian gradient and conjugate direction;

8: Select Polak-Ribiere $\gamma^{\alpha+1}$;

9: $\boldsymbol{p}^{\alpha+1} = -\boldsymbol{\kappa}^{\alpha+1} + \gamma^{\alpha+1}(\boldsymbol{p}^\alpha)^+$ ▷ the new conjugate direction;

10: $\alpha \leftarrow \alpha + 1$;

11: **Until** a stopping criterion triggers.

12: **end for**

so that, the total transmit power constraint can be guaranteed. The solution obtained by the proposed CMO algorithm satisfies the non-convexity constraints. However, the computational complexity of retraction step is quite high, especially when a large number of antennas are applied in the system. For this reason, it is important to develop a low-complexity hybrid D/A precoding scheme.

3.5.2 Particle Swarm Optimization (PSO)

In this subsection, a PSO-based algorithm named SPSO is proposed for the partially-connected antenna array system due to its simplicity, fast convergence speed, and robustness against local minima [52, 58, 59]. In SPSO, the coordinates of an agent represent the solution to the problem. Furthermore, in each iteration of SPSO, the direction of each agent is adjusted towards the best location and the best agent. For problem in (3.34), the M dimension vector $\bar{\boldsymbol{g}}_n$ needs to be optimized. Specifically, denote $\mathcal{E}$ as the set of initial agents with random positions $\bar{\boldsymbol{g}}_{n_1}(0), \bar{\boldsymbol{g}}_{n_2}(0), \cdots, \bar{\boldsymbol{g}}_{n_\mathcal{E}}(0)$. The position of the agent that maximizes (3.34) is denoted as $\boldsymbol{f}_{\text{best}}$ and is the global best agent. All positions are normalized to ensure that the power of n-th RF chain is unity. After that, the velocities of all the $\mathcal{E}$ agents $\boldsymbol{v}_{n_1}(0), \boldsymbol{v}_{n_2}(0), \cdots, \boldsymbol{v}_{n_\mathcal{E}}(0)$ are randomly initialized. The positions of the agents are used to evaluate (3.34). After initialization, update the velocity $\boldsymbol{v}_{n_e}$ and position $\bar{\boldsymbol{g}}_{n_e}$ of e-th agent as

$$\boldsymbol{v}_{n_e}(\alpha+1) = \boldsymbol{v}_{n_e}(\alpha) + c_1 \bar{\boldsymbol{w}}_1 \odot \left(\bar{\boldsymbol{g}}_{n_{e_{\text{best}}}}(\alpha) - \bar{\boldsymbol{g}}_{n_{e_{\text{current}}}}(\alpha)\right) + c_2 \bar{\boldsymbol{w}}_2 \odot \left(\boldsymbol{f}_{\text{best}} - \bar{\boldsymbol{g}}_{n_{e_{\text{current}}}}(\alpha)\right), \tag{3.41}$$

$$\bar{\boldsymbol{g}}_{n_e}(\alpha+1) = \bar{\boldsymbol{g}}_{n_e}(\alpha) + \boldsymbol{v}_{n_e}(\alpha+1), \qquad \mathrm{e} = 1, 2, \cdots, \mathcal{E}, \tag{3.42}$$

where $\bar{\boldsymbol{w}}_1$ and $\bar{\boldsymbol{w}}_2$ are uniformly distributed random vectors. $\bar{\boldsymbol{g}}_{n_{e_{\text{best}}}}$ is the best and $\bar{\boldsymbol{g}}_{n_{e_{\text{current}}}}$ is the e-th agent at the current iteration. c_1 and c_2 are positive acceleration coefficients. To guarantee the convergence of SPSO, parameters $\bar{\boldsymbol{w}}_1$ and $\bar{\boldsymbol{w}}_2$ elements should be chosen between 0 and 1 and $c_1 + c_2 < 3.8$, respectively [52, 53]. In the first iteration, $\bar{\boldsymbol{g}}_{n_{e_{\text{best}}}}(0) = \bar{\boldsymbol{g}}_{n_{e_{\text{current}}}}(0)$ for $e = [1, 2, \cdots, \mathcal{E}]$. After that, each agent keeps track of its own best position. Once the position of the e-th agent is updated, its fitness $C_n(\bar{\boldsymbol{g}}_{n_e}(\alpha+1))$ is evaluated. If the updated fitness of the agent is more than the previous best-fitness of the agent, then $\bar{\boldsymbol{g}}_{n_{e_{\text{best}}}}(\alpha) = \bar{\boldsymbol{g}}_{n_{e_{\text{current}}}}(\alpha)$. Similarly, the best-fitness of the e-th agent is set equal to its current-fitness. Finally, the fitness of all the agents and the agent which maximizes (3.34) is compared with the previous $\boldsymbol{f}_{\text{best}}$ and the one that maximizes (3.34) becomes the global best agent $\boldsymbol{f}_{\text{best}}$. When the iterative procedure is complete, $\bar{\boldsymbol{g}}_{n_{e_{\text{best}}}} = \boldsymbol{f}_{\text{best}}$.

The detailed procedures of SPSO are given in Algorithm 5, shown at the top of this page. Again, the joint precoder that is obtained from SPSO should be decomposed into $\bar{\boldsymbol{a}}_n$ and d_n. Unfortunately, we cannot choose $\bar{\boldsymbol{g}}_n$ freely, as $\bar{\boldsymbol{g}}_n$ does not satisfy the constraint (**T1**) and (**T2**). To find the feasible analog and digital precoders $\bar{\boldsymbol{a}}_n$ and d_n that are close to the optimal solution in terms of Euclidean distance, let us first normalize the obtained precoder as $\bar{\boldsymbol{g}}_n = \frac{\bar{\boldsymbol{g}}_{n_{e_{\text{best}}}}}{\sqrt{\bar{\boldsymbol{g}}^H_{n_{e_{\text{best}}}} \bar{\boldsymbol{g}}_{n_{e_{\text{best}}}}}}$. The optimization problem in (3.34) can then be reformulated as

$$\min \ \mathbb{E}\left\{\|\bar{\boldsymbol{g}}_n - d_n\bar{\boldsymbol{a}}_n\|_2^2\right\},$$

i.e.,

$$\min \ \mathbb{E}\left\{\bar{\boldsymbol{g}}_n^H\bar{\boldsymbol{g}}_n - d_n\mathbb{R}(\bar{\boldsymbol{a}}_n^H\bar{\boldsymbol{g}}_n + \bar{\boldsymbol{g}}_n^H\bar{\boldsymbol{a}}_n) + d_n^2\bar{\boldsymbol{a}}_n^H\bar{\boldsymbol{a}}_n\right\}. \tag{3.43}$$

As aforementioned in the system model, $\bar{\boldsymbol{a}}_n$ has M non-zero elements with the same amplitude, the optimal $\bar{\boldsymbol{a}}_n$ to maximize (3.34) is

$$\bar{\boldsymbol{a}}_n = e^{j\angle(\bar{\boldsymbol{g}}_n)}, \tag{3.44}$$

where $\angle(\bar{\boldsymbol{g}}_n)$ represents the phase vector of $\bar{\boldsymbol{g}}_n$ and all the elements of $\bar{\boldsymbol{a}}_n$ share the same amplitude. Substituting (3.44) into (3.43), d_n is obtained as

$$d_n = \mathbb{R}(\bar{\boldsymbol{a}}_n^H\bar{\boldsymbol{g}}_n + \bar{\boldsymbol{g}}_n^H\bar{\boldsymbol{a}}_n) = \frac{1}{M}\|\bar{\boldsymbol{g}}_n\|_1. \tag{3.45}$$

Note that $d_n\bar{\boldsymbol{a}}_n$ for all the partially-connected antenna array and the optimal solution $\bar{\boldsymbol{g}}_n$ have a similar form. After obtaining the precoder vector $\bar{\boldsymbol{g}}_n$ for the n-th RF chain, the same algorithm is applied to the other chains. After optimizing the last RF chain, the optimal precoding matrices $\boldsymbol{D}_{BB}, \boldsymbol{A}_{RF}$, and $\boldsymbol{G}$ are obtained. In SPSO, as $||\bar{\boldsymbol{g}}_n||_1 < ||\bar{\boldsymbol{g}}_n||_2^2 = 1$, the total transmit power constraint **C**1 is satisfied, i.e., $||\boldsymbol{G}||_F^2 = ||\text{blkdiag}\{\bar{\boldsymbol{g}}_1, \cdots, \bar{\boldsymbol{g}}_{N_{RF}^t}\}||_F^2 \leq N_{RF}^t$. In general, all the obtained vectors $\bar{\boldsymbol{g}}_n$ from CMO and SPSO that are used to construct $\boldsymbol{G}$ have the same amplitude as a result, the second constraint is also satisfied.

3.6 Hybrid D/A mm-Wave Design For Multi-User Large Antenna Arrary Systems

Now, we seek to design the hybrid precoders for the downlink multi-user system in which a BS equipped with large number of antennas N_t, but limited number of RF chains N_{RF}^t, serves K single-antenna users where $N_t >> K$. For such a multi-user system with hybrid precoding architecture at the BS, the optimization problem is formulated based on the

Algorithm 5 Particle Swarm optimization based partially-connected antenna array (SPSO)

for the n**-th RF chain** $(n \leq N_{RF}^{t})$ **do**

Require: E agents random positions $\bar{\boldsymbol{g}}_{n_1}(0), \bar{\boldsymbol{g}}_{n_2}(0), \cdots, \bar{\boldsymbol{g}}_{n_{\mathcal{E}}}(0)$

Require: the velocity of each agent $\boldsymbol{v}_{n_1}(0), \boldsymbol{v}_{n_2}(0), \cdots, \boldsymbol{v}_{n_{\mathcal{E}}}(0)$;

  Evaluate the fitness of candidates by using (3.34). Find the global best $\boldsymbol{f}_{\text{best}}$;

  for $\alpha = 1, 2, \cdots$; **do**

    for $e = 1, 2, \cdots, \mathcal{E}$; **do**

      Evaluate the velocity e by applying (3.41) and update the agent e utilizing (3.42);

      Evaluate the fitness of the agent e, via using (3.34);

      if (agent-current-fitness < agent-best-fitness);

      (agent-best-fitness = agent-current-fitness);

      $\boldsymbol{Q}_{e_{\text{best}}} = \boldsymbol{Q}_{e_{\text{current}}}$;

    end for

    Find the global best-weight, $\boldsymbol{Q}_{\text{best}}$.

  end for

end for

spectral efficiency expression for the multi-user can be expressed as

$$C_k = \max_{\{\boldsymbol{A}_{RF},\, \boldsymbol{D}_{BB}\}} \log(1 + \text{SINR}_k), \tag{3.46a}$$

$$\text{subject to} \quad |(\boldsymbol{A}_{RF})_{i,j}| = 1, \forall i, j, \tag{3.46b}$$

$$\|\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F^2 \quad = K = N_s, \tag{3.46c}$$

where SINR_k is the signal-to-interference-plus-noise ratio (SINR) of the user k which is presented as

$$\text{SINR}_k = \frac{|\boldsymbol{h}_k^H \boldsymbol{A}_{RF} \boldsymbol{d}_{bb_k}|^2}{\sum\limits_{i=1, i\neq k}^{K} |\boldsymbol{h}_i^H \boldsymbol{A}_{RF} \boldsymbol{d}_{bb_i}|^2 + \sigma^2}, \tag{3.47}$$

where $\boldsymbol{h}_k^H \in \mathbb{C}^{N_t}$, is the channel from the BS to the kth user and $\boldsymbol{d}_{bb_k}$ represents the lth column of the digital precoder $\boldsymbol{D}_{BB}$. The first constraints in (3.46b) are subjected to the non-convex constant modulus constraints, restricting all the non-zero elements to have the same amplitude in the analog domain. While, the second constraint (3.46c) represents the total normalized transmit power constraint. In the following, two sections aim to design optimal analog and digital precoders to maximize the sum rate in (3.46) of multi-user mm-Wave systems.

3.6.1 Design of Digital Precoder For Multi-User Systems

In this subsection, a next important step focus on designing the digital precoder for the multi-user system. Practically, to design the low dimension and asymptotically optimal fully digital preocders, here the digital precoder is proposed based on MMSE by exploiting the effective downlink analog channel gains which is expressed as

$$\boldsymbol{D} = D^{MMSE} \cdot \bar{\boldsymbol{H}}^H \boldsymbol{A}_{RF} \, (\boldsymbol{A}_{RF}^H \bar{\boldsymbol{H}}^H \bar{\boldsymbol{H}} \boldsymbol{A}_{RF} + K\sigma^2 \boldsymbol{A}_{RF}^H \boldsymbol{A}_{RF})^{-1}, \tag{3.48}$$

where $\bar{\boldsymbol{H}} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_k]$. Ultimately, in (3.48) D^{MMSE} is denoted as the normalization factors to ensure the power constraint $||\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}||_F^2 = K = N_s$, which is given as

$$D^{MMSE} = \frac{\sqrt{N_s}}{||\bar{\boldsymbol{H}}^H \boldsymbol{A}_{RF} \, (\boldsymbol{A}_{RF}^H \bar{\boldsymbol{H}}^H \bar{\boldsymbol{H}} \boldsymbol{A}_{RF} + K\sigma^2 \boldsymbol{A}_{RF}^H \boldsymbol{A}_{RF})^{-1}||_F} \tag{3.49}$$

By exploiting the effective downlink channel a low-dimension matrix can be generated, then MMSE is utilized to design the feasible digital precoder for multi-user hybrid precoding mm-Wave systems.

3.6.2 Design of Analog Precoder For Multi-User Systems

Now, we explore to design the analog precoder assuming the MMSE digital precoding as in (3.48). Fortunately, it can be observed that in (3.46) the analog precoder matrix $\boldsymbol{A}_{RF}$ is fixed, and also the sum rate with the MMSE digital precoder relies on the analog precoder only through the power constraint $||\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}||_F^2 = N_s$. Therefore, the design of the analog precoder problem can consider only the constant modulus constraints $|(\boldsymbol{A}_{RF})_{i,j}| = 1, \forall i, j$. In this case, the previous proposed algorithms ASO and APSO for fully-connected antenna array structures, CMO and SPSO for partially-connected antenna array structures, can be applied directly to the multi-user hybrid D/A system to design the analog precoder $\boldsymbol{A}_{RF}$. The optimization problem in this scenario can be re-formulated as

$$C_k = \max_{\{\boldsymbol{A}_{RF}\}} \log(1 + \mathrm{SINR}_k), \tag{3.50a}$$

$$\text{s.t.} \qquad |(\boldsymbol{A})_{n_t,n_{RF}}| = 1, \forall n_t, \forall n_{RF}. \tag{3.50b}$$

Unfortunately, the constraint (3.50b) in Problem (5.14a) is a non-convex constraint which can incur a high computational complexity to the system to solve since it cannot be relaxed. Therefore, the same novel efficient method to deal with this constraint that is used in (14) to relax the constant modulus constraint in (3.50b) into an additional unitary condition such that $\boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_{N_{RF}}$, resulting in, the reformulations of Problem (5.14a) can be expressed after replacing the analog precoder matrix $\boldsymbol{A}_{RF}$ with the matrix $\boldsymbol{X} \in \mathbb{C}^{N_t \times N_{RF}}$, as

$$C_k = \max_{\{\boldsymbol{A}_{RF}=\boldsymbol{X}\}} \log(1 + \mathrm{SINR}_k), \tag{3.51a}$$

$$\text{subject to} \quad \boldsymbol{X}^H\boldsymbol{X} = \boldsymbol{I}_{N_{RF}}. \tag{3.51b}$$

Now, the proposed algorithms can be used to design the matrix $\boldsymbol{X}$ with the same analog precoder algorithms steps that are applied in the Section III-B (i.e., ASO and APSO) and Section IV (i.e., CMO and SPSO). The prime objective here is the sum rate in (5.14a) which can be maximized directly. Note here, as the objective function of the spectral efficiency in (5.21a) is different from the one that is maximized in (10) for the single-user system, thus the gradient of the objective function has to be updated for ASO and CMO, which is expressed as

$$\nabla f(\boldsymbol{X}) = \left(\frac{1}{\ln(2)\left(1+\frac{|\boldsymbol{h}_k^H \boldsymbol{X} \boldsymbol{d}_{bb_k}|^2}{|\boldsymbol{h}_i^H \boldsymbol{X} \boldsymbol{d}_{bb_i}|^2+\sigma^2}\right)} \frac{\boldsymbol{X}(\boldsymbol{d}_{bb_k}\boldsymbol{h}_k^H\boldsymbol{h}_k+\boldsymbol{d}_{bb_k}^H\boldsymbol{h}_k\boldsymbol{h}_k^H)\cdot(|\boldsymbol{h}_i^H\boldsymbol{X}\boldsymbol{d}_{bb_i}|^2+\sigma^2)-|\boldsymbol{h}_k^H\boldsymbol{X}\boldsymbol{d}_{bb_k}|^2\cdot\boldsymbol{X}(\boldsymbol{d}_{bb_i}\boldsymbol{h}_i^H\boldsymbol{h}_i+\boldsymbol{d}_{bb_i}^H\boldsymbol{h}_i\boldsymbol{h}_i^H)}{(|\boldsymbol{h}_i^H\boldsymbol{X}\boldsymbol{d}_{bb_i}|^2+\sigma^2)^2} \right) \tag{3.52}$$

where $\nabla f(\boldsymbol{X})$ named the Euclidean gradient of (5.21a). In APSO and SPSO to design the analog precoder an objective function is demanded, thus in multi-user case the sum rate in (5.21a) represents the objective function that is utilized to find the $\boldsymbol{F}_{best}$.

Table 3.1: Computational complexity of fully-connected antenna array.

Algorithm	Number of operations per iterations: Additions	Multiplications	Iterations
MO-AltMin [19]	$2N_s(N_tN_{RF}^t - 1)N_t + N_tN_s(N_tN_{RF}^t - 1) + N_tN_{RF}^t(N_tN_s - 1) + 2N_tN_{RF}^t(N_tN_{RF}^t - 1)$	$2N_{RF}^tN_sN_t^2 + N_t^2N_sN_{RF}^t + N_t^2N_s + 2N_t^2N_{RF}^{t2} + N_tN_{RF}^t$	α
Proposed ASO	$N_{RF}^{t2}(N_s-1)+2N_tN_{RF}^t(N_{RF}^t-1)+2N_t^2(N_{RF}^t-1)+N_{RF}^t(N_t-1)+\frac{N_{RF}^{t3}}{6}$	$2N_t^2(N_{RF}^t+N_s)+N_{RF}^{t2}(N_t^2+N_s)+N_tN_sN_{RF}^{t2}+2N_{RF}^{t2}N_t+\frac{N_{RF}^{t3}}{6}$	α
OMP sparse precoding [10]	$\frac{N_{RF}^{t3}}{6}+N_{RF}^t(N_t-1)LN_s$	$N_sN_{RF}^tN_t^3L$	N_{RF}
GHP [41]	$2\frac{N_s^3}{6}+N_t(N_{RF}^t-1)NsL$	$2N_s^2LN_t+3N_sLN_t^2$	α
PE-AltMin [19]	$\frac{N_t^3}{6} + N_{RF}^t(N_s - 1)N_t + N_{RF}^t(N_t-1)N_s$	$2N_sN_tN_{RF}^t+\frac{N_t^3}{6}$	α
Proposed APSO	$7N_tN_{RF}^tN_sE$	$N_{RF}^tN_s(7N_tE+2E)$	α

3.7 Complexity and Energy Efficiency Analysis

3.7.1 Computational Complexity

3.7.1.1 Fully-connected antenna array

The complexity of the hybrid precoding mm-Wave system is evaluated by taking into account the number of multiplications and additions involved as shown in Table-I. Since the mm-Wave system uses sophisticated large antenna array, the computational complexity is dominated by the number of antennas. Similar to the complexity of MO-AltMin, our proposed ASO complexity consists of four parts: Euclidean gradient, Riemannian gradient, curvilinear search and Transport vectors. It requires $2N_t^2(N_{RF}^t + N_s) + N_{RF}^{t2}(N_t + N_s) + N_tN_sN_{RF}^t + 2N_{RF}^{t2}N_t + \frac{N_{RF}^{t3}}{6}$ multiplications and $N_{RF}^{t2}(N_s-1)+2N_tN_{RF}^t(N_{RF}^t-1)+2N_t^2(N_{RF}^t-1)+N_{RF}^t(N_t-1)+\frac{N_{RF}^{t3}}{6}$ additions. The inverse of matrix in (20) for ASO is derived with the help of cholesky decomposition. Notice that the number of antennas involved in ASO has a smaller impact on the complexity than MO-AltMin and OMP sparse precoding, as shown in Table-I, where L represents the number of multipaths in OMP sparse precoding [10]. APSO complexity consists of two parts, namely velocity and position of the agent. It takes $7N_tN_{RF}^tN_sE$ additions and $N_{RF}^tN_s(7N_tE+2E)$ multiplications per iteration, where E is the number of agents. APSO is preferred over the existing algorithms, due to its low complexity. This is will be proved in Section VI. Certainly, the number of iterations has the impact on the overall complexity of the algorithms. Therefore, the complexity of the whole hybrid D/A algorithms is magnified by the number of required iterations until convergence. Table-3.1 summarizes the computational complexity for the fully-connected antenna array.

3.7.1.2 Partially-connected antenna array

Table-3.2 provides the computational complexity comparison for the partially-connected antenna array. In general, M multiplications and M additions are required to determine $\bar{\boldsymbol{a}}_n$ and d_n respectively. Therefore, the complexity of the proposed CMO algorithm consists of three main steps, i.e., grad$f(\bar{\boldsymbol{g}_n})$, $\nabla f(\bar{\boldsymbol{g}}_n)$, and $\exp(\cdot)$. All these steps in each iteration takes $N_{RF}^t(5M-2+\frac{M^3}{6})$ additions and $N_{RF}^t(2M^2+M^3+\frac{M^3}{6})$ multiplications. According to Lemma 1 in [60], the complexity of the exponential is calculated as eigen-decomposition. SPSO includes $7ME$ additions and $7ME+2E$ multiplications per iteration to determine optimal $\bar{\boldsymbol{g}}_n$. The total complexity of the algorithm is $7ME+M$ additions and $7ME+2E+M$ multiplications. For the hybrid D/A BF system with N_{RF}^t chains, the complexity will be $N_{RF}^t(7ME+M)$ additions and $N_{RF}^t(7ME+2E+M)$ multiplications. In contrast, SVD-based hybrid D/A precoding system [21] takes $N_{RF}^{t^3}M+N_{RF}^{t^3}M^3$ additions and $N_{RF}^{t^3}M+N_{RF}^{t^3}M^3$ multiplications in each iteration.

3.7.2 Energy Efficiency Calculation

RF chain components consume up to 70% of the total transceiver power [14]. The practical energy efficiency of different types of hybrid D/A models have been discussed in [16–18]. Nevertheless, some circuit devices were not included in the hybrid D/A such as the filter

Table 3.2: Computational complexity of partially-connected antenna array.

Algorithms	Number of operations per iterations		Iterations
	Additions	Multiplications	
SVD-based hybrid D/A precoding [21]	$N_{RF}^{t3}M + N_{RF}^{t3}M^3$	$N_{RF}^{t3}M + N_{RF}^{t3}M^3$	α
Analog precoder [16]	$N_{RF}^{t}(M^2 - M)$	$N_{RF}^{t}M^2$	α
Proposed CMO	$N_{RF}^{t}(5M - 2 + \frac{M^3}{6})$	$N_{RF}^{t}(2M^2 + M^3 + \frac{M^3}{6})$	α
Proposed SPSO	$N_{RF}^{t}(7ME+M)$	$N_{RF}^{t}(7ME + 2E + M)$	α

P_{filter} and the base-band amplifier P_B. Therefore, we have taken the power consumption of these two devices into consideration, which are denoted by P_{filter} and P_B, respectively. As each power amplifier is connected to each antenna in our system, the power consumption of each RF chain is modelled as

$$P_{RF} = P_{mixer} + P_{LO} + P_{filter} + P_B, \tag{3.53}$$

where P_{mixer} is a power of mixer, and a P_{LO} is the power for local oscillator. From the energy perspective, a fully-connected antenna array consumes more power because the number of phase shifters scales linearly with the number of RF chains and the number of transmit antennas. As the phase shifters are independent of the number of RF chains in the partially-connected antenna array, it consumes less energy than the fully-connected antenna array. The total powers consumed by a hybrid precoding mm-Wave fully-connected antenna array and partially-connected antenna array systems are given by

$$\begin{aligned} P_{\text{total}}^{full} &= N_{RF}^t(P_{DAC} + P_{RF}) + N_t(P_{PA} + N_{RF}^t P_{PS}) + P_{BP}, \\ P_{\text{total}}^{sub} &= N_{RF}^t(P_{DAC} + P_{RF}) + N_t(P_{PA} + P_{PS}) + P_{BP}, \end{aligned} \tag{3.54}$$

respectively, where the power amplifier is modelled as $P_{PA} = P_t/\gamma_{PA}$, P_{DAC} is the power of the digital-to-analog converter, P_{BP} is the base-band processing power, and P_{PS} is the power consumed by the phase shifters (including the energy for the excitation and the energy for the compensation of insertion loss [61]). The energy efficiency is then given by

$$\gamma = \frac{C}{P_{\text{total}}}. \tag{3.55}$$

3.8 Simulation Results

In this section, simulation results are provided to evaluate the performance of our proposed algorithms. For the fully-connected antenna array, the existing algorithms such as MO-AltMin [19], GHP [41], PE-AltMin [19] and OMP [10] are used for comparison. An optimal fully digital system is also considered as a performance benchmark. For the partially-connected antenna array, an analog precoder system [16] and a hybrid D/A precoding in [18] are shown in the simulation results. In addition, a hybrid D/A precoding based on SVD method in [21] is also considered. A benchmark is applied based on hybrid D/A precoding partially-connected antenna array architecture [21]. The number of channel paths is set to $L = 13$ [13, 27, 30]. The transmitter antenna array is assumed to be a UPA with the antenna spacing $u_q = \frac{\lambda}{2}$ [30]. The AoAs and AoDs are generated independently from the uniform distribution with mean cluster angle uniformly distributed $[0, 2\pi]$, while the angle spread is set to 7.5 [10, 19, 30]. As shown in previous work on mm-Wave systems [10, 13, 19, 41, 48], most of our results in this section are tested with low and medium SNR levels.

Fig. 3.3 shows the convergence behavior of three different algorithms for the fully-connected antenna array system with SNR $= -15$dB. The spectral efficiency achieved is averaged over $10,000$ independent channel generations. It can be seen from Fig. 3.3 that our proposed ASO has faster convergence speed than other algorithms. The curvilinear search and BB enable our proposed ASO to converge within 12 iterations. MO-AltMin converges slower than the proposed ASO which requires about 20 iterations. The convergence performance of the APSO with the population sizes of 6 and 25 are sufficiently close, while the latter requires much higher complexity. Fig. 3.4 shows convergence behavior of the CMO algorithm and SPSO with different numbers of agents in partially-connected antenna array when SNR=0dB. It can be seen from Fig. 3.4 that CMO converges fast

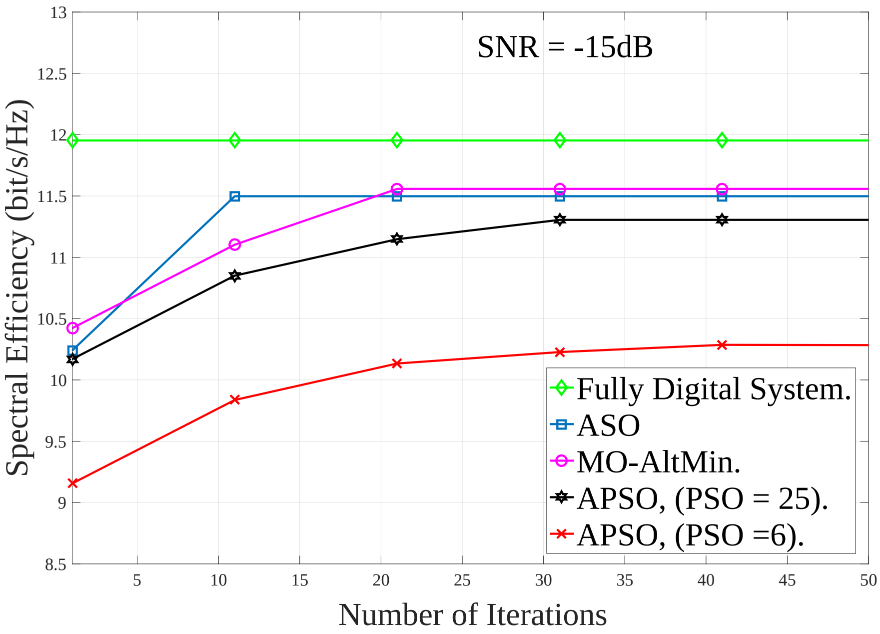


Figure 3.3: Learning curve of a mm-Wave communication system for Fully-connected antenna array algorithms.

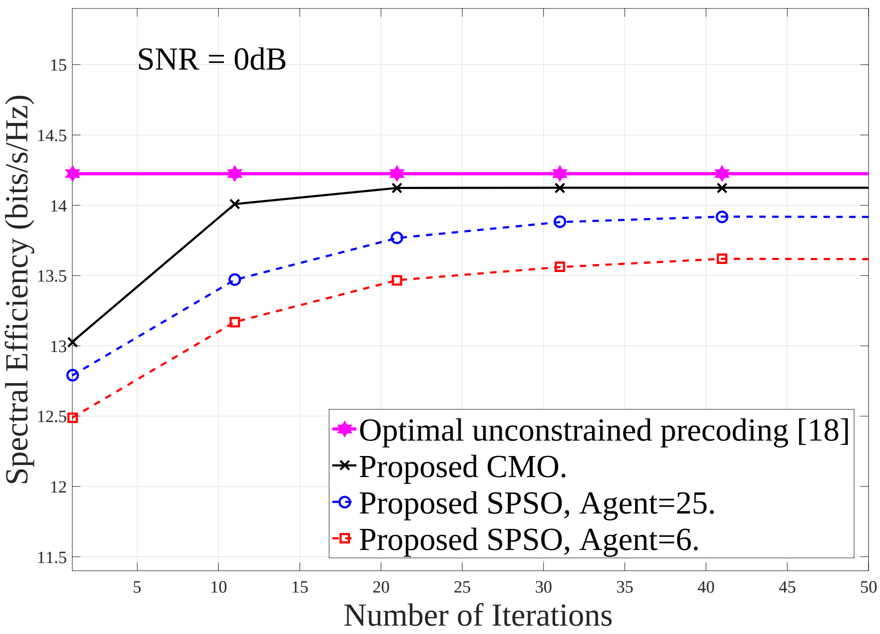


Figure 3.4: Learning curve of a mm-Wave communication system for Partially-connected antenna array algorithms.

within 22 iterations. Meanwhile, by increasing the number of agents, SPSO also converges faster (within 34 iterations) and can also achieve higher spectral efficiency. However, if we apply a lower number of agents, for e.g, when SPSO has 6 agents, it results in a lower spectral efficiency with slower convergence (40 iterations).

Fig. 3.5 plots the spectral efficiency performance for the fully-connected antenna array when $N_{RF}^t = N_{RF}^r = N_s$, as a function of SNR. From Fig. 3.5 it can be seen that the spectral efficiency achieved by OMP is always lower than those achieved by our proposed algorithms and the benchmark. The reason is that it is designed for a special scenario when $N_{RF}^t \geq 2N_s$. The spectral efficiency achieved by ASO is similar to that of MO-AltMin, which is close to the fully digital system. This is because ASO can accurately approximate the optimal precoder which outperforms OMP, GHP, PE-AltMin and APSO. With a lower number of agents, APSO can still achieve higher spectral efficiency than OMP and its spectral efficiency increases by increasing the number of agents. In Fig. 3.6, we investigate the effect of the number of RF chains on the spectral efficiency performance of different algorithms, for the fully-connected antenna array with a fixed numner of data streams. It can be seen that both ASO and MO-AltiMin achieve the same spectral efficiency as that of the fully digital system when $N_{RF} \geq 2N_s$. GHP and APSO need more RF chains to approach the performance of the fully digital system. As mentioned earlier, the PE-AltMin cannot take the advantage of extra RF chains and achieves constant spectral efficiency. OMP achieves a lower spectral efficiency for the case of $N_{RF} = N_s$, it is designed to take the advantage of extra RF chains in the system. The performance loss of ASO algorithm comes from the relaxation of the original problem (13) to be the problem (14). However, we have used the simulation tool as in Fig. 3.5 to show that this loss is small and this relaxation produces an unnoticeable gap from the optimal fully digital system. Moreover, the performance achievements of APSO depends on the relaxation problem and also on the number of agents, where with the high number of agents APSO can achieve close enough to the optimal solution.

Fig. 3.7 and Fig. 3.8 plot the spectral efficiency versus SNR for partially-connected antenna array hybrid precoding mm-Wave systems. To show effect of the number of antennas, we consider different numbers of antennas $N_{RF}^t = 2, N_t = 64, N_r = 16$ and

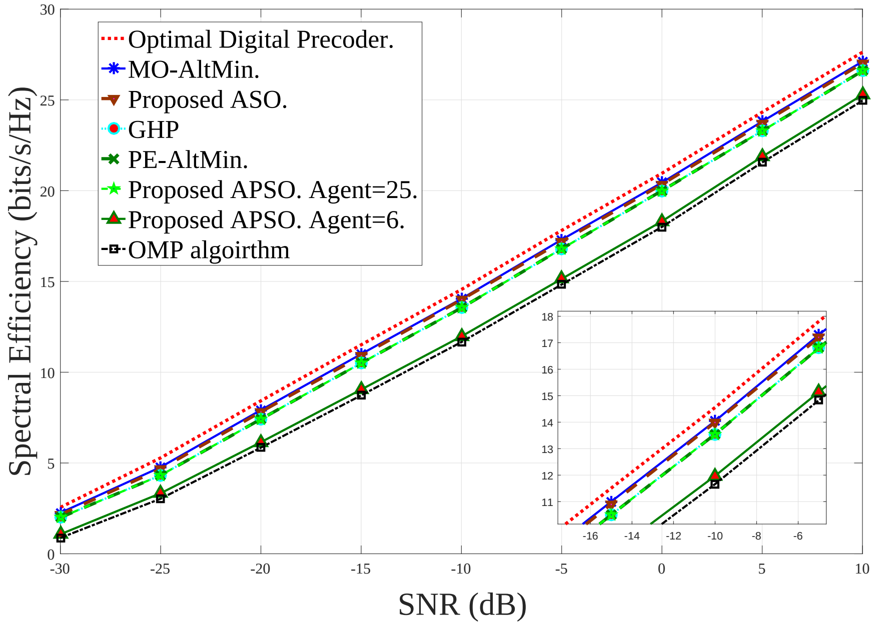


Figure 3.5: Spectral efficiency of fully-connected antenna array mm-Wave system $N_{RF}^t = N_{RF}^r = 2, N_t = 128, N_r = 32$

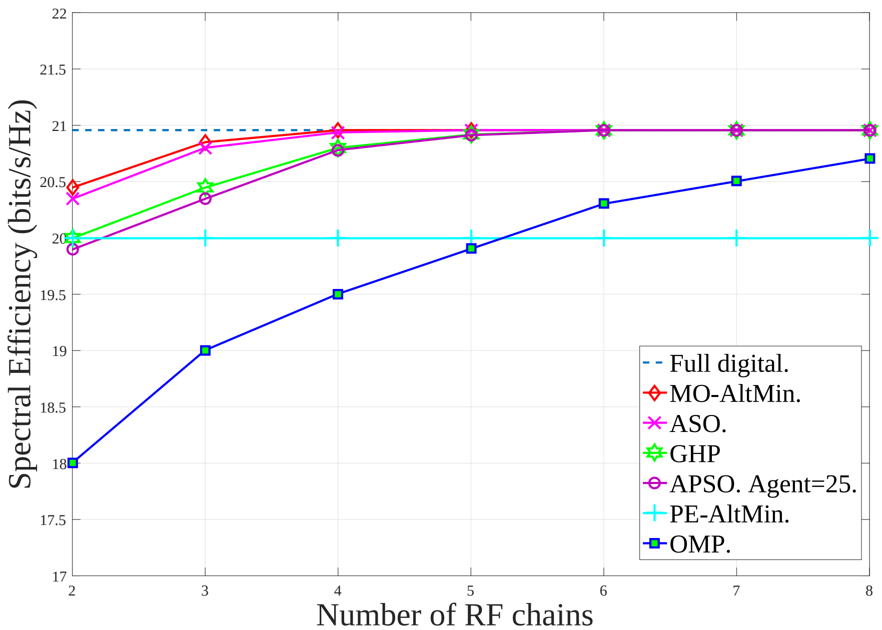


Figure 3.6: Spectral efficiency of partially-connected antenna mm-Wave system when different number of RF chains are applied, $N_t = 128, N_r = 32$

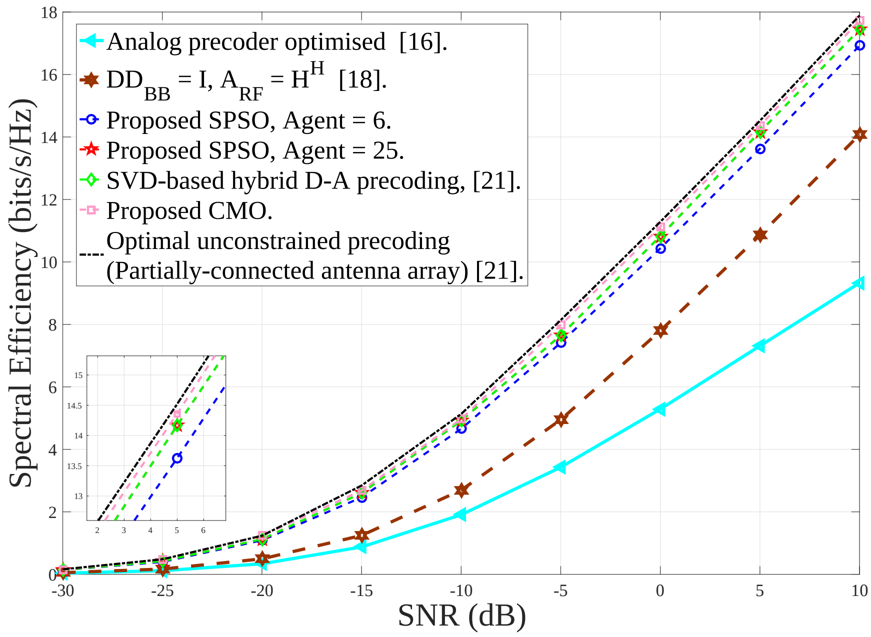


Figure 3.7: Spectral efficiency of partially-connected antenna mm-Wave system when different algorithms are used, $N_{RF}^t = N_{RF}^r = 2, N_t = 64, N_r = 16$

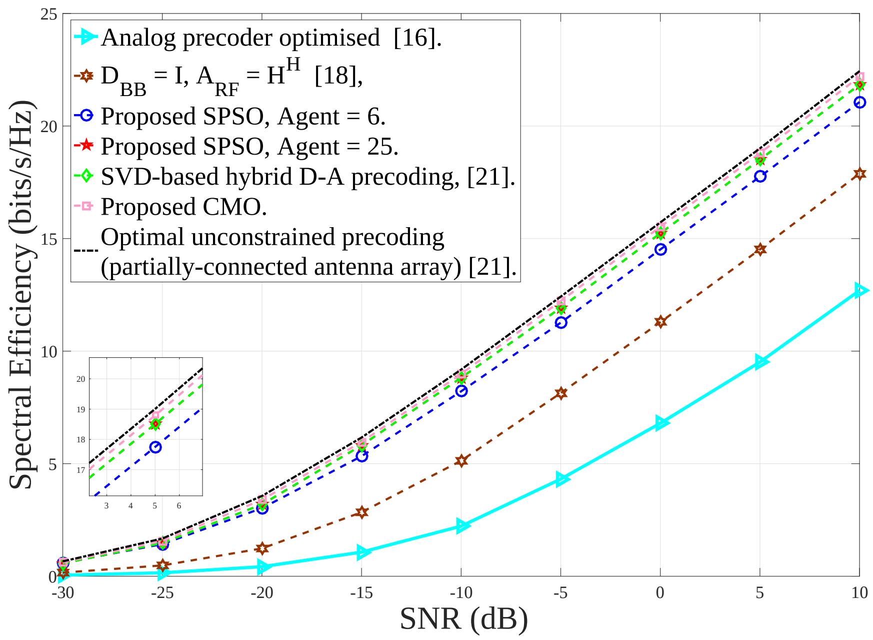


Figure 3.8: Spectral efficiency of partially-connected antenna mm-Wave system when different algorithms are used, $N_{RF}^t = N_{RF}^r = 2, N_t = 128, N_r = 32$

$N_{RF}^t = 2, N_t = 128, N_r = 32$. It can be seen that the spectral efficiency achieved by the analog precoder in [16] is the lowest. When SNR = 0dB the spectral efficiency gap between the analog precoder [16] and the proposed CMO and SPSO is about 4bits/s/Hz, while when SNR = 10dB, the spectral efficiency gap increases to about 9bits/s/Hz. Furthermore, the spectral efficiency achieved by hybrid D/A precoders [18] is lower than that of hybrid D/A precoders based on CMO. In addition, CMO achieves better spectral efficiency than SPSO and SVD-based hybrid D/A precoding. The reason is that CMO can approximate the optimal joint precoder more accurately. SPSO with 25 agents achieves similar spectral efficiency to the SVD-based hybrid D/A precoding [21]. By comparing results in Fig. 3.5 and Fig. 3.8, it is observed that the fully-connected antenna array achieves higher spectral efficiency as compared to partially-connected antenna array system.

Fig. 3.9 plots the sum rate performance for the fully-connected antenna array when $N_{RF}^t = N_{RF}^r = N_s$, as a function of SNR. A significant performance gain can be observed for the proposed algorithms compared to the existing algorithms, and the performance is close enough to the fully digital MMSE system. Fig. 3.9 shows that the sum rate achieved by the algorithm in [13] based ZF method is always lower than those achieved by our proposed algorithms and the fully digital MMSE. The reason is, the algorithm that is designed in [13] for a special scenario when $N_{RF}^t \geq 2N_s$. With a lower number of agents, APSO can still achieve higher sum rate than the algorithm in [13] and its sum rate increases by increasing the number of agents.

Fig. 3.10 illustrates the sum rate performance for the partially-connected antenna array when $N_{RF}^t = N_{RF}^r = N_s$ of the proposed algorithms. Without considering any interference in the system is simulated which is introduced as a benchmark and it is called free interference. It is observed that the proposed algorithms based CMO and SPSO still close enough to the free interference. SPSO with different number of agent is simulated where SPSO with low number of agents achieves the lowest curve as compared to the SPSO with higher number of agents.

Fig. 3.11 and Fig. 3.12 illustrate the impact of imperfect CSI on the spectral efficiency performance of various algorithms with $N_{RF}^t = 2$ and $N_t = 128$. From Fig. 3.11 and Fig. 3.12 it can be seen that OMP and the methods in [18] achieve very low spectral

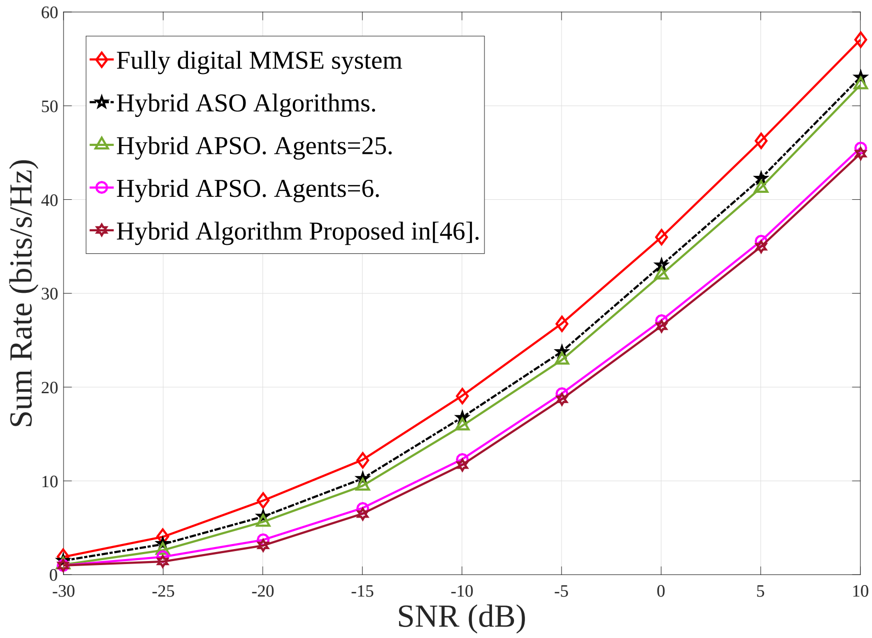


Figure 3.9: Sum rate achieved by different methods for fully-connected antenna array mm-Wave system, $N_{RF}^t = N_{RF}^r = 4, N_t = 128, N_r = 16$

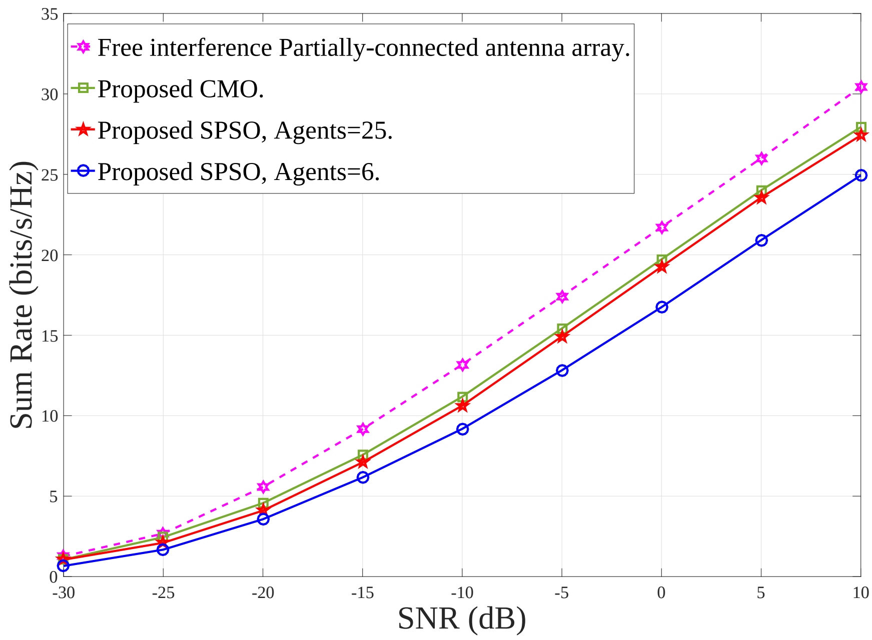


Figure 3.10: Sum rate achieved by different methods fo partially-connected antenna array mm-Wave system, $N_{RF}^t = N_{RF}^r = 4, N_t = 128, N_r = 32$

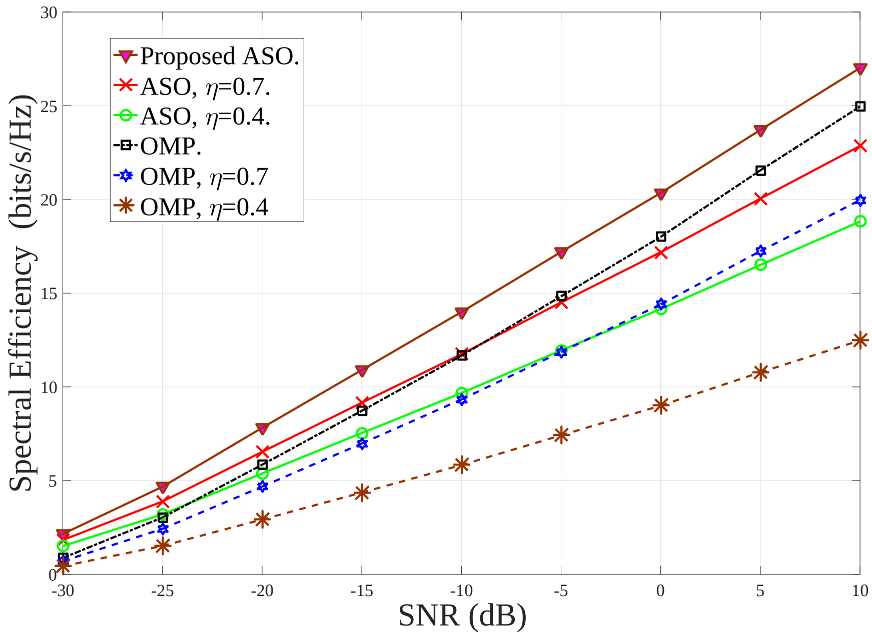


Figure 3.11: Spectral efficiency of hybrid D/A precoding mm-Wave system based on imperfect CSI for fully-connected antenna array algorithms

efficiency in the imperfect channel case with $\eta = 0.4$. The reason is that these two algorithms rely much on the channel estimation to design the analog precoding matrix. Our proposed algorithms have low spectral efficiency in the imperfect case with $\eta = 0.4$, but still can achieve relatively good spectral efficiency as compared to the perfect CSI case and the existing algorithms.

Fig. 3.13 plots the number of multiplications for full antenna array with respect to the number of antennas, as multiplications consume more computation than additions. Fully digital system has high complexity because the number of RF chains is proportional to the number of antennas. From the figure, it is seen that the complexity of MO-AltMin increases exponentially. The complexity of the proposed ASO is lower than that of MO-AltMin and OMP, while the complexity of the OMP increases exponentially but it is still lower than that of MO-AltMin. The complexity of the APSO scheme increases linearly. It is also verified that by increasing the number of agents for APSO, the complexity the system increases. Therefore, in the fully-connected antenna array system, APSO has lowest complexity than all algorithms, including MO-AltMin, ASO, GHP and OMP.

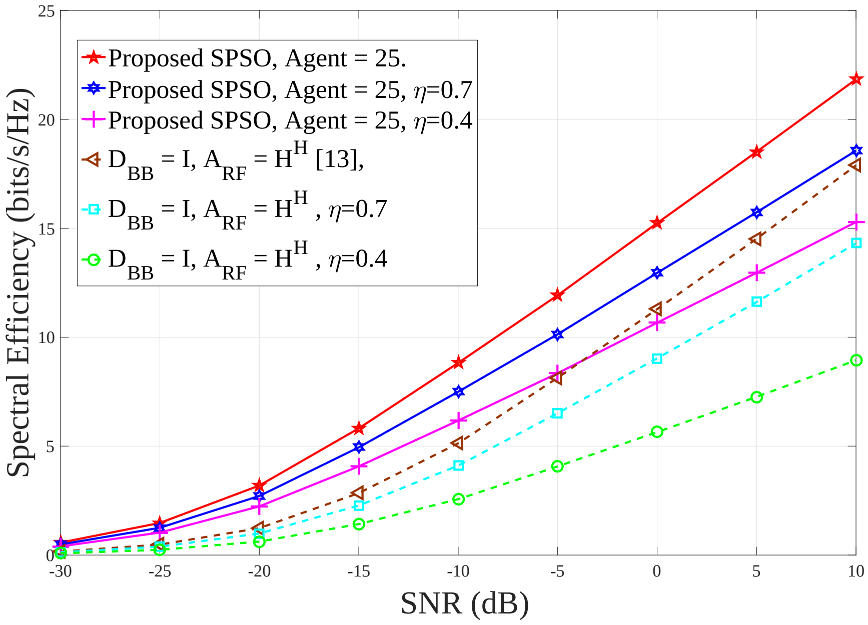


Figure 3.12: Spectral efficiency of hybrid D/A precoding mm-Wave system based on imperfect CSI for partially-connected antenna array algorithms

Fig. 3.14, shows the complexity for the number of selected area. In general, the partially-connected antenna array has lower complexity than the fully-connected antenna array. In the partially-connected antenna array structure, the complexities of CMO and SVD proposed in [21] increase exponentially, and CMO has lower complexity than SVD. The complexity of the SPSO and the analog precoding [16] scheme increases linearly. When SPSO adopts with a larger number of agents, the complexity of the system increases. Therefore, the complexity of the SPSO based method is a little higher than the analog precoding proposed in [16].

In terms of the running times of the proposed algorithms versus the existing algorithms, we report the following: the proposed ASO and CMO take around 35 and 27 second respectively. APSO and SPSO take around 0.0007 second which are the fastest algorithms amongst the proposed and the existing hybrid D/A algorithms. For fully antenna array, the existing algorithms MO-AltMin, OMP, GHP, and PE-AltMin consume 63, 0.075, 0.058, and 0.0011 seconds, respectively. However, the pure analog precoder [16] attains the fastest algorithm as compared to all other algorithms which are summarized in Table 3.3.

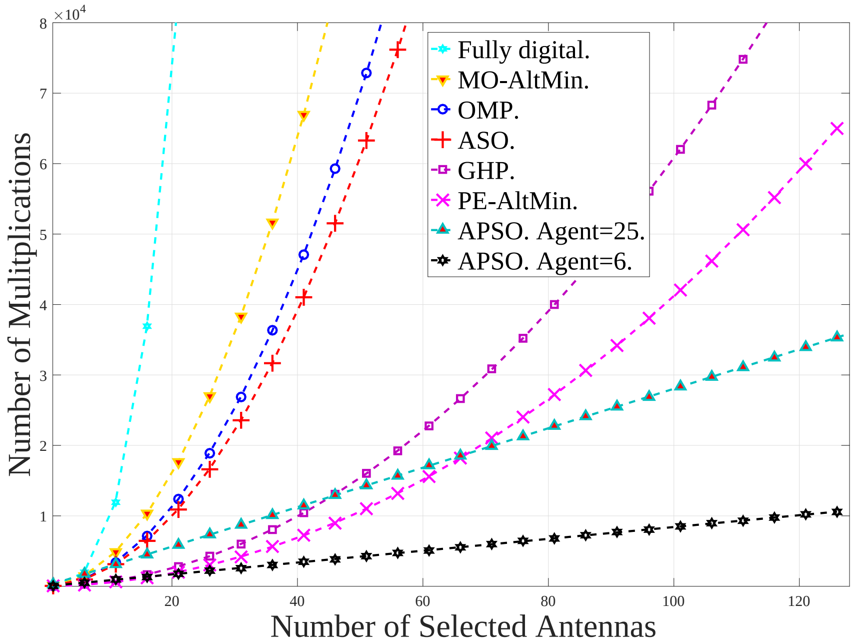


Figure 3.13: Number of multiplications versus the number of selected antennas for fully-connected antenna array complexity

Table 3.3: Running time of the proposed algorithms versus the existing algorithms.

Algorithms	running time	Algorithms	running time
Proposed ASO	35 second	MO-AltMin [19]	63 second
Proposed APSO	0.0007 second	OMP sparse precoding [10]	0.075 second
Proposed CMO	27 second	GHP [41]	0.058 second
Proposed SPSO	0.0007 second	PE-AltMin [19]	0.0011 second
SVD-based hybrid D/A precoding [62]	0.0025 second	Analog precoder [16]	0.0001 second

These running times are averaged over the 500 realizations, implemented in Mathworks' Matlab©.

Fig. 3.15 plots the energy efficiency versus the number of RF chains, N_{RF}^{t}, when $N_t = 128$ and SNR= 0dB. It is assumed that the average path loss is 100dB. The trans-

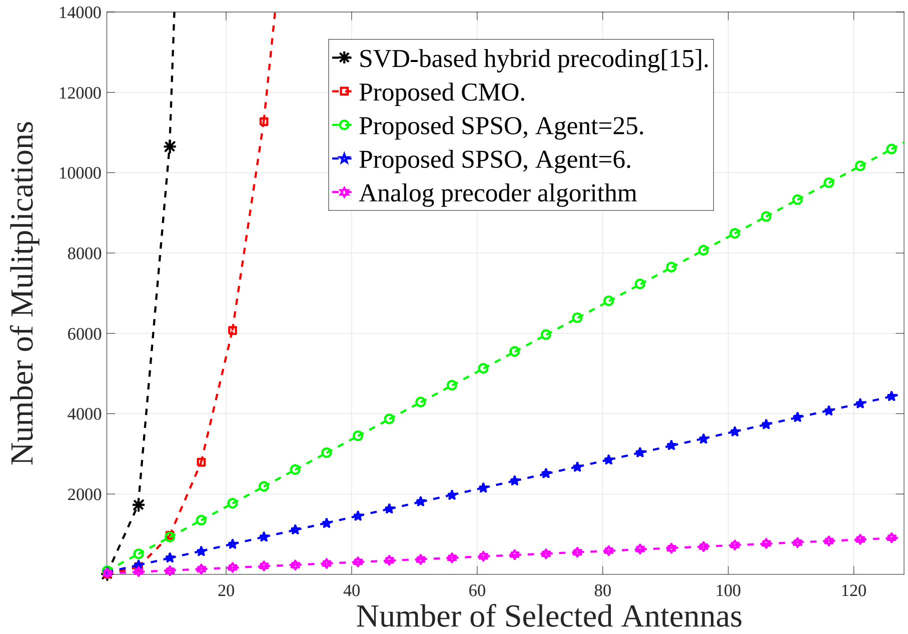


Figure 3.14: Number of multiplications versus the number of selected antennas for partially-connected antenna array complexity

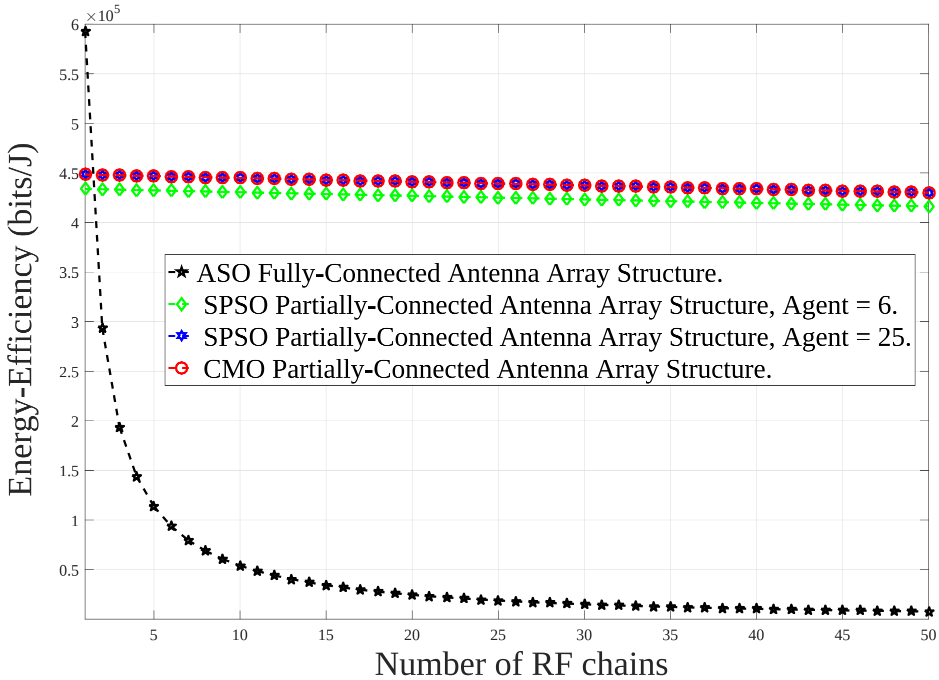


Figure 3.15: Energy efficiency for different antenna array structures as a function N^t_{RF} RF chains.

mission power and the efficiency of the power amplifier are $P_t = 1\text{W}$ and $\gamma_{PA} = 0.375$ respectively [17, 18]. The other system parameters are $P_{mixer} = 19\text{mW}$, $P_{LO} = 5\text{mW}$, $P_{filter} = 14\text{mW}$, and $P_B = 5\text{mW}$. In the hybrid D/A precoding mm-Wave system, the parameters are set as $P_{DAC} = 255\text{mW}$, $P_{BP} = P_{DAC}$ [61], and $P_{PS} = 30\text{mW}$. ASO algorithm for the fully-connected antenna array is considered to show the energy efficiency advantages of different hybrid D/A precoding architectures. CMO and SPSO are considered based on the partially-connected antenna array hybrid D/A precoding. The fully-connected antenna array consumes more power since the number of phase shifters scales linearly with N_{RF}^t and N_t. Conversely, the number of phase shifters is independent of N_{RF}^t in the partially-connected antenna array. It is seen from Fig. 3.15 that when the number of N_{RF}^t is small, ASO attains better energy-efficiency than the partially-connected antenna array algorithms due to the fact that the power consumption for RF chains is low and its achieved spectral efficiency is high. Nevertheless, by increasing N_{RF}^t, the spectral efficiency of the system will increase in the fully-connected antenna array along with the exponentially increased number of phase shifters that consumes excessive power. Subsequently, the energy efficiency of the fully-connected antenna array structure decreases gradually. The proposed algorithms CMO and SPSO for the partially-connected antenna array has higher energy efficiency and its performance decreases slowly with increasing N_{RF}^t. Moreover, SPSO with 6 agents is considered and attains less energy efficiency than SPSO with 25 agents.

***Remark* 3.** *A discussion based on practical aspects for OMP [10] and the proposed algorithms: From the aspect of the energy efficiency criteria OMP consumes more transmitted power than the proposed algorithms. The reason is, in Fig. 4(b), it is shown that hybrid precoding based on OMP to perfectly achieve the performance of fully digital precoding, the number of RF chains is required about* $N_{RF} \geq 10$ *under a practical condition where the number of RF chains same as the number of data streams. However, the proposed algorithms ASO and ASPO achieve the same spectral efficiency as that of the fully digital system when* $N_{RF} \geq 4$ *and* $N_{RF} \geq 5$*, respectively. This produces an extra problem, the number of phase shifters in the system is higher in OMP than the proposed algorithms, as OMP demands more RF chains than the proposed algorithms. To conclude,* *the proposed algorithms are not only bridging a spectral deficiency performance gap as shown in Fig4(a), the energy efficiency of the proposed algorithms is much higher than OMP, because the required RF chains and phase shifters by the proposed algorithms are lower than that required by OMP. In machine learning, the low-rank matrix is used widely in several applications such as communications, computer vision, recommended systems, and environmental science. We therefore often face optimization problems that require non-convex either high-rank or low-rank constraints as in the mm-Wave large antenna array systems. Therefore, one of the suitable solutions is to form the rank constraints problem as a smooth manifolds in* $\mathbb{R}^n$. *Because, intricate optimization problems based on manifolds can be solved more accurate than optimization problem based on for example Euclidean space [50, 56]. For APSO, it has been established widely in industry for designing industrial robots that are used in industrial manufacturing [**?**]. The wide use of PSO in practical systems, as well as the close optimal spectral efficiency performance that is attained by PSO for both structures makes it one of the perfect solutions for designing hybrid precoding mm-Wave systems.*

Finally, in this chapter, as a 3D BF gain is considered which means, the beampattern should be a 3D pattern. Furthermore, when N_t is large enough that will lead to negligible beampattern interference only in partially-connected antenna array. Therefore, the minimum angle for 3D beampatterns between two beams in partially-connected antenna array that can be distinguished and without interference to each other is investigated in this section. A beampattern function is equal to array factor (AF) that has been modelled in [33], and we may rearrange AF in the following form

$$\mathrm{AF} = \frac{\sin(\frac{N_{t_x}}{2}\Lambda_x)\sin(\frac{N_{t_y}}{2}\Lambda_y)}{(\frac{N_{t_x}N_{t_y}}{4}\Lambda_x\Lambda_y)}, \tag{3.56}$$

where $\Lambda_x = 2\pi\lambda^{-1}u_x\sin(\theta_l^i)\cos(\phi_l^i) + \beta_x, \quad \Lambda_y = 2\pi\lambda^{-1}u_y\sin(\theta_l^i)\sin(\phi_l^i) + \beta_y, \quad i \in [r\,,\,t]$, and $\beta x, \beta y$ is determined by user position in small cell. In essence, to find 3D beampattern nulls in the mainlobe based on (3.56), the first null determines the minimum angle between two beams that can be distinguished from each other. Second pair of nulls, determines the major property of the largest sidelobe, in general the maximum interference to adjacent users occurs here. Therefore, by observing (3.56) we can find nulls by zeroing either

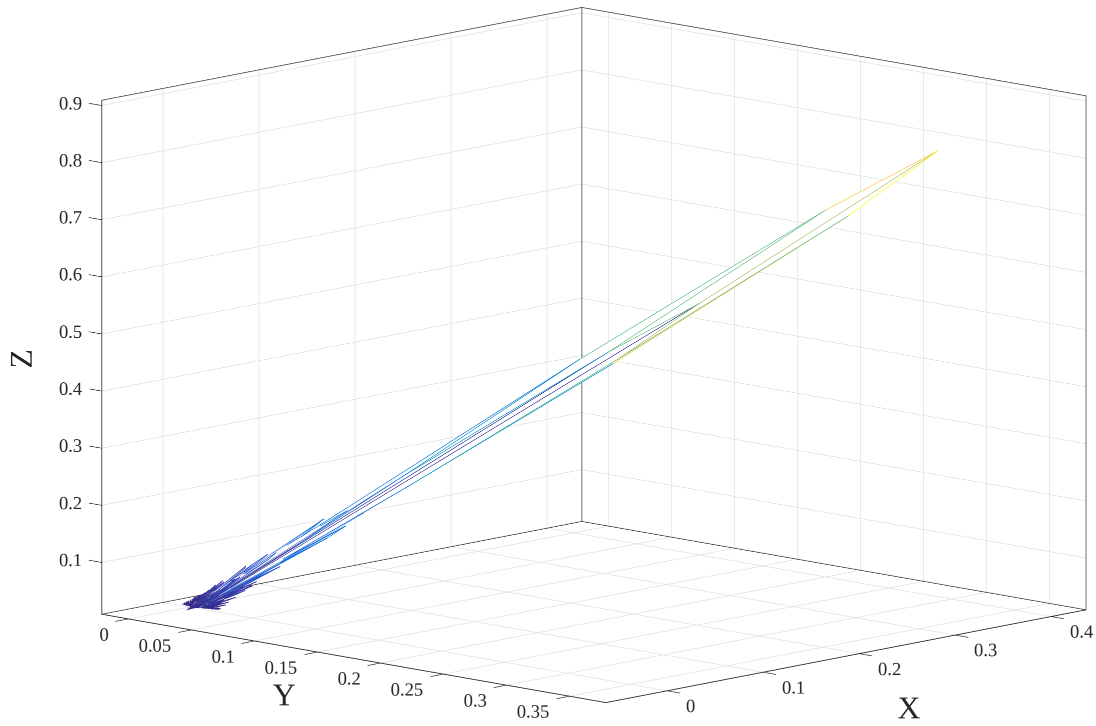


Figure 3.16: Beam pattern square array by using the proposed scheme APSO based on fully-connected antenna array for an array of size $N_t = 128, N_{RF}^t = 2$.

factors of its numerator

$$\frac{N_{t_x} N_{t_y}}{4} \Lambda_x \Lambda_y = \mu\pi, \tag{3.57}$$

where $\mu = 0, \pm 1, \pm 2, \cdots$. It is obvious from (3.57) that beampattern nulls depend on multiple parameters. Therefore, finding the 3D beampattern nulls over 2D scenario can be represented as

$$\theta_{null} = \sin^{-1}\left(\frac{\lambda}{2\pi u_x}\left(\pm\frac{2\pi N_{t_x}}{N_t} - \beta_x\right)\right), \tag{3.58}$$

$$\phi_{null} = \sin^{-1}\left(\frac{\lambda}{2\pi u_y}\left(\pm\frac{2\pi N_{t_y}}{N_t} - \beta_y\right)\right), \tag{3.59}$$

With

$$\left(\left\{ \left| \left(\frac{\lambda}{2\pi u_x}\left(\pm\frac{2\pi N_{t_x}}{N_t} - \beta_x\right)\right) \leq 1 \right| \right\} \cup \left\{ \left| \left(\frac{\lambda}{2\pi u_y}\left(\pm\frac{2\pi N_{t_y}}{N_t} - \beta_y\right)\right) \leq 1 \right| \right\} \right).$$

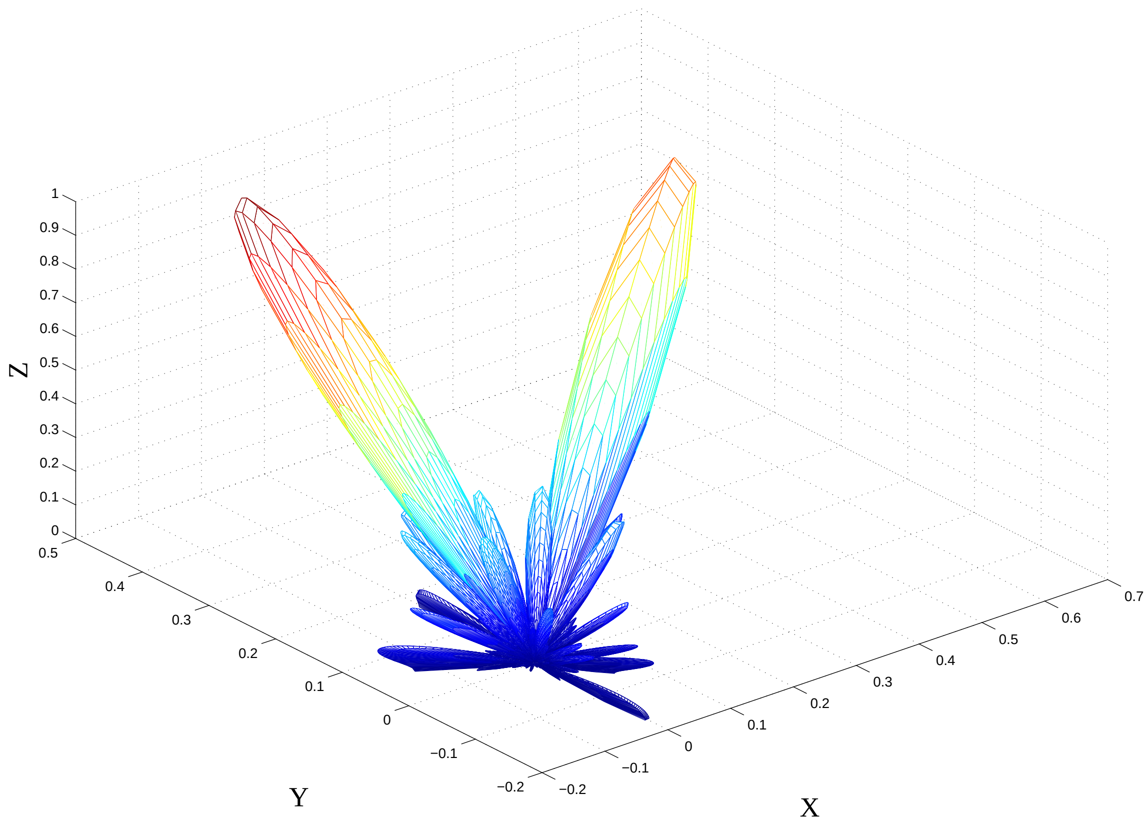


Figure 3.17: Beam pattern square array by using the proposed scheme SPSO based on partially-connected antenna array for an array of size $N_t = 64$ per RF chain, $N_{RF}^t = 2$.

where $\theta_{null}, \phi_{null}$ are the zenith and azimuth null angles, respectively. Notice that in Table-(3.3), the first null is approximately reciprocally proportional to the number of array antenna elements in the broadside of the antenna.

Fig. 3.16 and Fig. 3.17 shows that the beam pattern generated by a transmitter with a $N_t = 128$ and $N_{RF}^t = 2$ based on planar array. The beampatterns are generated by using APSO precoder for fully-connected antenna array and SPSO for partially-connected antenna array. It can be noticed that the beam pattern of fully-connected antenna array is much narrower than that of partially-connected antenna array. The reason is very obvious, in the fully-connected antenna array each RF chain is connected to all the antennas in the system which makes each RF chain use the whole antenna in the array $N_t = 128$ to create a beam. However, in the partially-connected antenna array each RF chain is connected to a sub antenna and create its own beam, in this case each RF chain is connected into

Table 3.4: Numerical Results

θ_{null}	ϕ_{null}	N_{t_x}	N_{t_y}
14.48	30	8	8

half of the number of antennas $N_t = 64$. In this case, the two beams can be separated by utilizing θ^t and ϕ^t directions. Furthermore, the fully-connected antenna array has lower sidelobes than the partially-connected antenna array, which results in sidelobes have lower impact on the main lobe in the fully-connected antenna array.

3.9 Conclusions

In this chapter, we have proposed novel hybrid D/A precoding algorithms for fully-connected antenna array and partially-connected antenna array mm-Wave systems. For the fully-connected antenna array structure, an efficient way was proposed uniquely to help reformulate the non-convex constraints and also to enable the low-complexity algorithm development. Based on this approach, ASO and APSO algorithms were proposed to satisfy the reformulated constraints. ASO can achieve high spectral efficiency at the cost of high complexity. APSO can achieve a relatively good spectral efficiency with lower complexity, related to the number of agents. CMO and SPSO were proposed for the partially-connected antenna array mm-Wave systems. Simulation results showed that our proposed algorithms were able to achieve higher spectral efficiency than the existing hybrid D/A precoding algorithms. Ultimately, the effect of the channel estimation was considered, which showed that our proposed algorithms are more robust to the estimation error than the existing algorithms. Finally, our simulation results showed that APSO and SPSO have lower complexity than the existing hybrid D/A precoding mm-Wave system algorithms.

Chapter 4

Energy-Efficient Design for Hybrid D-A Precoders in mm-Wave Systems

Contents

4.1 Overview

Hybrid D/A precoding is widely used in mm-Wave system to reduce power consumption. The RF chains consume maximum transmitted power, therefore, in this chapter, optimal number of RF chains are selected to achieve the desired energy efficiency (EE). The optimization problem is formulated as the fractional programming maximization. Therefore, the novelty of the proposed method is in two folds. First, the optimal number of RF chains are selected by the proposed bisection algorithm. In this work, the number of data streams equals the number of RF chains which results in the number of data streams is also optimized. Second, the optimal analog precoder and combiner are designed by eigenvalue decomposition, followed by the digital precoder and combiner which are designed based on applying the singular value decomposition to the proposed effective uplink and downlink channel gains matrix. Furthermore, the computational complexity of the hybrid D/A system is considered in this chapter as well. The proposed D/A precoders and combiners are designed carefully to attain a lower complexity than the existing D/A precoders and combiners algorithms simultaneously with achieving a good performance. The effect of the resolution of the phase shifters used in the analog precoder on EE is investigated. Simulation results show that the proposed algorithms outperform the existing algorithms in terms of the EE and the spectral efficiency. It also showed that the hybrid precoding with an optimal number of quantized bits of resolution attains the best EE performance. Finally, simulation results demonstrate that our proposed methods for designing D/A precoders and combiners have lower complexity than the existing D/A precoders and combiners algorithms.

4.2 Introduction

5G mobile communication is proposing mm-Wave bands (frequency ranges from 30 GHz to 300 GHz) to solve the spectrum scarcity issues [10, 19, 21, 38, 63]. However, there are a lot of challenges to implement 5G for mm-Wave systems [5, 54]. In conventional wireless communication systems, the number of RF chains is equal to the number of antennas and the optimal performance in terms of spectral efficiency is achieved by using the fully digital precoder. It is a well known fact that 70% of the transmitted power is consumed by the RF chains [10, 19, 21, 63]. However, if the fully digital precoder is adopted the power consumption will be extreme due to large number of RF chains. Therefore, a simpler approach is to use a single RF chain that is connected to a large number of antennas as proposed in [63] for mm-Wave system. However, the spectral efficiency performance of this approach is much lower than that of the prior system [10, 19, 21]. Therefore, a hybrid D/A precoding system has been studied to balance the spectral efficiency and power consumption [10, 19, 21] which consists of a low-dimension digital beamformer and a high-dimension analog beamformer implemented using phase shifters.

Sparse precoding based on OMP algorithm is proposed for the hybrid D/A mm-Wave system in [10]. The spectral efficiency of the hybrid D/A precoding system is maximized but the power consumption has not been considered in [10]. Hence, in Chapter 3 the computational complexity of designing the hybrid D/A system based on OMP is explained well and it is showed that OMP demands a high complexity to design the hybrid D/A system. Moreover, to address the power consumption issue in [19], the hybrid D/A precoding system is designed to achieve higher spectral efficiency than OMP, and also calculate the energy efficiency (EE) criteria. However, neither the number of RF chains is optimized nor reducing the computational complexity of the hybrid D/A is considered in [19]. Therefore, the number of RF chains in [64] is investigated, two new algorithms are proposed which optimized the number of RF chains while the hybrid D/A system is designed with high complexity based on conventional OMP algorithm. There are two issues with the approach in [64] firstly, finding the optimal number of RF chains which is equivalent to the number of antennas is quite difficult as the number of antennas is quite large. Secondly, the hybrid D/A precoding is not optimized for the selected number of RF chains. Hence, it is of paramount importance that the hybrid D/A mm-Wave system should be designed carefully to be more energy efficient.

As enumerated above, the existing hybrid D/A mm-Wave system neither the number of RF chains is optimized nor reducing the computational complexity of the hybrid D/A mm-Wave system is investigated. Therefore, the aim of this chapter is to develop new algorithms can design an energy efficiency hybrid D/A mm-Wave systems, where the number of RF chains is optimized simultaneously with designing low complexity D/A precoders and combiners. Motivated by this aforementioned discussions, the main contributions of this chapter are discussed as follows:

The optimization problem is formulated as the fractional programming maximization which helps in dividing the optimization problem into two sub-problems. The first sub-problem is formulated as a power allocation and then a bisection algorithm (BI) is proposed. All the positive and non-zero power values that are formulated based on a power allocation, are found by utilizing BI. After that, these power values are combined in a vector. Essentially, the optimal number of RF chains equals the number of all the positive and non-zero power values that are calculated via BI. Furthermore, in this model the number of data streams same as the number of RF chains. Subsequently, the number of data streams in the system is also optimized. The second sub-problem is focused on designing the D/A precoders and combiners with low complexity. To avoid multi-variable joint optimization, the literate works in [10, 19, 21] decouple the transmitter and receiver, then design separately, not jointly. To address the issue of the hybrid precoding design in the low energy consumption situation with the number of data streams equal to the number of RF chains, this chapter introduces the concept of effective channel gains to realize the joint design of transceiver. The analog beamformer under a non-convex constant modulus constraints is designed based on a method that exploits the potential of gaining the knowledge of the channel. This method is called the eigenvalue decomposition (EVD), which is theoretically based on calculating the eigenvector that corresponds to the largest eigenvalue of a matrix [65]. The main advantages of proposing EVD method is that: i) this scheme does not require calculating the gradient of the optimization function as shown in existing algorithms [10,19,63], which reduces the complexity of the system; ii) It also does not require a high number of iterations, as it calculates only the eigenvector that corresponds to the largest eigenvalue. After that, the digital precoder is designed by proposing a low dimension matrix that is generated based on the effective channel gains of the duality of the uplink and downlink channel gains. The optimal digital precoder can be obtained by applying SVD to the low dimension effective uplink and downlink channel gains matrix.

The computational complexity while designing the analog and digital precoder and combiner is precisely considered. The existing algorithm MO-AltMin [19] to design the analog part in each iteration have a high complexity which comes from three stages i.e., conjugate descent, find the local optimum, Kronecker product of large-sized matrices. For designing the digital part in [19] pseudo inverse of a matrix dimension equals to the number of antennas in the mm-Wave system is applied. This operation not only increases the computational complexity because of a large number of antennas, it also sacrifices performance when the analog beamformer irreversible. OMP in [10] when designing the analog precoder, the number of antennas is involved in addition to SVD operation is also applied for obtaining the fully digital precoder while the same digital precoder is used as in [19] which is designed based on least square (LS) method. Consequently, the computational complexity of hybrid D/A mm-Wave systems based on MO-AlitMin and OMP algorithms in [10, 19] is always high. Therefore, the proposed methods in this chapter managed to reduce the computational complexity of the hybrid D/A system. First, the analog part designed based on EVD that only involves the number of antennas without the need of the iterative optimization procedure and iterative SVD operation. Second, the digital part avoided the need of the large dimension matrix by proposing the low dimension matrix based on the effective uplink and downlink channel gains.

Since the power consumption increases with the resolution of the phase shifters of the analog precoder, a trade-off between the spectral efficiency and power consumption based on limited resolution is investigated. This essentially allows us to point out substantial conclusions about the EE of the intended hybrid D/A precoding when different numbers of quantized bits of precision are utilized.

Finally, the EE is compared with different existing systems such as a single RF chain beam steering algorithm [63], hybrid precoding based on fully-connected antenna array architecture OMP [10], MO-AltMin [19] and the fully digital system. Although in [19] two algorithms were proposed for fully-connected antenna array architecture, only MO-AltMin is compared here as it outperformed the other algorithms in [19]. The simulation results showed that the EE of the proposed hybrid D/A precoding algorithm is superior to all these precoders. Ultimately, the spectral efficiency of the proposed algorithm is simulated against the beam steering, OMP, MO-AltMin and the fully digital system. The simulation result also investigates the effects of utilizing few quantized bits on the spectral efficiency and EE. It also showed that the hybrid precoding with an optimal number of quantized bits of resolution attains the best EE performance. Meanwhile, numerical results show that the proposed hybrid D/A has much lower complexity than the existing hybrid D/A algorithm.

4.3 System Model

Fig. 1 shows a downlink mm-Wave system for hybrid D/A precoding based on fully-connected antenna array architecture where each RF chain is connected to all the antennas in the system. The BS sends N_s data streams via N_t transmit antennas to the receiver which consists of N_r receiver antennas. Note here, the receiver side is implemented symmetrically as the transmitter. Therefore, it is assumed that there are N^t_{RF} and N^r_{RF} RF chains at the BS and the receiver. The digital precoder is denoted as $\boldsymbol{D}_{BB}$ and the analog precoder is denoted as $\boldsymbol{A}_{RF}$. The analog precoder is designed under the condition that all the non-zero elements of $\boldsymbol{A}_{RF}$ have the same amplitude, but different phase shifts, while the digital precoder satisfies the total transmit power constraint by normalizing $\|\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|^2_F = N_s$. The dimensions of $\boldsymbol{D}_{BB}$ and $\boldsymbol{A}_{RF}$ are $N^t_{RF} \times N_s$ and $N_t \times N^t_{RF}$. Note that, both $\boldsymbol{D}_{BB}$ and $\boldsymbol{A}_{RF}$ may not be square matrices as $N_s \leq N^t_{RF}$ and $N_t >> N^t_{RF}$.

4.3.1 Signal Model

The received signal for all N_{RF} receive antennas $\boldsymbol{y} = [y_1, y_2, \cdots, y_{N_{RF}}]^T$ at the user is expressed as

$$\boldsymbol{y} = \sqrt{\rho_{\max}}\boldsymbol{W}^H_{BB}\boldsymbol{W}^H_{RF}\boldsymbol{H}\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\boldsymbol{s} + \boldsymbol{W}^H_{BB}\boldsymbol{W}^H_{RF}\boldsymbol{n}, \tag{4.1}$$

where $\rho_{\max}$ represents the average received power. The symbol vector $\boldsymbol{s} \in \mathbb{C}^{N_s}$ with $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^H] = \frac{1}{N_s}\boldsymbol{I}_{N_s}$. $\boldsymbol{n} \in \mathbb{C}^{N_r}$ is the zero-mean complex Gaussian noise vector with independent and identically distributed (i.i.d) elements with $\mathbb{E}[\boldsymbol{n}\boldsymbol{n}^H] = \sigma^2\boldsymbol{I}_{N_r}$. The analog

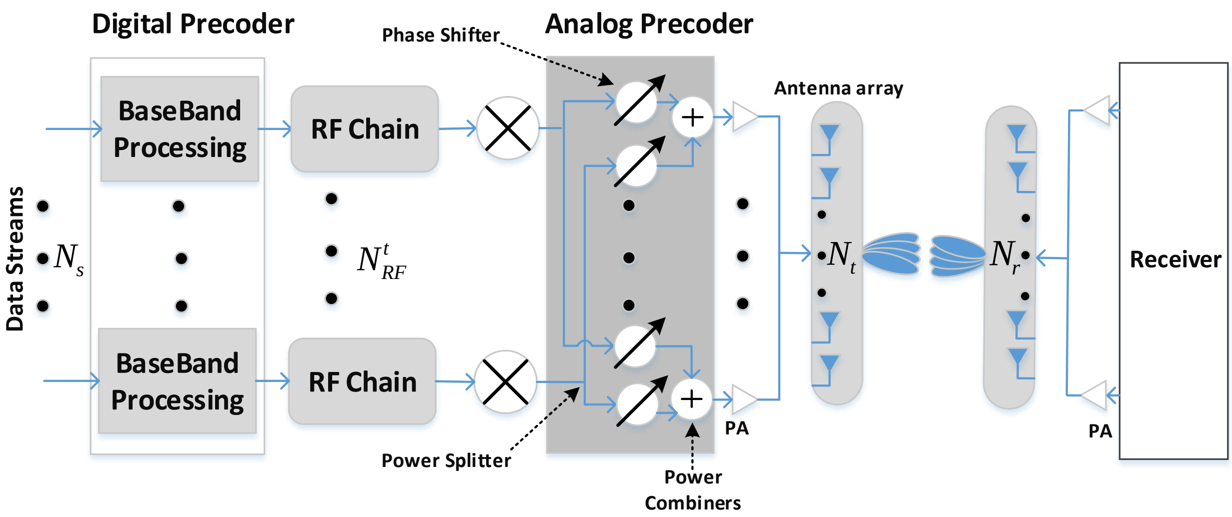


Figure 4.1: Block diagram of a hybrid D/A mm-Wave large antenna array system.

combiner is denoted as $\boldsymbol{W}_{RF}$ and the digital combiner is represented as $\boldsymbol{W}_{BB}$. The dimension of $\boldsymbol{W}_{RF}$ and $\boldsymbol{W}_{BB}$ are $N_r \times N_{RF}^r$ and $N_{RF}^r \times N_s$. The analog combiner is realized by the phase shifters which can only adjust the phase of the signal without changing its amplitude, therefore, all the non-zero elements of $\boldsymbol{W}_{RF}$ have the same amplitude. The objective is to design $\boldsymbol{A}_{RF}$, $\boldsymbol{D}_{BB}$, $\boldsymbol{W}_{RF}$ and $\boldsymbol{W}_{BB}$ to maximize the spectral efficiency C that is given by

$$C = \log_2 \left| \boldsymbol{I}_{N_s} + \frac{P_{\max}}{N_s \sigma^2} \boldsymbol{R}^{-1} \boldsymbol{W}_{BB}^H \boldsymbol{W}_{RF}^H \boldsymbol{H} \boldsymbol{A}_{RF} \boldsymbol{D}_{BB} \boldsymbol{D}_{BB}^H \boldsymbol{A}_{RF}^H \boldsymbol{H}^H \boldsymbol{W}_{RF} \boldsymbol{W}_{BB} \right|, \tag{4.2}$$

where $P_{\max}$ represents the maximum allocated power and $\boldsymbol{R}^{-1} = \boldsymbol{W}_{BB}^H \boldsymbol{W}_{RF}^H \boldsymbol{W}_{RF} \boldsymbol{W}_{BB}$ is the processed noise covariance matrix. Note in mm-Wave systems, the spectral efficiency can be improved by increasing the number of RF chains [10, 19, 21] which consumes up to 70% of the total transceiver power [21]. Therefore, EE criterion is required which is given as [65]

$$\gamma_{EE} = \frac{C}{P_{\text{total}}}, \tag{4.3}$$

where P_{total} is the total power consumed by the P_{Ttotal} transmitter hybrid D/A and P_{Rtotal} receiver hybrid D/A mm-Wave system and is given as

$$\begin{aligned} P_{\text{Ttotal}} &= \beta tr(\boldsymbol{G}\boldsymbol{G}^H) + N_{RF}^t(P_{DAC} + P_{RF}) \\ &+ N_t(P_{PA} + N_{RF}^t P_{PS}) + P_{BP}, \end{aligned} \tag{4.4}$$

$$\begin{aligned} P_{\text{Rtotal}} &= \beta tr(\boldsymbol{W}\boldsymbol{W}^H) + N_{RF}^r(P_{ADC} + P_{RF}) \\ &+ N_r(P_{PA} + N_{RF}^r P_{PS}) + P_{BP}, \end{aligned} \tag{4.5}$$

where $\boldsymbol{G} = \boldsymbol{A}_{RF}\boldsymbol{D}_{BB}$ and $\boldsymbol{W} = \boldsymbol{W}_{RF}\boldsymbol{W}_{BB}$, β denotes reciprocal of amplifier efficiency. The power amplifier is modelled as $P_{PA} = P_{n_t}/\gamma_{PA}$. P_{DAC} and P_{ADC} in (4.4) and (4.5) are the power of the digital-to-analog converter and the power of the analog-to-digital converter, respectively. P_{BP} is the base-band processing power, and P_{PS} is the power consumed by the phase shifters (including the energy for the excitation and the energy for the compensation of insertion loss [61]). γ_{PA} is the average transmitting power of each antenna, and is fixed to 0.375 [18, 21]. In (4.4) and (4.5), P_{RF} is the power consumption of each RF chain and is modelled as

$$P_{RF} = P_{mixer} + P_{LO} + P_{filter} + P_{BA}, \tag{4.6}$$

where P_{mixer} is the power consumed by the mixer and P_{LO} is the power required by the local oscillator. From (4.6), it can be observed that the mm-Wave system also considers the power consumed by the filter P_{filter} and the base-band amplifier P_{BA} which was not included in [19, 21] and [18] making our power consumption model more accurate and generalized.

4.3.2 Problem Formulation

Our objective is to maximize (4.3) when different constraints are applied on the hybrid D/A mm-Wave system. Notice here, the main focus of this work is on the transmitter side, this is because, the same work can be applied to the receiver side easily which helps in simplifying the optimization problem. In this case, for the transmitted side the spectral efficiency C can be reformulated as

$$C = \log_2 \left| \boldsymbol{I}_{N_s} + \frac{1}{\sigma^2}\boldsymbol{H}\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{H}^H \right|. \tag{4.7}$$

Therefore, the optimization problem is now formulated as:

$$\begin{aligned}
\underset{\boldsymbol{A}_{RF},\boldsymbol{D}_{BB},N_{RF}^{t}}{\arg\max} \quad & \gamma_{EE} = \frac{C}{P_{\text{total}}}, \\
\text{s.t.} \quad & C \geq v, \\
& tr(\boldsymbol{G}\boldsymbol{G}^{H}) \leq P_{\max}, \\
& |(\boldsymbol{A}_{RF})_{n_t}| = 1, \, \forall n_t, \\
& ||\boldsymbol{G}||_F^2 = N_s, \\
& 1 \leq N_{RF}^{t} \leq N_t,
\end{aligned} \tag{4.8}$$

where v is the minimum desired spectral efficiency. The non-convex constant modulus constraint is subjected to $\boldsymbol{A}_{RF}$ as restricting all the non-zero elements to have the same amplitude in the analog domain. The total power consumed P_{total} equals the total power consumed in the transmitter side P_{Ttotal} in (4.4). The total transmit power constraint and limited number of RF chains between 1 and the maximum number of antennas are considered respectively in (4.8). Undoubtedly, the optimization problem in this chapter targets the combination of different parameters $\boldsymbol{A}_{RF}, \boldsymbol{D}_{BB}$, and N_{RF}^{t} simultaneously which renders the problem difficult to be tractable. Moreover, the analog precoder $\boldsymbol{A}_{RF}$ constraints is non-convex and hard to deal with. Therefore, this problem can be partitioned into two sub-problems, initially the number of RF chains N_{RF}^{t} are optimized in the system, and finally the hybrid precoders $\boldsymbol{A}_{RF}, \boldsymbol{D}_{BB}$ are optimized. In this work, we assume $\boldsymbol{A}_{RF}^{opt}$, $\boldsymbol{D}_{BB}^{opt}$, and $N_{RF}^{t^{opt}}$ are the optimal solutions to the above mentioned problem. Let us now try to optimize the number of RF chains.

4.4 Optimized Number of RF Chains

It can be observed from (4.2) that the spectral efficiency of the system is a function of $\boldsymbol{H}$ and $\boldsymbol{G}$ where $\boldsymbol{G} = \boldsymbol{A}_{RF}\boldsymbol{D}_{BB}$. The spectral efficiency problem in (4.2) can be reformulated as a power allocation problem by taking SVD of $\boldsymbol{H}$ and $\boldsymbol{G}$ yielding [65]

$$C(N_{RF}^{t}, \{p_i\}) = \sum_{i=1}^{N_{RF}^{t}} \log_2\left(1 + \frac{1}{\sigma^2}\lambda_i^2 p_i\right), \tag{4.9}$$

where λ_i corresponds to the eigenvector of $\boldsymbol{G}$ and p_i denotes the power allocated to the i-th RF chain. The total power in (4.4) can be a function of this power allocation by replacing $tr(\boldsymbol{GG}^H)$ as

$$\begin{aligned} P_{\text{total}}(p_i) &= \beta \sum_{i=1}^{N_{RF}^t} p_i + N_{RF}^t(P_{DAC} + P_{RF}) \\ &+ N_t(P_{PA} + N_{RF}^t P_{PS}) + P_{BP}. \end{aligned} \tag{4.10}$$

From (5.25) and (4.10) the optimization function in (4.8) is redesigned to properly choose the optimized number of RF chains at the transmitter $N_{RF}^{t^{opt}}$ as

$$\begin{aligned} N_{RF}^{t^{opt}} = \quad & \arg\max_{N_{RF}^t} \frac{C(N_{RF}^t, \{p_i\})}{P_{\text{total}}(N_{RF}^t, \{p_i\})}, \\ & \text{s.t.} \quad C(N_{RF}^t, \{p_i\}) \geq v, \\ & \sum_{i=1}^{N_{RF}^t} p_i \leq P_{\max}, \\ & 1 \leq N_{RF}^t \leq N_t. \end{aligned} \tag{4.11}$$

The objective function in (4.11) denotes the EE that is obtained through finding the optimal number of RF chains $N_{RF}^{t^{opt}}$. This objective function can be solved by using primal fractional optimization programming which involves maximization of a ratio of two different functions [66,67]. (4.11) can be transformed using the Dinkelbach's method [67], and the optimal solution for a given φ can be achieved if and only if the function $E(\varphi) = 0$ which implies $\varphi = \varphi^{opt}$ and $\varphi^{opt} = \arg\max \frac{C}{P_{\text{total}}}$. The maximization of the ratio $\frac{C}{P_{\text{total}}}$ is mathematically identical to the maximization of the difference $C - \varphi P_{\text{total}}$ for a specific choice of the parameter φ [66]. Thus leveraging the theoretical results in [66, 67], the problem in (4.11) is transformed as the following parametric programming problem

$$\begin{aligned} E(N_{RF}^{t^*}, \{p_i^*\}, \varphi^*) = \quad & \arg\max_{N_{RF}^t} C(N_{RF}^t, \{p_i\}) - \varphi P_{\text{total}}(N_{RF}^t, \{p_i\}), \\ & \text{s.t.} \quad C(N_{RF}^t, \{p_i\}) \geq v, \\ & \sum_{i=1}^{N_{RF}^t} p_i \leq P_{\max}, \\ & 1 \leq N_{RF}^t \leq N_t. \end{aligned} \tag{4.12}$$

The above proposition implies that if we can find a parameter φ such that the optimal value of the objective function in (4.12) is zero i.e., $E(\varphi) = 0$, then this optimal solution can be directly applied to (4.11). However, from a mathematical point of view, the function $E(\varphi) = 0$ is a monotonically decreasing function over φ, in such case solving

Algorithm 6 Bisection Algorithm (BI) for selecting the optimal number of RF Chains.

1: Initialize $\varphi_{min} = 0$, $\varphi_{max} = \frac{C}{P_{\text{total}}}$ and small threshold value $\varpi > 0$;

2: Let $\varphi = \frac{\varphi_{min}+\varphi_{max}}{2}$. Solve $E(N_{RF}^{t*}, \{p_i^*\}, \varphi^*)$ in (4.12) with the given φ, and obtain the corresponding solutions;

3: **if** $E(N_{RF}^{t(m)}, \{p_i\}^{(m)}, \varphi^{(m)}) \geq 0$ **then**,

4: $\varphi_{min} = \varphi$; **Else** $\varphi_{max} = \varphi$.;

5: **if** the termination criteria $|\varphi_{max} - \varphi_{min}| \leq \varpi$, **then** Stop; **Else**, return to step 2

6: Consider $N_{RF}^{t*} = N_{RF}^{t}, \{p_i^*\} = \{p_i\}$ optimal solutions;

(4.12) through the Dinkelbach's method does not guarantee optimal solution to (4.11) [67]. Therefore, a bisection algorithm (BI) [68], that is demonstrated in Algorithm 6, perform better in this problem as compared to finding φ. In BI, it is important to decrease the search time of this procedure which is achieved by initializing the φ based on lower φ_{min} and upper bounds φ_{max}, such that $\varphi_{min} \leq \varphi \leq \varphi_{max}$. Intuitively, the lower and upper bounded are valued by $\varphi_{min} = 0$, and $\varphi_{max} = \frac{C}{P_{\text{total}}}$, respectively. Algorithm 6 at the top of this page provides the BI based solution to obtain the optimal number of RF chains $N_{RF}^{t^{opt}}$. The acquired combination of optimal power values $\{p_i^*\}$ can be represented as a power vector $\boldsymbol{p}_{opt}$. Importantly, by utilizing BI all the positive and non-zero power values are found. The optimal number of RF chains equal the number of all the positive and non-zero power values. Essentially, when the power value equal either a negative or zero, then we set it to zero and the algorithm start with a new RF chain until the BI finds all the optimal number of RF chains for the hybrid D/A design. In this system model, we have assumed that the number of data streams equal the number of RF chains, therefore, this method can target the optimal number of data streams as well. Theorem 3 in [69] provides the convergence properties of BI based algorithm. Based on the analysis in [69], BI converge to the optimal φ^*. Any solution φ, which satisfies $|\varphi_{max} - \varphi_{min}| \leq \varpi$ for a small $\varpi > 0$, can be found within at most α iterations, where α is the minimum integer such that $\alpha \geq \log_2(\frac{1}{\varpi} \cdot \frac{C}{P_{\text{total}}})$. Fortunately, the BI algorithm that is used to optimize the number of RF chains at the transmitted side can be applied directly to the receiver side.

At this point, we have had an efficient method to optimally find one of the main parameters that this chapter targets to optimize, the number of RF chains. Therefore, the remaining task is to obtain the precoding D/A design which will be discussed in the next section.

4.5 Precoders and Combiners Design for Hybrid D/A mm-Wave Large Antenna Array System

In this section, the main focus is to design hybrid D/A precoding and combining methods. Initially, the analog precoder and combiner $\boldsymbol{A}_{RF}$ and $\boldsymbol{W}_{RF}$ are designed based on the eigenvalue decomposition (EVD) followed by the digital precoder and combiner $\boldsymbol{D}_{BB}$ and $\boldsymbol{W}_{BB}$ that exploit the duality of the effective uplink and downlink channel gains where the non-convex constraint and the total transmit power constraints are satisfied. Although our proposed algorithms in Chapter 3 ASO and ASPO can be easily implemented to design the analgo precoder and combiner in this section, a new method is proposed here by avoiding the complexity of the optimization procedure. Furthermore, unlike the other algorithms proposed in literature [10, 19, 63] that are designed hybrid D/A precoders and combiners, the proposed methods based on EVD and the duality of the effective uplink and downlink channel gains while designing the hybrid D/A precoder and combiner can avoid some optimization steps which result in a low complex optimization procedure, these advantages points are summarized as

1. The proposed method does not require a gradient of an objective function to determine the optimal analog and digital precoder and combiner. This helps the system to reduce the complexity highly as compared to the existing algorithms in [10,19,63], especially when finding the optimal solutions for an optimization problem with large number of antennas.

2. Euclidean distance between the optimal fully digital system and the hybrid D/A system is required in the conventional algorithms [10, 19], this step is avoided in the proposed system with guarantee of attaining a near optimal spectral efficiency.

Let us first look into the design of analog precoders and combiners based EVD followed by desiging digital precoders and combiners based on the effective uplink and downlink channel gains matrix.

4.5.1 Analog Precoders and Combiners

The design of the analog precoder $\boldsymbol{A}_{RF}$ is proposed as follows. Suppose $\mathbf{P}_T = \boldsymbol{H}^H\boldsymbol{H}$ where $\mathbf{P}_T \in \mathbb{C}^{N_t \times N_t}$, is Hermitian and the diagonalizable matrix. Here, we only need to use EVD to decompose $\mathbf{P}_T$ and then select the eigenvector $\mathbf{u}_{N_t}$ that corresponds to the largest eigenvalue. Therefore, $\mathbf{P}_T = \boldsymbol{U}_T\boldsymbol{\Lambda}_T\boldsymbol{U}_T^H$, where $\boldsymbol{\Lambda}_T$ denotes a diagonal matrix with diagonal elements that represent the eigenvalues γ_{N_t}, and each column of $\boldsymbol{U}_T$ denotes as an eigenvector of $\mathbf{P}_T$. In this case $\mathbf{P}_T$ will consist of N_t linear independent eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_t}$ where $\mathbf{u}_{N_t} \in \mathbb{C}^{N_t \times 1}$ with corresponding eigenvalues of $\gamma_1, \gamma_2, \cdots, \gamma_{N_t}$, such that $\mathbf{P}_T\mathbf{u}_e = \gamma_e\mathbf{u}_e$ for $e = 1, \cdots, N_t$.

The analog precoder $\boldsymbol{A}_{RF}$ matrix can be designed with N_{RF}^t column vectors such that $\boldsymbol{A}_{RF} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{N_{RF}^t}]$ where $\boldsymbol{a}_i \in \mathbb{C}^{N_t \times 1}$ and i is the index of the RF-th chain. In this case, the first column of the analog precoder $\boldsymbol{a}_1$ is designed by the eigenvector corresponding to the largest eigenvalue that is selected from the output of EVD method. The element-wise normalization is satisfied for the non-convex constraint in (4.8) as

$$\boldsymbol{a}_1 = \frac{1}{\sqrt{N_t}} e^{j\arg[\boldsymbol{u}_1]}. \tag{4.13}$$

After obtaining $\boldsymbol{a}_1$, to design the next analog precoding vector $\boldsymbol{a}_2$, initially the first column vector of the matrix $\mathbf{P} = \boldsymbol{H}^H\boldsymbol{H}$ is replaced by the selected eigenvector that corresponds to the largest eigenvalue from previous EVD of $\boldsymbol{a}_1$, therefore, EVD method are repeated. Finally, the second analog vector $\boldsymbol{a}_2$ is designed by using equation (5.20). Similarly, all the columns of the optimal analog precoder matrix $\boldsymbol{A}_{RF}^{opt}$ are designed.

For designing the analog combiner matrix $\boldsymbol{W}_{RF} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_{N_{RF}^r}]$, the same steps can be followed which are utilized to design the analog precoder $\boldsymbol{A}_{RF}$ based on EVD. Therefore, suppose $\mathbf{P}_R = \boldsymbol{H}^H\boldsymbol{H}$ where $\mathbf{P} \in \mathbb{C}^{N_r \times N_r}$, is Hermitian and the diagonalizable matrix. Here, we only need to use EVD to decompose $\mathbf{P}_R$ and then select the eigenvector $\mathbf{u}_{N_r}$ that corresponds to the largest eigenvalue. Thus, the analog combiner is obtained as

$$\boldsymbol{w}_1 = \frac{1}{\sqrt{N_r}} e^{j\arg[\boldsymbol{u}_1]}. \tag{4.14}$$

After that to design the next analog combiner vector $\boldsymbol{w}_2$, initially the first column vector of the matrix $\mathbf{P}_R = \boldsymbol{H}^H\boldsymbol{H}$ is replaced by the selected eigenvector that corresponds to the largest eigenvalue from previous EVD of $\boldsymbol{w}_1$, therefore EVD method are repeated. Finally, the second analog vector $\boldsymbol{w}_2$ is designed by using equation (4.14). Similarly, all the columns of the optimal analog combiner matrix $\boldsymbol{W}_{RF}^{opt}$ are designed.

4.5.2 Quantized Analog Precoders and Combiners

In the previous subsection 4.5.1, the analog precoder and combiner are designed assuming that the resolution of the phase shifters (PSs) elements to be infinity ∞. However, this assumption affects proportionally the power consumptions, because the number of PSs scales linearly with the number of antennas in the system [19,21]. Utilizing low-resolution of PSs can enhance the EE of the hybrid D/A mm-Wave system. Therefore, to solve this issue, the phase ϑ of each elements of the analog precoder $\boldsymbol{A}_{RF}$ is quantized up to B bits of precision. The minimum Euclidean distance is formulated between the phase of nearest neighbour element to its quantized in order to accurately design the quantized phases of precision. The quantized phase of each elements of analog precoder $\boldsymbol{A}_{RF}$ and analog combiner $\boldsymbol{W}_{RF}$ is given as

$$\hat{\vartheta} = \frac{2\pi\hat{\tau}}{2^B}, \quad \hat{\tau} = \underset{\tau \in [0, 2\pi(\frac{1}{2^B}), \cdots, 2\pi(\frac{2^B-1}{2^B})]}{\arg\min} |\vartheta - \tau|, \tag{4.15}$$

where τ is the set of quantized phases. Finally, the analog precoder and combiner $\boldsymbol{A}_{RF}^{F^{opt}}(\hat{\vartheta})$ and $\boldsymbol{W}_{RF}^{F^{opt}}(\hat{\vartheta})$ are designed with quantized of each element.

4.5.3 Digital Precoders and Combiners

After having accomplished the designed of the analog beamformer, the joint design of the digital precoder and combiner is now our focus. We consider to design transceiver low complexity digital precoding and combining jointly based on the effective channel gains of the duality of the uplink and downlink channel gain.

In general, the existing algorithms [10, 19] use the well-known LSs solution, which is given by [10, 19]

$$\boldsymbol{D}_{BB} = (\boldsymbol{A}_{RF}^{F^{opt\,H}}(\hat{\vartheta})\boldsymbol{A}_{RF}^{F^{opt\,H}}(\hat{\vartheta}))^{-1}\boldsymbol{A}_{RF}^{F^{opt\,H}}(\hat{\vartheta})\boldsymbol{F}_{opt}, \tag{4.16}$$

where $\boldsymbol{F}_{opt}$ is the optimal fully digital precoder obtained through the SVD operation of the channel matrix. The solution in (4.16) employees pseudo inverse in the process of solving, this operation not only increases the computational complexity, it also sacrifices performance when the analog beamformer irreversible. In this case, to design low complexity and asymptotically optimal digital preocders and combiners, the optimal $\boldsymbol{D}_{BB}^{opt}$ and $\boldsymbol{W}_{BB}^{opt}$ are designed based on the duality of the effective uplink and downlink channel gains that is denoted as $\boldsymbol{H}^{EF}$, where $\boldsymbol{H}^{EF} = \boldsymbol{W}_{RF}^{F^{opt}H}(\hat{\vartheta})\boldsymbol{H}\boldsymbol{A}_{RF}^{F^{opt}}(\hat{\vartheta}), \boldsymbol{H}^{EF} \in \mathbb{C}^{N_{RF}^{r} \times N_{RF}^{t}}$. For this stage only, let ignore the power transmitted constraint, therefore, there is no restriction on $\boldsymbol{D}_{BB}$ and $\boldsymbol{W}_{BB}$. Based on our assumption, we can get the unconstrained digital precoder and combiner $\boldsymbol{D}_{BB}$ and $\boldsymbol{W}_{BB}$ as

$$\boldsymbol{D}_{BB} = \boldsymbol{V}^{EF}(1:N_s) = [\boldsymbol{v}_1^{EF}, \cdots, \boldsymbol{v}_{N_s}^{EF}], \tag{4.17}$$

$$\boldsymbol{W}_{BB} = \boldsymbol{U}^{EF}(1:N_s) = [\boldsymbol{u}_1^{EF}, \cdots, \boldsymbol{u}_{N_s}^{EF}], \tag{4.18}$$

where $\boldsymbol{V}^{EF}$ and $\boldsymbol{U}^{EF}$ are the unitary matrices that are derived by taking SVD of the effective channel gains $\boldsymbol{H}^{EF} = \boldsymbol{U}^{EF}\sum^{EF}\boldsymbol{V}^{EF^H}$. After obtaining the digital precoder and combiner from (4.17), now, we normalize $\boldsymbol{D}_{BB}^{opt} = \frac{\sqrt{N_s}\boldsymbol{D}_{BB}}{||\boldsymbol{A}_{RF}^{F^{opt}H}(\hat{\vartheta})\boldsymbol{D}_{BB}||_F}$. This step guarantees that the power transmitted constraint $\|\boldsymbol{A}_{RF}\boldsymbol{D}_{BB}\|_F^2 = N_s$ is satisfied. Although this work is focused on the transmitted side, the analog and digital combiners for the receiver side are designed jointly with the D/A precoders.

4.5.4 Computational Complexity of Proposed Hybrid D/A mm-Wave Large Antenna Array Systems

The complexity of the hybrid D/A mm-Wave system is calculated taking into account the number of multiplications and additions involved as shown in Table-(4.1) and Table-(4.2). Since the mm-Wave system uses sophisticated large antenna arrays, the computational complexity is dominated by the number of antennas. In this section, the the computational complexity of analog in Table-(4.1) and digital in Table-(4.2) precoders are discussed.

The complexity of our proposed method based on EVD for designing the analog precoder consists of two stages: multiplying two channel matrices and then decomposing the desired matrix $\mathbf{P}_T$. It requires N_t^3L multiplications and $N_t(N_r-1)+N_t(N_t-1)$ additions, where L represents the number of multipaths. The inverse of a matrix for any algorithm is derived with the help of cholesky decomposition. Notice that, the number of antennas involved in calculating the computational complexity of the proposed EVD method has a smaller impact on the system complexity as compared to the MO-AltMin and OMP algorithms when designing the analog precoder, as shown in Table-(4.1). Certainly, the number of iterations has the impact on the overall complexity of MO-AltMin and OMP algorithms. In Table-(4.1), we enumerate the calculation of the computational complexity per iteration. Thus, the overall computational complexity of each algorithm needs to be multiplied by the required number of iterations for each algorithm. In this case, suppose that the hybrid D/A mm-Wave system is designed based on $N_{RF}^t = N_s = 2, N_t = 144$ and $N_r = 64$, let now numerically calculate the computational complexity for MO-AltMin, OMP and the proposed method:

1. MO-AltMin algorithm requires 456480 number of multiplications to complete the first iteration. However, MO-AltMin algorithm needs 20 iterations to converge which makes the calculation of the number of multiplications increases highly to be 9129600 in order to design the analog precoder in the system [19].

2. In OMP algorithm, the number of multiplications is around 11943936 to complete the first iteration. In general, OMP needs approximately 5 to 10 iterations to converge, so let take 5 iterations to design the analog precoder in the system which implies that the number of multiplication after 5 iterations grows sharply to be around 59719680.

3. Fortunately, our proposed method does not require an iterative procedure. In this case, to design the analog precoder in the system based on the proposed EVD method, the number of multiplications is equal to 2985984. Undoubtedly, the number of multiplications that is demanded to perform the proposed method is significantly lower than that of MO-AltMin and OMP algorithms.

Note here, in the previous calculations the number of multiplications is only considered because the multiplications operation consume more computations than additions [70].

Table 4.1: Computational Complexity of Analog Precoders.

Algorithm	Number of operations per iterations	
	Additions	Multiplications
MO-AltMin analog precoding [19]	$2N_s(N_tN_{RF}^t - 1)N_t + N_tN_s(N_tN_{RF}^t - 1) + N_tN_{RF}^t(N_tN_s - 1) + 2N_tN_{RF}^t(N_tN_{RF}^t - 1)$	$2N_{RF}^tN_sN_t^2 + N_t^2N_sN_{RF}^t + N_t^2N_s + 2N_t^2N_{RF}^{t^2} + N_tN_{RF}^t$
OMP analog precoding [10]	$\frac{N_{RF}^{t^3}}{6} + N_{RF}^t(N_t - 1)LN_s$	$N_sN_{RF}^tN_t^3L$
Proposed analog precoding based on EVD	$N_t(N_r - 1) + N_t(N_t - 1)$	N_t^3L

Table 4.2: Computational Complexity of Digital Precoders.

Algorithm	Number of operations per iterations	
	Additions	Multiplications
LS solution for digital precoding in [10,19]	$N_{RF}^t(N_t - 1) + N_r(N_t - 1) + N_{RF}^tN_r(N_t^2 - 1)$	$N_{RF}^tN_t^2 + N_sN_t^2 + N_{RF}^tN_tN_s$
Proposed digital precoding based on SVD	$N_{RF}^tN_r(N_t - 1) + N_{RF}^t(N_r - 1)$	$N_{RF}^tN_rN_t + N_{RF}^{t^2}N_r + N_{RF}^{t^3}$

For designing the digital precoder, the proposed method based on SVD consists of three stages. Therefore, it requires $N_{RF}^tN_r(N_t - 1) + N_{RF}^t(N_r - 1)$ additions and $N_{RF}^tN_rN_t + N_{RF}^{t^2}N_r + N_{RF}^{t^3}$ multiplications. It can be noticed that the number of RF chains has more impact on the computational complexity of the proposed method. Essentially, in hybrid D/A mm-Wave systems the number of RF chains is roughly small as compared to the number of antennas, thus, the complexity of our proposed method attains lower than the LS solution. By seeking the same numerical example, the number of multiplications that are required in the proposed method based effective channel is 18696 and in the conventional LS solution is 83520, receptively.

To sum up, our proposed methods for the analog and digital precoders are preferred over the existing algorithms, due to its low complexity which is proved numerically. Furthermore, this is will be proved by simulation as well in Section 4.6. Ultimately, Table-(4.1) and Table-(4.2) summarizes the computational complexity for the hybrid D/A mm-Wave system.

Finally, after obtaining $\boldsymbol{A}_{RF}^{F^{opt}H}(\hat{\vartheta}), \boldsymbol{D}_{BB}^{opt}$, $N_{RF}^{t^{opt}}$, and $\boldsymbol{p}_{opt}$, the joint precoder is designed as $\boldsymbol{G}^{opt} = \boldsymbol{A}_{RF}^{F^{opt}H}(\hat{\vartheta})\boldsymbol{D}_{BB}^{opt}\text{diag}[\boldsymbol{p}_{opt}]^{\frac{1}{2}}$. Now, the spectral efficiency is determined by (4.7), total power by (4.4) and the EE by (4.3). The above proposition guides to an energy efficient hybrid D/A precoding system, and the performance of the proposed system is compared with the current known systems [10, 63] and [19] in the upcoming section.

4.6 Simulation Results

In this section, the EE performance of our desired hybrid precoding system is carried out. Our algorithm performance is compared with four known algorithms in literature. The first algorithm is based on beam steering which utilizes a single RF chain as proposed in [63]. Second and third algorithms OMP [10] and MO-AltMin [19] are also simulated with a fixed RF chains. The fourth algorithm is a fully digital system where the number of RF chains is equal to the number of antennas. Mm-Wave channel as described in Chapter 3 Section 3.3.1 is deployed and the number of multipaths is set to $L = 15$ [10, 19, 41]. The transmitter antenna array is assumed to be a UPA with the antenna spacing $u_q = \frac{\lambda}{2}$. The AoAs and AoDs are generated independently from the uniform distribution with mean cluster angle uniformly distributed $[0, 2\pi]$, while the angle spread is set to 7.5 [10, 19, 30]. Moreover, the minimum spectral efficiency is fixed as $v = 1$ bits/s/Hz. Table (4.3) provides the practical values of the power parameters that are used in the simulation results [19, 21, 64] and [18].

Fig. 4.2 illustrates the effect of the EE with respect to the number of transmit antennas when the SNR is fixed to 15 dB. In the simulation, the existing hybrid D/A precoding algorithms OMP and MO-AltMin utilize 6 number of RF chains. It is observed in Fig. 5.13 that the proposed hybrid precoding attain superior EE as compared to the other known solutions in literature [10, 63] and [19]. From the figure, it is observed that as the number

Table 4.3: The power expressions of different precoding systems where used in the simulation results [19, 21, 64] and [18].

Power Type	Values
P_{mixer}	19mW
P_{LO}	5mW
P_{filter}	14mW
P_{BA}	5mW
P_{PS}	10mW
$P_{DAC} = P_{BP}$	255mW

of transmit antennas increase the EE of all the known solution decreases linearly in performance except the proposed one. The proposed solution, enhances the EE by searching the optimal number of RF chains and then implementing the optimal precoders D/A in the system. In the fully digital system the number of transmit antennas are equal to the number of RF chains, therefore, as the number of antennas increases, the EE performance decreases as the RF chain consumes more power. Both the OMP and MO-AltMin hybrid precoding losses the EE linearly because the number of RF chains both have not been optimized. The simplified beam steering neither achieves the highest nor the lowest EE because there is no effort for optimizing the number of transmitted antennas.

Fig. 4.3 shows the EE variation for different precoding solutions with respect to SNR. For a 64×16 mm-Wave system, it is very clear that the proposed EE hybrid precoding system based algorithms outperforms all the existing precoding solutions in [10, 63] and [19] with increase in SNR. The reason is obvious as the proposed hybrid precoding designed based on finding the optimal number of RF chains and then applying the optimal precoders D/A in the system. The hybrid precoding based on OMP and MO-AltMin algorithms achieve a reasonable performance because the main focus is to find only the optimal D/A precoders in the system. Finally, the fully digital system and the beam steering have the lowest performance as compared to the hybrid D/A precoding system.

In Fig. 4.4 the EE of the different algorithms based on different systems with respect to the transmission power is tested for $N_t = 64 \times N_r = 48$, and the number of RF

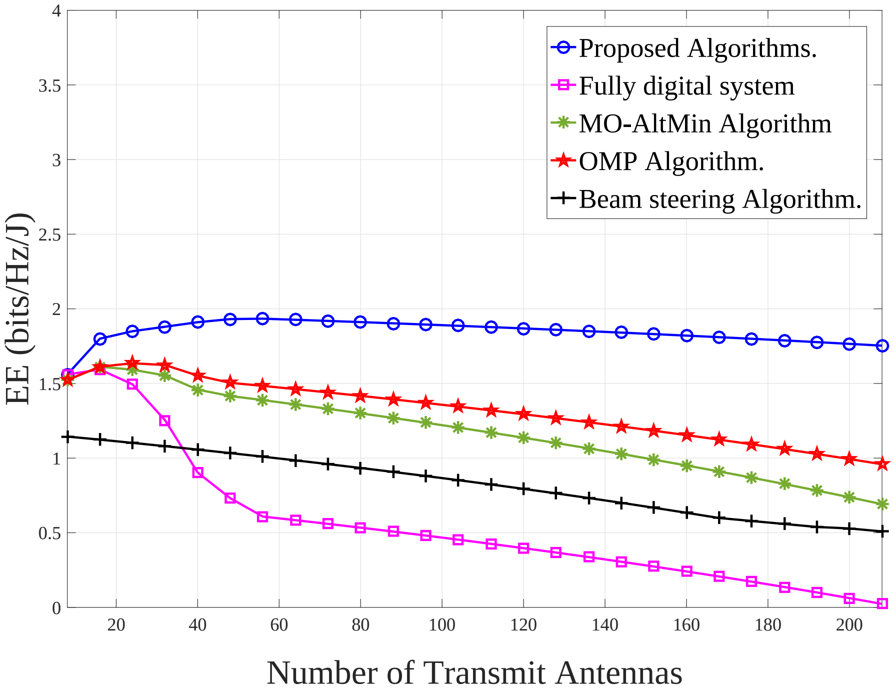


Figure 4.2: EE performance versus the number of transmitter antennas at SNR = 15 dB.

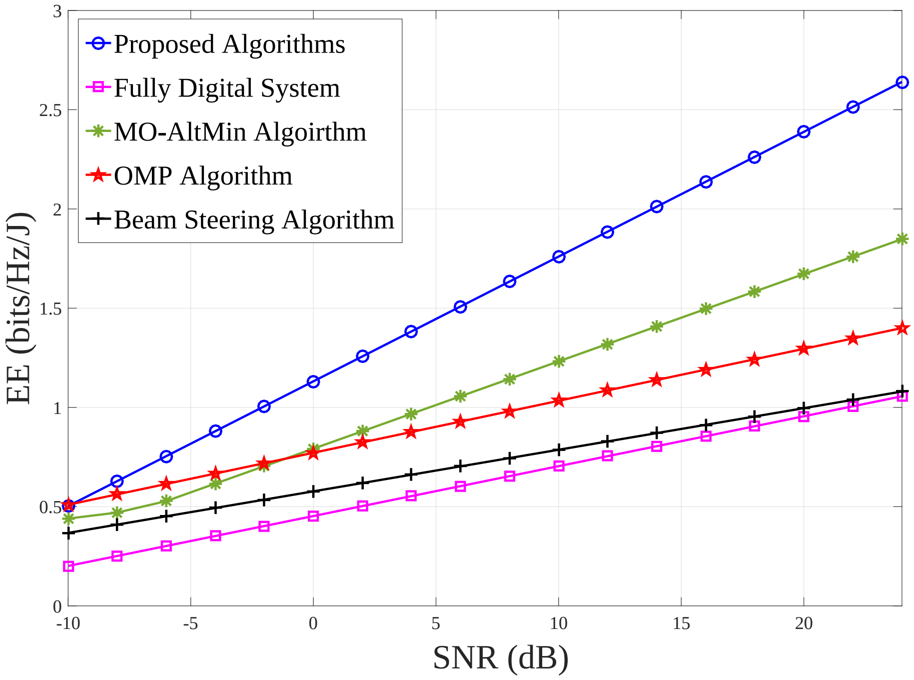


Figure 4.3: EE performance versus SNR when $N_t = 64 \times N_r = 16$, B=∞.

chains is fixed as $N_{RF}^t = N_s = 4$ only for the proposed algorithms, MO-AltMin [19], and OMP [10] based on hybrid D/A precoding system. As depicted when the transmission power is higher than 8W, the proposed algorithm is more energy efficient than the existing algorithms. This is due to, the beam steering algorithm [63] have only single RF chain, thus for power up to 8W, the gains of the spectral efficiency are becoming higher as compared to the power cost. However, as the power increases above that value, the beam steering performance decreases sharply as the spectral efficiency deteriorates as compared to the power consumption, and at some point it performs even worse than the other algorithms. Furthermore, MO-AltMin and OMP algorithms achieve lower EE than that of the proposed algorithm, since the both algorithms focus on the precoding solutions only, without any efforts on optimizing the total power of the hybrid D/A system which makes the power costs having more impact on their EE. Moreover, for OMP algorithm which is based on codebook, the analog precoder performance is limited which affects directly the gains of the spectral efficiency as compared to the power costs. On the other hands, the proposed algorithm always achieves superior EE, because the spectral efficiency is tightly close to the fully digital system and the total power is optimized which balances the EE performance as compared to the power consumptions. In general, as the power consumption grows linearly all the algorithms decreases, since the achievable capacity has lower gains as compared to the power consumption.

Fig. 4.5 plots the spectral efficiency performance for the hybrid D/A mm-Wave system when $N_{RF}^t = N_{RF}^r = N_s$, as a function of SNR. Since the beam steering [63] employs a single RF chain, the spectral efficiency achieved by beam steering is always lower than the spectral efficiency achieved by our proposed algorithms and the benchmark which is the fully digital system. OMP [10] is designed to achieve a spectral efficiency very close to the fully digital system only for a special scenario when $N_{RF}^t \geq 2N_s$. Therefore, when this case $N_{RF}^t = N_{RF}^r = N_s$ is applied, the spectral efficiency performance is not close to the fully digital system. It is observed that the spectral efficiency achieved by our proposed precoder based on EVD and LS methods outperforms the OMP and the beam steering and is slightly higher than MO-AltMin performance, which is close to the fully digital system. The reason is that the proposed precoders EVD and LS methods can accurately

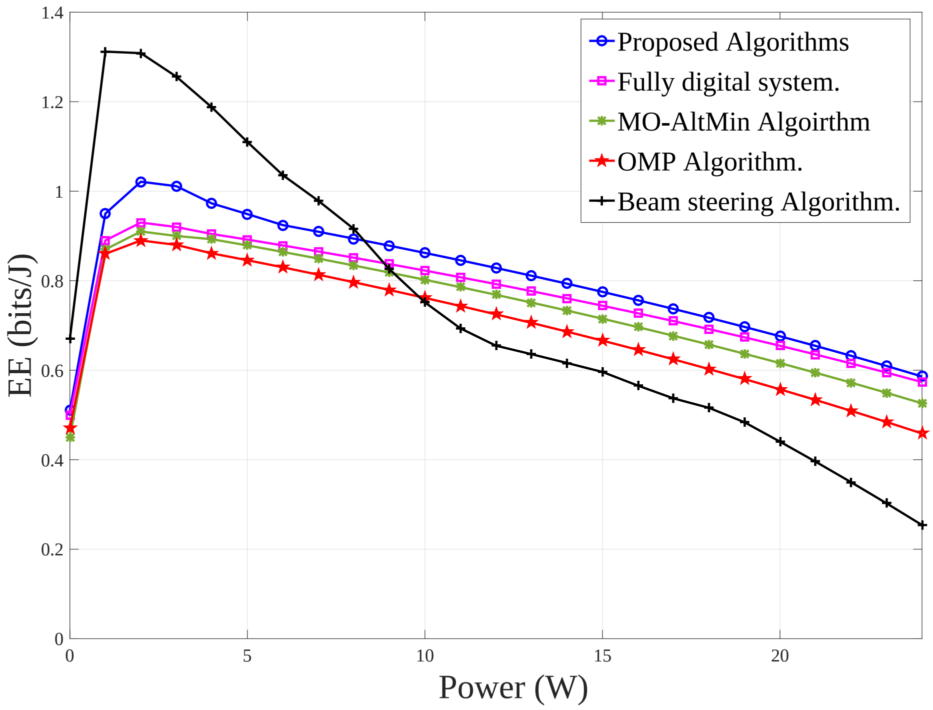


Figure 4.4: EE performance versus transmission power when $N_t = 64 \times N_r = 48$, with fixed RF chains $N_{RF}^t = N_s = 4$, B=∞.

design the optimal precoders which maximizes the spectral efficiency of the system and therefore, outperforms the existing algorithms.

Fig. 4.6 presents the spectral efficiency and the EE with respect to the number of quantized bits of the resolution of the PSs. Initially, as expected, the spectral efficiency of the finite-bit grows with resolution. It can be observed that the proposed hybrid precoding based EVD method for B=8 bits similar performance is achieved as B=∞. This implies that increasing the number of bits of precision do not always provide much gain as compared to the few-finite-bits. Based on this observations, we lay insight into the optimal number of quantization bits required the proposed hybrid precoding based algorithm perspective. From simulation results it can be observed that when the quantization B =4-5 bits, the maximum EE is achieved. This is because, the hybrid precoding at very-low-resolution (B =1-2 bits) suffers from a severe spectral efficiency lack, while at the high resolution (B =7-8 bits) the power consumption increases linearly.

Fig. 4.7 shows also the convergence behavior of the BI algorithm that searches for the optimal number of RF chains with SNR = 0dB. The spectral efficiency achieved is

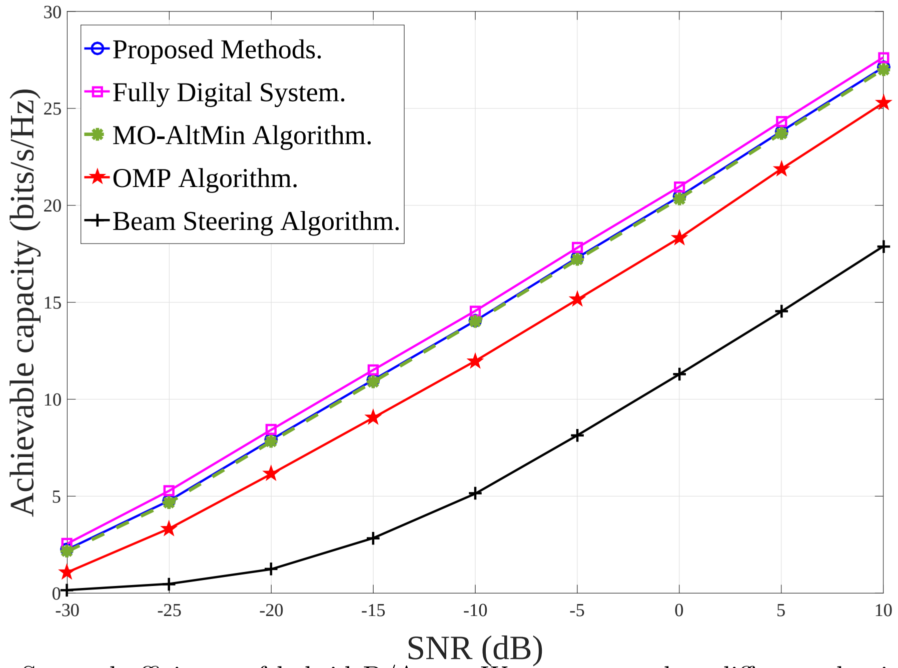


Figure 4.5: Spectral efficiency of hybrid D/A mm-Wave system when different algorithms are used and $N_{RF}^{t} = N_{RF}^{r} = 2, N_t = 128, N_r = 32$, B=$\infty$.

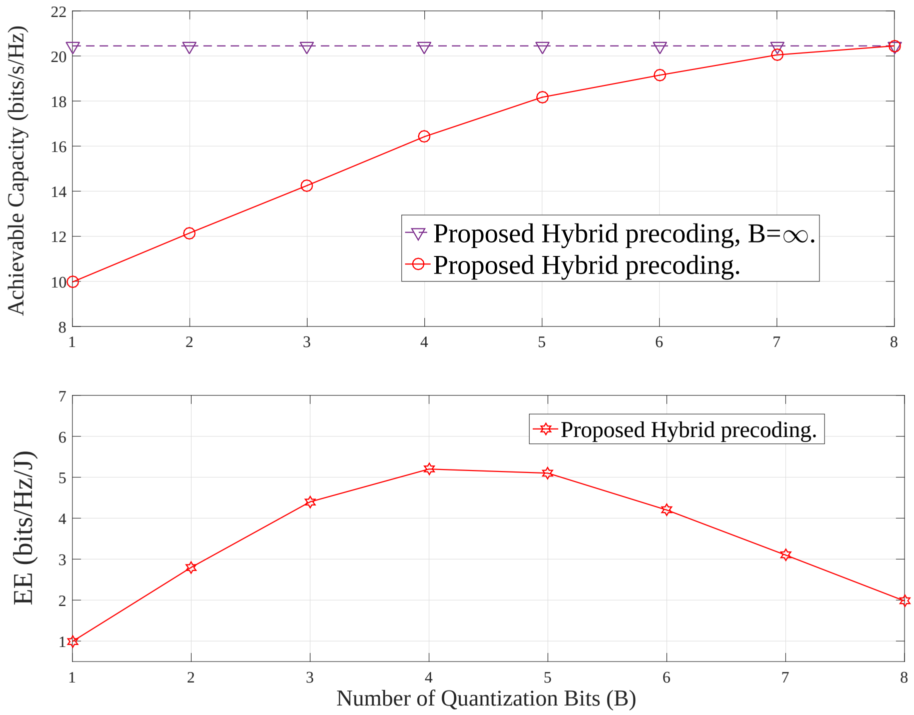


Figure 4.6: Spectral efficiency and EE versus the number of quantization bits, where $N_t = 128$, SNR = 0 dB.

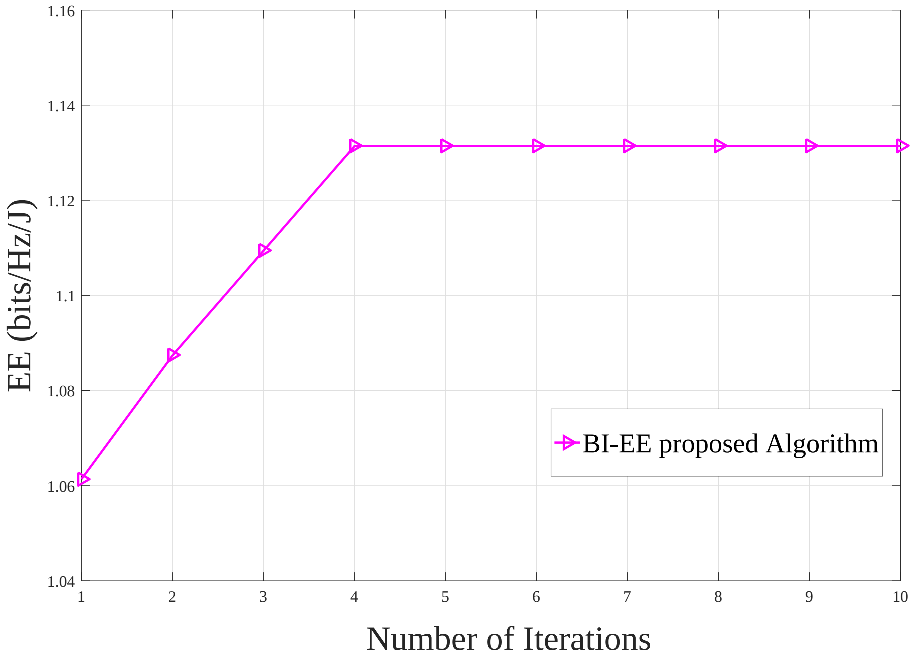


Figure 4.7: Convergence of the proposed algorithm BI in terms of the EE.

averaged over $10,000$ independent channel generations. It can be seen that BI converges very fast and only 4 iterations are required. The reason is that BI has a very simple structure and it is bounded to find the solutions quickly.

In both Fig. 4.8 and Fig. 4.9 the structure of the hybrid D/A mm-Wave system is based on $N_{RF}^{t} = N_s = 2$, and $N_r = 16$. Fig. 4.8 plots the number of multiplications for designing the analog precoder for the hybrid D/A mm-Wave system with respect to the number of antennas, as multiplications consume more computation than additions. Fully digital system has high complexity because the number of RF chains is proportional to the number of antennas. From the figure, it is seen that the complexity of OMP with 5 iterations increases exponentially. The complexity of MO-AltMin with 20 iterations is lower than that of OMP when the number of antennas increases. While, the complexity of the the proposed scheme based EVD increases linearly. It is also verified that by increasing the number of of antennas, the complexity of the system increases. Therefore, in the analog precoder that is designed for the hybrid D/A mm-Wave system, the proposed scheme based EVD has lowest complexity than all algorithms, including MO-AltMin, and

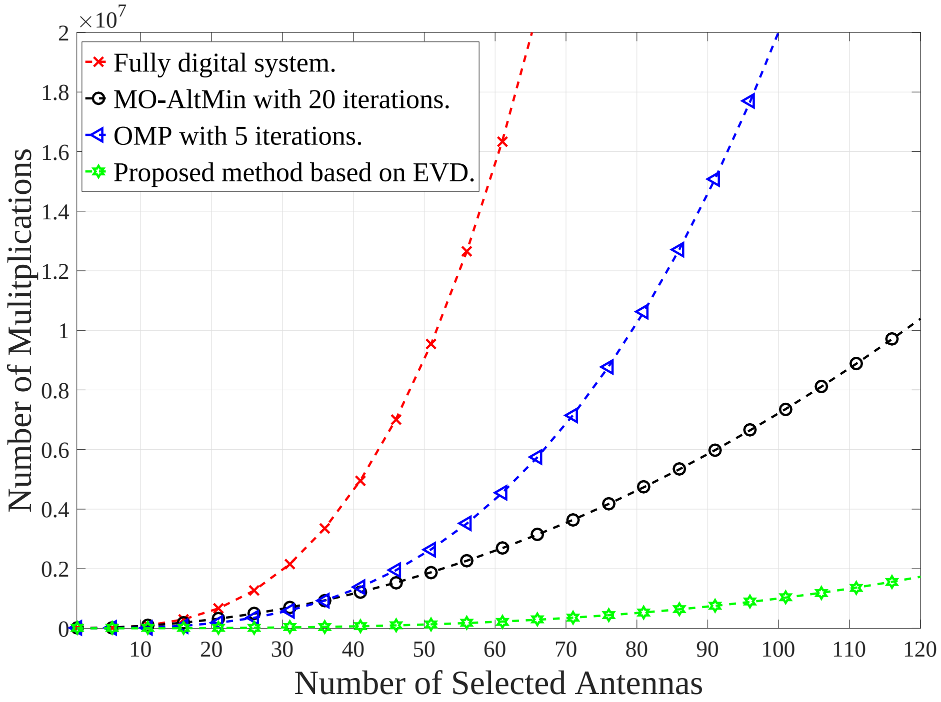


Figure 4.8: Number of multiplications versus the number of selected antennas for Analog precoder complexity

OMP.

Fig. 4.9 demonstrates the computational complexity for designing the digital precoder for the hybrid D/A mm-Wave system as a function of the number of multiplications versus the number of antennas. From the figure, it is seen that the complexity of the LS solution grows exponentially. This is because, when calculating the computational complexity of LS solution the number of antennas involved in the calculation is doubled, as shown in Table (4.2). While, the complexity of the the proposed scheme based SVD increases linearly. It is shown that at very low number of antennas (> 22) the proposed scheme has slightly higher complexity than the LS solution. The reason, as our proposed method is designed the digital precoder jointly which involves the number of received antennas in the calculation. However, as the number of received antennas keeps constant and the number of transmitted antennas increases above (< 22) the computational complexity of the proposed method attains much lower complexity than the LS solutions. Therefore, in the digital precoder that is designed for the hybrid D/A system, the proposed scheme based SVD and effective channel has lowest complexity than conventional LS solution [10, 19].

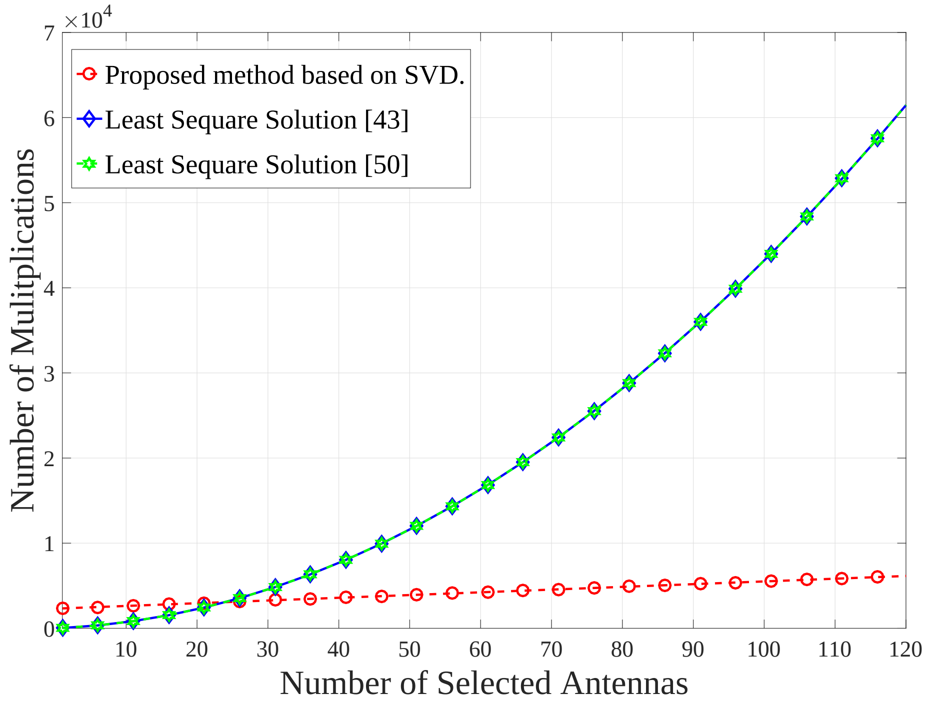


Figure 4.9: Number of multiplications versus the number of selected antennas for Digital precoder complexity.

4.7 Conclusions

In this chapter EE hybrid precoding mm-Wave system is designed. The optimization problem was separately aimed to initially optimize the number of RF chains in the system where BI algorithm was proposed. This algorithm also optimizes the number of data streams in the system. After that, the optimal analog precoder is designed with the help of EVD method without the needs of finding the gradient of the cost function. While, the digital precoder was designed based on LS solution. Ultimately, the proposed hybrid precoding based on limited resolution of the phase shifters was explored to show a potential EE gain as compared to the hybrid precoding having infinite bits of resolution. One of the main concluded points is that neither increasing highly the number of quantized bits of the resolution results in more spectral efficiency nor does a few-finite-bits (i.e., 1-2 bits) achieves the best EE for hybrid precoding systems. Simulation results showed significant EE performance advantages of our proposed algorithms over the existing systems.

Chapter 5

Designing Hybrid Precoding for Uplink Multi-User Millimetre Wave Massive MIMO Systems

Contents

5.1 Overview

In this chapter, we develop novel precoders and combiners for multi-user hybrid D/A beamforming for mm-Wave massive MIMO systems, where the number of RF chains is much less than the number of antennas equipped with each transceiver. Due to some propagation uplink paths of multiple users are transmitted via physical common scatters regime, some of these paths may arrive with overlapped angle of arrivals (AoAs) at the BS. In this scenario, the correlation between the channel vectors also increases highly which affects the achievable uplink rate severely. Therefore, the intrinsic focus is on substantially maximizing the desired signal while reducing the system interferences. The analog precoders and combiners are designed by using the power iteration and the Riemannian optimization method based on Stiefel manifold, algorithms respectively. The digital combiner adapts the MMSE by judiciously exploiting the effective uplink analog channel gains. Further, the estimation channel is investigated by designing a two-step procedure. The strongest power point in each scattering point is formulated and the accuracy is improved by employing an angler domain scheme. Ultimately, simulation result show that the proposed algorithms surpass the existing algorithms in hybrid D/A paradigms in terms of the achievable uplink rate, and also attain a smaller error estimated channels.

5.2 Introduction

The high prevalence of smartphones has generated a continuing demand for increasing data rate in wireless systems. One solution for substantially increasing the data rate in 5G mobile communication is to combine mm-Wave bands (frequency ranges from 30 GHz to 300 GHz) [11, 38, 71] with MIMO [5, 72]. Therefore, massive MIMO systems operating at mm-Wave bands are anticipated to become very popular [12]. In fact, there are plenty of challenges to implement mm-Wave massive MIMO communication systems. As known, it is definite that in conventional wireless systems, the number of antennas is equal to the number of RF chains, so that the highest performance gain is achieved by using the fully digital precoder scheme. However, if the fully digital precoder scheme is adopted in the mm-Wave massive MIMO systems, the associated large number of RF chains will consume extremely high level of power and the hardware complexity will increase dramatically [10–12, 19]. Therefore, a hybrid D/A precoding and combining system has been studied to balance the achievable rate, power consumption, and the implementation complexity [10, 19, 24–27]. Fundamentally, the hybrid D/A structure consists of a small-number of RF chains called the digital part and a large-number of antennas called the analog part. Hence, the optimal hybrid D/A precoding and combining design has the challenge of solving an intractable non-convex constant modulus constraint optimisation problem [10,19,24–27].

[10] has been focused on designing hybrid D/A for single-user mm-Wave massive MIMO systems, where OMP based approach was proposed to procure the analog and digital solution. Later on, MO-AltMin and PE-AltMin based on SVD as iterative algorithms were proposed in [19] to further improve the achievable rate of the OMP approach for single-user scenarios.

To extend to the multi-user case in order to improve the maximum achievable rate of mm-Wave massive MIMO systems, various hybrid D/A precoding and combining schemes have been proposed [24–27]. For instance, a hybrid D/A system was designed by considering the weighted sum mean square error (WSMSE) based on the theory of compressed sensing in [24]. The number of RF chains in the system [24] were assumed to be half of the number of antennas in the system. To reduce the number of RF chains in [25] the precoder and combiner problem are proposed to split into two stages. In stage one, a beamsteering based algorithm was proposed for the analog part and in stage two, ZF solution was developed for the digital part [25]. Though simple, such a system sacrifices the achievable rate when a realistic multipath channel is applied to the system. To resolve this issue in multipath channels, maximum ratio combining (MRC) was proposed to design the analog precoding and then ZF was applied for the digital precoding [26]. However, MRC [26] did not consider the effects of large interference in the system which consequently spoils the achievable rate. In [27], the analog precoding was designed by using an extension algorithm based on a conventional OMP and MMSE was applied for the digital precoding. The elements of analog precoding and combining matrices were selected from a pre-defined dictionary which depends on angles of arrivals (AoAs) and angles of departures (AoDs) [27]. However, in mm-Wave massive MIMO systems with 5G high user density, there is a noticeable achievable rate gap between the hybrid D/A based on OMP and the fully digital system.

All the algorithms in [24–27] have assumed that the propagation paths corresponding to different users have independent channel vectors which results in non-overlapping AoAs at the BS. However, it has been demonstrated in recent research that the rough surfaces of buildings and the rifts between small buildings can cause diffused scattering in mm-Wave channels (short distance propagation) [11, 38, 39]. This makes the propagating wave attenuate significantly as the wavelength shrinks. Practically, in urban areas (i.e., shopping malls, stadiums, city centres), the number of unforeseen scatterings increases considerably so that mm-Wave signal channels may not be necessary to have isolated AoAs at BS. For example, measurements in Daejeon City, Korea, based on ray-tracing simulation [40] showed that the propagation paths of multiple users were spotted under the impact of non-negligible common scattering scenarios, thus the arrived signals tend to overlap at the BS. With high probability, in this scenario we intuitively observe that some propagation paths of multiple users sharing some common scatterers, therefore these different propagation paths will enforce to have the same AoAs at the BS. This phenomenon generates a non-negligible severe interference in the system. Indubitably, with the hybrid D/A mm-Wave massive MIMO system, this type of interference cannot be eliminated by the digital processing domain only, because the digital domain is restricted with fewer RF chains. In this case, if the existing precoders and combiners [24–27] are applied, this serves to severely limit the achievable rate in the current hybrid D/A mm-Wave massive MIMO systems. Hence, it is of paramount importance that the hybrid D/A precoders and combiners should be designed precisely to mitigate this type of interference.

[73] was developed to manage the interference due to signals having the overlapped AoAs at the BS. Initially the Gram-Schmidt (GS) algorithm is formed for the analog stage and then applying the MMSE scheme at the digital stage. However, mathematically GS

based on iterate vector by vector, in the end, these vectors do not completely eradicate the interference at the analog stage. Undoubtedly, this results in an illustrious performance gap away from the fully digital system. Ultimately, despite of the massive MIMO diversified advantages, algorithms in [24–27, 73] have been applied only for multiple-input single-output (MISO) systems.

Channel estimation for hybrid D/A mm-Wave massive MIMO systems is crucial. The main disputes arise from the certainty that the digital part estimator is positioned to the rear of the analog part. This makes the digital part only access to the effective analog channel gains with lower dimensionality. This limitation motivated researchers to develop new channel estimation algorithms. A promising scheme based on compressed sensing theory [48] was proposed which relies heavily on the mm-Wave channels using sparse channel matrices. This limits the algorithm practicality, as it may not be feasible when there are more scattering in mm-Wave channels (i.e., shopping malls, city centre). To overcome the previous drawbacks, a recent developed work based on Hierarchical beamforming training algorithm (HBT) was designed in [74]. For estimating the channel angles, HBT begins with dividing the whole estimated angle spaces, number of sectors, and number of antennas into hierarchical levels. The main constraint is that the number of estimated angles and sectors must double that of the corresponding number of antennas at each level. In view of prior work, as the number of sectors increases linearly, HBT suffers severely from the sectors' overlapped interference problem. Therefore, the number of sectors is compulsorily small which affects the accuracy of choosing the estimated angles in HBT.

Motivated by aforementioned discussions, the existing hybrid D/A mm-Wave massive MIMO system has clearly assumed that there is always non-overlapped AoAs case, or that there is no precise decrease in the overlapped AoAs interference which sacrifices the system's achievable rate. Furthermore, the current estimation channel algorithms are neither practical in 5G dense areas with several scatters environments nor do they find accurate estimated angles due to overlapped sectors' interference. Therefore, the prime target of this Chapter is to maximize the desired channel gain concurrently mitigating the interferences that are generated frontally when the signals are transmitted through some common scatters which produce overlapped AoAs, and backwardly multi-streams of data transmission, arrive at the BS. To clarify the aim, a low dimension matrix based on combining all the precoders, channels, and combiners is formulated as an effective uplink analog channel gains. In this matrix, the diagonal elements represent the desired user gain and the off-diagonal elements denote the interference of the system. Ultimately, a method is proposed that can estimate the mm-Wave channels under the common scatters scenarios with a feature providing a more accurate estimation of the channel angles. Specifically, the contributions of this chapter are summarized as follows:

1. The analog precoder is designed based on an algorithm that exploits the potential of gaining the knowledge of the channel at the user. This algorithm is called Power Iteration (PI) based on calculating the eigenvector that corresponds to the largest eigenvalue of a matrix. In this design, to satisfy the non-convex constant modulus constraint, the optimal analog precoder is designed by extracting the phases of the eigenvector that corresponds to the largest eigenvalue which is optimally found by PI. Conceptually, this algorithm is similar to the optimal SVD operation, but importantly avoids the high complexity of SVD especially in large matrices. Although PI is used previously in Google [75], in rare cases we may not find an exact solution which leads to iterate infinitely. Therefore, in this chapter a developed version is proposed that avoids previous convergence problems.

2. After designing the analog precoder optimally, the analog combiner at the BS needs to be proposed with the aim of maximizing the received power of the desired signal simultaneously with suppressing the off-diagonal elements impact, using a Riemannian optimisation method based on Stiefel manifold (ROSt) under orthogonal constraints. The non-convex constant modulus constraints is ignored at the begging steps until find the optimal solution based on the cost function by using ROSt. Then in order to guarantee the constant modulus constraint, the analog combiner will extract the phases of the optimal solution at the last step. In contrast to [24, 25, 27], our proposed approach to find the analog combiners aims to optimize the achievable uplink rate of the hybrid D/A mm-Wave massive MIMO system directly. One of the other advantages of proposing ROSt is that it removes the need for any pre-defined

candidate set for analog part (e.g. OMP).

3. In an enterprise to magnify the achievable uplink rate, further reduce the interference in the effective uplink analog channel gains. The digital combiner will be designed jointly by exercising MMSE based on uplink analog channel gains.

4. We propose a two-step multi-user channel estimation scheme for mm-Wave channels. In the first step, detecting the strongest power point among multi-power points in the scattering scenarios. As each scatterer point in the channel consist of multi-power points where there is a point that is defined as the most of the power focuses around in that scatterer point. Therefore, instead of estimating jointly all the combination of multiple paths and angles as in [74], the proposed method can find directly the strongest power point in each scatter point. In the second step, the accuracy of the estimation process is considered by proposing a search method based on an angular domain that aims in searching for the strongest points for all the antenna array. This allows our method to avoid the drawback of lacking the accuracy, as the sectors' overlapped interference problem does in [74].

Simulation results show that the proposed PI and ROSt algorithms based hybrid D/A under practical circumstances can outperform the existing algorithms [24–27, 73] in terms of achievable uplink rate. Ultimately, the proposed estimation method attains lower error estimated channels than HBT [74].

5.3 Uplink Model

Consider a single cell uplink multi-user hybrid D/A mm-Wave massive MIMO system as illustrated in Fig. 5.1. A BS is equipped with a large-scale antenna array of N_r antennas and N_{RF} RF chains, where each RF chain is connected to all N_r antennas. It is assumed that the BS communicates with K users, each of which has N_t antennas and one single RF chain. Therefore, each user can transmit one data stream at one time. The number of users K is assumed to be $K \leq N_{RF}$ in the system. K users are allowed to share the same spectrum. Then, the received signal at BS is given by

$$\boldsymbol{y}_{\text{ul}} = \sum_{k=1}^{K} \boldsymbol{H}_k \boldsymbol{j}_{RF_k} s_k + \boldsymbol{n}, \tag{5.1}$$

where the matrix $\boldsymbol{H}_k \in \mathbb{C}^{N_r \times N_t}$ represents the uplink channel for the user k, s_k is the transmitted data symbol of the user k. $\boldsymbol{j}_{RF_k}$ is the analog precoder which is implemented by analog components like phase shifters, i.e., an $N_t \times 1$ vector with constant modulus entries. $\boldsymbol{n}$ is the noise vector that follows $\boldsymbol{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{N_r \times 1})$. It is assumed that the data symbols for users are independent with unit power, i.e., $\mathbb{E}[s_k s_i] = 0, k \neq i$, and $\mathbb{E}[|s_k|^2] = 1$. For data-transmission, beamforming is implemented at the BS to decode users' data streams. Specifically, at the BS, by using combing vector $\mathbf{g}_k$ to decode user k's signal, we have

$$\mathbf{g}_k^H \boldsymbol{y}_{\text{ul}} = \mathbf{g}_k^H \boldsymbol{H}_k \boldsymbol{j}_{RF_k} s_k + \mathbf{g}_k^H \sum_{i \neq k} \boldsymbol{H}_i \boldsymbol{j}_{RF_i} s_i + \mathbf{g}_k^H \boldsymbol{n}, \tag{5.2}$$

where $\mathbf{g}_k = \boldsymbol{A}\mathbf{d}_k \in \mathbb{C}^{N_r \times 1}$ denotes the hybrid D/A combiner at the BS. The analog combiner matrix at the BS is denoted as $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{N_{RF}}] \in \mathbb{C}^{N_r \times N_{RF}}$, where $\boldsymbol{a}_i \in \mathbb{C}^{N_r \times 1}$, is a phase array of adjustable phase shifters connecting to the ith RF chain. In other words, each analog combiner $\boldsymbol{a}_i$ has the same amplitude but different phase shifters. The data symbol will be further separated by applying an $N_{RF} \times K$ digital combiner (without the constant modulus constraint on its elements) $\mathbf{D} = [\mathbf{d}_1, \cdots, \mathbf{d}_K] \in \mathbb{C}^{N_{RF} \times K}$, where $\mathbf{d}_k \in \mathbb{C}^{N_{RF} \times 1}$.

5.3.1 Channel Model

The mm-Wave channel can be portrayed with the popular multipath channel model [38, 39, 76], in which a number of propagation paths are associated with a few scatterers points. To consolidate this phenomena, in this chapter the extended 3D Saleh-Valenzuela geometry-statistical model is adopted [76]. The channel of user k can be written as

$$\boldsymbol{H}_k = \frac{1}{\sqrt{N_{cl}}} \sum_{m=1}^{N_{cl}} \sqrt{c_{km}} \boldsymbol{H}_{km}, \tag{5.3}$$

where $\boldsymbol{H}_{km}$ denotes the 3D mm-Wave channel from user k to the BS via cluster m. In general, when the signals hit an object (e.g., wall), this phenomena will originate l paths

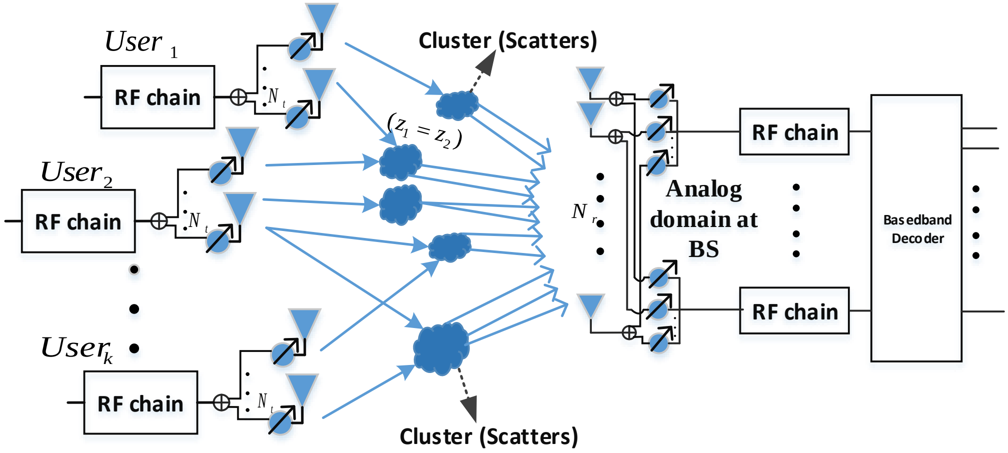


Figure 5.1: Diagram of K users that applies only analog precoder simultaneously transmit to the BS using hybrid D/A combiner mm-Wave massive MIMO system.

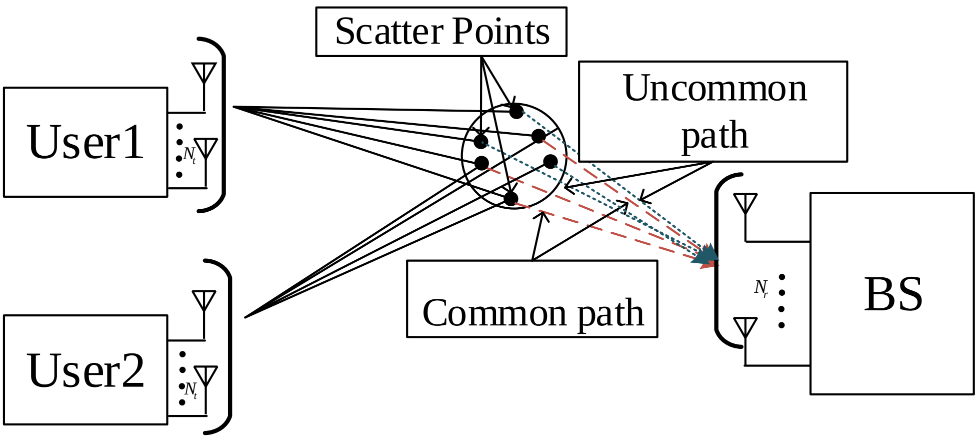


Figure 5.2: Example of a scenario with two users are transmitting through common scatter points to the BS.

from a few physical scatters close to each other forming a cluster. Therefore, in such case where the number of scatters has a significantly large physical size, it is meaningful to contemplate the entire scatterer as common scatters for different l paths and different users as shown in Fig. 5.2. Extraordinarily, when the wavelength shrink, it is possible to have l paths from each scatter point. However, mm-Wave channel measurements [38, 39, 76] showed that most of these paths carrying a very low power which can be ignored and the dominant power is concentrated merely on one or two paths, thus each scatter point produce one path in this work. Even though, it can be also observed from Fig. 5.2 that there is still autonomous lth path which is represented as uncommon path. In $\boldsymbol{H}_k$, the total number of clusters is denoted as N_{cl}. c_{km} denotes the significance of cluster m for user k, which measures the percentage of energy carried by cluster m with $\sum_{m=1}^{N_{cl}} c_{km} = N_{cl}$ [76]. $\boldsymbol{H}_{km}$ is represented as

$$\boldsymbol{H}_{km} = \sqrt{\frac{N_t N_r}{L_{km}}} \sum_{l=1}^{L_{km}} e^{j\Delta_{kz_{l_{km}}}} h^{3D}_{z_{l_{km}}} \boldsymbol{sv}_r(\theta^r_{z_{l_{km}}}, \phi^r_{z_{l_{km}}}) \boldsymbol{sv}_t^H(\theta^t_{z_{l_{km}}}, \phi^t_{z_{l_{km}}}), \tag{5.4}$$

where $\mathbb{L}_{km} = [1, \cdots, L_{km}]$ is the propagation paths index set of user k from cluster m. $\Delta_{kz_{l_{km}}} \in [-\pi/2, \pi/2]$ is the uniformly distributed random phase for l_{km} propagation path and cluster m. $h^{3D}_{z_{l_{km}}}$ is independent and identically distributed (i.i.d) $\mathcal{CN}(0, \delta_k^2)$ random variables with δ_k^2 representing the user k data channel's propagation attenuation. The scatter index set is denoted as $\mathbb{Z}$ where $z_{l_{km}} \in \mathbb{Z}(\forall l_{km} \in \mathbb{L}_{km})$ represents the scatter index of l_{km} paths. $\theta^r_{z_{l_{km}}}$ ($\phi^r_{z_{l_{km}}}$) and $\theta^t_{z_{l_{km}}}$ ($\phi^t_{z_{l_{km}}}$) represent AoA and AoD, respectively. As the uniform planar antenna (UPA) structure [19] is a preferred choice for 3D channel, in this case, $N_t = N_{t_x} \times N_{t_y}$, and $N_r = N_{r_x} \times N_{r_y}$, where (N_{t_x}, N_{r_x}) and (N_{t_y}, N_{r_y}) represent the numbers of antennas in the x-axis and y-axis. In the following notation, the superscript (or subscript) t and r are used to represent transmitter and receiver, respectively. $\boldsymbol{sv}_b(\theta^b_{z_{l_{km}}}, \phi^b_{z_{l_{km}}})$ in (5.4) represent the normalized receive and transmit array response vectors associated with AoA and AoD, where $b \in \{r, t\}$ represents arrival at the receiver (r) or departure at the transmitter (t), is given by [10, 19]

$$\boldsymbol{sv}_b(\theta^b_{z_{l_{km}}}, \phi^b_{z_{l_{km}}}) = \text{vec}[\boldsymbol{sv}_{b_x}(\Psi) \boldsymbol{sv}^H_{b_y}(\Phi)], \quad b \in [r, t], \tag{5.5}$$

where the directions of the array responses $\boldsymbol{sv}_{b_x}$ and $\boldsymbol{sv}_{b_y}$ are represented by $\boldsymbol{sv}_{b_q}$ and devised as

$$\boldsymbol{sv}_{b_q}(\Omega) = \frac{1}{\sqrt{N_{b_q}}} \left[1, e^{j\Omega}, \cdots, e^{j(N_{b_q}-1)\Omega}\right]^T, \quad q \in [x, y], \tag{5.6}$$

where $\Omega \in \{\Psi, \Phi\}$ are the values of Ψ and Φ are given by

$$\Psi = -2\pi\lambda^{-1} u_x \sin(\theta^b_{z_{l_{km}}}) \cos(\phi^b_{z_{l_{km}}}), \quad \Phi = -2\pi\lambda^{-1} u_y \sin(\theta^b_{z_{l_{km}}}) \sin(\theta^b_{z_{l_{km}}}), \tag{5.7}$$

where λ is the wavelength, and u_x and u_y are the inter-element distances in the x-axis and y-axis, respectively.

5.3.2 Cases of Common Scatterers (Overlaped AoA) on the Propagation signal

It has been substantiated that in mm-Wave channels (short distance propagation), the rough surfaces of buildings and small building rifts can cause diffused scattering, which make the propagating wave attenuate considerably as the wavelength shrinks [38, 39]. Therefore, utilizing mm-Wave communications in a dense area (i.e., shopping malls, stadiums, city centres) will force multiple users paths to share the same scatterers. Thus uplink signal transmission from different users in these scenarios produce the same AoAs at the BS as depicted in Fig. 5.1 and Fig. 5.2. Here, substantially, under these scenarios different cases need to be investigated in multi-user hybrid D/A mm-Wave massive MIMO systems. First, the extreme case when all the paths of several users are shared the same common scatterers which arrive with completely overlapped AoAs at the BS. Second, the most likely case to occur when some of the propagation paths are overlapped which are well spotted when two users share some of the common scatterers, i.e., there exists at least two paths $l_1 \in \mathbb{L}_{k_1}, l_2 \in \mathbb{L}_{k_2}$ sharing the same common scatterers points $z_1 = z_2$. Third, the perfect case is defined as several users paths arrive with non-overlapped AoAs at the BS. Our main analysis in this section is stated in Lemma 1.

Lemma 1. *By using $\boldsymbol{sv}_b(\theta^b_{z_{l_{km}}}, \phi^b_{z_{l_{km}}})$ (as in (5.4)), and the assumption of UPA structure, the array responses of different users propagated through common scatters then we have the following result:*

$$\boldsymbol{sv}_b^H(\theta_1^b, \phi_1^b)\boldsymbol{sv}_b(\theta_2^b, \phi_2^b),$$

$$\begin{cases} = 1, & \textbf{if}(\theta_1^b = \theta_2^b)\,\textbf{and}\,(\phi_1^b = \phi_2^b) \\ = \mathrm{v}_2(\theta_1^{\mathrm{b}}, \phi_1^{\mathrm{b}}, \theta_2^{\mathrm{b}}, \phi_2^{\mathrm{b}}), & \begin{bmatrix} \textbf{if}(\theta_1^b \neq \theta_2^b)\,\textbf{and}\,(\phi_1^b = \phi_2^b) \\ \textbf{if}(\theta_1^b = \theta_2^b)\,\textbf{and}\,(\phi_1^b \neq \phi_2^b) \end{bmatrix} \\ = \mathrm{v}_1(\theta_1^{\mathrm{b}}, \theta_2^{\mathrm{b}})\mathrm{v}_2(\theta_1^{\mathrm{b}}, \phi_1^{\mathrm{b}}, \theta_2^{\mathrm{b}}, \phi_2^{\mathrm{b}}), & else \end{cases} \tag{5.8}$$

where $\mathrm{v}_2(\theta_1^b, \phi_1^b, \theta_2^b, \phi_2^b) = \left[\frac{1-e^{jkdn_{b_x}(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}}{1-e^{jkd(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}}\right]$, and $\mathrm{v}_1(\theta_1^b, \theta_2^b) = \left[\frac{1-e^{jkdn_{b_y}(\cos(\phi_1^b)-\cos(\phi_2^b)}}{1-e^{jkd(\cos(\phi_1^b)-\cos(\phi_2^b)}}\right]$
For simplicity and making the analysis clear, the number of paths and shared scatters is assumed to be equal one, where (θ_1^b, ϕ_1^b) and (θ_2^b, ϕ_2^b) are angles for two different users, respectively.

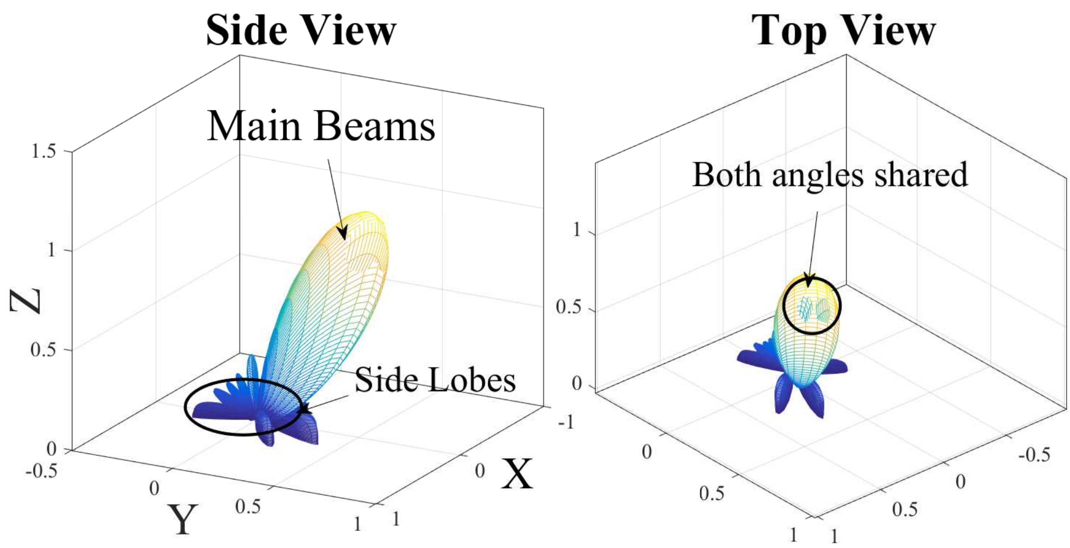


Figure 5.3: Case1: Fully overlapped AoAs where $(\theta_1^b = \theta_2^b)$ and $(\phi_1^b = \phi_2^b)$.

***Proof*:** From (5.4), we have

$$\boldsymbol{sv}_b^H(\theta_1^b, \phi_1^b)\boldsymbol{sv}_b(\theta_2^b, \phi_2^b),$$

$$= \frac{1}{N_b}\sum_{n_{b_x}=1}^{N_{b_x}}\sum_{n_{b_y}=1}^{N_{b_y}} e^{jkd(n_{b_x}-1)(\cos(\theta_1^b)-\cos(\theta_2^b))} \times e^{jkd(n_{b_y}-1)(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}$$

$$= \frac{1}{N_b}\left[\sum_{n_{b_x}=1}^{N_{b_x}} e^{jkd(n_{b_x}-1)(\cos(\theta_1^b)-\cos(\theta_2^b))}\right] \times \left[\sum_{n_{b_y}=1}^{N_{b_y}} e^{jkd(n_{b_y}-1)(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}\right]. \quad (5.9)$$

In the most influential case when the AoAs of different users are completely overlapped due to all the propagation paths and angles of two users are completely shares the same common scatters $(\theta_1^b = \theta_2^b)$ and $(\phi_1^b = \phi_2^b)$, the right hand side of (5.9) is equal to

$$\frac{1}{N_b}[N_{b_x}][N_{b_y}] = \frac{1}{N_b}[N_b] = 1.$$

Importantly, this extrema case means that we cannot differentiate the signals from different users by constructing array response since they share the same AoAs. Thus one of the key concepts of using the hybrid D/A for the mm-Wave massive MIMO system is not practically useful in the system any more, i.e., beamforming. The resulting array response of this case is shown in Fig. 5.3. Note that, in Fig. 5.3, Fig. 5.4, and Fig. 5.5 the system parameters are set as $K = 2$ and $N_r = 16$, so there should be two main spatially beams along with their own side lobes. It can be observed from Fig. 5.3 that it is impossible to distinguish between the two users spatial beams, as it looks as a one beam from the top view of the BS.

Interestingly, in the most likely case where $(\theta_1^b \neq \theta_2^b)$ and $(\phi_1^b = \phi_2^b)$, or $(\theta_1^b = \theta_2^b)$ and $(\phi_1^b \neq \phi_2^b)$, the right hand side of (5.9) is equal to

$$\frac{1}{N_b}[N_{b_y}]\left[\frac{1-e^{jkdn_{b_x}(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}}{1-e^{jkd(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}}\right]=\frac{1}{N_{b_x}}\mathrm{v}_2(\theta_1^{\mathrm{b}},\phi_1^{\mathrm{b}},\theta_2^{\mathrm{b}},\phi_2^{\mathrm{b}}).$$

The array response in this case is shown in Fig. 5.4. It is seen from Fig. 5.4 that the propagation shape has shared an angle whenever either the zenith or azmiuth due to share some scatterers. However, a spatial beam with its own side lobes can be detected by the receiver in Fig. 5.4 as either the zenith or azimuth AoAs does not share common scatterers. For the last case where $(\phi_1^b \neq \phi_2^b)$ and $(\theta_1^b \neq \theta_2^b)$, again we take the right hand side of (5.9) which gives

$$\frac{1}{N_b}\left[\frac{1-e^{jkdn_{b_y}(\cos(\phi_1^b)-\cos(\phi_2^b)}}{1-e^{jkd(\cos(\phi_1^b)-\cos(\phi_2^b)}}\right]\left[\frac{1-e^{jkdn_{b_x}(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}}{1-e^{jkd(\sin(\theta_1^b)\sin(\phi_1^b)-\sin(\theta_2^b)\sin(\phi_2^b))}}\right]=$$

$$\frac{1}{N_b}\mathrm{v}_1(\phi_1,\phi_2)\mathrm{v}_2(\theta_1^{\mathrm{b}},\phi_1^{\mathrm{b}},\theta_2^{\mathrm{b}},\phi_2^{\mathrm{b}}). \tag{5.10}$$

Fig. 5.5 shows the case when AoAs of different user from side and top views are well separated $(\theta_1^b \neq \theta_2^b)$ and $(\phi_1^b \neq \phi_2^b)$, i.e., the users do not share any common scatterers, which allows them to align their beams in a specific direction. Therefore, two spatially separated beams are formed in Fig. 5.5, this perfect case is named as non-overlapped AoAs. □

Lemma 1 shows that the multi-user hybrid D/A mm-Wave massive MIMO system can easily lose its effectiveness in applying the beamforming technique for transmission when the propagation signal (wave) of different users passes through common scatterers in 5G scenarios.

In the existing proposed multiuser hybrid D/A precoding [24–27] assumed that as a large number of antennas is implemented, the channel vector of a desired user will tend to be independent or more asymptotically orthogonal to the channel vector of a randomly

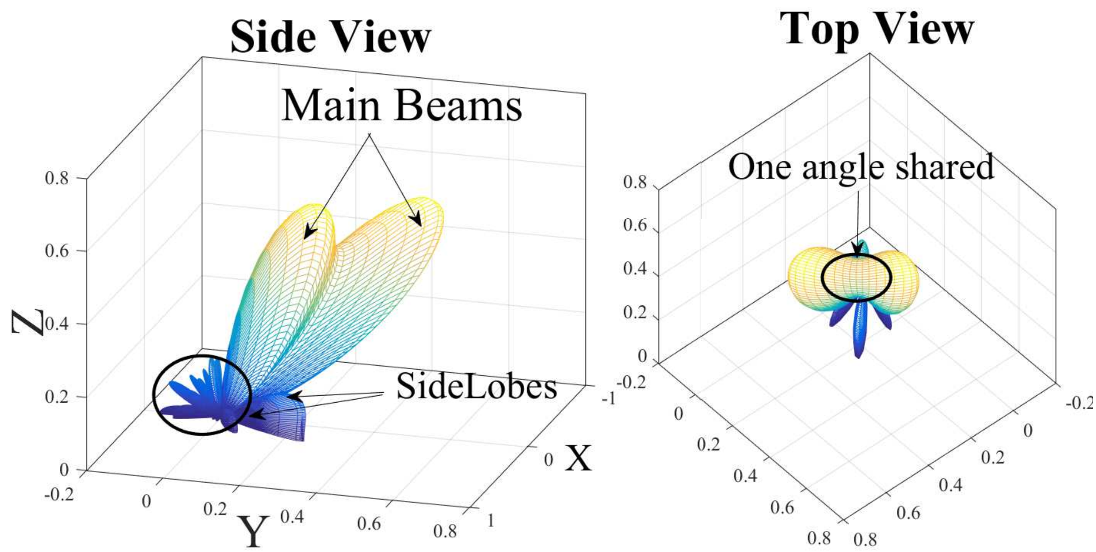


Figure 5.4: Case2: Partially overlapped AoAs where ($\theta_1^b \neq \theta_2^b$) and ($\phi_1^b = \phi_2^b$), or ($\theta_1^b = \theta_2^b$) and ($\phi_1^b \neq \phi_2^b$).

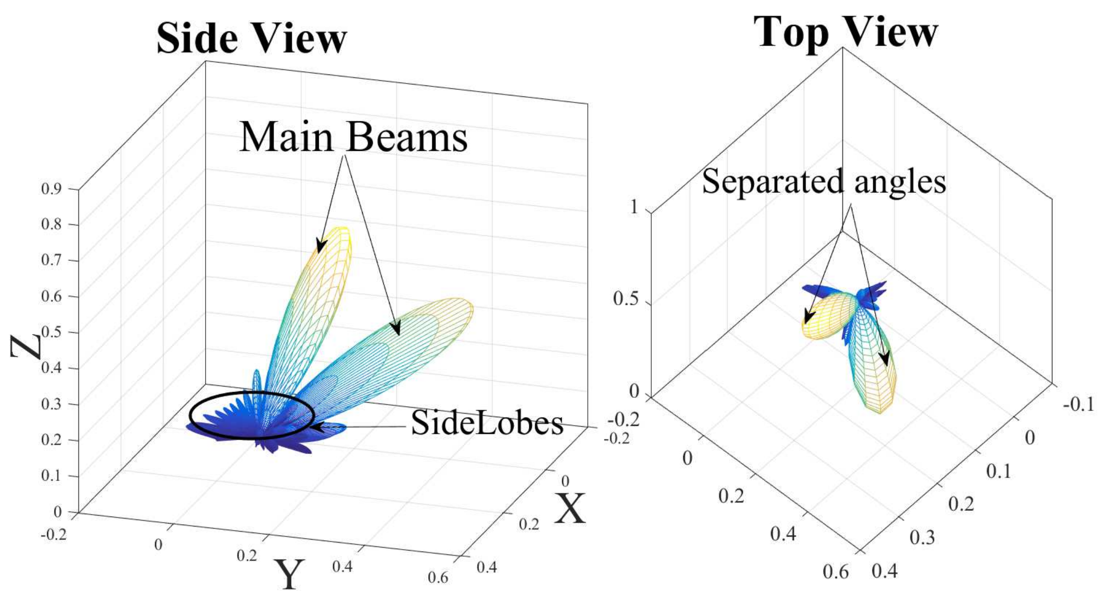


Figure 5.5: Case3: Non-overlapped AoAs where ($\theta_1^b \neq \theta_2^b$) and ($\phi_1^b \neq \phi_2^b$).

selected interfering user most of the time. However, in reality, when different users transmitting through common scattering environments the channel vectors for different users are correlated, or not asymptotically orthogonal any more [77]. Intuitively, the correlation between the channel vectors of different users in mm-Wave system defined in (5.11), is a good option to measure the orthogonality of these channel vectors. The correlation $\Pi_{i,k}$ of the magnitude of the inner product between the channel vectors is defined as

$$\Pi_{i,k} \triangleq \frac{|\boldsymbol{sv}_b^H(\theta_i^b, \phi_i^b)\boldsymbol{sv}_b(\theta_k^b, \phi_k^b)|}{||\boldsymbol{sv}_b(\theta_i^b, \phi_i^b)||_2||\boldsymbol{sv}_b(\theta_k^b, \phi_k^b)||_2}. \tag{5.11}$$

To show the impact of the channel correlation (i.e., when different users propagate through the common scattering regime) on the achievable uplink rate, a simulation result is formed between two users in Fig. 5.6. The conventional algorithms GS [73], OMP [27] and MRC [26] based on hybrid D/A mm-Wave massive MIMO systems is applied in this result at SNR -10 dB. The channel parameters are set as 12 scatters, 4 clusters and $L_{km} = 3$, also the BS is equipped with $N_{RF} = K = 2, N_r = 128$, and each user is built with $N_t = 16$. Importantly, it is seen that all the algorithms loss their performance sharply when there is a high correlation between the channel vectors of the two users especially after -20 dB. In other words, most the propagation paths of the measured users share most of the common scatters in the channel which arrive with highly overlapped AoAs at the BS. However, in a low correlation when the propagation paths not share common scatters all the algorithms achieve similar and high achievable uplink rate. Therefore, theoretically there are two correlation regions; low correlation region and high correlation region. As our main encourage in this chapter is to maximize the achievable uplink rate in the common scattering environment, therefore the high correlation region is the target when designing the precoders and combiners for the hybrid D/A system. Undoubtedly, the conclusion of this important result is contrastingly to the law of large numbers of antenna can provide independent channel vectors, and also show the existing algorithms cannot perform good in the practical environments.

***Remark* 4.** *The key finding from Fig. 5.3, Fig. 5.4, Fig. 5.5 and Fig. 5.6, is that the propagation signals of different users that arrive at the BS through common scatterers with different cases of overlapped AoA can interfere with each other. In [24–27], this type of impractical interference was not considered which spoils the achievable uplink rate of multi-user hybrid D/A mm-Wave massive MIMO systems. Another impractical assumption is, when there are high number of antennas that are equipped in the system, the AoAs of different users can be spatially separated well as having a very low correlation. However,*

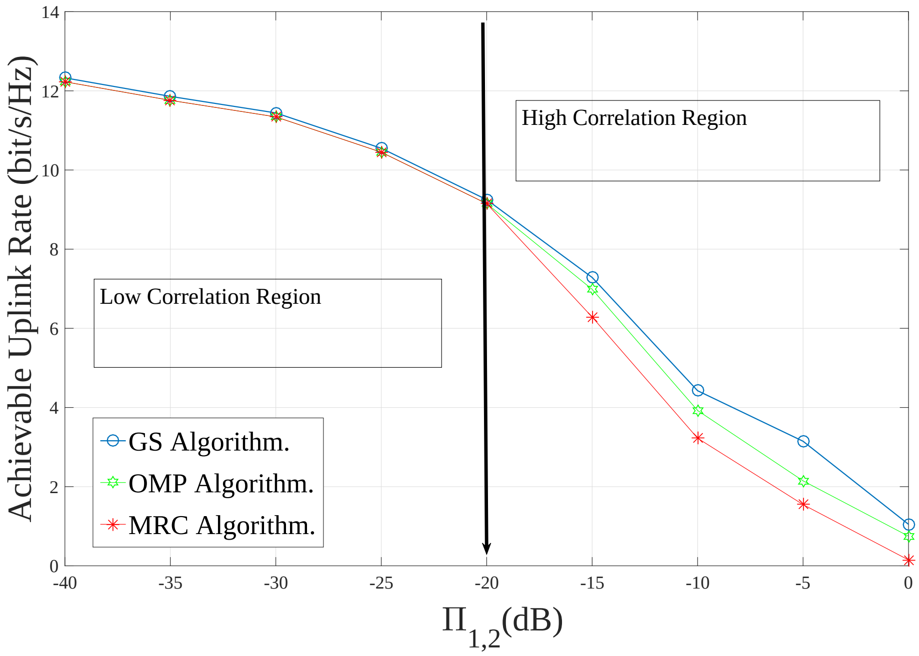


Figure 5.6: Achievable uplink rate versus correlation between two users.

our analysis and results in this section have proved that these assumptions are neither realistic nor valid for multi-user hybrid D/A mm-Wave massive MIMO systems.

5.4 Problem Formulation

For the scenarios considered in this chapter where mm-Wave transmission is utilized in a dense area (i.e., shopping malls, stadiums, city centres), multiple users transmit their signals through common scatterers. This results in the AoA spreading and the propagation multipath for the desired user being confined to a region of space where interfering paths from other users are very likely to exist. If this is the case, this will cause severe interference at the BS as discussed in Section 5.3.2. This problem constitutes a challenge that exhausts the achievable uplink rate of multi-user hybrid D/A mm-Wave massive MIMO systems. Moreover, hybrid D/A with a limited number of RF chains is utilized in 5G, thus interference due to the propagation paths that are transmitted through common scatterers prevails and cannot be eliminated directly by using traditional interference cancellation such as MMSE and ZF at the digital part only [24–27]. Therefore, the BS becomes incapable of precisely eliminating and separating those uplink signals from different users that share the same AoAs even when applying the existing hybrid D/A precoders and combiners in [24–27].

Previously, to mitigate the risk of this problem, a null space method solution has been proposed in [78], principally based on capturing the desired user paths from the interfering paths while requiring the fully digital system. The role of this solution is assembled based on the assumption that the orthogonal signal subspaces can be derived from the covariance matrix $\boldsymbol{R}$, where $\boldsymbol{R}$ is obtained separately from the desired and interference channels. This provides the ability to abstract the desired signal from the interfering signals by making all the paths from different users fall into the null space of its covariance matrix. Consequently, an ideal role is assumed that, *the possibility of different propagation paths to have overlapped AoAs at the BS is almost zero*. This result is rooted in the law of the fully digital system, because such a solution needs to be spatially applied in the system which requires that the number of RF chains equals the number of antennas. However, this method is impossible to be directly applied to hybrid D/A mm-Wave massive MIMO systems, as fewer RF chains are utilized in the system. One of the limiting points of employing the null space solution, inaccurately, it is also suggested that the BS is always situated in a high building, and the AoA spread of different user paths that impinges at the BS, are well separated. Nevertheless, this is not the case in 5G short distance propagation i.e., mm-Wave channels [38, 39].

Recent research in [24–27, 79] focused on multi-user hybrid D/A mm-Wave massive MIMO systems based on the same assumption that, *the possibility of different propagation paths to have overlapped AoAs at the BS is almost zero*. However, this role cannot be contemplated as a straightforward solution for 5G environments when multi-user hybrid D/A mm-Wave massive MIMO systems are utilized because of the aforementioned reasons.

In this chapter, the user and BS have knowledge of channel $\boldsymbol{H}_k$. In contrast to the literature [24, 25, 27], our proposed approach to find the hybrid D/A precoders and combiners aim to maximize the achievable uplink rate and eliminating interference due to the same AoAs of the multi-user hybrid D/A mm-Wave massive MIMO system directly. The achievable uplink rate of the system is given by

$$R^{UL} = \sum_{k=1}^{K} \log(1 + \mathrm{SINR}_k), \tag{5.12}$$

where SINR_k is SINR of the user k which is presented as

$$\mathrm{SINR}_k = \frac{|\mathbf{g}_k^H \boldsymbol{H}_k \boldsymbol{j}_{RF_k}|^2}{\sum\limits_{i=1, i\neq k}^{K} |\mathbf{g}_k^H \boldsymbol{H}_i \boldsymbol{j}_{RF_i}|^2 + \sigma^2 ||\mathbf{g}_k||_2^2}. \tag{5.13}$$

The prime target of this section below is to maximize the achievable uplink rate of the hybrid D/A precoding and combing system synchronously with reducing the system interferences that are combinationally generated due to preliminary the sharing of the same AoAs and then the multi-streams at the BS. To accomplish our target, we are interested in optimizing the analog precoder $\boldsymbol{j}_{RF_k}$, combiner $\boldsymbol{A}$, and the digital combiner $\mathbf{d}_k$. Therefore, the optimization problem is formulated as

$$R^{UL} = \max_{\{\boldsymbol{j}_{RF_k},\, \boldsymbol{A},\, \mathbf{d}_k\}} \sum_{k=1}^{K} \log(1 + \mathrm{SINR}_k), \tag{5.14a}$$

$$\text{s.t.} \quad |(\boldsymbol{j}_{RF})_{n_t}| = 1, \forall n_t, \tag{5.14b}$$

$$|(\boldsymbol{A})_{n_r, n_{RF}}| = 1, \forall n_r, \forall n_{RF}, \tag{5.14c}$$

where the first and second constraints in (5.14b), and (5.14c), are subjected to the non-convex constant modulus constraints, restricting all the non-zero elements to have the same amplitude in the analog domain of the user and BS sides. Unfortunately, this makes the Problem in (5.14a) more intractable. Practically, harvesting the potential gains of nulling the overlapped AoAs interference in the analog beamformers aids to reduce the quantization noise concurrently relieves the degrees of freedom necessity at the digital combiner. Thus technically decreases the required number of RF chains in the hybrid D/A system. In this situation, the analog beamformers needs to be precisely designed so as to alleviate interference due to comparable AoAs as much as possible, as well as correspondingly adjust the analog precoder.

In the following section, the analog and digital precoders and combiners are designed for multi-user hybrid D/A mm-Wave massive MIMO systems for attaining the target of the optimization Problem in (5.14a).

5.5 Proposed Hybrid D/A Precoders and Combiners

We present the proposed schemes in this section, where the aim is to design the hybrid D/A schemes that help in maximizing the effective gain for a particular user and concurrently reducing the system interference in both the analog and digital precoders and combiners. The analog precoder $\boldsymbol{j}_{RF_k}$ for each user k is designed based on knowledge of $\boldsymbol{H}_k$, and then the analog combiner $\boldsymbol{A}$ is designed at the BS. Effectively, to conceive the effective uplink analog channel coefficient between the receiver and transmitter in a consolidated matrix that is denoted as $\boldsymbol{H}^{EF}$. Primarily, let combine all the analog precoder vectors in a matrix as $\boldsymbol{J} = \text{diag}[\boldsymbol{j}_{RF_1}, \boldsymbol{j}_{RF_2}, \cdots, \boldsymbol{j}_{RF_K}]$ and then the channel of each user in a matrix that is represented as $\boldsymbol{H} = [\boldsymbol{H}_1^T, \boldsymbol{H}_2^T, \cdots, \boldsymbol{H}_K^T]^T$, thus $\boldsymbol{H}^{EF}$ is expressed as

$$\boldsymbol{H}^{EF} = \boldsymbol{A}^H \boldsymbol{H} \boldsymbol{J}. \tag{5.15}$$

Considerably, the diagonal entries of $\boldsymbol{H}^{EF}$ denote the desired effective uplink analog channel gains, while the off-diagonal entries unwieldy consist of the analog interference due to the same AoAs and also the digital multi-stream interference. For this reason, in what follows, the designing of the analog precoders and combiners must be precise, for initially maximizing the diagonal elements of $\boldsymbol{H}^{EF}$, then sequentially vanishing the off-diagonal elements. Importantly, the digital combiner will be designed wisely to further decrease the off-diagonal entries effects. For more clarification, the k-th diagonal term of $\boldsymbol{H}^{EF}$ can be represented as

$$\boldsymbol{H}^{EF}(N_{RF}, k) = \boldsymbol{a}_{N_{RF}}^H \boldsymbol{H}_k \boldsymbol{j}_{RF_k}. \tag{5.16}$$

Note here, the analog precoder and combiner are designed in the system separately, the aim of doing this is to reduce the complexity of the system and also evade any extra information that can be exchanged between the receiver and transmitter. The proposed framework of the steps and algorithms that enhance the receive SNR as well as nulling the interferences three key steps. Firstly, the analog beamformers $\boldsymbol{j}_{RF_k}$ and $\boldsymbol{A}$ are designed respectively under a non-convex constant modulus constraints (5.14b), and (5.14c), for each user intending at maximizing particularly the corresponding effective uplink analog channel gains. Secondly, to tractably mitigate the interference from the same AoAs, Power Iteration (PI) is proposed for forming the analog precoder, then a Riemannian optimization method based on Stiefel manifold (ROSt) under orthogonal constraints is imposed to design the analog combiner. Thirdly, in the sequel to further diminish the effect of the interference in the system, MMSE solution is adopted which aims in designing the digital combiner based on the effective uplink analog channel gains. Therefore, the upcoming subsections representative the design of the analog precoder, combiner and the digital combiner that are discussed in details.

5.5.1 Analog Precoders Design

The design of the analog precoder $\boldsymbol{j}_{RF_k}$ is proposed as follows. As defined previously, $\boldsymbol{j}_{RF_k}$ is the analog precoder of each user k which is implemented by phase-only. To exploit the potential gains that each user posses its channel $\boldsymbol{H}_k$, an algorithm is proposed to calculate only the eigenvector that corresponds to the largest eigenvalue of a matrix. This algorithm is called Power Iteration (PI), which is used robustly in Google for the web page rank, as the web page, nodes and their hyperlinks are represented as a large matrix [75]. PI algorithm involves multiplying the diagonalize matrix by a randomly chosen vector, and iteratively normalizing and multiplying the matrix by the normalized vector from the previous step. After efficiently computing the largest eigenvalue and denoting the corresponding eigenvector of a diagonalizable matrix, we then extract the phase of the eigenvector that is corresponding to the largest eigenvalue to precisely create the analog precoder $\boldsymbol{j}_{RF_k}$ at each user. Suppose $\mathbf{P} = \boldsymbol{H}_k^H \boldsymbol{H}_k$ where $\mathbf{P} \in \mathbb{C}^{N_t \times N_t}$, is Hermitian and the diagonalizable matrix. Therefore, $\mathbf{P}$ consist of N_t linear independent eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_t}$ where $\mathbf{u}_{N_t} \in \mathbb{C}^{N_t \times 1}$ with corresponding eigenvalues of $\gamma_1, \gamma_2, \cdots, \gamma_{N_t}$, such that $\mathbf{P}\mathbf{u}_e = \gamma_e \mathbf{u}_e$ for $e = 1, \cdots, N_t$. For the convergence analysis, those eigenvalues are organized in a way that $|\gamma_1| \geq |\gamma_2| \geq \cdots \geq |\gamma_{N_t}|$, where γ_1 is the dominant eigenvalue with a corresponding eigenvector of $\mathbf{u}_1$. In PI, suppose that a random given vector $\mathbf{v}^0 \in \mathbb{C}^{N_t \times 1}$ be an approximation to an eigenvector of $\mathbf{P}$ with $||\mathbf{v}^0|| = 1$ and ${\mathbf{v}^0}^H \mathbf{u}_1 \neq 0$, continually multiply by $\mathbf{P}$, producing a sequence of vectors $\mathbf{v}^1, \mathbf{v}^2, \cdots$ defined as

$$\mathbf{v}^\alpha = \mathbf{P}\mathbf{v}^{(\alpha-1)} = \mathbf{P}^\alpha \mathbf{v}^0, \tag{5.17}$$

where $\alpha = 1, 2, \cdots$ is denoted as the number of iterations. Fortunately, as $\mathbf{P}$ is diagonalizable matrix, thus each vector is a linear combination of the eigenvectors which is represented as

$$\mathbf{v}^0 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_{N_t}\mathbf{u}_{N_t}, \tag{5.18}$$

where $c_1 \neq 0$ is a normalizing constant to avoid $\mathbf{u}_e$ from getting too large or too small. Then, we have

$$\begin{aligned}
\mathbf{v}^1 = \mathbf{P}\mathbf{v}^0 &= c_1\mathbf{P}\mathbf{u}_1 + \cdots + c_{N_t}\mathbf{P}\mathbf{u}_{N_t} \\
&= c_1\gamma_1\mathbf{u}_1 + \cdots + c_{N_t}\gamma_{N_t}\mathbf{u}_{N_t} \\
&= \gamma_1\left[c_1\mathbf{u}_1 + \left(\frac{\gamma_2}{\gamma_1}\right)c_2\mathbf{u}_2 + \cdots + \left(\frac{\gamma_{N_t}}{\gamma_1}\right)c_{N_t}\mathbf{u}_{N_t}\right] \\
\mathbf{v}^2 = \mathbf{P}^2\mathbf{v}^1 &= \gamma_1^2\left[c_1\mathbf{u}_1 + \left(\frac{\gamma_2}{\gamma_1}\right)^2 c_2\mathbf{u}_2 + \cdots + \left(\frac{\gamma_{N_t}}{\gamma_1}\right)^2 c_{N_t}\mathbf{u}_{N_t}\right] \\
&\vdots \\
\mathbf{v}^\alpha = \mathbf{P}^\alpha\mathbf{v}^{\alpha-1} &= \gamma_1^\alpha\left[c_1\mathbf{u}_1 + \left(\frac{\gamma_2}{\gamma_1}\right)^\alpha c_2\mathbf{u}_2 + \cdots + \left(\frac{\gamma_{N_t}}{\gamma_1}\right)^\alpha c_{N_t}\mathbf{u}_{N_t}\right].
\end{aligned} \tag{5.19}$$

As $|\gamma_1| \geq |\gamma_e|$ for $e = 2, \cdots, N_t$, it inferences that the coefficients of $\mathbf{u}_e$ where $e = 2, \cdots, N_t$, converge to zero as $\alpha \to \infty$. Therefore, as α increases the vector $\mathbf{P}^\alpha\mathbf{v}^0$ tends to assume an increasingly significant component in the direction of converges eigenvector $\mathbf{u}_1$, while its components in other directions $\mathbf{u}_{N_t}$ decrease. This conclude to an exceedingly simple method of evaluating an eigenvalue and eigenvector. In PI, the simplest criterion that is proposed to find the accuracy of an approximate eigenpair $\mathbf{v}^\alpha, \gamma^\alpha$, is to compute the norm of the so called residual vector

$$\mathbf{r} = \mathbf{P}\mathbf{v}^\alpha - \gamma^\alpha\mathbf{v}^\alpha.$$

This information can in turn be helpful in selecting a stopping criterion in PI. Therefore, the stopping criterion in PI is introduced as

$$|\gamma^\alpha - \gamma^{\alpha-1}| \leqslant ||\mathbf{r}||_2.$$

Another challenge, the vector $\mathbf{v}^\alpha$ may either underflow in the case that $||\mathbf{P}|| < 1$ or overflow when $||\mathbf{P}|| > 1$ especially for large α. Moreover, the convergence rate of

Algorithm 7 Power Iteration based algorithm (PI) for designing the analog precoder.

1: Initialize a starting vector randomly with $||\mathbf{v}^0|| = 1$ and $\mathbf{v}^{0^H}\mathbf{u}_1 \neq 0$;
2: **for** $\alpha = 1, 2, \cdots$ **do**
3: $\quad\alpha = \alpha + 1$;
4: $\quad\mathbf{v}^\alpha = \mathbf{P}\mathbf{v}^{(\alpha-1)}$;
5: $\quad\mu_\alpha = ||\mathbf{P}\mathbf{v}^{\alpha-1}||_2$;
6: $\quad\mathbf{v}^\alpha = \mathbf{v}^\alpha/\mu_\alpha$;
7: $\quad\gamma^\alpha = \mathbf{v}^{\alpha^H}\mathbf{P}\mathbf{v}^\alpha$
8: $\quad$**if** $|\gamma^\alpha - \gamma^{\alpha-1}| \leqslant ||\mathbf{r}||_2$, **then** stop
9: **end for**

PI depends on $|\frac{\gamma_2}{\gamma_1}|$, significance that the distance between $\mathbf{v}^\alpha$ and a vector $\mathbf{u}_1$ decreases roughly by this factor from iteration to iteration. Practically, this may leads PI to converge slowly when γ_2 is almost as large as γ_1. Thus in PI that is structured in Algorithm 7, both problems are remedied by normalizing the vector $\mathbf{v}^\alpha$ at each iteration to avoid PI having underflow, overflow and slower convergence. Thus the normalization step is defined as

$$\mathbf{v}^{\alpha+1} = \frac{\mathbf{P}\mathbf{v}^\alpha}{||\mathbf{P}\mathbf{v}^\alpha||_2}.$$

Now, after finding the eigenvector and eigenvalue by utilizing PI in Algorithm 7, the analog precoder $\boldsymbol{j}_{RF_k}$ is then proposed as the eigenvector which corresponds to the largest eigenvalue with an element-wise normalization to satisfy the non-convex constraint (5.14b), of phase shifters, represented as

$$\boldsymbol{j}_{RF_k}(e) = \frac{1}{\sqrt{N_t}} e^{j \ \arg[\varrho_e]}, \tag{5.20}$$

where $\arg(x)$ generates an element containing the phase of the entries of the element x, which results in, ϱ_e is the phase of the eth element in $\mathbf{u}_1$. Hereafter efficiently designing the optimal first part in the diagonal elements $\boldsymbol{a}^H_{N_{RF}}\boldsymbol{H}_k\boldsymbol{j}_{RF_k}$ of $\boldsymbol{H}^{EF}$, the next important step in the following subsection is to design the analog combiner in order to shed further light on the significance of maximizing the diagonal while handling interference of off-diagonal of $\boldsymbol{H}^{EF}$.

5.5.2 Analog Combiners Design

In this subsection, the essential challenge now is to design the analog combiner that simultaneously satisfies the non-convex constraint in (5.14c), helps in maximizing the diagonal of $\boldsymbol{H}^{EF}$ and to actively put down the off-diagonal elements value that represents the interference of the system. Unfortunately, the constraint (5.14c) in Problem (5.14a) is a non-convex constraint which can incur a high computational complexity to the system to solve since it cannot be relaxed. Therefore, a novel efficient method to deal with this constraint is to convert this constraint into an additional unitary condition such that $\boldsymbol{W}^H\boldsymbol{W} = \boldsymbol{I}_{N_{RF}}$, resulting in, the reformulations of Problem (5.14a) can be expressed after replacing the analog combiner matrix $\boldsymbol{A}$ with the matrix $\boldsymbol{W} \in \mathbb{C}^{N_r \times N_{RF}}$, as

$$R^{UL} = \max_{\{\boldsymbol{A}=\boldsymbol{W}\}} \sum_{k=1}^{K} \log(1 + \mathrm{SINR}_k), \tag{5.21a}$$

$$\text{subject to} \quad \boldsymbol{W}^H\boldsymbol{W} = \boldsymbol{I}_{N_{RF}}, \tag{5.21b}$$

where $\boldsymbol{W}$ is the optimization parameter that needs to be optimized. Note here, the constraint (5.14b) has not imposed in Problem (5.21a), because it is satisfied in the previous Section 5.5.1. The constraint set $\boldsymbol{W}^H\boldsymbol{W} = \boldsymbol{I}_{N_{RF}}$ is a submanifold of $\mathbb{C}^{N_r \times N_{RF}}$ called the Stiefel manifold. The optimization Problem in (5.21a) is defined over a novel ROSt based on Stiefel manifold which is of interest in the study of geometry, topology, and analysis [56, 57]. The standard notation of Stiefel manifold is noted as $\mathcal{S}_{N_{RF}}^{N_r}$. The reason for selecting the Stiefel manifold is that the (compact) Stiefel manifold is the Riemannian submanifold of orthogonal matrices [56, 57]. Therefore, we only use orthogonal Stiefel manifold matrices $\boldsymbol{W} \in \mathcal{S}_{N_{RF}}^{N_r}$ to represent subspaces. An optimal $\boldsymbol{W}$ is required to be selected in the tangent plane of the Stiefel manifold which can maximize the cost function. The feasible set of Stiefel manifold $\mathcal{S}_{N_{RF}}^{N_r}$ is represented as $\mathcal{S}_{N_{RF}}^{N_r} = \left\{\boldsymbol{W} \in \mathbb{C}^{N_r \times N_{RF}}, \boldsymbol{W}^H\boldsymbol{W} = \boldsymbol{I}_{N_{RF}}\right\}$. Algorithm 8 based on ROSt with their three stages in each iteration is proposed to find the optimal $\boldsymbol{W}$. Following steps of Algorithm 8 based on ROSt will be carried out; 1) the tangent space refinement; 2) Riemannian gradient; 3) retraction. Let us discuss the steps in details.

Step-1 *Tangent Space Refinement of a Manifold Structure:*

The ROSt based on Stiefel manifold admits a tangent space at each point. Each point

$\boldsymbol{W}$ in an embedded sub-manifold has a tangent space which is represented by $\mathcal{T}_{\boldsymbol{W}}\mathcal{S}_{N_{RF}}^{N_r}$. The tangent space consists of a tangent vector ξ of the curves which is represented by a matrix $\boldsymbol{\xi} \in \mathbb{C}^{N_r \times N_{RF}}$, thus, the tangent space at point $\boldsymbol{W}$ can be depicted as

$$\mathcal{T}_{\boldsymbol{W}}\mathcal{S}_{N_{RF}}^{N_r} = \{\boldsymbol{\xi} \in \mathbb{C}^{N_r \times N_{RF}} : \mathbb{R}\{\boldsymbol{\xi} \odot \boldsymbol{W}^H\} = \mathbf{0}_{N_r \times N_{RF}}\}. \tag{5.22}$$

In a manifold, each tangent space is given with an inner product. In this algorithm, Riemannian metric is used as the inner product which varies smoothly from point to point and is represented as

$$\forall \boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathcal{T}_{\boldsymbol{W}}\mathcal{S}_{N_{RF}}^{N_r}, \langle \boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \rangle = \text{trace}(\boldsymbol{\xi}_1^H \boldsymbol{\xi}_2). \tag{5.23}$$

Step-2 *The Riemannian Gradient:* The computation of the gradient descent direction is based on the Riemannian gradient structure of the Stiefel manifold. Generally, the Riemannian gradient of f at point $\boldsymbol{W}$ is simply the linear projection of the ordinary gradient of f onto the tangent space. Therefore, the Riemannian gradient of the function at $\boldsymbol{W}$ is defined by computing the gradient of $\text{grad}f(\boldsymbol{W})$, in usual way, then projecting the Euclidean gradient $\triangledown f(\boldsymbol{W})$ firstly onto the tangent space $\mathcal{T}_{\boldsymbol{W}}\mathcal{S}_{N_{RF}}^{N_r}$, and then to the Stiefel manifold

$$\text{grad}f|_{\mathcal{S}_{N_{RF}}^{N_r}}(\boldsymbol{W}) = \mathbf{Proj}_{\boldsymbol{W}}^{\mathcal{S}} \triangledown f(\boldsymbol{W}) = (\boldsymbol{I}_{N_{RF}} - \boldsymbol{W}\boldsymbol{W}^H) \triangledown f(\boldsymbol{W}), \tag{5.24}$$

where $\triangledown f(\boldsymbol{W})$ named the Euclidean gradient of (5.21a), which is derived as

$$\triangledown f(\boldsymbol{W}) = \left(\frac{1}{\ln(2)\left(1 + \frac{|\mathbf{g}_k^H \boldsymbol{H}_k \boldsymbol{j}_{RF_k}|^2}{|\mathbf{g}_k^H \boldsymbol{H}_i \boldsymbol{j}_{RF_i}|^2 + \sigma^2 ||\mathbf{g}_k||_2^2}\right)} \cdot \frac{\mathbf{g}_k^H \boldsymbol{H}_k \boldsymbol{j}_{RF_k} \boldsymbol{j}_{RF_k}^H \boldsymbol{H}_k^H (|\mathbf{g}_k^H \boldsymbol{H}_i \boldsymbol{j}_{RF_i}|^2 + \sigma^2 ||\mathbf{g}_k||_2^2) - |\mathbf{g}_k^H \boldsymbol{H}_k \boldsymbol{j}_{RF_i}|^2 (\mathbf{g}_k^H \boldsymbol{H}_i \boldsymbol{j}_{RF_i} \boldsymbol{j}_{RF_i}^H \boldsymbol{H}_i^H + \sigma^2 \mathbf{g}_k^H)}{(|\mathbf{g}_k^H \boldsymbol{H}_i \boldsymbol{j}_{RF_i}|^2 + \sigma^2 ||\mathbf{g}_k||_2^2)^2} \right), \tag{5.25}$$

where $\mathbf{g}_k = \boldsymbol{W}\mathbf{d}_k$, in this stage only the digital combiner is assumed to be an identity matrix $\boldsymbol{I}_{N_{RF}}$.

Step-3 *The Retraction:* Retraction $\mathrm{R}_{\boldsymbol{W}}(\boldsymbol{\xi})$ on $\mathcal{S}_{N_{RF}}^{N_r}$ is another concept which maps every vector from a given point $\boldsymbol{W}$ along a prescribed direction $\boldsymbol{\xi}$ while remaining on the manifold, and is expressed as

$$\mathrm{R}_{\boldsymbol{W}}(\boldsymbol{\xi}) = \text{polar}(\boldsymbol{W} + \boldsymbol{\xi}), \tag{5.26}$$

where polar$(\boldsymbol{W}+\boldsymbol{\xi}) \in \mathcal{S}_{N_{RF}}^{N_r}$ designates the $N_r \times N_{RF}$ orthonormal factor of the polar decomposition of $(\boldsymbol{W}+\boldsymbol{\xi})$. For example, if $\boldsymbol{\mho} = (\boldsymbol{W}+\boldsymbol{\xi})$, then to compute the polar of $\boldsymbol{\mho}$, a thin SVD is used which means $\boldsymbol{\mho} = \boldsymbol{\Xi\Sigma\Gamma}^T$, and polar$(\boldsymbol{\mho}) = \boldsymbol{\Xi\Gamma}^T$. In order to guarantee that the objective function is non-decreasing at each iteration α, Armijo backtracking line step of length ε and Polak-Ribiere parameter [57] are defined and tracked in Algorithm 8. Essentially on manifolds, it is necessary to compare vectors belonging to different tangent spaces. Ideally, this happens in the case where the gradient at the current iteration wants to combine with the search direction followed at the previous iteration. Consequently to combine them, a mathematical tool capable of computing the transport vectors [56] of the tangent vector $\boldsymbol{\xi}$ at $\boldsymbol{W}^{\alpha}$ to $\boldsymbol{W}^{\alpha+1}$ is applied, which is represented as

$$\text{Transp}_{\boldsymbol{W}^{\alpha}\rightarrow\boldsymbol{W}^{\alpha+1}} = (\boldsymbol{I}_{N_{RF}} - \boldsymbol{W}^{\alpha+1}(\boldsymbol{W}^{\alpha+1})^H) \bigtriangledown f(\boldsymbol{W}). \tag{5.27}$$

(5.27) is applied as given in step 6 of Algorithm 8 which provides the architecture of ROSt algorithm based on Stiefel manifold optimization. For making certain that Algorithm 8 based on the ROSt converge, suppose $\boldsymbol{W}^{\alpha}$ be an infinite sequence of iterations originated by Algorithm 8. The resulting each accumulation point is a critical point of the defined cost function. Therefore, in theorem 4.3.1 in [56] and section 4.1 in [80], it has been proved that the convergence on manifolds is guaranteed.

Notice that, one of the main points of selecting (5.21a) to be the Euclidean gradient is when designing $\boldsymbol{W}$ instead of only focusing on the desired user (i.e., the diagonal elements of $\boldsymbol{H}^{EF}$), is also to effectively design an optimal $\boldsymbol{W}$ that helps in keeping the off-diagonal of the effective uplink analog channel elements in low values as an another target as well. To clarify this issue conceptually, the receiver has to apply the same analog beamforming matrix to combine all K users' data. Fortunately, the optimal $\boldsymbol{W}$ already consist of K columns of optimal vectors $[\boldsymbol{w}_1^*, \boldsymbol{w}_2^*, \cdots, \boldsymbol{w}_K^*]$, where $\boldsymbol{w}_k \in \mathbb{C}^{N_r}$, which can help in guaranteeing the fairness among all K users. If this is the case then mathematically each vector can simultaneously maximize its corresponding gain $\boldsymbol{w}_k^* = \underset{\boldsymbol{w}_k}{\text{argmax}}\,||\boldsymbol{w}_k^H\boldsymbol{H}_k\boldsymbol{j}_{RF_k}||_F^2$. Mutually the target of decreasing the interference, when different users with different effective uplink analog channel coefficients for instance $\boldsymbol{w}_k^H\boldsymbol{H}_k\boldsymbol{j}_{RF_k}$ and $\boldsymbol{w}_i^H\boldsymbol{H}_i\boldsymbol{j}_{RF_i}$ are highly correlated, the vectors $\boldsymbol{w}_k$ and $\boldsymbol{w}_i$ in the analog domain $\boldsymbol{W}$ attains this condition ${\boldsymbol{w}_k^*}^H\boldsymbol{w}_i^* \approx 0$ which can actively put down the interference due to the correlated vectors.

Algorithm 8 Riemannian optimization methods based on Stiefel manifold (ROSt).

1: Initialize $\alpha = 0$ and $\boldsymbol{W}^0 \in \mathcal{S}_{N_{RF}}^{N_r}$;

2: $\boldsymbol{\xi}^0 = -\mathrm{grad} f(\boldsymbol{W}^0)$

3: **Repeat**

4: $\varepsilon^\alpha = \mathrm{Linesearch}(\boldsymbol{W}^\alpha, \boldsymbol{\xi}^\alpha, \boldsymbol{W}^{\alpha+1})$ ▷ Armijo backtracking step size;

5: $\boldsymbol{W}^{\alpha+1} = \mathrm{R}_{\boldsymbol{W}}(\boldsymbol{\xi})$ ▷ search for the next point;

6: $\boldsymbol{\kappa}^{\alpha+1} = \mathrm{grad} f(\boldsymbol{W}^{\alpha+1})$ ▷ Determine the gradient;

7: $(\boldsymbol{\xi}^\alpha)^+ = \mathrm{Transp}_{\boldsymbol{W}^\alpha \to \boldsymbol{W}^{\alpha+1}}$ ▷ find the transport;

8: Select Polak-Ribiere $\gamma^{\alpha+1}$;

9: $\boldsymbol{\xi}^{\alpha+1} = -\boldsymbol{\kappa}^{\alpha+1} + \gamma^{\alpha+1}(\boldsymbol{\xi}^\alpha)^+$ ▷ the new conjugate direction;

10: $\alpha \leftarrow \alpha + 1$;

11: **Until** a stopping criterion triggers.

So far, the obtained $\boldsymbol{W}$ from Algorithm 8 is the solution that only satisfies the orthogonal constraints (5.21b), but it does not satisfy the constant modulus constraint in (5.14c). Since matrix $\boldsymbol{W}$ has N_{RF} orthonormal columns, and matrix $\boldsymbol{A}$ needs constant modulus, i.e., $\frac{1}{\sqrt{N_r}}$, then the minimum value can be achieved when $\boldsymbol{A}$ has the same phase components as $\boldsymbol{W}$, i.e.,

$$\boldsymbol{A} = \frac{1}{\sqrt{N_r}} e^{j\ \arg(\boldsymbol{W})}. \tag{5.28}$$

where $\arg(\boldsymbol{X})$ generates a matrix containing the phases of the entries of the matrix $\boldsymbol{X}$. It is confirmed that phases of the analog combiner $\boldsymbol{A}$ are extracted from the phases of $\boldsymbol{W}$. Ideally, as all the elements in $\boldsymbol{A}$ have a constant modulus, the phases of all the elements in $\boldsymbol{A}$ are the same as the elements of the corresponding orthonormal $\boldsymbol{W}$. This point is proven by utilizing the properties of complex numbers [46].

Note here, a metric is defined based on the diagonally dominant matrix condition $|\boldsymbol{H}_{k,k}^{EF}| > \sum_{i \neq k} |\boldsymbol{H}_{k,i}^{EF}|$ [81], in the aim of illustrating the robustness of the proposed algorithms PI and ROSt. This metric is introduced as a Ratio of the magnitude of the diagonal element in each row to the sum of the magnitudes of all the off-diagonal elements in that

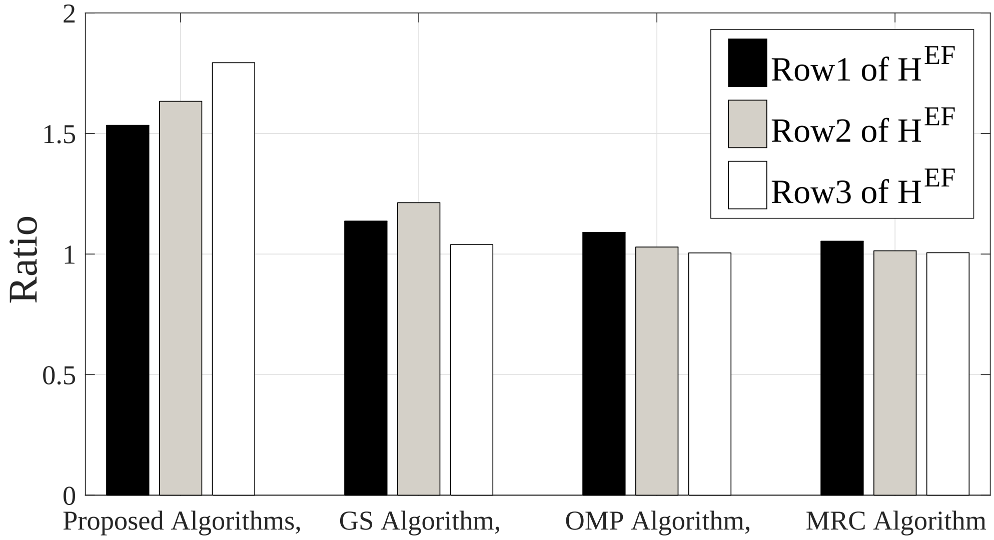


Figure 5.7: Illustration of a Ratio between the diagonal and off diagonal elements of $\boldsymbol{H}^{EF}$.

row which is represented as

$$\text{Ratio} = \frac{|\boldsymbol{H}_{k,k}^{EF}|}{\sum\limits_{i\neq k} |\boldsymbol{H}_{k,i}^{EF}|}, \tag{5.29}$$

where $\boldsymbol{H}_{k,k}^{EF}$ is the diagonal element in each row, and $\boldsymbol{H}_{k,i}^{EF}$ is the off-diagonal element in the same row of $\boldsymbol{H}^{EF}$, respectively. Suppose that there are 3 users transmitting through common scattering scenario, where the number of RF chains equals the number of users i.e., $N_{RF} = K = 3$ and $N_t = 4$ and $N_r = 16$. Now, for comparison, several $\boldsymbol{H}^{EF}$ are generated by different algorithms, initially the proposed algorithms (PI and ROSt), GS [73], OMP [10], and MRC [26] algorithms, respectively. Fig. 5.7 compares the Ratio of 4 different algorithms where each of which designs its $\boldsymbol{H}^{EF}$, and each bar represents a row of $\boldsymbol{H}^{EF}$. It can be noticed that the diagonal entries of $\boldsymbol{H}^{EF}$ that are generated by our proposed algorithms attain higher values as compared to the off-diagonal elements. Consequently, superior Ratio based on the proposed algorithms is attained than that of the other compared algorithms, as expected. Intuitively, it can be also recognized that the off-diagonal values of the other compared algorithms are close enough to the diagonal values which affects the Ratio directly and making it always close to one.

Ultimately, ROSt that is summarized in Algorithm 8, is achieved the target of designing the analog combiner $\boldsymbol{A}$ for maximizing the desired user and helping in putting all the off-diagonal elements effects down. As mentioned earlier, the off-diagonal elements of $\boldsymbol{H}^{EF}$ represent the interference entries, thus in the upcoming subsection to further eliminate their effects on the system, the digital combiner $\boldsymbol{D}$ will be designed for multi-user hybrid D/A mm-Wave massive MIMO systems.

5.5.3 Design of Digital Combiners

In this subsection, after having accomplished the design of the analog precoder and combiner, a next important step focus on designing the digital combiner. Practically, here the digital combiner is proposed based on MMSE by exploiting the effective uplink analog channel gains. Note here, for each user k, its effective uplink analog channel gains can be defined as $\bar{\boldsymbol{h}}_k = \boldsymbol{H}_k \boldsymbol{j}_{RF_k} \in \mathbb{C}^{N_r}$. Therefore, the optimal MMSE combiner which helps in further decreasing the off-diagonal elements is designed as [27]

$$\boldsymbol{D} = \bar{\boldsymbol{H}}^H \boldsymbol{A} \, (\boldsymbol{A}^H \bar{\boldsymbol{H}}^H \bar{\boldsymbol{H}} \boldsymbol{A} + \frac{K\sigma^2}{\eta} \boldsymbol{A}^H \boldsymbol{A})^{-1}, \tag{5.30}$$

where $\bar{\boldsymbol{H}} = [\bar{\boldsymbol{h}}_1, \cdots, \bar{\boldsymbol{h}}_k]$. Ultimately, the digital combiner $\boldsymbol{D}$ that is obtained from (5.30) based on the effective uplink analog channel can efficiently applied for making the off-diagonal elements have lower impact on the desired diagonal elements.

To sum up, the hybrid D/A precoders and combiners for the multi-user mm-Wave massive MIMO system that considers 5G dense area were designed properly in a way to mitigate the system interference in both analog and digital precoders and combiners.

5.6 AoA Estimation for Hybrid D/A mm-Wave systems

The aim of this section is to show how the proposed two-steps method can estimate the AoAs for the mm-Wave signals by exploiting an idea of detecting the point that represents the most power focuses around in the scattering scenario. This power point can be represented as $(q_{z_{l_{km}}}, r_{z_{l_{km}}})$, where illustrated in Fig. 5.8. In [82], it has been discussed that each scattering point consist of a group of power points. Interestingly, it is also shown that one of those power points, most probably the centre point, has the

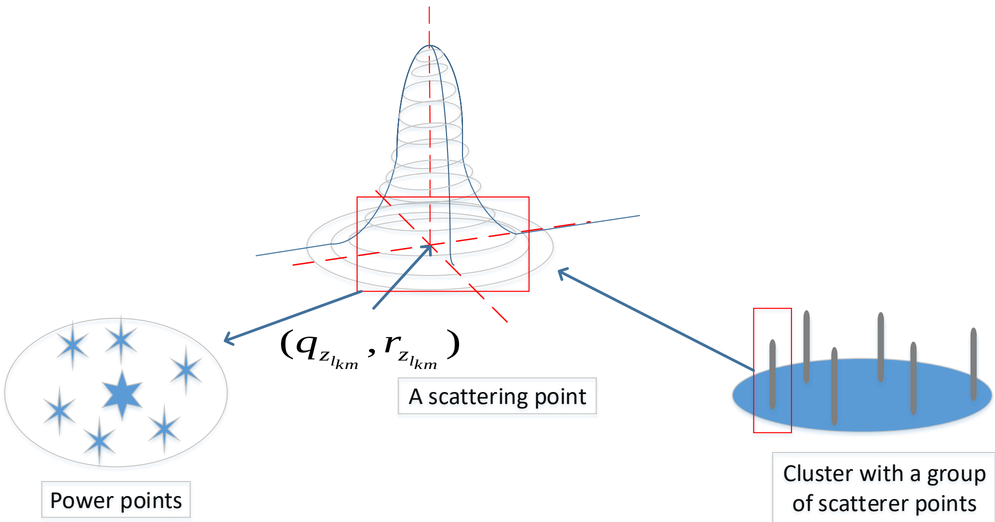


Figure 5.8: Illustration of a group of power points inside a scattering point in our scenario.

most power around which produces the strongest lth path as depicted clearly in Fig. 5.8. Authors in [82] exploited the point that has most of the power around to idealize the scattering function in frequency-wavenumber space in order to evaluate the performance of detectors. Specifically, this theory is employed in this section for a different purpose. This theory is proposed as one of the main stages that are formulated for helping in estimating AoAs of the mm-Wave signals. In this case in order to formulate the point that has the most power around, thus let us show the (q,r)th component of the channel estimation matrix $[\hat{\boldsymbol{H}}_{km}]_{q,r}$ which is defined as

$$\begin{aligned}
[\hat{\boldsymbol{H}}_{km}]_{q,r} &= \frac{1}{\sqrt{N_b}} h^{3D}_{z_{l_{km}}} \sum_{n_{b_x}=1}^{N_{b_x}} \sum_{n_{b_y}=1}^{N_{b_y}} e^{-j(\frac{2\pi}{N_{b_x}} n_{b_x} r + \frac{2\pi}{N_{b_y}} n_{b_y} q - 2\pi d/\lambda (n_{b_y} \sin\hat{\theta}^b_{z_{l_{km}}} \cos\hat{\phi}^b_{z_{l_{km}}} + n_{b_x} \sin\hat{\theta}^b_{z_{l_{km}}} \sin\hat{\phi}^b_{z_{l_{km}}}))} \\
&= \frac{1}{\sqrt{N_b}} h^{3D}_{z_{l_{km}}} e^{-j\frac{N_{b_y}-1}{2}(\frac{2\pi}{N_{b_y}} n_{b_y} q - 2\pi d/\lambda(\sin\hat{\theta}^b_{z_{l_{km}}} \cos\hat{\phi}^b_{z_{l_{km}}})} e^{-j\frac{N_{b_x}-1}{2}(\frac{2\pi}{N_{b_x}} n_{b_x} r - 2\pi d/\lambda(\sin\hat{\theta}^b_{z_{l_{km}}} \sin\hat{\phi}^b_{z_{l_{km}}})} \\
&\cdot \frac{\sin\left[\frac{N_{b_y}}{2}(\frac{2\pi}{N_{b_y}} q - 2\pi d/\lambda(\sin\hat{\theta}^b_{z_{l_{km}}} \cos\hat{\phi}^b_{z_{l_{km}}}))\right]}{\sin\left[\frac{1}{2}(\frac{2\pi}{N_{b_y}} q - 2\pi d/\lambda(\sin\hat{\theta}^b_{z_{l_{km}}} \cos\hat{\phi}^b_{z_{l_{km}}}))\right]} \cdot \frac{\sin\left[\frac{N_{b_x}}{2}(\frac{2\pi}{N_{b_x}} r - 2\pi d/\lambda(\sin\hat{\theta}^b_{z_{l_{km}}} \sin\hat{\phi}^b_{z_{l_{km}}}))\right]}{\sin\left[\frac{1}{2}(\frac{2\pi}{N_{b_x}} r - 2\pi d/\lambda(\sin\hat{\theta}^b_{z_{l_{km}}} \sin\hat{\phi}^b_{z_{l_{km}}}))\right]}.
\end{aligned} \tag{5.31}$$

It can be easily noticed that, when $N_{b_x} \rightarrow \infty$ and $N_{b_y} \rightarrow \infty$, then $[\hat{\boldsymbol{H}}_{km}]_{q_{z_{l_{km}}}, r_{z_{l_{km}}}} = \sqrt{N_b} h^{3D}_{z_{l_{km}}}$ that possess only a point where all the power will concentrate around which is represented as $(q_{z_{l_{km}}}, r_{z_{l_{km}}})$. Therefore, the estimation AoAs in both angles (i.e., zenith and azimuth) is formulated as

$$\hat{\theta}^b_{z_{l_{km}}} = \arcsin\left(\lambda/d\sqrt{\left(\frac{q_{z_{l_{km}}}}{N_{b_x}}\right)^2 + \left(\frac{r_{z_{l_{km}}}}{N_{b_y}}\right)^2}\right), \quad \hat{\phi}^b_{z_{l_{km}}} = \arctan\frac{\left(\frac{q_{z_{l_{km}}}}{N_{b_x}}\right)}{\left(\frac{r_{z_{l_{km}}}}{N_{b_y}}\right)}, \tag{5.32}$$

where $\hat{\theta}^b_{z_{l_{km}}}$ and $\hat{\phi}^b_{z_{l_{km}}}$ are the estimated zenith and azimuth angles. Importantly, the power point is defined as

$$q_{z_{l_{km}}} = \frac{N_{b_y}d}{\lambda(\sin\varpi^b_{z_{l_{km}}}\cos\omega^b_{z_{l_{km}}})}, r_{z_{l_{km}}} = \frac{N_{b_x}d}{\lambda(\sin\varpi^b_{z_{l_{km}}}\sin\omega^b_{z_{l_{km}}})}.$$

By doing this, the first stage of formulating the estimated angles is done. It is confirmed that this point $(q_{z_{l_{km}}}, r_{z_{l_{km}}})$ plays a main role in entire of the estimation process, where each of which depends on a $\varpi^b_{z_{l_{km}}}$ and $\omega^b_{z_{l_{km}}}$. Therefore, to find $\varpi^b_{z_{l_{km}}}$ and $\omega^b_{z_{l_{km}}}$ in this chapter, a search method called angular domain search is proposed as the second step. In this step that is expedited in estimating the AoAs, a wide search in the angular domain ranged from 0 to π and from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ are performed linearly with a step size of $\frac{\pi}{\Omega}$ and $\frac{\pi/2}{\Omega}$ that is used as an angle step size search. Practically, $\Omega \approx \frac{2N_b}{1.7}$ [82] is denoted as the minimum step search that relies on the number of antennas N_b employed for the AoAs search. The main concept of using such a step is to search among a range of angles and finding which angle can achieve the minimum estimation error. Where for each antenna in both axis the angular domain search is represented as $\varpi^b_{z_{l_{km}}} = (n_{b_x} - 1)\frac{\pi}{\Omega}$, where $n_{b_x} = [1, \cdots, N_{b_x}]$, and $\omega^b_{z_{l_{km}}} = (n_{b_y} - 1)\frac{\pi/2}{\Omega}$, where $n_{b_y} = [1, \cdots, N_{b_y}]$. The normalized estimation error criteria between the estimation channel that is obtained based on the estimated angles and the given channel, is applied as a performance metric which is denoted as E_{rr} and then formulated as

$$E_{rr} = 10\log_{10}\left(\frac{\sum_{l=1}^{L_{km}} ||\bar{\boldsymbol{H}}_{km} - \boldsymbol{H}_{km}||_F^2}{\sum_{l=1}^{L_{km}} ||\boldsymbol{H}_{km}||_F^2}\right). \tag{5.33}$$

Lastly, the two angles that achieve the minimum estimation error between the estimation channel and the given channel are chosen as the estimated angles for the mm-Wave signals.

With the aforementioned criterion, we provide an estimation approach for the mm-Wave signals that were suffered also from the overlapped AoAs. Basically, the strongest power point will be first found and formulated and then an accurate search method is also involved in this approach. The both stages are applied to strongly find the estimated angles that constructively minimize the normalized estimation error criteria between an estimation channel and the given channel.

5.7 Simulation Results

In this section, simulation results are provided to evaluate the performance of our proposed algorithms. For simplicity, the number of users equals the number of RF chains at the BS. The simulation parameters are set as follows. Both BS and users are equipped with UPA, having antenna spacing $u_\delta = \frac{\lambda}{2}$, where $\delta \in [x, y]$ [30]. In this result, the BS is equipped with $N_{RF} = K = 6, N_r = 128$, and each user is built with $N_t = 16$. For fair comparison, any ZF solution in literature algorithms is replaced by an MMSE solution. As shown in previous work on mm-Wave systems [10, 13, 19, 41, 48], most of our results in this section are tested with low and medium SNR levels.

Fig. 5.9 illustrates the effect of the correlation between the channel vectors of different users in mm-Wave system on the achievable uplink rate when SNR is set as -10 dB. There is a strong relation between the scattering environment and the orthogonality between the channel vectors of users. It is seen in Fig. 5.9 that all the algorithms achieve lower performance when the correlation reach to 0 dB. This is because, as most of the propagation paths of the two users share most of the scattering points, then their signals overlapped at the BS, the correlation between the channel vectors reach to the worse level by losing the orthogonality feature. It can be concluded that our proposed algorithms attain higher achievable uplink rate than the other algorithms at the practical regime when there is a high correlation between users staring from -20 dB until 0 dB.

In Fig. 5.10 and Fig. 5.11, we investigate the achievable uplink rate as a function of SNR for two users $K = 2$, where the correlation is fixed at -10 dB and -5 dB, respectively. It can be seen when the correlation high at -5 dB in Fig. 5.11 all the algorithm perform lower achievable uplink rate as compared to the Fig. 5.10 when the correlation value is set as -10 dB at all SNR values. Intuitively, when the correlation of the channel vectors of these users increase due to the propagation paths of the two users sharing most of the

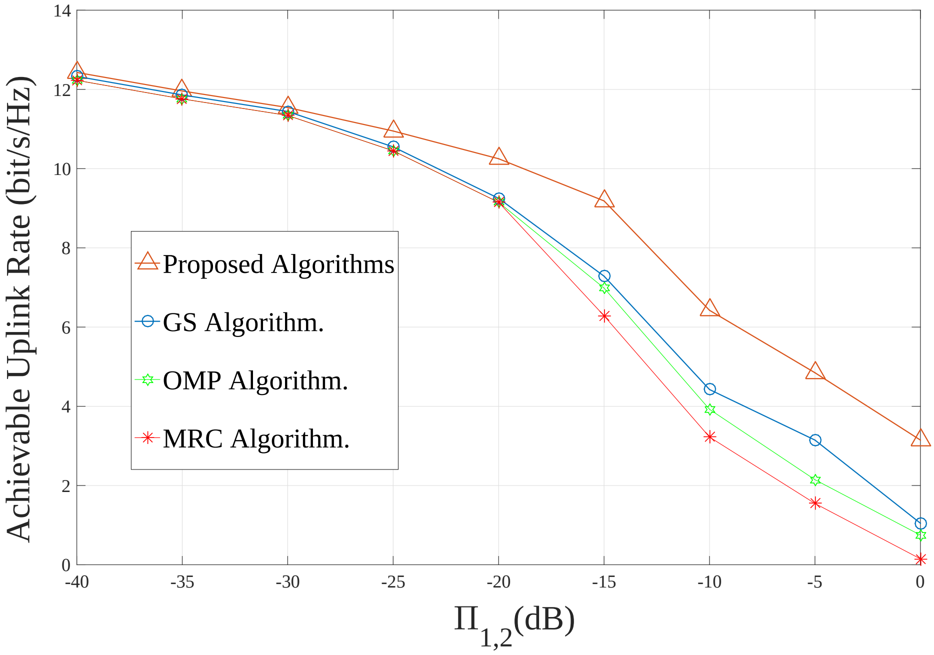


Figure 5.9: Achievable uplink rate versus channel correlation between $K = 2$B.

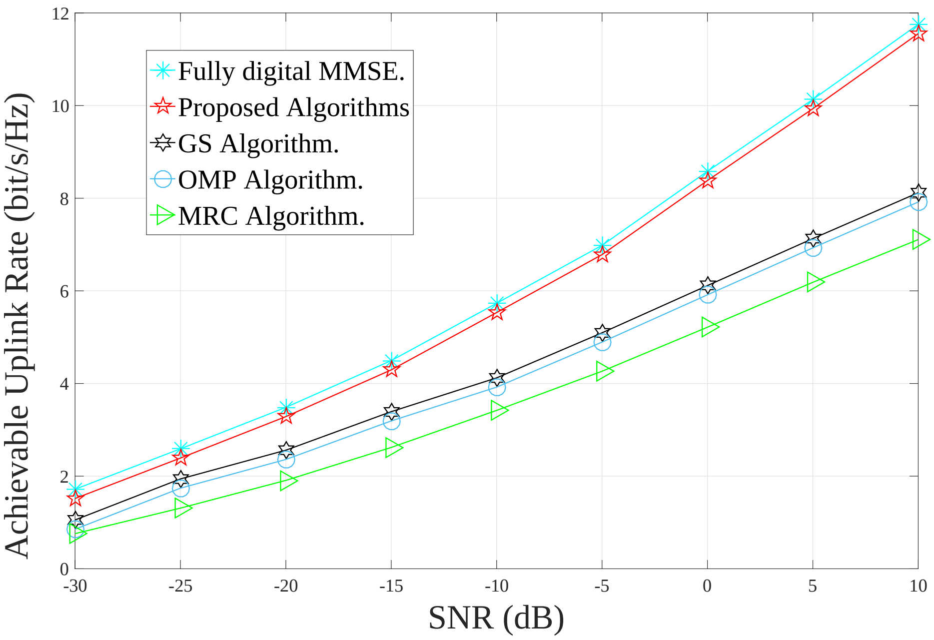


Figure 5.10: Achievable uplink rate versus SNR when $K = 2$ and correlation $= -10$ dB.

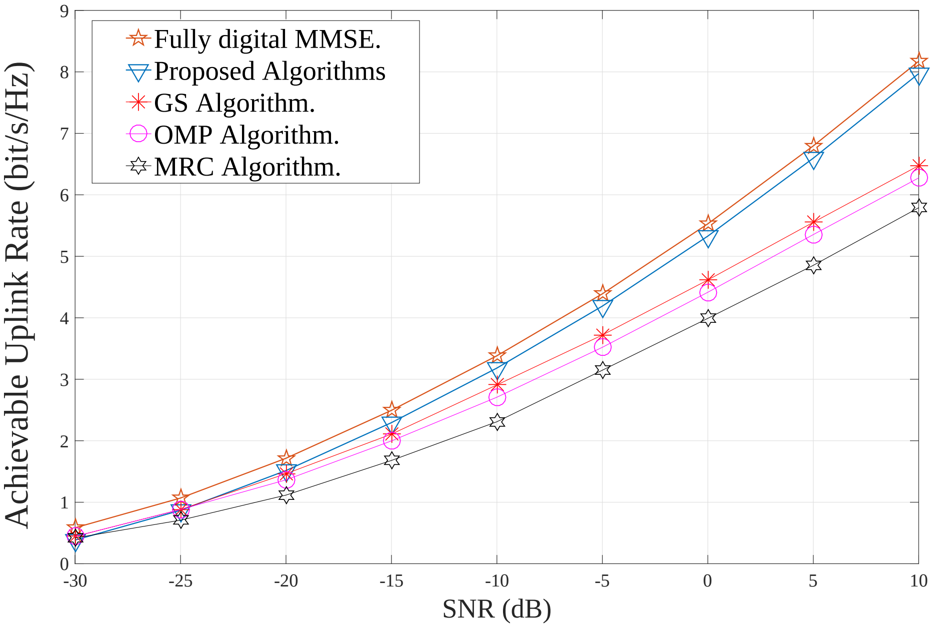


Figure 5.11: Achievable uplink rate versus SNR when $K = 2$ and correlation $= -5$ dB.

common scattering point in the regime and arrive with overlapped AoAs, the interference at the BS increases sharply which spoils the performance of the system. In both figures the proposed algorithms superior than the existing algorithms, as the analog domain is designed directly to decrease the correlation effect which acquires more achievable uplink rate.

Fig. 5.12 reveals the effect of the interference due to some of the propagation paths share some common scatters on the achievable uplink rate of hybrid D/A mm-Wave massive MIMO systems. In the practical regime the correlation between users is varying from user to user in the system, for instance when two users are shared most of the common scatters then their correlation maybe -6 dB, but in the case that there is no sharing between these users then the correlation maybe -18 dB. Thus in this figure, we have $K = 6$ users and the correlation between each two users randomly selected, but within the high correlation range $(-20 \sim 0)$ dB. Further, the number of scatterers is set as 12 and there are 4 clusters and $L_{km} = 13$ [13, 27, 30]. We randomly select $(3 \sim 5)$ scatterers from an existing set of scatterers with a pre-defined probability that users share

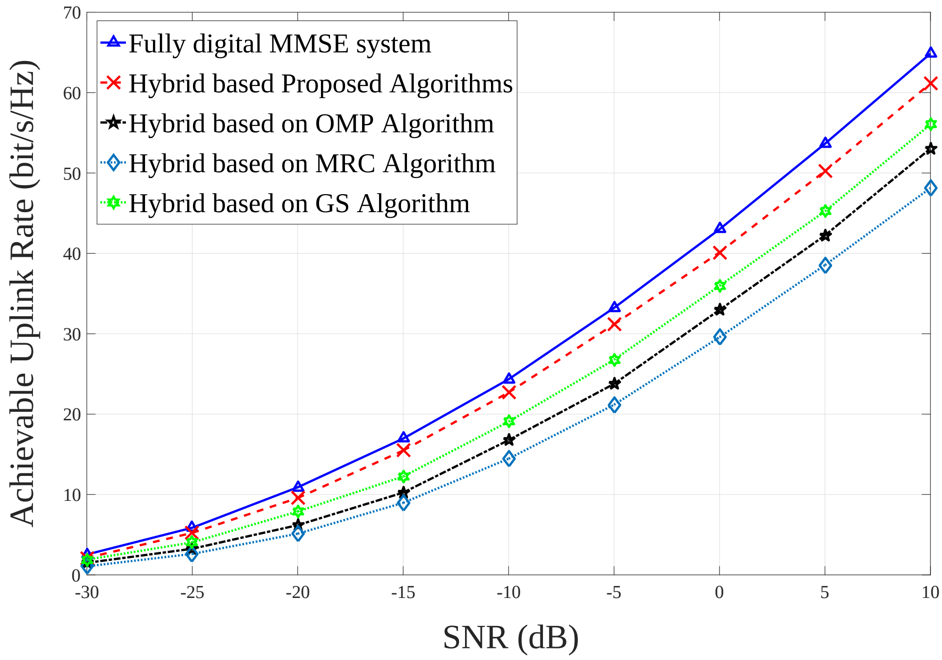


Figure 5.12: Achievable uplink rate under the overlapped AoAs scenario versus SNR, $K = 6$..

common scatterers, also each scatterer has a unit average power. The fully digital MMSE system and the hybrid D/A based on GS [73], MRC [26] and OMP [27] algorithms are applied as comparisons. Noticeably the hybrid D/A based on the proposed algorithms gain higher achievable uplink rates than the existing hybrid D/A based on GS [73], MRC [26] and OMP [27] algorithms as expected. Although, GS algorithm is built to decrease the overlapped AoAs interference, unacceptably it has a noticeable gap from the fully digital MMSE system. This can prove that the correlation interference has not been efficiently removed. Furthermore, MRC [26] and OMP algorithms [27] have a high gap from our proposed algorithms and the fully digital MMSE because the interference that is created due to the common scattering fades their performance improperly. Fig. 5.12 represents the practical scenario, as users physically move which affects the correlation values either increase or decrease, but among all these affect parameters our designed algorithms attain reasonable performance.

To investigate more, let us test the system when the achievable uplink rate as a function of the number of users in Fig. 5.13. In the dense area the number of users is high and

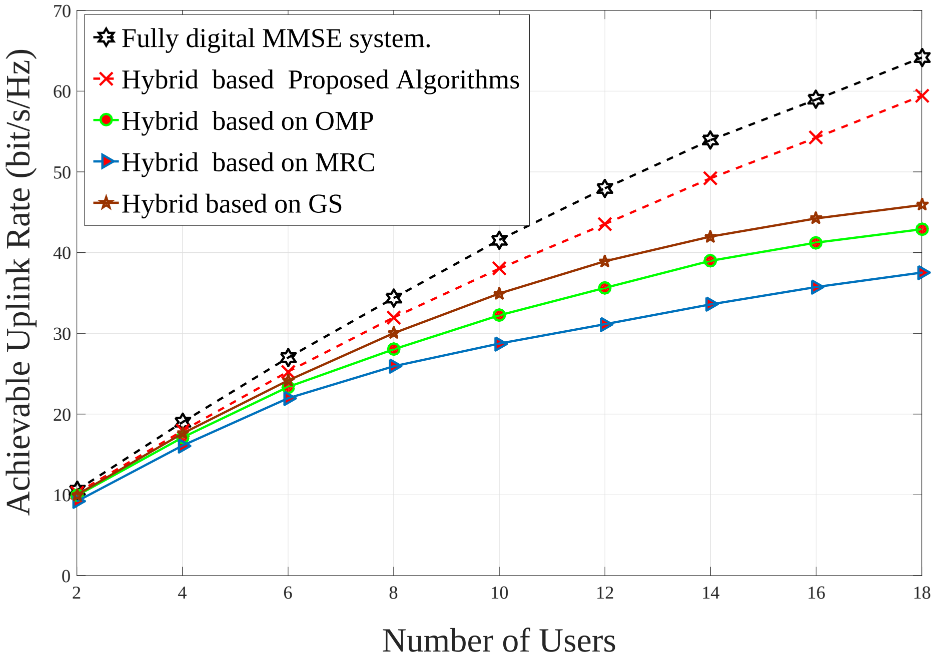


Figure 5.13: Achievable uplink rate the overlapped AoAs scenario versus K.

varying between spot to spot which influences the users correlation values. From previous figure, the correlation between each two users in Fig. 5.13 is also guaranteed to be in the high correlation range ($-20 \sim 0$) dB at SNR -10 dB. It can be observed that the achievable uplink rates of our proposed PI and ROSt algorithms significantly outperform GS [73], MRC [26] and OMP [27] algorithms, especially when the number of users is large. The reason is, as the number of users grows excessively, more users share a small number of scatterers resulting in the correlation increases between each two users which generates high interference in the system. It is obvious that GS algorithm suffers from the high number of users which makes GS curve bend toward OMP curve. Furthermore, the existing MRC [26] and OMP [27] algorithms have not designed their hybrid D/A to reduce the effect of the interference due to common scattering in the system. Therefore, in Fig. 5.13 the curves of GS, MRC, and OMP algorithms almost saturate.

Fig. 5.14 plots the achievable uplink rate versus SNR for the hybrid D/A mm-Wave massive MIMO system in the low correlation region which implies that the channel vectors are more independent. The number of channel parameters are set as 4 clusters, 8 scatters

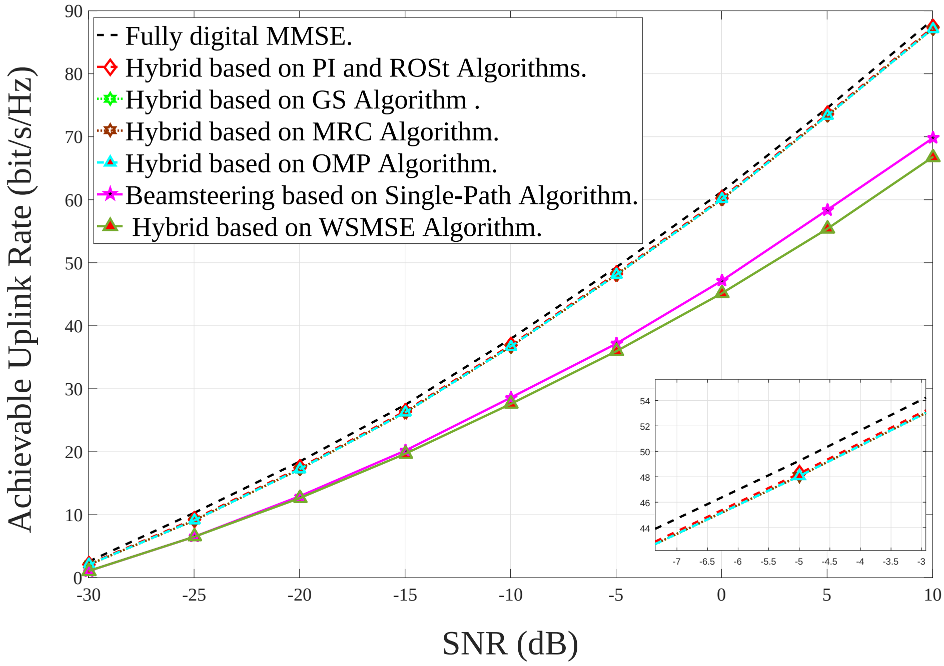


Figure 5.14: Achievable uplink rate of non-overlapped AoAs versus SNR, $K = 6$.

and $L_{km} = 13$ [13, 27, 30]. The AoAs and AoDs are taken independently from the uniform distribution within $[0, 2\pi]$, while the angle spread is set to 7.5 [10, 19, 30]. The fully digital MMSE system achieves the highest achievable uplink rate which is selected as a benchmark for the hybrid D/A mm-Wave massive MIMO system. Specifically, it is observed in Fig. 5.14 that, the hybrid D/A based on proposed algorithms PI and ROSt always attain an achievable uplink rate that is close to the fully digital MMSE system, and slightly higher than the existing algorithms in [26, 27, 73]. Further, the achievable uplink rate of our proposed algorithms surpasses that of the beamsteering and WSMSE in [24, 25] algorithms.

In Fig. 5.15, the estimation error performance as a function of the estimation channel and the given channel versus the number of antennas at the BS is illustrated. As shown here, the performance of the proposed estimation method improves significantly when the number of antennas increases from 10 to 20 as compared to HBT. The reason is that, in our proposed estimation method the search stage with more antennas can improve quickly the accuracy of the estimation channel by examining all antennas in both axes which acquires

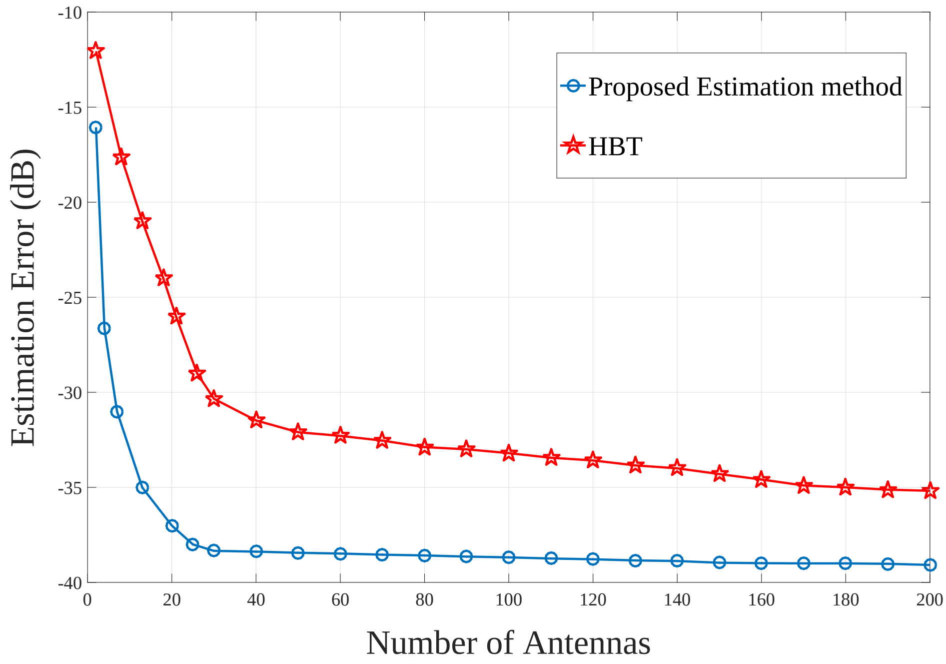


Figure 5.15: Estimation channel error performance.

better estimation. On the other hand, AoA estimation in HBT suffers from the overlapped sectors problem thus lower number of sectors in the space is always recommended. As a consequence, the performance for estimating the AoA will be inaccurate, having high estimation errors. Even though later on, as the number of antennas grows, this can help the sector space increase with less overlapped impact which aids the slow improvement in performance of HBT.

The normalized mean square error (MSE) of estimation channel for both the proposed method and HBT is presented in Fig. 5.16, where the MSE is defined as MSE $= \frac{1}{K}\sum_{k=1}^{K}\frac{||\hat{\boldsymbol{H}}_{km}-\boldsymbol{H}_{km}||^2}{||\boldsymbol{H}_{km}||^2}$. The configurations in which both the transmitter and the receivers are equipped with $N_t = 8, 32$ and $N_r = 64, 128$ antennas are compared. The channel model is set as 4 clusters, 10 scatters and $L_{km} = 13$ [13,27,30]. As displayed in Fig. 5.16, for all number of antennas, the MSE achievement of the proposed method are all superior to those of HBT. Essentially, because the proposed method can provide higher accuracy as compared to HBT. It can also seen that by enlarging the number of antennas, the MSE performance diminishes for both methods. This is because the accuracy of the both

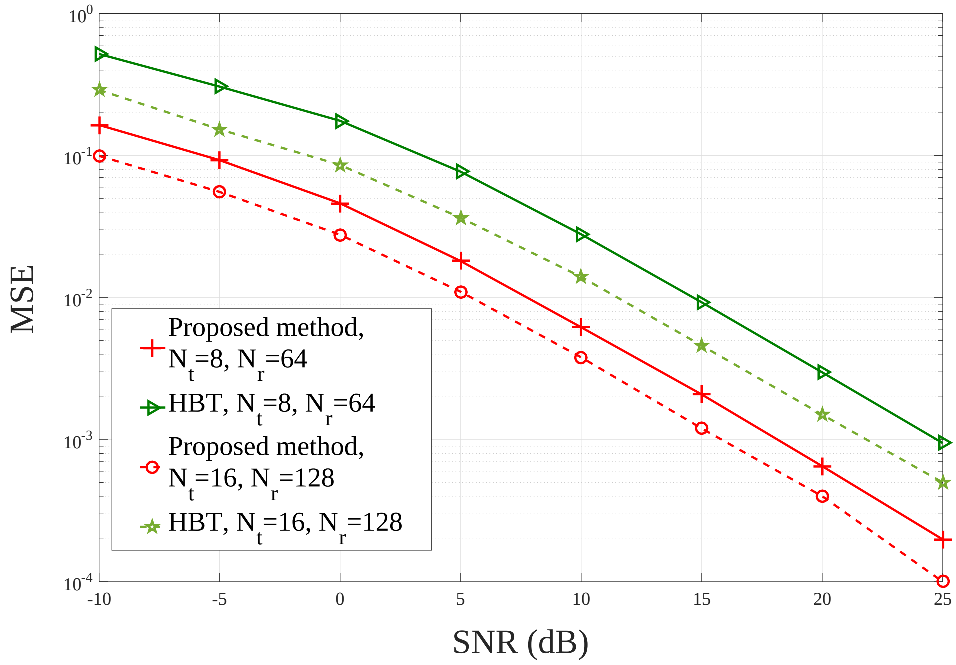


Figure 5.16: The uplink MSE performance comparison.

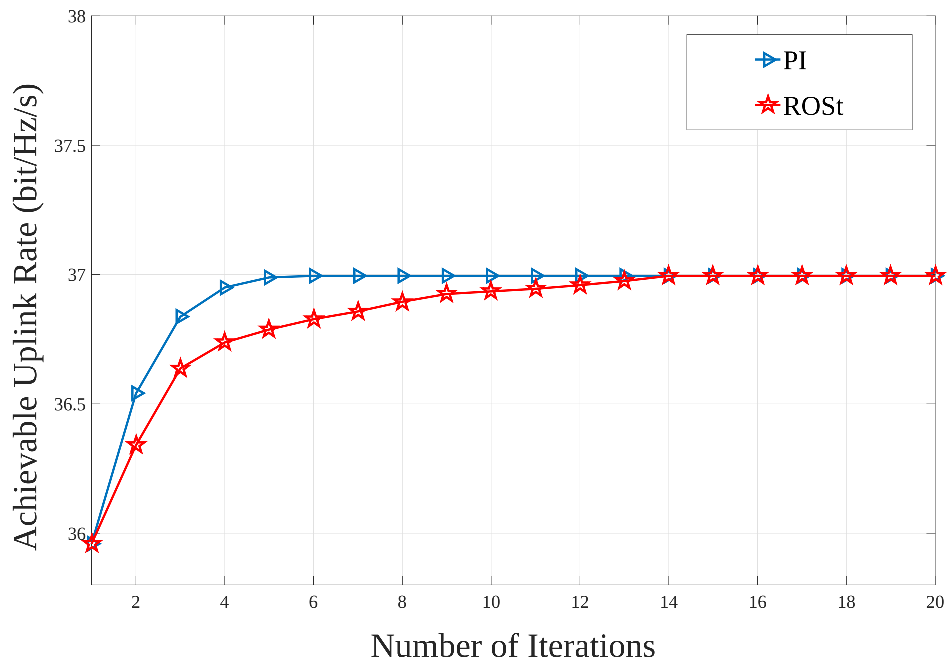


Figure 5.17: Learning curve of a our proposed algorithms for designing the analog precoder and combiner of the hybrid D/A mm-Wave large antenna array system.

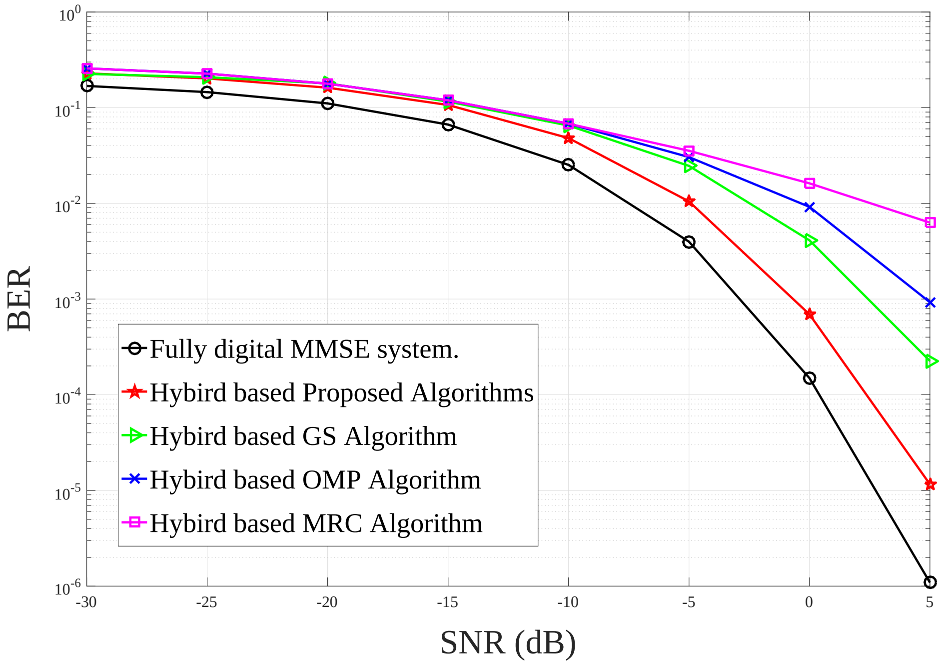


Figure 5.18: BER performance of hybrid D/A mm-Wave large antenna array systems.

methods grows significantly when the number of antennas increases in the system. Take SNR= 16 dB as an example, the MSE of 128 antennas is around 38% better than that of 64 antennas in our proposed method.

Fig. 5.17 shows the convergence behaviour of the ROSt algorithm and PI algorithm at 0 dB SNR. The channel model parameters are set as 4 clusters, 8 scatters and $L_{km} = 13$ [13, 27, 30]. The performance achieved is averaged over $10,000$ independent realisations of the channel. The curves in the figure show that the PI algorithm converge within 6 iterations which is considerably faster than ROSt. The simple framework and the normalized step of PI algorithm are actively enhanced the convergence speed to reach a steady state swiftly. On the other hand, ROSt algorithm converges within 16 iterations which is slower than PI algorithm. The retraction step needs more mathematical operations to perform and then move to the next step, this was predicted in our analysis and discussions.

In Fig. 5.18, we turn to consider binary phase-shift keying (BPSK) transmission to illustrate the bit-error-rate (BER) performance versus SNR when $N_{RF} = K = 4$, $N_t = 16$ and $N_r = 128$. The parameters of the channel are described as 4 clusters, 10 scatterers, and $L_{km} = 13$ [13, 27, 30]. Our proposed hybrid D/A based on ROSt and PI algorithms has better BER performance over the hybrid D/A based on the existing algorithms. The observation results agreed with our previous analysis and simulations where our proposed algorithms have focused on eliminating the interference that are initialized due to the overlapped AoAs and the multi-streams from the off-diagonal elements of $\boldsymbol{H}^{EF}$. This results in attaining a reasonable BER performance as compared to the fully digital MMSE system. As anticipated the existing algorithms GS, OMP, and MRC curves always have a lower performance as compared to our proposed algorithms.

5.8 Conclusions

The addressed issue in this chapter was to investigate the effect of the overlapped AoAs scenario on the achievable uplink rate of hybrid D/A mm-Wave massive MIMO systems. Due to the uplink signals transmitted via the common scatterers, users were spotted suffer from a non-negligible severe interference in the system as the correlation between channel vectors increases highly. The hybrid D/A was proposed for the aim of primarily maximize the desired signal while minimizing the system interference that was generated from overlapped AoAs at the BS during simultaneous uplink transmission. Therefore, the novel precoders and combiners were designed in this chapter based on the trust able PI algorithm at the user and then ROSt algorithm at the BS for the analog part. Then, the digital combiner was formed jointly applying MMSE by exploiting the effective uplink analog channel gains. Simulation results showed significant achievable uplink rate performance advantages in our proposed PI and ROSt algorithms, over the existing algorithms GS, OMP, and MRC algorithms for overlapped AoAs scenarios. Importantly, the channel estimation for hybrid D/A was examined by proposing a two-steps procedure. The strongest power point in each scattering point was formulated. Then, the accuracy of this procedure was actively improved by proposing a search method based on the angler domain scheme. Ultimately, the proposed estimation method attains lower error estimated channels than the recent existing HBT scheme.

Chapter 6

Conclusions and Future Works

Contents

Large antenna array systems which are called massive MIMO for mm-Wave systems with conventional digital precoders and combiners are one of the promising technologies to offer high data rates. Nevertheless, such systems are very costly and also power hungry due to the associated large number of RF chains that are demanded to connect the large number of antennas to the baseband. Since the associated phase shifters are significantly cheaper and consume lower power as compared to RF chains, hybrid D/A precoders and combiners are considered as an enabling technology to attain the promises of the existing of large antenna arrays systems. To do this, the design of hybrid D/A beamformers is a challenging task this is because the non-convex nature of the optimization problem.

This thesis addressed some of the key challenges in the design of hybrid D/A beamformers with large antenna array for mm-Wave systems. In this concluding Chapter, to begin with the summary and conclusions of this research will be presented. After that, considering the state-of-the-art survey on hybrid D/A precoders and combiners that was presented in Chapter 2, future directions of this research will be introduced in this Chapter.

6.1 Summary and Conclusions

This thesis focused on developing precoding and combining solutions that address key challenges in large antenna array systems, namely, the hardware constraints, the precoding and combining design complexity, the energy efficient precoding and combining, and the channel estimation, for single user and multi-user mm-Wave large antenna array systems.

Firstly, in this thesis, we provided a comprehensive literature survey to identify the gaps in the existing works for hybrid D/A beamformers mm-Wave large antenna array systems. To the best of our knowledge, there were no closed-form expressions for the hybrid D/A beamformers mm-Wave large antenna array system and also for the spectral efficiency that was achieved by the existing researchers. Moreover, most of the existing researchers were concentrated on sparse scattering channels only without considering the applicability of other scenarios such as rich scattering channels. High computational complexity associated with most of the state-of-the-art techniques and algorithms was another limiting factor for the implementation of hybrid D/A precoding in large antenna array scenarios. Further, the use of digital phase shifters with full resolution can result in a dramatically high complexity system and poor energy efficiency. Last but not the least, the power consumption increases with the resolution of the phase shifters of the analog precoder, which can be very high. Therefore, a trade-off between the spectral efficiency and the power consumption based on finite number of resolution require to be discussed. Having investigated these challenges, the contributions of this thesis are summarized as following:

In Chapter 3, we proposed efficient low complexity algorithms for hybrid D/A precoding and combining mm-Wave large antenna array systems. Two types of hybrid D/A architectures namely, fully-connected antenna array and partially-connected antenna array structures, have been adopted in the literature. To this end, considering the fact that each structure has its own practical advantages, in this thesis, novel hybrid D/A precoding and combining algorithms that can achieve high system spectral efficiency with low computational complexity were proposed for both structures. For the fully-connected antenna array structure, the precoding and combining design was formulated as an optimization problem to minimize the Euclidean distance between the hybrid D/A and fully digital system. The objective function of the approximation problem have a non-convex constraint, however, much less complex than the original one. This formulation is an efficient way to find the optimal precoders and combiners that achieve close performance to the single user fully digital system. Moreover, the optimization problem aimed to design optimally two different parameters each of which has its own constraint, non-convex constant modulus constraint and transmit power constraint. Alternating minimization approach was utilized to solve the optimization problem based on the Euclidean distance. We were mainly focused on designing the analog precoder with its non-convex constant modulus. Therefore, two algorithms were proposed in Chapter 3. Firstly, ASO based on Stiefel manifold which achieves always high spectral efficiency which close to the fully digital system with a reasonable computational complexity as compared to the existing algorithms. Secondly, to reduce the complexity of the proposed ASO algorithm, PSO with alternative minimization was utilized to propose APSO which aimed to design the low complexity analog part for fully-connected antenna array structures.

For the partially-connected antenna array structure, the hybrid D/A precoding and combining was formulated as a joint D/A optimization problem to maximize the spectral efficiency of the system directly. The first algorithm CMO was based on the principle of complex circle manifold optimization which can directly satisfy the constant modulus constraint by re-writing the precoding matrix as a vector and laying the elements of the vector into a complex circle. Subsequently, CMO algorithm achieves higher spectral efficiency than the existing algorithms. In order to find the optimal solution in CMO, Riemannian gradient and retraction were preformed. The second algorithm SPSO based on PSO was again proposed here to reduce the complexity of CMO. The attained simulation results illustrated that our proposed algorithms were able to achieve higher spectral efficiency than the existing hybrid D/A precoding algorithms. Ultimately, the effect of the channel estimation was considered, which showed that our proposed algorithms were more robust to the estimation error than the existing algorithms. Finally, our simulation results demonstrated that APSO and SPSO have lower complexity than the existing hybrid D/A precoding mm-Wave large antenna array system algorithms.

In Chapter 4, one of the main points of proposing the hybrid D/A precoding and combining for the mm-Wave large antenna array system was to reduce the power consumption of the conventional fully digital mm-Wave large antenna array system. However, RF chains component in the hybrid D/A system consume up to 70% of the total transceiver power. Therefore, in this thesis an energy efficient hybrid D/A precoding and combining mm-Wave large antenna array system was designed through optimizing the number of RF chains and the number of data streams while optimizing the D/A precoder and combiner matrices. To achieve the target of this optimization problem, we proposed simple methods for initially, finding the optimal number of RF chains and then, designing the low complexity D/A precoder and combiner matrices based on fully-connected antenna array architecture. The optimization problem was separately aimed to pramarelly optimize the number of RF chains in the system where BI algorithm was proposed. This algorithm also optimizes the number of data streams in the system. After that, the optimal analog precoder and combiner were designed with the help of the EVD method without the needs of finding the gradient of the cost function which reduced the complexity of the optimization problem. The digital precoder and combiner was designed based on applying SVD operation into the proposed low dimension effective channel gains matrix. Consequently, the computational complexity of the proposed digital precoder and combiner for the hybrid D/A system was much lower than that of the existing designed system based on LS solution.

Ultimately, the proposed hybrid precoding based on limited resolution of the phase shifters was explored to show a potential EE gain as compared to the hybrid precoding having infinite bits of resolution. One of the main concluding remarks was that neither the highly increase of the number of quantized bits of the resolution results in more spectral efficiency nor does a few-finite-bits (i.e., 1-2 bits) achieve the best EE for hybrid precoding systems. The simulation results showed significant EE performance advantages of our proposed algorithms over the existing systems, such as fully digital system, MO-AltMin and OMP algorithms based on hybrid precoding and beam steering based on a single RF chain in the system. Further, we analyzed the proposed hybrid precoding and combining strategy under a finite number of quantized bits in the analog precoding and combining to show the impacts of limited number of bits on EE performance.

In Chapter 5, the addressed issue of the effect of the overlapped AoAs scenario on the achievable uplink rate of hybrid D/A mm-Wave massive MIMO systems was investigated. Due to the uplink signals transmitted via the common scatterers, users were spotted suffer from a non-negligible severe interference in the system as the correlation between channel vectors increases highly. The hybrid D/A was proposed for the aimed of primarily maximizing the desired signal while minimizing the system interference that was generated from overlapped AoAs at the BS during simultaneous uplink transmission. Therefore, the novel precoders and combiners were designed in Chapter 5 based on the trust able PI algorithm at the user and then ROSt algorithm at the BS for the analog part. Then, the digital combiner was formed jointly applying MMSE by exploiting the low dimension effective uplink analog channel gains matrix. The simulation results showed significant achievable uplink rate performance advantages in our proposed PI and ROSt algorithms, over the existing algorithms GS, OMP, and MRC algorithms for overlapped AoAs scenarios.

Importantly, the channel estimation for hybrid D/A was examined by proposing a two-steps procedure. The strongest power point in each scattering point was formulated. Ultimately, the accuracy of this procedure was actively improved by proposing a search method based on the angler domain scheme. Finally, the proposed estimation method attained lower error estimated channels than the recent existing HBT scheme.

6.2 Future Works

The hybrid D/A precoding and combining approaches investigated in this research were developed and evaluated under certain assumptions such as the availability of perfect CSI, no RF impairments and ideal lossless hardware. Nevertheless, in order to integrate hybrid D/A precoders and combiners into practical mm-Wave large antenna array systems, the effect of these important parameters should be designed carefully. Under more realistic scenarios the hybrid D/A system requires solving many more research problems including imperfections CSI that can degrade the performance. In this section, we only highlight the research topics that are associated with channel estimation.

Designing Hybrid Precoding for mm-Wave Large Antenna Array Systems : The multiplicative structure of the proposed D/A precoding framework in Chapter 3 makes it feasible for hybrid architecture implementations. Although, we highlighted one main idea for this implementation, several open problems still require further investigations. For example, it is important to develop novel and optimized solutions for implementing the inter-cell interference avoidance stage in the analog precoding and combining. One initial way to do that is by employing antenna elements that have electrically controlled directional patterns. This, however, needs further research in the optimization and the actual implementation.

Another important challenge with hybrid D/A precoding mm-Wave large antenna array system is the channel acquisition. All of the proposed algorithms rely on the assumption of perfect CSI at the transmit end of the large antenna array link. However in reality, CSI cannot magically be assumed to appear at the transmitter, and this needs to be realistically taken into account in the study. In Chapter 3 the estimated channel is a noisy version of the true channel which is a simple method to check the robustness of the proposed algorithms. However, the amount of CSI imperfections is directly a function of how much time, or bandwidth was spent after training, as well as the electrical power utilized used in the training phase of transmission. These are not being taken into account because of the difficulties of designing the D/A precoders. Therefore, considering these parameters to be part of the channel estimation analysis is very curial when designing D/A precoders for mm-Wave large antenna array systems.

Energy-Efficient Design for Hybrid D/A Precoders in mm-Wave Systems : In this thesis, we proposed several combinations of RF chains and phase shifters to reduce the power consumption of hybrid D/A precoding and combining systems. The RF chains consume maximum transmitted power, therefore, in Chapter 4 optimal number of RF chains was selected to achieve the desired EE. Several challenges, though, face the extension of this aim into multi-user systems, which makes it non-trivial. This is mainly due to finding the optimal number of RF chains in a system that is redirected with the number of RF chains is equivalent to the number of data streams and equal to the number of users, is quite difficult as the number of users is large in the next mobile wireless communications system. Therefore, if the number of users and data streams are larger than the number of RF chains, the hybrid D/A precoding system might be unable to optimize the selected number of RF chains. Hence, it is of paramount importance that the hybrid D/A mm-Wave system should be designed carefully to adaptively search for the optimal number of RF chains in the system for such different cases. Moreover, the hybrid D/A precoding system should be designed to support more data streams per RF chain in some extrema cases.

Impact of Overlapped AoAs on the Achievable Uplink Rate for mm-Wave Systems : In Chapter 5, we developed novel precoders and combiners for multi-user hybrid D/A beamforming for mm-Wave massive MIMO systems, where the number of RF chains was much less than the number of antennas equipped with each transceiver. Channel mm-Wave measurements in Daejeon City, Korea, showed that the propagation paths of multiple transmitted signal users were spotted under the impact of non-negligible common scattering scenarios. Thus the arrived signals tend to overlap at the BS. With high probability, in this scenario we intuitively observe that some propagation paths of multiple users sharing some common scatterers, therefore these different propagation paths were enforced to have the same AoAs at the BS. This phenomenon generates a non-negligible severe interference in the system. Indubitably, with the hybrid D/A mm-Wave massive MIMO system, this type of interference cannot be eliminated by the digital processing domain only, because the digital domain is restricted with fewer RF chains. In this case, if the existing precoders and combiners are applied, this serves to severely limit the achievable rate in the current hybrid D/A mm-Wave massive MIMO systems. Hence, with high number of users the overlapped signals will block the BS to distinguish between these users which results in designing inaccurate combiners. Therefore importantly, it is of paramount importance that the analog precoders and combiners of the hybrid D/A system should be designed precisely to help the digital domain in mitigating this type of interference.

In real wireless communications systems, the channel changes over the time quickly, in this case the channel estimation must repeat after the coherence time of the channel is over. Investigating the tradeoffs between the tolerable estimation error, the signalling overhead and data transmission within the coherence time of the channel are necessary steps towards the implementation of hybrid D/A precoding and combining mm-Wave systems. In other words, the target of these steps are to optimize the signalling and data transmission time given the performance of the hybrid D/A precoding and combining mm-Wave which is directly related to the accuracy of the estimated channel. The problem in the mm-Wave large antenna array system is represented in the hardware constraints which make it challenging to derive closed-form recovery guarantees. This, therefore, requires developing new analytical tools to overcome the challenges and to enable formulating solid statements on the mm-Wave channel estimation performance under large number of antennas at the transmitters and receivers. One difficulty of doing that comes from the mm-Wave geometric channel model that leaves a small gap for tractable analysis. This opens new rooms for future work.

It would also be interesting to develop impairment-aware precoding and channel estimation solutions. Based on the insights that will be obtained from analysing hybrid D/A structures under practical circuit models, it might be possible to design new optimized hybrid D/A structures, precoding and combining, and channel estimation solutions that take these insights into discussion. Moreover, promoting such solutions that are more robust to the hardware impairments can boost the actual spectral and energy efficiency performance of hybrid D/A structures under practical conditions. Thereby, improving such hybrid D/A structures and the associated signal processing solutions while analysing their performance are very important for future mm-Wave and large antenna array systems.

Appendix A

Appendices

Contents

A.1 Proof of Sherman-Morrison-Woodbury (SMW) (3.21)

We show that the Sherman-Morrison-Woodbury (SMW) formula can be used to simplify (19) to obtain (20). To prove it, we present the following:

Lemma 2. *Postulate that* $\boldsymbol{W} = \bigtriangledown f(\boldsymbol{X})\boldsymbol{U}^H - \boldsymbol{U} \bigtriangledown f^H(\boldsymbol{X})$, *where,* $\boldsymbol{U} = -\boldsymbol{X}$. *Rewrite* $\boldsymbol{W} = \boldsymbol{K}\boldsymbol{P}^H$ *for* $\boldsymbol{K} = [\bigtriangledown f(\boldsymbol{X}), \boldsymbol{U}]$ *and* $\boldsymbol{P} = [\boldsymbol{U}, -\bigtriangledown f(\boldsymbol{X})]$. *If* $(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}$ *is invertible,* (19) *is*

$$\boldsymbol{X}^{\alpha+1}(\tau) \quad = \quad \boldsymbol{X}^{\alpha} - \tau^{\alpha}\boldsymbol{K}(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\boldsymbol{P}^H\boldsymbol{X}^{\alpha}.$$

Proof : By applying the Sherman-Morrison-Woodbury (SMW) formula [55], one obtains:

$$(\boldsymbol{B} + \alpha\boldsymbol{K}\boldsymbol{P}^H)^{-1} = \boldsymbol{B}^{-1} - \alpha\boldsymbol{B}^{-1}\boldsymbol{K}(\boldsymbol{I} + \alpha\boldsymbol{P}^H\boldsymbol{B}^{-1}\boldsymbol{K})^{-1}\boldsymbol{P}^H\boldsymbol{B}^{-1}. \tag{A.1}$$

Letting $\boldsymbol{B} = \boldsymbol{I}$, and, substituting into (A.1), one obtains

$$(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{W})^{-1} = \boldsymbol{I} - \frac{\tau}{2}\boldsymbol{K}(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\boldsymbol{P}^H, \tag{A.2}$$

where $\boldsymbol{W} = \boldsymbol{K}\boldsymbol{P}^H$, then

$$
\begin{aligned}
(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{K}\boldsymbol{P}^H)^{-1}(\boldsymbol{I} - \frac{\tau}{2}\boldsymbol{K}\boldsymbol{P}^H) \quad &= \quad (\boldsymbol{I} - \frac{\tau}{2}\boldsymbol{K}(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\boldsymbol{P}^H)\,(\boldsymbol{I} - \frac{\tau}{2}\boldsymbol{K}\boldsymbol{P}^H) \\
&= \quad \boldsymbol{I} - \frac{\tau}{2}\boldsymbol{K}(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\boldsymbol{P}^H - \frac{\tau}{2}\boldsymbol{K}\boldsymbol{P}^H \\
&+ \quad \frac{\tau^2}{4}\boldsymbol{K}(\boldsymbol{I}\frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\boldsymbol{P}^H\boldsymbol{K}\boldsymbol{P}^H \\
&= \quad \boldsymbol{I} - \frac{\tau}{2}\boldsymbol{K}(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\left[\boldsymbol{I} + (\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K}) - \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K}\right]\boldsymbol{P}^H \\
&= \quad \boldsymbol{I} - \tau\boldsymbol{K}(\boldsymbol{I} + \frac{\tau}{2}\boldsymbol{P}^H\boldsymbol{K})^{-1}\boldsymbol{P}^H. \qquad \text{(A.3)}
\end{aligned}
$$

□

Bibliography

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2016-2021 white paper." `https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html`, 2017. [Online; accessed 28-March-2017].

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 742–758, Oct 2014.

[3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, pp. 40–60, Jan 2013.

[4] J. Zhang, X. Huang, V. Dyadyuk, and Y. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Trans. on Wireless Communi.*, vol. 22, pp. 79–87, Feb. 2015.

[5] F.-L. Luo and C. Zhang, *Massive MIMO for 5G: Theory, Implementation and Prototyping*. Wiley-IEEE Press, 2016.

[6] . E. G. Larsson, "Very large MIMO systems: Opportunities and challenges." `:http://www.kth.se/polopoly-fs/1303070/Menu/general/column-content/attachment/Large-MIMO.pdf`, 2012. [Online; accessed 2012].

[7] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3590–3600, November 2010.

[8] J. Wang, H. Zhu, and N. J. Gomes, "Distributed Antenna Systems for Mobile Communications in High Speed Trains," *IEEE J. Sel. Areas in Commun.*, vol. 30, pp. 675–683, May. 2012.

[9] T. S. Rappaport, R. W. H. Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Upper Saddle River, NJ, USA: Prentice-Hal, Sep. 2013.

[10] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 13, pp. 1499–1513, Mar. 2014.

[11] J. Zhang, X. Ge, Q. Li, M. Guizani, and Y. Zhang, "5G Millimeter-Wave Antenna Array: Design and Challenges," *IEEE Wireless Commun.*, vol. PP, pp. 2–8, Oct. 2016.

[12] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Commun. Surveys Tutorials*, vol. 18, pp. 1617–1655, thirdquarter 2016.

[13] F. Sohrabi and W. Yu, "Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays," *IEEE J. Sel. Topics in Signal Processing*, vol. 10, pp. 501–513, Apr. 2016.

[14] P. V. Amadori and C. Masouros, "Low RF-Complexity Millimeter-Wave Beamspace-MIMO Systems by Beam Selection," *IEEE Trans. on Commun.*, vol. 63, pp. 2212–2223, Jun. 2015.

[15] Y. Y. Lee, C. H. Wang, and Y. H. Huang, "A Hybrid RF/Baseband Precoding Processor Based on Parallel-Index-Selection Matrix-Inversion-Bypass Simultaneous Orthogonal Matching Pursuit for Millimeter Wave MIMO Systems," *IEEE Trans. on Signal Processing*, vol. 63, pp. 305–317, Jan. 2015.

[16] O. El Ayach, R. W. Heath Jr., S. Rajagopal, and Z. Pi, "Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays," in *Proc. IEEE GLOBECOM, 2013*, pp. 3476–3480, Dec. 2013.

[17] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, pp. 186–194, Jan. 2015.

[18] S. Han, C.-L. I., Z. Xu, and S. Wang, "Reference Signals Design for Hybrid Analog and Digital Beamforming," *IEEE Commun. Lett.*, vol. 18, pp. 1191–1193, Jul. 2014.

[19] X. Yu, J. C. Shen, J. Zhang, and K. B. Letaief, "Alternating Minimization Algorithms for Hybrid Precoding in Millimeter Wave MIMO Systems," *IEEE J. Sel. Topics in Signal Processing*, vol. 10, pp. 485–500, Apr. 2016.

[20] L. Dai, X. Gao, J. Quan, S. Han, and C.-L. I, "Near-optimal hybrid analog and digital precoding for downlink mmWave massive MIMO systems," in *Proc. IEEE ICC*, pp. 1334–1339, Jun. 2015.

[21] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-Efficient Hybrid Analog and Digital Precoding for MmWave MIMO Systems With Large Antenna Arrays," *IEEE Jour.l on Sel. Areas in Commui.*, vol. 34, pp. 998–1009, Apr. 2016.

[22] S. Han, C.-L. I., C. Rowell, Z. Xu, S. Wang, and Z. Pan, "Large scale antenna system with hybrid digital and analog beamforming structure," in *Proc. IEEE ICC*, pp. 842–847, Jun. 2014.

[23] C. Rowell and S. Han, "Practical Large Scale Antenna Systems for 5G cellular networks," in *Proc. IEEE IWS*, pp. 1–4, Mar. 2015.

[24] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, pp. 122–131, Dec. 2014.

[25] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited Feedback Hybrid Precoding for Multi-User Millimeter Wave Systems," *IEEE Trans. on Wireless Commun.*, vol. 14, pp. 6481–6494, Nov. 2015.

[26] D. Ying, F. W. Vook, T. A. Thomas, and D. J. Love, "Hybrid structure in massive MIMO: Achieving large sum rate with fewer RF chains," in *IEEE ICC*, pp. 2344–2349, Jun. 2015.

[27] D. H. N. Nguyen, L. B. Le, T. Le-Ngoc, and R. W. Heath, "Hybrid MMSE Precoding and Combining Designs for mmWave Multiuser Systems," *IEEE Access*, vol. 5, pp. 19167–19181, 2017.

[28] F.-L. Luo and C. Zhang, *5G Standard Development: Technology and Roadmap*, p. 616. Wiley-IEEE Press, 2016.

[29] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Transactions on Communications*, vol. 61, pp. 1436–1449, April 2013.

[30] 3rd Generation Partnership Project (3GPP), "Study on channel model for frequencies from 0.5 to 100 GHz," Tech. Rep. 36.331, 2017. Version 38.901.

[31] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-Cell Interference in Noncooperative TDD Large Scale Antenna Systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 192–201, February 2013.

[32] X. Rao and V. K. N. Lau, "Distributed Compressive CSIT Estimation and Feedback for FDD Multi-User Massive MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 62, pp. 3261–3271, June 2014.

[33] C. Balanis, *Antenna theory analysis and design*. John Wiley and Sons, 2013.

[34] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, pp. 436–453, April 2016.

[35] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, "MIMO for millimeter-wave wireless communications: beamforming, spatial multiplexing, or both?," *IEEE Communications Magazine*, vol. 52, pp. 110–121, December 2014.

[36] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks-With a Focus on Propagation Models," *IEEE Transactions on Antennas and Propagation*, vol. 65, pp. 6213–6230, Dec 2017.

[37] M. Rumney, *LTE Introduction*, pp. 648–. Wiley Telecom, 2013.

[38] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Trans. on Commun.*, vol. 63, pp. 3029–3056, Sep. 2015.

[39] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, pp. 172–179, June 2013.

[40] S. Hur, S. Baek, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, "Proposal on Millimeter-Wave Channel Modeling for 5G Cellular System," *IEEE J. Sel. Areas in Signal Processing*, vol. 10, pp. 454–469, April 2016.

[41] R. Méndez-Rial, C. Rusu, N. González-Prelcic, and R. W. Heath, "Dictionary-free hybrid precoders and combiners for mmWave MIMO systems," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 151–155, June 2015.

[42] IEEE802.11ad, "IEEE 802.11ad standard draft D0.1." `http://www.ieee802.org/11/Reports/tgad-update.html.`, 2014. [Online; accessed 2014].

[43] "Ericsson energy and carbon report, 2014-2021 white paper." `https://www.ericsson.com/res/docs/2014/ericsson-energy-and-carbon-report.pdf`, 2014. [Online; accessed 2014].

[44] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint Precoding and RRH Selection for User-Centric Green MIMO C-RAN," *IEEE Trans. on Wireless Commun.*, vol. 16, pp. 2891–2906, May 2017.

[45] H. Zhu, "Performance Comparison Between Distributed Antenna and Microcellular Systems," *IEEE J. Sel. Areas in Commun.*, vol. 29, pp. 1151–1163, Jun. 2011.

[46] G. H. Golud and C. F. V. Loan, *MATRIX COMPUTATIONS*. The Johns Hopkins University Press, 2013.

[47] A. Kammoun, H. Khanfir, Z. Altman, M. Debbah, and M. Kamoun, "Preliminary results on 3D channel modeling: from theory to standardization," *IEEE J. Sel. Areas in Commun.*, vol. 32, pp. 1219–1229, Jun. 2014.

[48] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 831–846, Oct 2014.

[49] D. D. Lee and H. S. Seung, "Algorithms for Non-negative Matrix Factorization," tech. rep., In NIPS, 2000.

[50] U. Shalit, D. Weinshall, and G. Chechik, "Online Learning in The Manifold of Low-Rank Matrices," in *Advances in Neural Information Processing Systems 23*, pp. 2128–2136, 2010.

[51] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2839–2848, Nov 2003.

[52] Q. Z. Ahmed, S. Ahmed, M. S. Alouini, and S. Aïssa, "Minimizing the Symbol-Error-Rate for Amplify-and-Forward Relaying Systems Using Evolutionary Algorithms," *IEEE Transactions on Communications*, vol. 63, pp. 390–400, Feb. 2015.

[53] F. van den Bergh, "An Analysis of Particle Swarm Optimizers," tech. rep., PhD thesis, Department of Computer Science, University of Pretoria, 2002.

[54] O. Alluhaibi, Q. Z. Ahmed, J. Wang, and H. Zhu, "Hybrid digital-to-analog precoding design for mm-wave systems," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, May 2017.

[55] Z. Wen and W. Yin, "A Feasible method for Optimization with Orthogonality Constraints," tech. rep., Rice University, 2010.

[56] R. S. P.-A. Absil, R. Mahony, "Optimization Algorithms on Matrix Manifolds," tech. rep., Princeton University Press, 2008.

[57] N. Boumal, "Optimization and estimation on manifolds," tech. rep., PhD thesis, Department of Computer Science, University of Pretoria, 2014.

[58] O. Alluhaibi, Q. Z. Ahmed, C. Pan, and H. Zhu, "Capacity Maximisation for Hybrid Digital-to-Analog Beamforming mm-Wave Systems," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec 2016.

[59] O. Alluhaibi, Q. Z. Ahmed, C. Pan, and H. Zhu, "Hybrid Digital-to-Analog Beamforming Approaches to Maximise the Capacity of mm-Wave Systems," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, June 2017.

[60] D. Klain, "The Matrix Exponential," tech. rep., Linear Algebra II - Spring, 2006.

[61] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Computationally Efficient Vector Perturbation Precoding Using Thresholded Optimization," *IEEE Trans. on Commun.*, vol. 61, pp. 1880–1890, May. 2013.

[62] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath Jr., "Energy-Efficient Hybrid Analog and Digital Precoding for mmWave MIMO Systems with Large Antenna Arrays," *CoRR*, vol. abs/1507.04592, Jul. 2015.

[63] O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *IEEE 13th International Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, pp. 100–104, Jun. 2012.

[64] R. Zi, X. Ge, J. S. Thompson, C. Wang, H. Wang, and T. Han, "Energy Efficiency Optimization of 5G Radio Frequency Chain Systems," *CoRR*, vol. abs/1604.02665, 2016.

[65] A. Goldsmith, *Wireless Communications.* New York, NY, USA: Cambridge University Press, 2005.

[66] H. P. Benson, "Fractional programming with convex quadratic forms and functions," *European Journal of Operational Research*, vol. 173, no. 2, pp. 351 – 369, 2006.

[67] I. Stancu-Minasian, *Fractional Programming: Theory, Methods and Applications.* Springer Netherlands, 1997.

[68] S. Boyd and L. Vandenberghe, *Convex Optimization.* New York, NY, USA: Cambridge University Press, 2004.

[69] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Trans. on Commun.*, vol. 58, pp. 545–554, Feb. 2010.

[70] Q. Z. Ahmed, K. H. Park, M. S. Alouini, and S. Aissa, "Linear Transceiver Design for Nonorthogonal Amplify-and-Forward Protocol Using a Bit Error Rate Criterion," *IEEE Trans. on Wireless Commun.*, vol. 13, pp. 1844–1853, Apr. 2014.

[71] S. L. H. Nguyen, K. Haneda, J. Jarvelainen, A. Karttunen, and J. Putkonen, "On the Mutual Orthogonality of Millimeter-Wave Massive MIMO Channels," in *IEEE 81st VTC Spring*, pp. 1–5, May. 2015.

[72] J. Wang, H. Zhu, L. Dai, N. J. Gomes, and J. Wang, "Low-Complexity Beam Allocation for Switched-Beam Based Multiuser Massive MIMO Systems," *IEEE Trans. on Wireless Communi.*, vol. 15, pp. 8236–8248, Dec 2016.

[73] J. Li, L. Xiao, X. Xu, and S. Zhou, "Robust and Low Complexity Hybrid Beamforming for Uplink Multiuser MmWave MIMO Systems," *IEEE Commun. Letters*, vol. 20, pp. 1140–1143, June 2016.

[74] G. Wang, S. A. Naeimi, and G. Ascheid, "Low complexity channel estimation based on DFT for short range communication," in *IEEE (ICC)*, pp. 1–7, May 2017.

[75] A. N. Langville and C. D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings.* Princeton, NJ, USA: Princeton University Press, 2012.

[76] J. Poutanen, F. Tufvesson, K. Haneda, V. M. Kolmonen, and P. Vainikainen, "Multi-Link MIMO Channel Modeling Using Geometry-Based Approach," *IEEE Trans. on Antennas and Propagation*, vol. 60, pp. 587–596, Feb. 2012.

[77] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The Multicell Multiuser MIMO Uplink with Very Large Antenna Arrays and a Finite-Dimensional Channel," *IEEE Trans. on Communi.*, vol. 61, pp. 2350–2361, Jun. 2013.

[78] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A Coordinated Approach to Channel Estimation in Large-Scale Multiple-Antenna Systems," *IEEE J. Sel. Areas in Commun.*, vol. 31, pp. 264–273, Feb. 2013.

[79] G. Zhu, K. Huang, V. K. N. Lau, B. Xia, X. Li, and S. Zhang, "Hybrid Beamforming via the Kronecker Decomposition for the Millimeter-Wave Massive MIMO Systems," *IEEE J. Sel. Areas in Commun.*, vol. PP, pp. 1–1, Jun. 2017.

[80] E. F. C. Quispe and E. A. P. Quiroz, "An implementation of the steepest descent method using retractions on riemannian manifolds," 2015.

[81] F. Farid, "Notes on matrices with diagonally dominant properties," *Linear Algebra and its Applications*, vol. 435, no. 11, pp. 2793 – 2812, 2011.

[82] H. L. Van Trees, *Optimum Array Processing, Detection and Other Topics.* John Wiley and Sons, Inc., 2002.