Hybrid Analog/Digital Signal Processing for mmWave Massive-MIMO Communications

Alireza Morsali

Department of Electrical & Computer Engineering
McGill University
Montréal, Canada

November 2020

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

To my lovely wife:

Ailar

&

To my beloved parents:

Reza and Nasrin

Abstract

Massive multiple-input multiple-output (MIMO) and millimeter-wave (mmWave) communications are established as key technologies for fifth generation (5G) and beyond (5G&B) networks. However, the practical implementation of mmWave massive-MIMO systems remains challenging. Conventional MIMO systems are implemented using the fully-digital (FD) architecture, in which signal processing is performed in the digital domain by means of dedicated processors and/or digital circuitry. At the transmitter, the digital baseband output signals are then converted to analog signals for transmission, which requires a dedicated radio frequency (RF) chain per antenna element. For the large-scale antenna arrays envisaged for massive-MIMO systems, however, the FD architecture is impractical due to the huge power consumption and production costs.

One the most effective solutions to this problem is hybrid analog/digital (A/D) beamforming (HBF). In this approach, an additional signal processing layer in the analog domain, referred to as analog beamformer, is added between the RF chains and the antenna elements. In effect, by properly designing the analog beamformer, it becomes possible to reduce the number of RF chains while achieving a performance comparable to the FD architecture.

There are three parts to this thesis all of which have a common goal, which is to achieve the performance of FD systems with HBF. In the first part, we consider HBF at the transmitter side and study the minimum number of required RF chains for realizing a given FD precoder with the HBF architecture. We further investigate HBF designs based on the single RF chain architecture for mmWave massive-MIMO systems. We present three novel beamformer designs which achieve the performance of FD precoding systems. Finally, we extend these results to MIMO-OFDM systems.

The second part studies HBF at the receiver. Particularly, we propose a novel hybrid structure for realizing a given FD combiner with the minimum number of required RF chains. We then focus on a more practical scenario where phase-shifters can realize a finite number of phase angles. Accordingly, we propose a modified hybrid structure by introducing an additional degree of freedom, i.e., phase-offset between the finite-resolution phase-shifts and optimize this parameter via close approximations. Robust hybrid combiners are then studied for the case of imperfect channel knowledge at the receiver.

In the final part of this thesis, we explore the hybrid A/D structure as a general frame-

work for signal processing in massive and ultra-massive-MIMO systems. To exploit the full potential of the analog domain, we first focus on the analog signal processing (ASP) network. We investigate a mathematical representation suitable for any arbitrarily connected feed-forward ASP network comprised of the common RF hardware elements in the context of hybrid A/D systems, i.e., phase-shifter and power-divider/combiner. A novel ASP structure is then proposed which is not bound to the unit modulus constraint, thereby facilitating the hybrid A/D systems design. We then study MIMO transmitter and receiver designs to exploit the full potential of digital processing as well. An optimization model based on the proposed structure is presented that can be used for hybrid A/D system design. Finally, precoding and combining designs under different conditions are discussed as examples.

Sommaire

Les communications massives à entrées multiples et sorties multiples (MIMO) et les ondes millimétriques (mmWave) sont désormais établies comme des technologies clés pour les réseaux de cinquième génération (5G). Néanmoins, la mise en œuvre pratique des systèmes MIMO massifs mmWave reste difficile. Les systèmes MIMO conventionnels sont mis en œuvre en utilisant l'architecture entièrement numérique (FD), dans laquelle le traitement du signal est effectué dans le domaine numérique au moyen de processeurs dédiés et/ou de circuits numériques. Dans le scénario de liaison descendante, les signaux de sortie numériques en bande de base sont ensuite convertis en signaux analogiques pour la transmission, ce qui nécessite une chaîne radio-fréquence (RF) dédiée par élément d'antenne. Cependant, pour les réseaux d'antennes à grande échelle envisagés pour les systèmes MIMO massifs, une architecture FD n'est pas pratique en raison de l'énorme consommation d'énergie et des coûts de production.

L'une des solutions les plus efficaces à ce problème est l'hybride formation de faisceaux analogique/numérique (HBF). Dans cette approche, un couche de traitement du signal dans le domaine analogique, référencée en tant que formateur de faisceau analogique, est ajouté entre le RF chaînes et les éléments d'antenne. En effet, en concevant correctement le formateur de faisceaux analogique, il devient possible de réduire le nombre de chaînes RF tout en atteignant des performances comparables à l'architecture FD.

Cette thèse se compose de trois parties qui ont toutes un objectif commun qui est d'atteindre les performances de systèmes entièrement numériques avec HBF. Pour cela, nous considérons d'abord la formation de faisceaux hybride au niveau de l'émetteur et étudions le nombre minimum de chaînes de RF requises pour réaliser un précodeur FD MIMO arbitraire avec l'architecture HBF. Nous étudions ensuite des conceptions hybrides de formation de faisceau basées sur l'architecture à chaîne RF unique pour les systèmes MIMO massifs mmWave. Nous présentons trois nouvelles conceptions de formateurs de faisceau qui permettent d'obtenir les performances des systèmes de précodage entièrement numériques. Nous explorons plus en détail les applications de ces conceptions pour un précodage optimal dans les scénarios mono-utilisateur et multi-utilisateur. Enfin, nous étendons ces résultats aux systèmes MIMO-OFDM.

La deuxième partie étudie l'utilisation de l'architecture HBF au niveau du récepteur. En particulier, nous proposons une nouvelle structure hybride pour réaliser tout combina-

teur FD avec le nombre minimum de RF requis par une structure HBF pour correspondre aux performances d'un formateur de faisceau FD. Nous nous concentrons ensuite sur un scénario plus pratique où les déphaseurs peuvent réaliser un nombre fini d'angles de phase. En conséquence, nous proposons une structure hybride modifiée en introduisant un degré de liberté supplémentaire, c'est-à-dire un décalage de phase entre les déphasages à résolution finie et optimisons ce paramètre via des approximations serrées qui permetted d'obtenir la même solution que la recherche exhaustive. Nous développons également une nouvelle technique d'estimation de canal basée sur un pilote qui peut atteindre les mêmes performances que l'estimation linéaire optimale FD.

Dans la dernière partie de cette thèse, nous explorons la structure hybride analogique/numérique en tant que cadre général pour le traitement du signal dans les systèmes MIMO massifs et ultra-massifs. Pour exploiter tout le potentiel du domaine analogique, nous nous concentrons d'abord sur le réseau de traitement du signal analogique (ASP). Nous étudions une représentation mathématique adaptée à tout réseau ASP à anticipation, connecté arbitrairement, et composé des éléments matériels RF communs dans le contexte de systèmes hybrides, c'est-à-dire déphaseur et diviseur/combinateur de puissance. Une nouvelle structure ASP est alors proposée qui n'est pas liée à la contrainte de module unitaire, facilitant ainsi la conception de systèmes hybrides. Nous étudions ensuite les conceptions d'émetteurs et de récepteurs MIMO afin d'exploiter également tout le potentiel du traitement numérique. Un modèle d'optimisation basé sur la structure proposée est présenté qui peut être utilisé pour la conception de systèmes hybrides. Plus précisément, la conception de précodeurs et combinateurs est dicustée sous différentes conditions à titre d'exemples.

Acknowledgements

First and foremost, I would like to express my deepest appreciation and gratitude to my supervisor, Prof. Benoit Champagne, for his insightful guidance and generous support at every stage of my Ph.D. study. Without his technical support and constant help, the completion of this thesis would have not been possible. I would also like to thank him for his endless patience and great effort in editing and improving my academic papers. I would also like to express my gratitude towards my supervisory committee members Prof. Le-Ngoc, Prof. Zhu and Prof. Shepherd for their time and efforts to evaluate my work and provide valuable feedback.

I am deeply grateful for the generous funding support from the McGill Engineering Doctoral Awards (MEDA). I am also grateful for the financial support provided by my supervisor, through research grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and by InterDigital Inc, without which this thesis would not have been possible. I would like to thank the technical staff of InterDigital Inc, specifically, Dr. Afshin Haghighat for his constructive inputs, comments and ideas.

I would like to acknowledge my colleagues and friends, who made my PhD the most memorable experience of my life. In particular, I would like to thank Ali Baghaki, Seyyed Saleh Hosseini, Sara Norouzi, Alec Poulin, Jiaxin Yang, and Hao Li. I would like to deeply thank my parents, Reza and Nasrin, for their endless love and support through my life. Last but not least, I owe my deepest thanks to my lovely wife, Ailar, for her endless love, countless sacrifices and timely encouragements.

Preface and Contributions of the Author

The research presented in this dissertation was carried out in the Department of Electrical and Computer Engineering (ECE) of McGill University from January 2016 to November 2020. This dissertation is the result of my original work under the supervisions of Prof. Benoit Champagne from the same department. The research has led to several journal and conference publications. In all of these publications, I proposed the idea, formulated the problem, developed the algorithms, implemented simulations and prepared the first draft of the manuscripts. Prof. Benoit Champagne, who is co-author of all these publications, reviewed the works and assisted with the editing and writing of the final manuscripts. All the other co-authors provided constructive comments to improve the quality of manuscripts. Our collaboration with the industry also resulted in two US patent applications with the help and support of Dr. Afshin Haghighat. The full list of my publications during my PhD studies is given below:

- **Patents**

  1. A. Morsali, A. Haghighat, and B. Champagne, "Efficient implementation of hybrid beamforming," *U.S. Patent,* 10,790,889, issued September 29, 2020.
  2. A. Morsali, A. Haghighat, and B. Champagne, "Efficient receiver combining for hybrid analog/digital beamforming ," *U.S. Patent,* PCT/US2020/024318, March 23, 2020.

- **Journal Papers**

  1. A. Morsali, A. Haghighat and B. Champagne, "Generalized framework for hybrid analog/digital signal processing in massive and ultra-massive-MIMO Systems," *IEEE Access*, vol. 8, pp. 100262–100279, 2020.
  2. A. Morsali, A. Haghighat, and B. Champagne, "Realizing fully digital precoders in hybrid A/D architecture with minimum number of RF chains," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2310-2313, Oct. 2017.
  3. A. Morsali, S. S. Hosseini, B. Champagne and X. Chang, "Design criteria for omnidirectional STBC in massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1435-1439, Oct. 2019.

- **Conference Papers**

  1. A. Morsali and B. Champagne, "Achieving fully-digital performance by hybrid analog/digital beamforming in wide-band massive-MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, 2020, pp. 5125-5129.
  2. A. Morsali and B. Champagne, "Single RF chain hybrid analog/digital beamforming for mmWave massive-MIMO," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP),* Ottawa, Canada, 2019.
  3. A. Morsali and B. Champagne, "Robust hybrid analog/digital beamforming for uplink massive-MIMO with imperfect CSI," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC),* Marrakesh, Morocco, 2019.

Contents

List of Figures

List of Tables

List of Acronyms

B5G	Beyong fifth generation
4G	Forth generation
ADC	Analog to digital converter
BS	Base station
CSI	Channel state information
DAC	Digital to analog converter
FD-MIMO	Full dimension MIMO
FD	Full(y) digital
FDP	Fully-digital precoding
FS	Frequency selective
HBF	Hybrid analog/digital beamforming
HSP	Hybrid analog/digital signal processing
IMT	International mobile telecommunication
IOT	Internet of things
LTE-A	Long term evolution advanced
LSAS	Large scale antenna system
mmWave	Millimeter-wave
MIMO	Multiple-input-multiple-output
MU	Multi-user
MMSE	Minimum mean squared error
MSE	Mean square error
SU	Single-user
UE	User equipments
THz	Terahertz

Chapter 1

Introduction

In this chapter, we first introduce the main features of the next generation of wireless cellular networks, i.e., fifth generation (5G) and beyond 5G (B5G). We then point out some of the major challenges of the proposed communication techniques for B5G and present the concepts of hybrid architecture as a key framework for physical layer 5G. We then state the detail objectives and contributions of the thesis. Finally, we present its organization along with some mathematical notations.

1.1 Massive-MIMO mmWave Communications for 5G

A huge growth, in the range of 10-100 times, is anticipated for the international mobile telecommunication (IMT) traffic from 2020 to 2030 [1]. As shown in Fig. 1.1, mobile traffic is projected to grow at an annual rate of around 55% in next decade. The global mobile traffic per month is also estimated to reach about 600 EB (Exabyte=10^{18}B) in 2025 and 5000 EB in 2030. The official introduction and early implementation of 5G cellular networks happened in early 2020 while its widespread deployments is expected by 2025. Compared to fourth generation (4G), also known as long-term evolution (LTE) and LTE-advanced (LTE-A), several objectives have been set for 5G, including in particular: 1000 times higher mobile data traffic per geographic area, 10 times more connected user equipments (UEs), 100 times higher user data rate, 10 times lower network energy consumption and 5 times less latency [3, 4].

Existing communication systems, nevertheless, are reaching the Shanon limit and many optimal communication techniques have been already explored and even implemented.

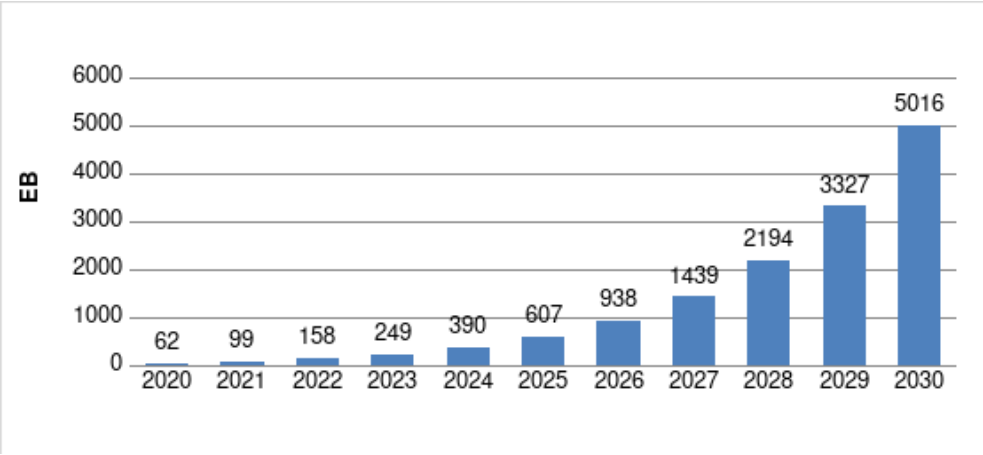


Fig. 1.1: Estimations of global mobile traffic from 2020 to 2030 [1].

Thus, realizing the ambitious 100-fold increase in data rebate and 10-fold increase in user capacity of 5G compared to legacy networks is a major challenge which only seems possible by taking advantage of: (1) much larger bandwidth (available at higher frequencies); (2) much larger number of antennas; and (3) ultra dense networks. The first avenues are complementary to each other, leading to the implementation of large scale arrays in millimeter-wave (mmWave) spectrum, i.e., frequency bands in the rage from say, 10 to 100 Ghz. Moreover, their integration will allow to make advances along the third avenue [5].

Deployment of multiple antennas at both transmitter and receiver, i.e., multiple-input multiple-output (MIMO) was undoubtedly a huge step up for wireless communication systems. While bandwidth is limited and increasing signal-to-noise ratio (SNR) increases the capacity only logarithmically, it has been shown that MIMO can linearly increase the capacity by increasing the number of antennas [6]. Although, this is true if the channel matrix coefficients are independent and identically distributed which is not always the case, especially, in mmWave systems, as the number of antennas grows larger, this condition is met in asymptotic limit according to random matrix theory. From an information theoretic point of view, it is shown that as the number of antennas becomes extremely large, capacity can increase linearly with the minimum number of employed antenna at transmitter or receiver. This conceptual extension of MIMO has been presented under different names in the literature, i.e.: massive-MIMO, large scale antenna system (LSAS), full dimension MIMO (FD-MIMO) and hyper MIMO [7–9].

Conventional LTE cellular networks operate in frequency bands between 2 to 6 GHz, i.e., microwave which provides a bandwidth of up to 2 GHz. Clearly, the congested microwave can no longer support the increasing demand for additional bandwidth and we are at the dawn of the mmWave era with carrier frequencies extending up to 100 GHz and offering potential bandwidths on the order of 20 GHz. The mmWave communications is not a new technology; it has already been proposed for indoor and fixed outdoor communications which provide data rates on the order of Gbps [10, 11]. Unfortunately, mmWave signals experience high path loss, penetration loss and atmospheric absorption compared with microwave signals, which explains why mmWave was not considered for cellular networks until now. Recent advances in mmWave hardware and more importantly beam-steering capabilities of massive-MIMO have brought a new life to mmWave communications. Specifically, directional communication and the ability to shape sharp beams by the means of LSAS will help to overcome the severe path loss experienced by mmWave signals [12].

Massive-MIMO and mmWave communications are both compelling technologies when considered separately, although each one suffers from a number of shortcomings in practice. Surprisingly, their integration, not only accentuates their capabilities but also resolves some of their shortcomings. Massive-MIMO offers highly directional communications which help combat the open air path loss of mmWave transmissions. In turn, operation at mmWave frequencies allows a reduction in the dimension of antenna elements which facilitates implementation of LSAS [9].

Consequently, Massive-MIMO communications with LSAS at the mmWave is the prime candidate technology for B5G cellular networks [7, 8, 11, 13]. In fact, base-stations (BS) with 64 antennas have been recently deployed for commercial use in some countries [14]. Moreover, an extensive theory for massive MIMO has been developed in recent years, including capacity and spectral efficiency analysis, system design for high energy efficiency, pilot contamination, etc. However, implementation of such systems faces many technical difficulties, and to this day remains very challenging and costly [9, 15].

1.2 Hybrid Analog/Digital Signal Processing

In conventional fully-digital (FD) MIMO systems, each antenna element requires a dedicated radio frequency (RF) chain. Hence, the direct FD implementation for massive-MIMO systems is not practical and efficient due to the ensuing high production costs and more im-

portantly, huge power consumption. Hybrid analog/digital (A/D) signal processing (HSP) is an effective approach to overcome this problem by cascading an analog signal processing (ASP) network with the baseband digital signal processor [16, 17]. While in conventional FD MIMO transmitters [18–20], each antenna element is directly controlled by the digital processor, in an HSP-based transmitter, the digital processor generates a low-dimensional RF signal vector, whose size is then increased by analog circuitry for driving the large-scale antenna array. Similarly, in an HSP-based receiver, the size of the high-dimensional vector of antenna signals is reduced by an ASP network, whose outputs are then converted to the digital domain for baseband processing by means of RF chains.

There are practical constraints in the implementation and design of ASP networks and only a few types of RF components are commonly used in practice. Specifically, the power-divider (splitter), power-combiner (adder), and phase-shifter are the key analog components of the ASP design [21–29]. In the existing hybrid beamforming (HBF) structures, due to the particular configuration of the aforementioned analog components, a constant modulus constraint is imposed to the analog beamformer weights which turn the overall transceiver design into an intractable non-convex optimization [21, 22].

1.3 Contributions

The premise of HBF is the efficient implementation of massive-MIMO systems while achieving the performance of the FD systems, which has been the main focus of our contributions in this thesis. The latter have been grouped into three main parts: we first consider HBF at the transmitter, i.e., hybrid precoding; we then focus on the receiver, i.e., hybrid combining; and finally we present a generalized framework for HSP. The key contributions for each part are summarized below.

Part I, Hybrid A/D Precoding: We mainly focus on realizing any given fully-digital precoder (FDP) with the hybrid architecture while minimizing the required number of RF chains. Main contributions of Part I are summarized as follows:

- We first introduce a generalized system model where the baseband digital precoder is viewed as an abstract transformation, and then, we formulate a constrained optimization problem to minimize the number of required RF chains such that the FDP and HBF outputs are identical.

- We show that feasible optimum solutions exist for only two RF chains; however, optimal solutions are not unique, which allows for a freedom in the design. Based on the optimum solution so obtained, we propose a simple yet novel HBF design for the realization of any given FDP with a single RF chain, which does not require extra hardware or computational complexity in its implementation.

- The realization of any given FDP with the HBF architecture in wide-band massive-MIMO-OFDM systems is investigated by inroducing appropriate system and channel models.

- We then delve into the time-domain reconstruction of OFDM signals via HSP and take the extra step towards realizing any FDP, by discussing the required conditions for the most general case.

- Next we introduce one specific case and derive a systematic design for arbitrary number of RF chains. In particular, the trade-off between number of RF chains and the system bandwidth is discussed. Finally, we present the design for two RF chains are present along with the simulation results.

Part II, Hybrid A/D Combining: Extending our hybrid precoding design to the receiver side is not trivial: indeed, while at the transmitter side, the input of the RF beamformer is known at each time slot, at the receiver, we only have statistical knowledge of the RF beamformer input. Consequently, in Part II, we investigate HBF structures that can achieve FD performance. Our main contributions can be summarized as follows:

- We develop new HBF realization that can match the FD combiner performance where the number of RF chains is equal to the number of data streams. The new structure is not bound to unit modulus constraint and thus facilitates the design process of HSP systems.

- We then consider a more practical scenario where only finite-resolution phase-shifters are used to implement the analog beamformer and optimize the proposed hybrid design. A modified hybrid structure is presented to improve the performance. Taking advantage of available degrees of freedom in our design, we minimize the maximum error between the RF beamformer and its FD counterpart.

- We also present a novel channel estimation method based on the proposed hybrid structure which achieves the performance of optimal linear FD channel estimation, but with reduced number of RF chains.
- Next, we investigate the design of hybrid A/D beamformers for uplink connection in massive-MIMO systems under imperfect channel state information (CSI) for both single user (SU) and multi-user (MU) scenarios.
- In the SU scenario, we consider the computational capabilities of the user for designing the beamformer. For users with limited computational resources, a hierarchical optimization schemes is presented which only puts the burden of robust calculations on the BS.
- For users with extended computational capabilities, joint precoder/combiner design is proposed. We then extend these optimization techniques to propose a new robust hybrid combiner design for MU.

Part III, Generalized Hybrid A/D Signal Processing: In Part III, our goal is to investigate and exploit the full potential of HBF in massive-MIMO systems. Aiming at this challenge, we can summarize our contributions as follows:

- We first explore the degrees of freedom in the analog domain by developing a compact mathematical representation for any given feed-forward ASP network with arbitrary connections of any number of RF components, i.e., phase-shifters, power dividers and power combiners.
- Based on the above generalization, a simple and novel ASP architecture is conceived out of the above RF components, which is not bound to the constant modulus constraint. Removing this constraint facilitates system design as non-convex optimizations are difficult to solve and global optimality of the solutions cannot usually be guaranteed.
- The transmitter and receiver sides are then studied separately by exploiting the newly proposed ASP architecture and generalizing the digital processing. Specifically, the optimization problem for the HSP beamformer is reformulated within the new representation framework, which facilitates its solution under a variety of constraints and requirements for the massive MIMO system.

- The realization of optimal FD by HBF and the problem of RF chain minimization are presented as guideline examples to illustrate potential applications of the proposed theoretical framework.

1.4 Thesis Outline and Notations

Except for the introductory and concluding Chapters 1 and 8, respectively, the remaining chapters of this thesis are presented in the same chronological order as the research itself, and are organized as follows:

- In Chapter 2, an overview of the relevant fundamental material on hybrid beamforming for mmWave massive-MIMO communication is presented followed by a brief literature survey as well as a generic mmWave massive-MIMO system model.

Part I, Hybrid A/D Precoding

- In Chapter 3, the massive-MIMO transmitters with hybrid A/D architecture sretudied. We first present a single RF chain architecture (SRCA) for efficient implementation of HBF for massive-MIMO systems. Then, we present three novel beamformer schemes which eliminate the defects of SRCA, and can achieve the performance of FD beamforming schemes.

- In Chapter 4, we extend our design to wideband massive-MIMO mmWave communications by introducing the notion of time domain signal reconstruction of OFDM signals. We then present the general solution for any arbitrary number of RF chains and as a specific example, we consider a design for only two RF chains.

Part II, Hybrid A/D Combining

- In Chapter 5, the massive-MIMO receivers with hybrid A/D architecture are studied. We present a novel HBF structure for massive-MIMO communication systems that matches the performance of any given FD combiner. We also focus on finite-resolution phase-shifters and present a modified hybrid structure with phase-offset as an additional degree of freedom in order to improve the system performance. We then find a closed-form solution for the phase-offset which minimizes the error between FD combiner and its analog realization. A novel hybrid channel estimation is also presented based on the proposed hybrid structure that achieves FD performance.

- In Chapter 6, we study robust hybrid combiner design for imperfect CSI. A norm-bounded channel error model is used to capture the imperfect CSI conditions and the objective function is formulated based on the worst-case robustness. For the SU scenario, we present hierarchical optimization as well as joint optimization (based on UE capabilities to perform extra calculations) for robust transceiver design. We then propose a robust hybrid combiner design for MU uplink connections.

Part III, Generalized Hybrid A/D Signal Processing

- Finally Chapter 7, we investigate the hybrid A/D structure as a general framework for signal processing in massive-MIMO systems. We first explore the ASP network in detail by developing a mathematical representation for any arbitrarily connected feed-forward ASP network comprised of phase-shifters, power-dividers and power combiners. Then, a novel ASP structure is proposed which is not bound to the unit modulus constraint. Subsequently, we focus on the transmitter and receiver sides by exploiting the newly proposed ASP architecture and generalizing the digital processing. Specifically, the optimization problem for the HSP beamformer is reformulated within the new representation framework, which facilitates its solution under a variety of constraints and requirements for the massive MIMO system.

Notations

Throughout this thesis, we use bold capital and lowercase letters to represent matrices and vectors, respectively. Superscripts $(\cdot)^H$ and $(\cdot)^T$ indicate Hermitian and transpose, respectively. Kronecker product is denoted by $\otimes$. $\mathbf{I}_n$ and $\mathbf{1}_n$ denotes an identity matrix of size $n \times n$ and a column vector of all ones with size n, respectively. The element on the p^{th} row and q^{th} column of matrix $\mathbf{A}$ is denoted by $A(p,q)$, while the p^{th} element of vector $\mathbf{x}$ is denoted by x_p. $\mathrm{Tr}(\mathbf{A})$ and $\|\mathbf{A}\|_F$ denote the trace and Frobenius norm of matrix $\mathbf{A}$, respectively. $\mathbf{A} = \mathrm{diag}(a_1, a_2, \ldots, a_n)$ represents a diagonal matrix, in which $a_1, a_2, \ldots, a_n$ are placed on the main diagonal. A complex $n \times 1$ Gaussian random vector $\mathbf{x}$ with mean vector $\mathbf{m} = \mathbb{E}\{\mathbf{x}\}$, where $\mathbb{E}\{\cdot\}$ stands for expectation, and covariance matrix $\mathbf{R} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ is denoted by $\mathcal{CN}(\mathbf{m}, \mathbf{R})$. The absolute value and phase of a complex number z is denoted by $|z|$ and $\angle z$, respectively, and $\mathbb{C}$ stands for complex field. $\mathbf{A} = \mathrm{bd}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n)$ represents a block-diagonal matrix, in which $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$ are the diagonal blocks of $\mathbf{A}$. By $\mathbf{x}_1 \overset{\pi}{=} \mathbf{x}_2$, it is meant that there exists a permutation matrix

$\mathbf{P}_\pi$ such that $\mathbf{x}_1 = \mathbf{P}_\pi \mathbf{x}_2$. The greatest (least) integer less (greater) than or equal to x is denoted by $\lfloor x \rfloor$ ($\lceil x \rceil$). Moreover, $x = a \bmod n$ denotes the remainder of the division of a by n.

Chapter 2

Background and Literature Review

In this chapter, background study and literature review of hybrid beamforming for mmWave massive-MIMO communications are provided in preparation for the technical chapters presented in the thesis. We first outline the distinctive features and advantages of mmWave spectrum, and then review the potential applications of massive-MIMO mmWave communications. Afterwards, we present a brief literature survey of the key research studies in the topical area of this thesis. Finally, a simplified, generic system model of FD and HBF massive-MIMO mmWave systems is introduced.

2.1 Massive-MIMO and mmWave Communications

The undesirable characteristics of high-frequency propagation prevent the key wireless technologies used at the sub-6 GHz frequencies to be directly applied in mmWave spectrum. The main distinctive features which must be considered for modelling mmWave transmission can be summarized as follows:

- Pathloss and Signal Power: The transmitted signal in all wireless communication systems undergoes distance-based attenuation. The isotropic path loss increases in direct proportion to the inverse of the squared wavelength, specifically, we have

$$P_R = \frac{P_T G_T G_R \lambda^2}{(4\pi d)^2}, \tag{2.1}$$

  where d is the distance between the transmitter and receiver, λ is the wavelength,

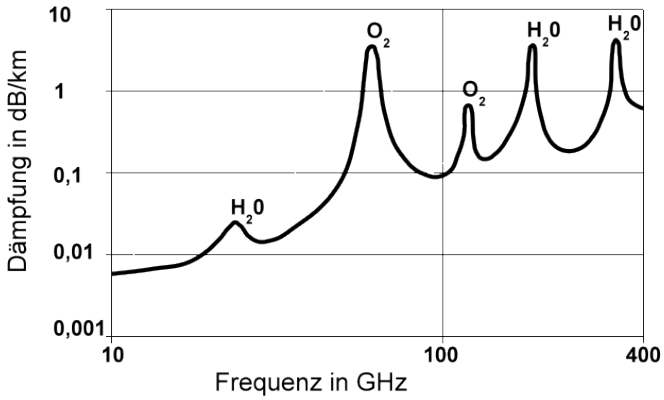


Fig. 2.1: Atmospheric attenuation where peaks in absorption at specific frequencies are due to atmosphere constituents such as water vapour (H2O) and molecular oxygen (O2).

G_T, and G_R are the transmitter and receiver gains, P_R and P_T are the receive and transmit power, respectively [30]. Consequently, for fixed antenna gains, mmWave signals experiences higher pathloss compared to microwave ones due to their shorter wavelength. Hence, other conditions being equal, the received power is much less compared to lower frequency bands. However, directional transmissions with higher gain antennas can compensate for this high path loss [31]. In practice, the directivity gain G of an antenna (whether used as a transmitter or receiver) can be expressed as a function of the antenna effective area A and the operating wavelength, i.e. $G \propto \frac{A}{\lambda^2}$. Hence, by using antenna arrays along with beamforming, we can effectively increase the aperture size and the antenna gain. [2]

- Sparse Scattering Channel: Compared to microwave, mmWave signals are more sensitive to blockages because most obstacles in the propagation environment such as buildings, cars, etc, are made up of non-translucent surfaces larger than the wavelength of mmWave communications signals. However, reflection and scattering can facilitate the transmission between the transmitter and receiver if steerable antennas are used to locate the objects that reflect or scatter or avoid those that absorb the signal. In Fig. 2.1, the atmospheric attenuation in of the mmWave spectrum is illustrated. In addition, since many objects in the surrounding will absorb the mmWave signals, only a small number of scatterers or reflectors will contribute to the

propagation. Consequently, the channel impulse response at these frequencies will tend to be sparse [32, 68].

In spite of the above difficulties, the mmWave band offers remarkable advantages compared to lower frequency bands, some of which are summarized as follows:

- Higher Data Rate: Clearly the first benefit of mmWave communications is a very large bandwidth (on the order of 20 GHz) which directly translates to higher data rates.

- Dense Networks: In addition to the increased channel bandwidth, mmWaves can be leveraged to reduce coverage areas, i.e., establish more densely packed communication links and exploit spatial reuse to provide increased capacity gains.

Moreover, mmWave communications facilitate implementation of massive-MIMO systems by allowing the deployment of a larger number of smaller antenna elements which results in additional performance gains. Interference suppression, which primarily relies on beamforming at the transmitter and receiver (respectively known as precoding and combining), plays an important role in multi-user communications. Aside from beamforming, MIMO can be used for spatial diversity and multiplexing. In addition to other sources of diversity such as time and frequency, spatial diversity can be exploited at both ends to defeat the impact of fading in wireless channels. MIMO can be also used to transmit parallel data streams without using additional bandwidth or power which increases the number of spatial dimensions for communication.

Although there are certain advantages in operating at mmWave frequencies as explained above, a number of challenges must be addressed to exploit these benefits which are explained below.

- Doppler Spread: Due to a large doppler spread, the coherence time in the mmWave range is very small which becomes a major challenge with user mobility [34]. Consequently, the rapidly chaining behaviour of the wireless channel must be taken into account when designing the channel estimation, modulation and coding schemes in mmWave cellular networks.

- RF Circuitry: RF design has been always a bottleneck in reaching the theoretical limits of wireless communication systems. Particularly, when designing RF integrated

circuits for mmWave systems, non-linear distortions in the power amplifiers, phase noise and IQ imbalance become more dramatic due to the high carrier frequencies and wide bandwidth.

- Power Consumption: Massive-MIMO is a double-edged sword for mmWave communications. While a larger number of antennas enables directional beamforming with higher resolution, a direct consequence of this is an increase in hardware complexity and power consumption. Specifically, since each antenna requires a dedicated RF chain which includes data converters, mixers, and power amplifiers, the power consumed by RF chains scale up with increasing number of antennas. Consequently, an alternative signal processing technique, i.e., hybrid A/D processing is required for mmWave massive-MIMO systems, as further listed below.

2.2 Applications of Massive-MIMO mmWave Communications

Although the mmWave spectrum has been in use for military, radar and wireless backhaul applications, the first commercial use case of mmWave was indoor communications [10]. In fact, the first commercial applications of mmWave technology appeared in the context of wireless personal area networks and wireless local area networks (WLAN) standards for operation in the unlicensed bands around 60 GHz. WirelessHD is one of the most popular standards that uses mmWave technology to provide high-bandwidth wireless links as a replacement for wired connections, and capable of handling uncompressed high-definition multi-media content. IEEE 802.11ad which is an amendment to the ubiquitous IEEE 802.11 protocol is another standard for WLAN operating in the 60 GHz band [35]. These early and successful applications of mmWave paved the way for many potential applications which are discussed below.

- B5G communication systems: While it has been demonstrated that mmWave transmission can provide gigabit data rates in indoor communications and point-to-point backhaul links, conventional system design techniques call for expensive hardware components in order to provide reliable long-ranged communication links. However, with the advent of massive-MIMO, low-cost mmWave technologies and advanced signal processing techniques, mmWave transmission can now be used as well for outdoor

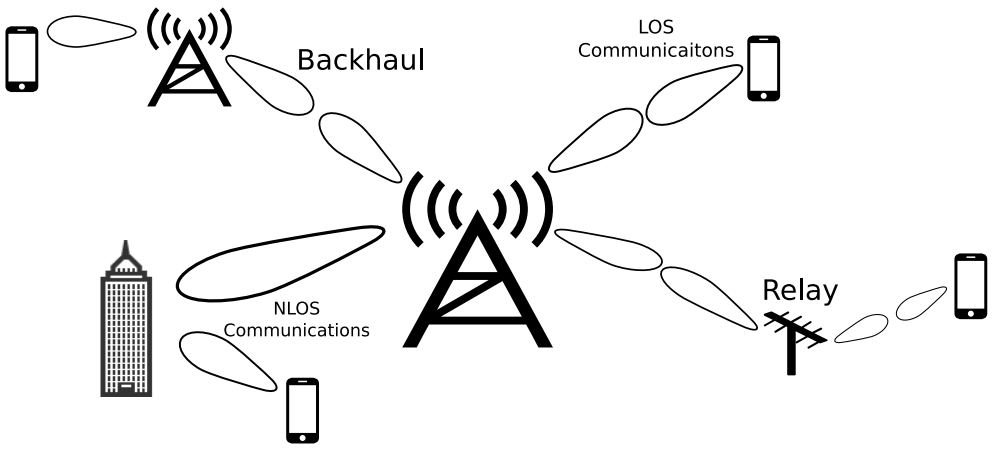


Fig. 2.2: Illustration of a mmWave cellular network [2].

communication networks [36]. The architecture of mmWave cellular networks however is likely to be much different than in microwave systems, as illustrated in Fig. 2.2.

- P2P networks: For many years, cellular and personal networks were the main focus of both research and industry. Recently, there has been a growing interest in peer-to-peer (P2P) networks. Specifically, vehicular applications such as autonomous vehicles, vehicle-to-vehicle communications, and vehicle-to-infrastructure have received considerable attention [37]. Because of all the aforementioned benefits of mmWave and the fact that this technology is already established in automotive radar applications [38], we can foresee that mmWave will play a significant role in vehicular networks.

- Personal networks: Advanced wireless technology is capable of establishing high-speed connection via low-power and small-scale devices. One direct use case is in wearable networks for providing wireless connections for smart devices such as cell-phones, tablets, watches, activity-tracking devices, virtual-reality headsets and glasses. Space limitation is one key aspect why mmWave is of interest to these networks, with the added benefit of high data rate and very low latency communication which is also another requirement in this case [39].

- IoT networks: The advent of internet of things (IoT) is indeed dependent on machine-to-machine communications which provides networking connectivity between physical objects of various sizes. This may eventually give rise to huge networks comprising

millions of connections; fortunately, mmWave offers a massive spectrum and has the capacity to locally support large number of connections at high data rates. This explains why mmWave is a prime candidate for the physical wireless infrastructure of IoT [40].

2.3 Literature Review

At the beginning of each one of the subsequent technical chapters, a comprehensive literature review of relevant works to the chapter's main topic is presented. Nevertheless, in this section, a brief survey of the key works in mmWave communications, massive-MIMO and HSP is provided. The mmWave communications frameworks and then fundamental works on massive-MIMO are presented. Finally, a brief review of the major works on HSP is provided.

mmWave Communications: One of the early applications of mmWave communications was for indoor communications enabling multigigabit wireless links; this technology was even standardized in commercialized systems [10, 35, 41, 42]. Since the densely crowded microwave spectrum cannot meet the requirements of 5G, hence, the mmWave bands between 30 to 100 GHz with abundant available bandwidth of more than 20 GHz is the new horizon of next generation of cellular wireless systems [33, 43–45].

The first step towards mmWave communications is indeed a profound understanding of propagation effects and channel modelling. The fundamentals of mmWave propagation in fact, had already been studied for indoor communications [46]. Nevertheless, mmWave MIMO channel modeling have been revisited in recent years and many works have been published in this area [41, 47–49].

The biggest drawback of mmWave communications for outdoor systems is the severe pathloss, resulting from combined effects of free-space pathloss, penetration loss, atmospheric absorption and rain attenuation [9, 50]. It was shown however that the deployment of a larger number of antennas at both transmitter and receiver can effectively defeat pathloss by means of beamforming [17, 51]. Essentially, outdoor mmWave communications are only possible in massive-MIMO systems. Besides, due to the form factor restrictions, implementation of massive-MIMO is also only logical in mmWave bands. This complementarity makes the integration of these techniques a fortunate opportunity for future

B5G wireless networks.

Massive-MIMO: The concept and main theories of massive-MIMO [52] were developed following the success of MIMO techniques in wireless local area and cellular systems [6, 53]. However, it took more than a decade for massive-MIMO to be considered as a practical solution [7–9, 51]. There have been many research studies dedicated to massive-MIMO systems and there are still many questions to be answered [54, 54]. While massive-MIMO is now established as a key technology for B5G [55], different aspects, challenges and opportunities are still being explored. In this thesis, we specifically focus on the hybrid architecture for implementation of massive-MIMO systems.

Hybrid A/D Signal Processing: In MIMO systems precoding can be performed when the radio channel descriptors are known at the transmitter [56]. In massive-MIMO, particularly, precoding plays a more critical role due to the severe pathloss effects of mmWave transmission. Conventional FD implementation of massive-MIMO is in fact inefficient because in such systems, driving large number of antennas require one dedicated RF chain for each antenna element which increases the overall system cost and power consumption [21].

Hybrid analog/digital systems, were introduced to simplify and reduce the cost and power consumption of massive-MIMO systems [17]. In fact, by cascading an analog (RF) precoder after the baseband digital precoder, it is possible to reduce the number of required RF chains. In [16], the first attempt was made to realize any FDP by means of hybrid A/D precoding with two RF chains for the case of a single data stream, i.e., one symbol per transmission. In [22, 23], multi-stream hybrid designs were presented in which the number of RF chains must be set equal to the number of symbols per transmission. The HBF has an intricate structure where the entries of the RF precoder matrix satisfy a constant modulus constraint associated to the use of phase-shifters. Since the ensuing precoder optimization is non-convex, many works have alternatively focused on designing HBF directly using heuristic iterative algorithms or reconstruction algorithms [21, 22]. Additional practical power constraints for the analog precoder designs are discussed in [57].

MIMO-OFDM Hybrid Beamforming: Orthogonal frequency division multiplexing (OFDM) is standardized in 4G [58] and in many other wireless systems such as WiFi [59]. The simplicity and flexibility of OFDM makes it the dominant multi-carrier (MC)

technique, even in next generation of wireless systems. However, implementation of OFDM in massive-MIMO systems has been even more challenges than for narrow-band systems. While the hybrid architecture is a powerful method in massive-MIMO-OFDM to reduce the number of RF chains, it makes the design of RF signal processing network more intricate. Since the RF precoder must have a similar response for all subcarriers, designing hybrid precoders is much harder than in the narrow-band scenario. There have been a few attempts to design wide-band hybrid precoders [60, 61]; however, this research is still too young and many questions remain to be answered [24, 62–65].

mmWave massive-MIMO Channel Estimation: Above, we reviewed recent works on designing hybrid beamforming for both narrow-band and wide-band transmission. However, all these techniques require an accurate estimation of the underlying channel parameters. Unfortunately, when the hybrid architecture is used for implementation of the massive-MIMO system, the conventional FD channel estimation techniques cannot be directly adapted; consequently, HBF-based channel estimation methods are required. The algorithm presented in [66] was one of the first works to consider channel estimation, along with HBF structure in narrow-band transmission; this work was later extended to the wide-band scenario in [67]. The estimation of channel related parameters for mmWave MIMO-OFDM systems is further studied in [68–71]. In all these works, the sparsity of the mmWave channels is exploited in the HSP-based channel estimators; however, there are still many other possibilities that can be explored and further studied.

2.4 System model

Herein, we present a simplified system model of a point-to-point mmWave system where the transmitter and receiver are equipped with large scale antenna arrays. This generic system model serves as a basis and sets the stage for designing new hybrid structures in the following chapters of the thesis.

2.4.1 Transmitter Beamforming

We consider a massive-MIMO transmitter with N_T antennas and N_{RF} RF chains where $N_T \gg N_{RF}$. The BS transmitter performs precoding on the symbol vector $\mathbf{s} = [s_1, s_2, ..., s_K]^T \in$

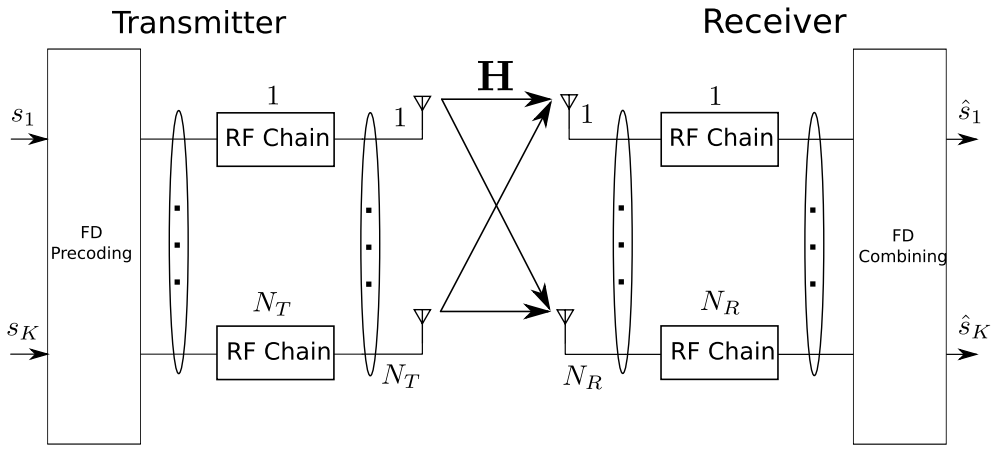


Fig. 2.3: Point-to-point massive-MIMO system with FD architecture.

$\mathcal{A}^K$, where each s_i is taken from a discrete constellation $\mathcal{A}$, such as M-QAM or M-PSK and K is the number of symbol streams. This representation applies to both SU and MU scenarios: in the former case, the K symbols are intended to a single multi-antenna user; while in the latter case, the K symbols are intended to K single-antenna users. In the sequel, we focus on the SU case.

FD Precoding: In Fig. 2.3, a point-to-point SU massive-MIMO system with FD beamforming at both ends is depicted. In a FD architecture, each antenna element is connected to a dedicated RF chain, i.e., $N_{RF} = N_T$. Thus, the transmitted signal can be written as

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}} = \mathbf{P}_{\mathrm{FD}}\mathbf{s}. \tag{2.2}$$

where $\mathbf{P}_{\mathrm{FD}}$ is the FD precoding (FDP) matrix of size $N_T \times K$.

Hybrid Precoding The massive-MIMO transmitter in HBF architecture is shown in Fig. 2.4, where $N_{RF} < N_T$ is assumed [16, 23]. Since only a limited number of RF chains is available, the symbol vector is first precoded with the digital beamformer and then analog beamforming is performed by the means of analog circuitry. The transmitted signal is therefore given as

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{A}}\mathbf{P}_{\mathrm{D}}\,\mathbf{s}, \tag{2.3}$$

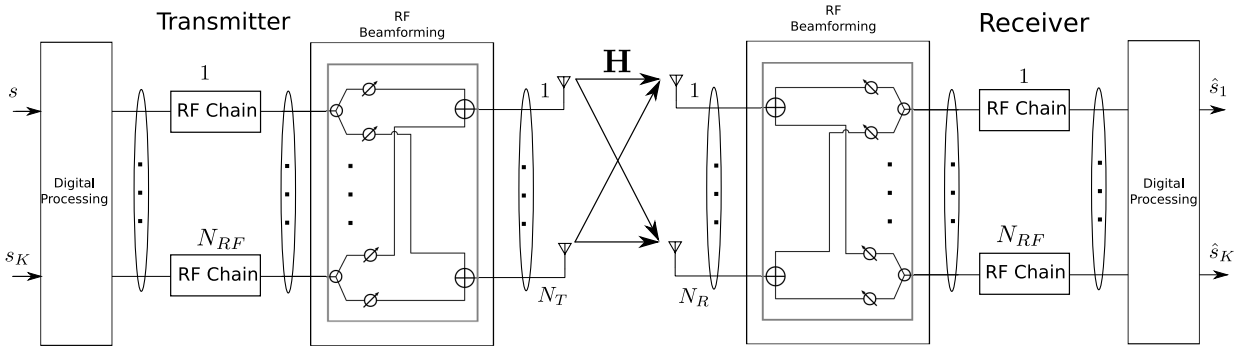


Fig. 2.4: Point-to-point massive-MIMO system with hybrid architecture.

where $\mathbf{P}_{\mathrm{D}} \in \mathbb{C}^{N_{RF} \times K}$ and $\mathbf{P}_{\mathrm{A}} \in \mathbb{U}^{N_T \times N_{RF}}$ are the digital and analog precoders, respectively, with $\mathbb{U}$ being the set of complex numbers with unit norm:

$$\mathbb{U} = \{z \in \mathbb{C} \colon |z| = 1\}. \tag{2.4}$$

2.4.2 Receiver Beamforming

Now, we consider the uplink connection of a massive-MIMO system where the BS is equipped with N_R receive antennas and N_{RF} RF chains. The received signal at the BS is given by

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} + \mathbf{n}, \tag{2.5}$$

where $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the point-to-point mmWave MIMO channel matrix, ρ is the average transmit power and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_R})$ is an additive white Gaussian noise (AWGN) vector.

FD Combining In FD systems (Fig. 2.3), since there are no limitations on the number of RF chains, i.e., $N_{RF} = N_R$, the transmitted symbols can be directly estimated a the digital combiner $\mathbf{W}_{\mathrm{FD}} \in \mathbb{C}^{N_R \times K}$ as

$$\hat{\mathbf{s}}_{\mathrm{FD}} = \mathbf{W}_{\mathrm{FD}}^{H}\mathbf{y}. \tag{2.6}$$

Hybrid Combining In the hybrid combining structure (Fig. 2.4), to reduce the dimension of the received signal so that $N_{RF} < N_R$, the signal first goes through the analog

combiner $\mathbf{W}_A \in \mathbb{U}^{N_R \times N_{RF}}$

$$\mathbf{x} = \mathbf{W}_A^H \mathbf{y}. \tag{2.7}$$

The RF chains then convert the analog vector $\mathbf{x}$ to a digital which in turn is used to estimate the transmitted symbols as

$$\hat{\mathbf{s}}_{\text{HY}} = \mathbf{W}_D^H \mathbf{x}, \tag{2.8}$$

where $\mathbf{W}_D \in \mathbb{C}^{N_{RF} \times K}$ is the digital combiner.

2.5 Conclusion

In this chapter, we presented background study and literature review on the main topic of this thesis, i.e., hybrid beamforming for mmWave massive-MIMO communications. We first covered the basics of massive-MIMO mmWave communications while discussing the distinctive features and advantages of mmWave spectrum. We then reviewed the applications of massive-MIMO mmWave systems. Moreover, we presented a brief literature survey of the key technologies studied in this thesis. Finally, we presented system model of FD and HBF massive-MIMO mmWave systems. In the next two chapters, which constitute the firs part of the thesis, we explore the design of hybrid A/D transmitter with minimum number of RF chain.

Part I

Hybrid A/D Precoding

Chapter 3

Massive-MIMO Transmitter with Hybrid A/D Architecture

In this Chapter[1], the minimum number of required RF chains for realizing a given FDP with the HBF architecture is investigated from a new perspective. First, in order to take full advantage of the digital domain, an extended system formulation of the hybrid structure is presented in which the digital precoder is taken from an abstract transformation space. Then, a constrained optimization problem is formulated, where the aim is to minimize the number of RF chains subject to an equality constraint on the multi-stream FDP and HBF outputs. We show that non-unique optimum solutions exist for only two RF chains. Based on this result, we further develop a single RF chain design. The implementation of the latter does not require extra hardware or computational complexity. We compare the proposed HBF design to recent works that require larger number of RF chains and demonstrate that significant improvement in performance can be achieved over direct HBF design.

[1]Parts of the work presented in this chapter have been patented, presented and published in [72–74]:

- A. Morsali, A. Haghighat, and B. Champagne, "Efficient Implementation of Hybrid Beamforming," *U.S. Patent,* 10,790,889, issued September 29, 2020.
- A. Morsali, A. Haghighat, and B. Champagne, "Realizing Fully Digital Precoders in Hybrid A/D Architecture with Minimum Number of RF Chains," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2310-2313, Oct. 2017.
- A. Morsali and B. Champagne, "Single RF Chain Hybrid Analog/Digital Beamforming for mmWave Massive-MIMO, " in *Proc IEEE Global Conf. Signal Inf. Process. (GlobalSIP).* Ottawa, ON, Canada, 2019.

3.1 Introduction

Massive-MIMO empowered by mmWave technology is emerging as a strong candidate for 5G and future wireless systems. Precoding is undoubtedly a fundamental and powerful technique in the design of MIMO systems. In massive-MIMO, precoding plays a more critical role due to the severe path loss affecting mmWave communications. However, because of the very large number of antennas in massive-MIMO systems, the conventional implementation of FDPs is not efficient. Indeed, FDP-based techniques require the use of one RF chain per antenna element, which increases the overall system cost and power consumption.

Recently, HBF have attracted considerable interest as a practical solution to this problem. Indeed, by cascading an analog (RF) precoder after the baseband digital precoder, it is possible to reduce the number of required RF chains. In [16], the first attempt was made to realize any FDP by means of the HBF structure with two RF chains for the case of a single data stream, i.e., one symbol per transmission. In [22, 23] multi-stream hybrid designs were presented in which the number of RF chains must be set equal to the number of symbols per transmission.

The HBF has an intricate structure where the entries of the RF precoder matrix satisfy constant modulus constraint (i.e., phase-shifters). Since the ensuing precoder optimization is non-convex, many works have alternatively focused on designing HBF directly using heuristic iterative algorithms or reconstruction algorithms [28, 75]. In recent years, several optimization techniques have been introduced to directly design the analog and digital beamforming matrices of an HBF system [17, 25, 26].

Since for any HBF, there exists an equivalent FDP, HBF cannot improve FDP. Moreover, FDPs have been studied for a long time and optimal designs for various scenarios already exist. In this chapter, motivated by these considerations, we focus on realizing any given FDP with the HBF architecture while minimizing the required number of RF chains. To this end, we first introduce a generalized system model where the baseband digital precoder is viewed as an abstract transformation. Then, we formulate a constrained optimization problem to minimize the number of required RF chains such that the FDP and HBF outputs are equal. We show that feasible optimum solutions exist for only two RF chains; however, optimal solutions are not unique, which allows for freedom in design. Based on the optimum solution so obtained, we then propose a simple yet novel HBF design for the realization of

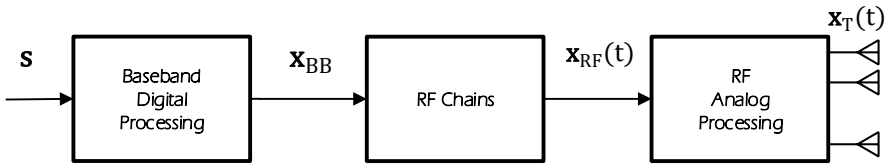


Fig. 3.1: Block diagram of the transmitter modules.

any given FDP with single RF chain architecture (SRCA), which does not require extra hardware or computational complexity in its implementation. We further propose three HBF schemes which facilitate the practical implementation of the SRCA concept. While these designs can be employed for the realization of various FD techniques, our discussions focus on the optimal precoding for single-user and multi-user MIMO scenarios. Finally, we compare the proposed HBF design to recent works that require a larger number of RF chains and show that significant performance improvement can be achieved by realizing FDP with the proposed HBF architecture.

3.2 System Model

Consider a generic massive-MIMO transmitter with N_T antennas and N_{RF} RF chains. Within each transmission (time slot), there are K users to be served and for each user, D symbols are transmitted.

Let $\mathbf{s}_i = [s_{i,1}, s_{i,2}, ..., s_{i,D}]^T$ be the symbol vector of the ith user, where $s_{i,j}$ is taken from a discrete constellation $\mathcal{A}$ (such as M-QAM or M-PSK). Hence, the symbol vector $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, ..., \mathbf{s}_K^T]^T \in \mathcal{A}^{N_s}$, where $N_s = DK$, must be precoded and then transmitted. As shown in Fig. 3.1 which depicts the block diagram of the transmitter, vector $\mathbf{s}$ is first processed by a baseband digital precoder whose output vector is denoted as $\mathbf{x}_{\mathrm{BB}} \in \mathbb{C}^{N_{RF}}$. Then, N_{RF} parallel RF chains are used to convert the baseband digital vector signal $\mathbf{x}_{\mathrm{BB}}$ into the bandpass modulated RF vector signal $\mathbf{x}_{\mathrm{RF}}(t)$. The latter is next fed to the analog precoder which consists of relatively simple RF circuitry. The resulting output signal $\mathbf{x}_{\mathrm{T}}(t)$ is then transmitted by means of an antenna array. Since the notation of analog signals adds unnecessary complication, in the sequel we let $\mathbf{x}_{\mathrm{RF}} \in \mathbb{C}^{N_{RF}}$ and $\mathbf{x}_{\mathrm{T}} \in \mathbb{C}^{N_T}$ represent the baseband sampled version of the analog signals $\mathbf{x}_{\mathrm{RF}}(t)$ and $\mathbf{x}_{\mathrm{T}}(t)$, respectively, and drop the

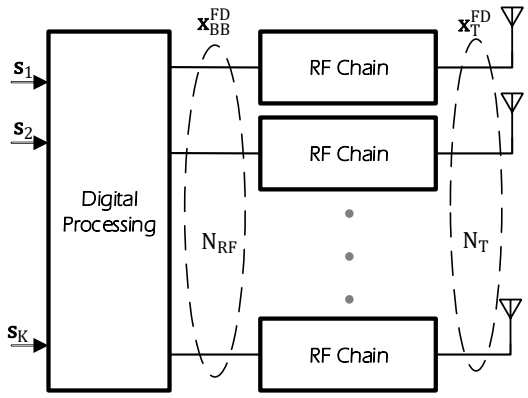


Fig. 3.2: Massive-MIMO transmitter with FD architecture.

time dependence without loss of generality (see also [21]). This mathematical abstraction allows us to identify the baseband vector $\mathbf{x}_{\mathrm{BB}}$ with its RF counterpart $\mathbf{x}_{\mathrm{RF}}$, as they contain the same information.

3.3 Generalized Precoder Architecture

In this section, first the conventional baseband FDP system and then the generalized architecture of HBF are presented.

Fig. 3.2 illustrates a conventional FDP where each antenna element is directly connected to an RF chain, so that $N_{RF} = N_T$. Thus, the transmitted signal of each antenna element can be processed and controlled by the digital baseband processor. Hence, RF precoding is not required and we can write: $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}} = \mathbf{x}_{\mathrm{RF}}^{\mathrm{FD}} = \mathbf{x}_{\mathrm{BB}}^{\mathrm{FD}} \in \mathbb{C}^{N_T}$.

Letting $\mathbf{P}_{\mathrm{FD}} \in \mathbb{C}^{N_T \times N_s}$ denote the digital precoder matrix, the output of the FDP can be written as [21],

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}} = \mathbf{P}_{\mathrm{FD}}\mathbf{s}. \tag{3.1}$$

Fig. 3.3 illustrates a massive-MIMO transmitter with an HBF implementation. In this

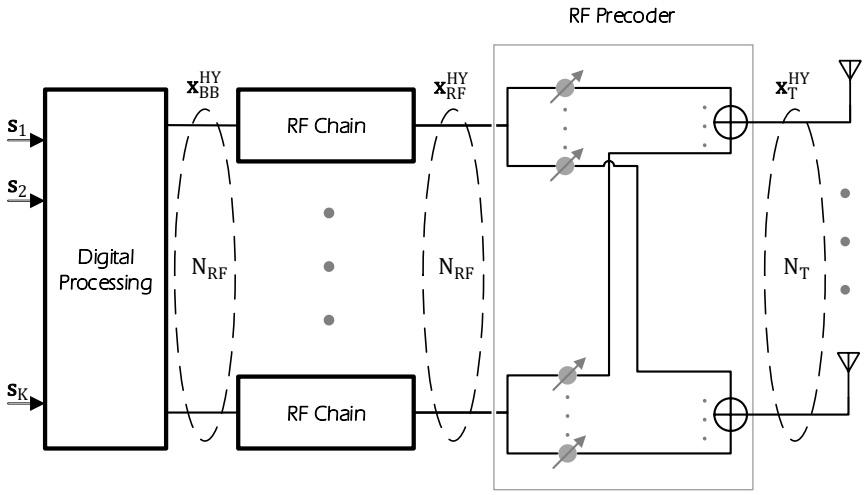


Fig. 3.3: Massive-MIMO transmitter with hybrid architecture.

case, due to inherent feature of an HBF design, we have that $N_{RF} \ll N_T$. In this work, we consider a general framework in which the output of the digital baseband processor can be written as,

$$\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = \mathcal{P}(\mathbf{s}) \in \mathbb{C}^{N_{RF}}. \tag{3.2}$$

where $\mathcal{P} : \mathcal{A}^{N_s} \rightarrow \mathbb{C}^{N_{RF}}$ is a mapping from the given symbol vector $\mathbf{s}$ to desired output $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$. This is in contrast to the previous works in the literature, e.g. [16, 21–23, 57], where the digital precoder is limited to a linear transformation, i.e.,

$$\mathcal{P}(\mathbf{s}) = \mathbf{P}_{\mathrm{D}}\mathbf{s}, \tag{3.3}$$

where $\mathbf{P}_{\mathrm{D}} \in \mathbb{C}^{N_{RF} \times N_s}$ is a digital precoder matrix with $N_{RF} \geq N_s$.

Comparing (3.1) and (3.3), we note an important difference between matrices $\mathbf{P}_{\mathrm{FD}}$ and $\mathbf{P}_{\mathrm{D}}$: while the former represents the $N_T \times N_s$ precoder for FDP where $N_{RF} = N_T$, the latter represents the $N_{RF} \times N_s$ digital precoder for HBF with $N_s \leq N_{RF} \ll N_T$. The output of the analog processing module can be written as,

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{A}}\mathbf{x}_{\mathrm{RF}}^{\mathrm{HY}}, \tag{3.4}$$

where matrix $\mathbf{P}_{\mathrm{A}} \in \mathbb{C}^{N_T \times N_{RF}}$ represents a linear transformation implemented via RF circuitry. Therefore, the overall input-output relationship of the system is

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{A}} \mathcal{P}(\mathbf{s}). \tag{3.5}$$

It can be observed from Fig. 3.3 that in the adopted baseband notation, $\mathbf{x}_{\mathrm{RF}}^{\mathrm{HY}} = \mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} \in \mathbb{C}^{N_{RF}}$ since the digital and analog modules are connected via N_{RF} dedicated RF chains. The analog precoding is performed in the RF domain via analog circuitry. Specifically, the analog precoder for mmWave massive-MIMO communications only consists of phase shifters and adders.

In the HSP literature [16,21–23,57], entries of analog precoder matrix $\mathbf{P}_{\mathrm{A}}$ are constrained to have equal magnitude, where in [57] further practical power constraints are discussed. For simplicity, we assume that these entries all have unit norm. However, for any other given magnitude constraint $\kappa > 0$, as dictated by practical RF considerations, the RF precoder matrix $\mathbf{P}_{\mathrm{A}}$ and baseband digital output vector $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ designed with the proposed approach can be scaled by κ and $1/\kappa$ respectively, while maintaining the same overall analog output. Thus, without loss of generality, we model the analog precoder as a linear transformation, represented by matrix $\mathbf{P}_{\mathrm{A}} \in \mathbb{U}^{N_T \times N_{RF}}$ where $\mathbb{U}$ is the set of complex number with unit norm:

$$\mathbb{U} = \{z \in \mathbb{C} \colon |z| = 1\}. \tag{3.6}$$

Using this information, the output of the analog precoder can be expressed as,

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{A}} \mathbf{x}_{\mathrm{RF}}^{\mathrm{HY}}. \tag{3.7}$$

Hence, from (3.2) and (3.7), the output of the generalized HBF can be written as,

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{A}} \mathcal{P}(\mathbf{s}). \tag{3.8}$$

From (3.3) and (3.8), we observe that the conventional hybrid structure in [16, 21–23, 57] is a special case of (3.8) and has the following form:

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{A}} \mathbf{P}_{\mathrm{D}}\, \mathbf{s}. \tag{3.9}$$

3.4 Realizing FDP in HBF architecture

In the recent literature, based on the model in (3.9), the HBF design problem is reduced to the design of the matrix factors $\mathbf{P}_{\mathrm{A}}$ and $\mathbf{P}_{\mathrm{D}}$. Consequently, for a given symbol vector $\mathbf{s}$, the effective precoder determines the output of the HBF:

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{eff}}\,\mathbf{s} = \mathbf{P}_{\mathrm{A}}\mathbf{P}_{\mathrm{D}}\,\mathbf{s}. \tag{3.10}$$

Hence, designing $\mathbf{P}_{\mathrm{A}}$ and $\mathbf{P}_{\mathrm{D}}$ such that $\mathbf{P}_{\mathrm{eff}} = \mathbf{P}_{\mathrm{FD}}$, where $\mathbf{P}_{\mathrm{FD}}$ is the given FDP, under the constraint $N_{RF} \ll N_T$ is the main goal in [16, 22, 23]. Extending the design of [16] to multi-stream results in utilizing $N_{RF} = 2N_s$ RF chains for realizing any FD precoder in hybrid structure. In [23] and [22], realizations of any FDP in hybrid architecture with $N_{RF} = N_s$ RF chains are proposed.

Here, we investigate the minimum number of required RF chains for realizing any FDP in hybrid structure. To this end, we use the generalized digital baseband precoding model (3.2) and seek to realize the underlying given FDP by means of the proposed generalized HBF structure, that is,

$$\mathbf{P}_{\mathrm{A}}\mathcal{P}(\mathbf{s}) = \mathbf{P}_{\mathrm{FD}}\,\mathbf{s}. \tag{3.11}$$

Accordingly, we formulate the following optimization problem for minimizing the number of RF chains while realizing any arbitrary FDP $\mathbf{P}_{\mathrm{FD}}$ with HBF:

$$\min_{\mathbf{P}_{\mathrm{A}},\mathcal{P}(.)} N_{RF} \tag{3.12a}$$

$$\text{subject to} \quad \mathbf{P}_{\mathrm{A}}\mathcal{P}(\mathbf{s}) = \mathbf{P}_{\mathrm{FD}}\,\mathbf{s} \tag{3.12b}$$

$$\mathbf{P}_{\mathrm{A}} \in \mathbb{U}^{N_T \times N_{RF}} \tag{3.12c}$$

This optimization is clearly not convex as $\mathbb{U}$ is a non-convex set as shown in (3.6); moreover, the generality of $\mathcal{P}$ makes it difficult to solve this problem. To overcome this limitation, we first replace $\mathcal{P}(\mathbf{s})$ by the auxiliary signal $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ in (3.2) and then solve the problem for $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ for the given $\mathbf{s}$. Subsequently, we shall seek $\mathcal{P}$ such that $\mathcal{P}(\mathbf{s}) = \mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$. Hence, we have,

$$\min_{\mathbf{P}_{\mathrm{A}}, \mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}} N_{RF} \tag{3.13a}$$

$$\text{subject to} \quad \mathbf{P}_{\mathrm{A}} \mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{FD}}\, \mathbf{s} \tag{3.13b}$$

$$\mathbf{P}_{\mathrm{A}} \in \mathbb{U}^{N_T \times N_{RF}} \tag{3.13c}$$

which means minimizing N_{RF} while finding RF precoder $\mathbf{P}_{\mathrm{A}}$ and vector $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ (3.2) such that $\mathbf{P}_{\mathrm{A}}\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{FD}}\mathbf{s}$. The existence of a solution to this problem is asserted by Theorem 1 below, whose proof relies on the following lemma [16, 22, 23].

Lemma 3.1. *For positive real numbers β_1 and β_2, any complex number z where $|\beta_1 - \beta_2| \leq |z| \leq |\beta_1 + \beta_2|$ can be written as: $z = \beta_1 e^{j\theta_1} + \beta_2 e^{j\theta_2}$ where $\theta_1, \theta_2 \in [0, 2\pi]$ and θ_1, θ_2 might be non-unique.*

Proof. Different proofs have been presented in [16, 22, 23]. Here we present an alternative proof. We first show that for positive real numbers β_1 and β_2, the set

$$\{w = \beta_1 e^{\theta_1} + \beta_2 e^{\theta_2} \mid \ \theta_1, \theta_2 \in [0, 2\pi)\} \tag{3.14}$$

spans an annulus in the complex plane delimited by circles with radii $r_1 = |\beta_1 - \beta_2|$ and $r_2 = |\beta_1 + \beta_2|$ as illustrated in Fig. 3.4. With simple manipulations we can write w in polar form, i.e., $w = re^{j\theta}$ where,

$$r = \sqrt{\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2 \cos(\theta_1 + \theta_2)} \tag{3.15a}$$

$$\theta = \text{angle}(\beta_1 \sin\theta_1 + \beta_2 \sin\theta_2, \beta_1 \cos\theta_1 + \beta_2 \cos\theta_2). \tag{3.15b}$$

For fixed β_1 and β_2, it can be observed from (3.15a) that by varying θ_1 and θ_2 from 0 to 2π, values of r will cover the range $[r_1, r_2]$ while θ will cover the range $[0, 2pi)$. Hence the transformation in (3.14) establishes a subjective map from $[0, 2\pi]^2$ onto the above annulus, as illustrated by the shaded area in Fig. 3.4. Hence, any point in the complex plane with absolute value between $|\beta_1 - \beta_2|$ and $|\beta_1 + \beta_2|$ can be presentation as: $z = \beta_1 e^{\theta_1} + \beta_2 e^{\theta_2}$, although the representation is in general not unique. □

From Lemma 3.1, we can see that for any $z \in \left[|\beta_1 - \beta_2|, |\beta_1 + \beta_2|\right]$, θ_1, θ_2 are not unique

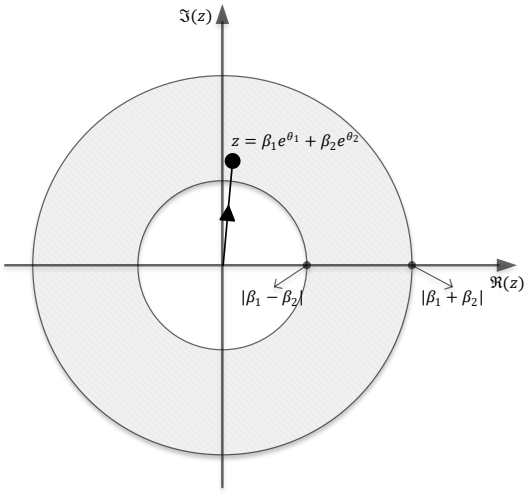


Fig. 3.4: Annulus in complex plane

and different solutions have been presented in [16,22,23]. Nevertheless, it means that there exist degrees of freedom for designing $\mathbf{P}_{\mathrm{A}}$ and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$.

Theorem 3.1. *There exist non-unique solutions for* $\mathbf{P}_A$ *and* $\mathbf{x}_{BB}^{HY}$ *which minimize* (3.13) *with optimum objective function value* $N_{RF} = 2$*.*

Proof. Since $N_{RF} \in \{1, 2, 3, \dots\}$ is a natural number, in order to solve (3.13), we can take the simple approach of starting from the least possible value for N_{RF} and show whether the constraints can be satisfied or not; if not, we move on to the next number and so on until the constraints are satisfied. As $\mathbf{P}_{\mathrm{FD}}$ and $\mathbf{s}$ are given, the right hand side of (3.13b), i.e., $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}} = \mathbf{P}_{\mathrm{FD}}\mathbf{s}$ in (3.1), is an arbitrary complex vector of size N_T.

$N_{RF} = 1$**:**

In this case, $\mathbf{P}_{\mathrm{A}}$ is a vector of size N_T in $\mathbb{U}^{N_T \times 1}$ and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ is a complex scalar. Since $\mathbf{P}_{\mathrm{A}}$ has unit modulus entries, $\mathbf{P}_{\mathrm{A}}\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ would be a vector of size N_T where all the entries have the same magnitude. Therefore, because $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}}$ is an arbitrary complex vector, $\mathbf{P}_{\mathrm{A}}\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = \mathbf{x}_{\mathrm{T}}^{\mathrm{FD}}$ is not generally satisfied. Thus, $N_{RF} = 1$ is not the minimum of (3.13) under the given constraints.

$N_{RF} = 2$**:**

Let $x_i = |x_i|e^{j\vartheta_i}$ denote the polar representation of the ith entry of the vector $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}}$. Define $|x|_{min}$ and $|x|_{max}$ as the minimum and maximum values of $|x_i|$ for $i = 1, \ldots, N_T$, respectively. In the case $N_{RF} = 2$, $\mathbf{P}_{\mathrm{A}}$ and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ can be explicitly written as,

$$\mathbf{P}_{\mathrm{A}} = \begin{bmatrix} e^{j\phi_{1,1}} & e^{j\phi_{2,1}} & \ldots & e^{j\phi_{N_T,1}} \\ e^{j\phi_{1,2}} & e^{j\phi_{2,2}} & \ldots & e^{j\phi_{N_T,2}} \end{bmatrix}^T \tag{3.16a}$$

$$\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix}^T. \tag{3.16b}$$

From Lemma 3.1, there exist non-unique angles $\phi_{i,1}$ and $\phi_{i,2}$ such that,

$$x_i = \alpha_1 e^{j\phi_{i,1}} + \alpha_2 e^{j\phi_{i,1}}, \tag{3.17}$$

where α_1 and α_2 are positive real numbers satisfying,

$$|\alpha_1 - \alpha_2| \leq |x|_{min} \quad \text{and} \quad |x|_{max} \leq |\alpha_1 + \alpha_2|. \tag{3.18}$$

Thus, in the case $N_{RF} = 2$, we have shown that for any given vector $\mathbf{s}$ and FDP matrix $\mathbf{P}_{\mathrm{FD}}$, there exist optimal solutions for $\mathbf{P}_{\mathrm{A}}$ (3.16a) and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ (3.16b) such that the constraint $\mathbf{P}_{\mathrm{A}}\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{FD}}\,\mathbf{s}$ is satisfied. □

Theorem 3.1 indicates that in order to realize an arbitrary FDP with an HBF architecture, each antenna requires only two phase shifters and therefore only two RF chains are needed to provide $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = [\alpha_1, \alpha_2]^T$ to the two sets of phase shifters. That is α_1 is connected to the phase shifters of $\mathbf{P}_{\mathrm{A}}$'s first column and α_2 to the second column. From Lemma 3.1, however, we can see that the optimal solutions for $\mathbf{P}_{\mathrm{A}}$ and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ are not unique and there exists degrees of freedom for designing $\mathbf{P}_{\mathrm{A}}$ and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$. In particular, if there exists an optimal solution where $\alpha_1 = \alpha_2 = \alpha$, i.e., $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} = [\alpha, \alpha]^T$, then, only one RF chain is sufficient to realize any FDP in the HBF system, as depicted in Fig. 3.5. In the following theorem, an optimal solution of (3.13) is presented under the constraint $\alpha_1 = \alpha_2$ which leads to a novel HBF design with a single RF chain.

Theorem 3.2. *Any FDP can be realized by a single RF chain HBF as shown in Fig. 3.5*

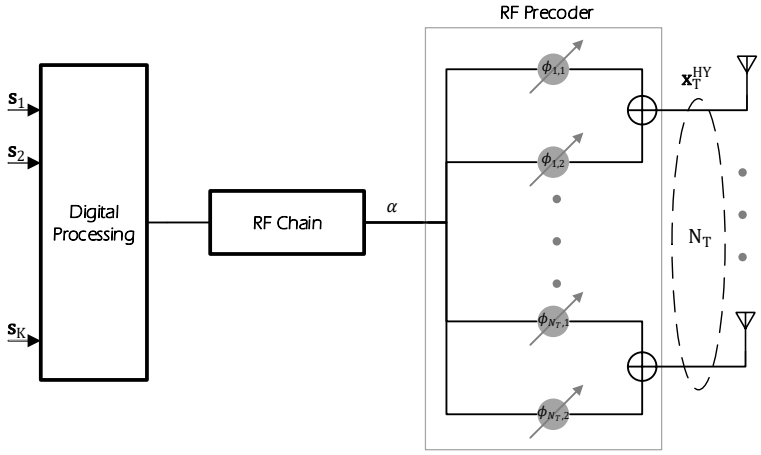


Fig. 3.5: HBF with single RF chain.

with the following parameters:

$$\alpha \geq \frac{1}{2}|x|_{max} \tag{3.19a}$$

$$\phi_{i,1} = \vartheta_i - \cos^{-1}\left(\frac{|x_i|}{2\alpha}\right) \tag{3.19b}$$

$$\phi_{i,2} = \vartheta_i + \cos^{-1}\left(\frac{|x_i|}{2\alpha}\right). \tag{3.19c}$$

Proof. For $\alpha \geq \frac{1}{2}|x|_{max}$, setting $\alpha_1 = \alpha_2 = \alpha$ satisfies the necessary condition (3.18) because $|x|_{min} \geq |\alpha_1 - \alpha_2| = 0$ and $|x|_{max} \leq |\alpha_1 + \alpha_2| = 2\alpha$. Also, since $0 \leq |x|_{max} \leq 2\alpha$, it follows that the argument of the $\cos^{-1}$ functions in (3.19b) and (3.19c) satisfies $0 \leq \frac{|x_i|}{2\alpha} \leq 1$. Under this condition, using the principal value of $\cos^{-1}$ in $[0, \pi/2]$, it can be verified through algebraic manipulations that the representation (3.17) is satisfied for every x_i, with $i = 1, \ldots, N_T$. □

From Theorem 3.2, one can infer that any complex number x can be written as: $x = \alpha(e^{j\phi_1} + e^{j\phi_2})$ for some $\alpha \geq \frac{1}{2}|x|$, which provides a generalization of Theorem 1 in [23].

The only problem left for realizing any FDP in HBF architecture with minimum N_{RF} is to find a mapping $\mathcal{P}$ such that $\mathcal{P}(\mathbf{s}) = \mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ which at first, might seem to be a challenging task. However, since our design requires only one RF chain and, more importantly, the

Algorithm 1 Realizing any FDP in HBF

Given: $\mathbf{s}$, $\mathbf{P}_{\mathrm{FD}}$
Output: $\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}$
1. Calculate the desired transmit signal by: $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}} = \mathbf{P}_{\mathrm{FD}}\,\mathbf{s}$.
2. Choose α as (3.19a).
3. Set $\mathbf{P}_{\mathrm{A}}$ in (3.16a) using (3.19b) and (3.19c).
4. Feed α to the RF chain.
Note: The output of the HBF, i.e., $\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}$, is identical to $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}}$.

output of the baseband digital processing module is already determined, i.e., the positive number α, this problem can be solved by taking advantage of the digital processing. In other words, since we know the desired output of $\mathcal{P}(\mathbf{s})$, it is not necessary to implement the mapping $\mathcal{P}(.)$ and then apply it on $\mathbf{s}$ to obtain $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ which is available from the solution of (3.13). In effect, we simply define $\mathcal{P}(.)$ as

$$\mathcal{P} : \mathbf{s} \mapsto \mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}} \tag{3.20}$$

Algorithm 1 summarizes our proposed SRCA design for realizing any FDP in HBF with a single RF chain, i.e., SRCA. Note that, since $\mathbf{s}$ and $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ are given, the function $\mathcal{P}$ is well defined with input of $\mathbf{s}$ and corresponding image of $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$. Now, one can see that generalizing $\mathbf{P}_{\mathrm{D}}$ to $\mathcal{P}(.)$ did not result in more complicated transformation; actually, $\mathcal{P}(.)$ is simply represents the produce desired $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$ which is required in (3.13).

3.4.1 Implementation Aspects

Except for its reduced number of RF chains, the proposed scheme is characterized by an HBF architecture similar to [16,21–23,57,76,77]. However, certain important considerations need to be taken into account for its practical implementation as explained below.

RF precoder update rate: In a conventional HBF system, to allow for accurate tracking of the wireless channel conditions, the RF precoder coefficients are updated according to the channel coherence time, T_c. In the proposed scheme however, the RF precoder coefficients are also affected by the transmit symbol vector $\mathbf{s}$ and as such, they must be updated according to the symbol duration T_s. Since $T_s < T_c$ in slowly-varying channels, this means that in this case, the RF precoder needs be updated at a higher rate, i.e., by a factor T_c/T_s.

However, recent advances in RF semiconductor and mmWave hardware [76,77] suggest that it will be possible to accommodate such requirements in near future. Alternatively, at the price of increasing latency, the update rate of the RF precoder can be reduced by increasing the symbol duration T_s, i.e. the length of the symbol vector.

Computational complexity: The proposed HBF design does not introduce any major additional complexity. Compared to conventional HBF methods, only the digital baseband processing is modified and extended, while the RF processing has been much simplified through reduction of the number of RF chains. There is also no major computational burden introduced to the digital processing module. For a chosen FDP method, our proposed design provides its HBF realization in two phases. Every T_c, the desired FDP matrix must be calculated according to the selected algorithm, which is also the case for other HBF methods [22, 23]. Since FDP techniques usually admit closed form solutions, less computational complexity is required compared with iterative HBF designs. For instance, the optimal eigenmode transmission FDP can be obtained via singular value decomposition and water filling power allocation, whereas the direct HBF designs in [21, 22] are based on iterative algorithms where complex calculations are performed in each iteration, while the required number of iterations itself varies. Having designed the FDP matrix, within each T_s, a new symbol vector must be precoded. To this end, the HBF designs in [21–23] require calculations with complexity order of $\mathcal{O}(N_{RF}N_s)$. In our design, step 1 of the algorithm is a simple matrix multiplication with complexity order $\mathcal{O}(N_T N_s)$ while from (3.16) and (3.19), the second and thirds steps only require $\mathcal{O}(N_T)$ complex operations or function evaluations. Thus the overall complexity order is dominated by the first step. However, this only represents a small fraction of the total computations needed for the FDP calculation or HBF design in iterative approaches [21, 22]. The calculation of the FDP matrix $\mathbf{P}_{\mathrm{FD}}$ by water filling requires singular value decomposition whereas the hybrid designs in [21, 22] are based on iterative algorithms which have much higher complexity order required for FDP calculation.

Digital mapping $\mathcal{P}$: Although the digital processing is extended to the mapping $\mathcal{P}(\mathbf{s})$ in (3.2), as opposed to the linear transformation $\mathbf{P}_{\mathrm{D}}\mathbf{s}$, the realization of the proposed HBF scheme is similar to the conventional HBF ones because we neither need to implement the mapping nor calculate it. As shown in Algorithm 1, first the desired precoded signal

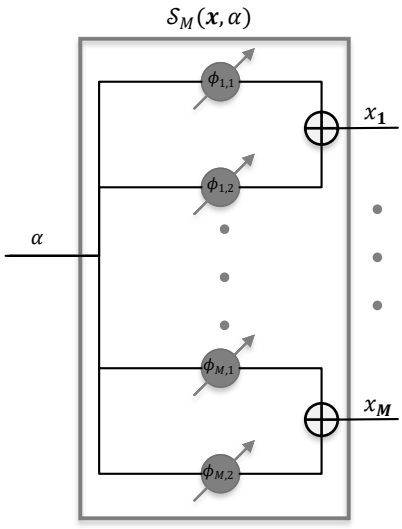


Fig. 3.6: Block diagram of the $\mathcal{S}_M(\boldsymbol{x},\alpha)$ module.

$\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}}$ is computed, and then α is determined and assigned as the output of the digital domain processor to drive the RF precoder. In fact, since the desired output of $\mathcal{P}(\mathbf{s})$ is already calculated, i.e., $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$, we do not need to implement the nonlinear mapping $\mathcal{P}$ nor apply it on $\mathbf{s}$ to obtain $\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}$. In other words, there is no need for implementation of non-linear mapping.

Number of served users: Since increasing the number of RF chains may not be feasible in an already deployed system, the number of precoded symbols per transmission in conventional HBF-based systems [16, 21–23, 57] is limited by N_{RF} (i.e., $N_s \leq N_{RF}$), which makes a direct migration of such system to massive-MIMO unlikely. In contrast, migration of a system based on the proposed scheme does not require any change on the RF chain; only the RF precoder needs to be expanded according to N_T since the proposed HBF does not limit the number of precoded symbols per transmission.

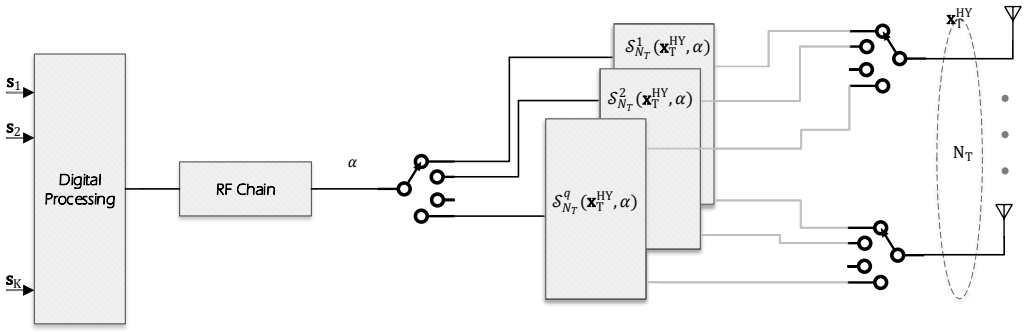


Fig. 3.7: SRCA with phase-shifter bank.

3.5 HBF Schemes Based on SRCA

The SRCA developed in the previous section enables us to generate any desired signal vector $\boldsymbol{x}$ of size M in the RF domain with one RF chain. Let us first introduce $\mathcal{S}_M(\boldsymbol{x}, \alpha)$ as a primary building block, shown in Fig. 3.6, which can be configured to generate a given signal $\boldsymbol{x}$ in this way. In what follows, three additional HBF schemes are presented based on the $\mathcal{S}_M(\boldsymbol{x}, \alpha)$ module. For the single user case, we further let $D = 1$, i,e,. $N_s = K$ for simplicity.

3.5.1 Phase-Shifter Bank

As discussed earlier, direct implementation of SRCA for realizing FDP realization requires analog precoder to be updated at each symbol period T_s. In certain systems however, meeting this requirement may be difficult; hence, we present the following alternative design based on phase-shifter banks. Assuming the minimum update period of the chosen phase-shifter is T_p, this architecture requires a phase-shifter bank of size $q \geq \lceil \frac{T_p}{T_s} \rceil$, as shown in Fig. 3.7. Specifically, the output of the RF chain is connected to an analog switch (multiplexer) which selects each of the analog precoder in turn. In the same way, each antenna is connected to an analog switch which selects the active analog precoder. The switches are synchronized, i.e., operate simultaneously, and are controlled by the digital processor. Assuming the q consecutive symbol vectors are precoded to obtain the desired output signals $\mathbf{x}_T^1, \mathbf{x}_T^2, ..., \mathbf{x}_T^q$, each vector $\mathbf{x}_T^i$ for $i = \{1, 2, ..., q\}$ is then used in turn to configure $\mathcal{S}_M^i(\boldsymbol{x}, \alpha)$ based on the SRCA and the multiplexers are then switched to state i to

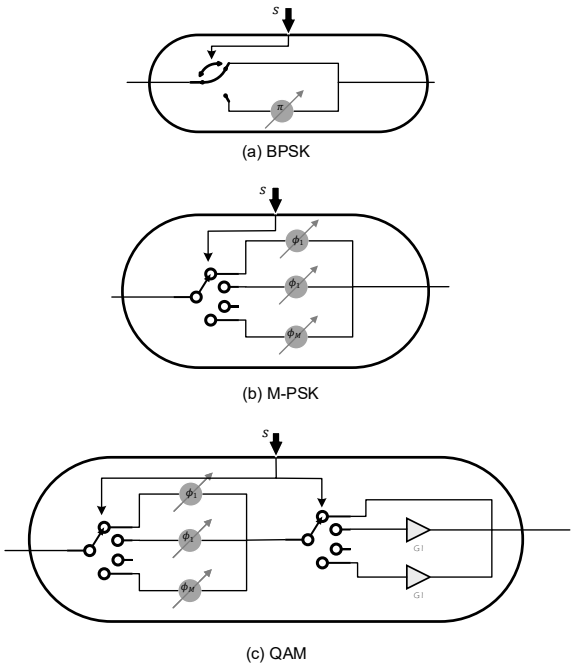


Fig. 3.8: AC block for different constellations.

transmit the precoded signal. The step-by-step FDP realization by the phase-shifter bank scheme can be summarized as follows:

- Calculate the desired transmit signal by: $\mathbf{x}_{\mathrm{T}}^{\mathrm{FD}} = \mathbf{P}_{\mathrm{FD}}\, \mathbf{s}^{i}$.
- Choose α as in (3.19a).
- Set $\mathcal{S}_{N_T}^{i}(\boldsymbol{x}, \alpha)$ using (3.19).
- Feed α to the RF chain.
- Set all switches to the ith position

Note that this scheme does not require the use of additional memory or delay as each symbol vector is precoded and transmitted immediately. Furthermore, since there are $q \geq \lceil \frac{T_p}{T_s} \rceil$ RF blocks available, adequate time is available to update the ith block when it needs to be reconfigured.

3.5.2 Analog Constellation

To realize a given FDP in HBF, the output of the hybrid beamformer must be equal to the output of the FDP, i.e.,

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{FD}}\mathbf{s}, \tag{3.21}$$

where $\mathbf{s} = [s_1, s_2, ..., s_{N_s}]^T$. Let us express the FDP matrix as,

$$\mathbf{P}_{\mathrm{FD}} = [\mathbf{p}^1, \mathbf{p}^2, ..., \mathbf{p}^{N_S}], \tag{3.22}$$

where $\mathbf{p}^j$ denote the j-th column of the precoder matrix. By vectorizing $\mathbf{P}_{\mathrm{FD}}$ as

$$\mathbf{p} = [\mathbf{p}^{1^T}, \mathbf{p}^{2^T}, ..., \mathbf{p}^{{N_S}^T}]^T, \tag{3.23}$$

and defining

$$\mathbf{S} = [s_1\mathbf{I}_{N_T}, s_2\mathbf{I}_{N_T}, ..., s_{N_s}\mathbf{I}_{N_T}] \in \mathbb{C}^{N_T \times N_T N_s}, \tag{3.24}$$

(3.21) can be also written as

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{S}\mathbf{p}. \tag{3.25}$$

Using SRCA, vector $\mathbf{p}$ can be generated by one RF chain and matrix $\mathbf{S}$ can be implemented by analog constellation (AC) blocks as illustrated in Fig. 3.8. The AC blocks are designed based on the constellation (such as QAM or PSK) and are comprised of a switch and fixed phase-shifters and gains, as shown in Fig. 3.8. For instance, it can be observed that BPSK only requires a single switch and a phase-shifter which is just a negative sign. Step-by-step FDP realization by means of AC can be summarized as follows:

- Construct $\mathbf{S}$ from (3.24).
- Update all AC modules according to $\mathbf{S}$.
- Calculate $\mathbf{P}_{\mathrm{FD}}$ and construct $\mathbf{p}$ from (3.23).
- Set $\mathcal{S}_M(\mathbf{p}, \alpha)$ using (3.19).

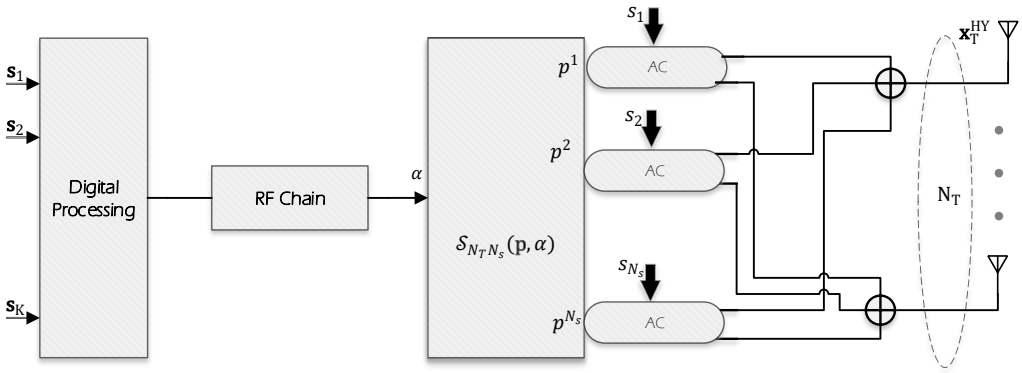


Fig. 3.9: SRCA with analog constellation.

3.5.3 Hybrid Beamformer with RF Multiplier

Using RF multipliers, we present a new architecture for HBF which relaxes the unit modulus constraints of the analog precoder. RF multipliers are not useful in conventional FDP and even HBF designs because the implementation of conventional FDP with RF multipliers requires more RF chains than baseband FDP, i.e., $(N_T+1)K$ RF chains. However, using RF multipliers in SRCA simplifies the RF precoder design. We introduce a general architecture which can be used for designing various hybrid signal processing systems. This technique also relaxes the unit modulus constraint of analog precoders. Fig. 3.10 illustrates the system architecture for this technique.

The output of the system is,

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}} = \mathbf{P}_{\mathrm{HY}} \mathbf{x}_{\mathrm{RF}}^{\mathrm{HY}} \tag{3.26}$$

where $\mathbf{x}_{\mathrm{RF}}^{\mathrm{HY}}$ is the output of RF chains and,

$$\mathbf{P}_{\mathrm{HY}} = [\mathbf{p}_{\mathrm{HY}}^1, \mathbf{p}_{\mathrm{HY}}^2, ..., \mathbf{p}_{\mathrm{HY}}^{N_S}] \in \mathbb{C}^{N_T \times N_{RF}}, \tag{3.27}$$

is the new analog precoder. By vectorizing the matrix $\mathbf{P}_{HY}$, we have,

$$\mathbf{p}_{\mathrm{HY}} = [{\mathbf{p}_{\mathrm{HY}}^1}^T, {\mathbf{p}_{\mathrm{HY}}^2}^T, ..., {\mathbf{p}_{\mathrm{HY}}^{N_s}}^T]^T, \tag{3.28}$$

which can be generated using SRCA. Thus we can set $\mathbf{P}_{\mathrm{HY}} = \mathbf{P}_{\mathrm{FD}}$ and $\mathbf{x}_{\mathrm{RF}}^{\mathrm{HY}} = \mathbf{s}$. However,

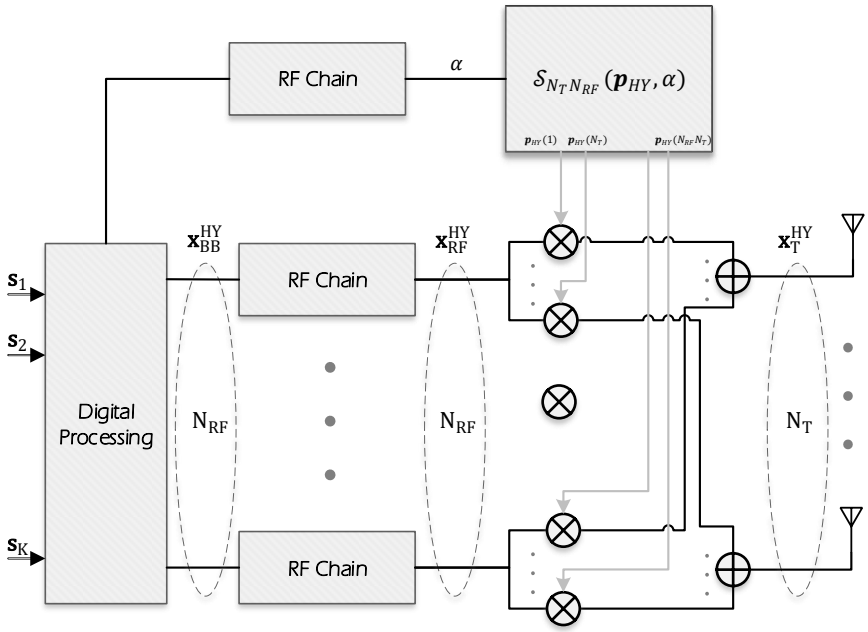


Fig. 3.10: SRCA with RF multipliers.

more sophisticated optimizations and/or decompositions can be used to minimize the number of RF chains. Step-by-step FDP realization SRCA scheme with RF multipliers can be summarized as follows:

- Feed symbol vector $\mathbf{s}$ to RF chains number 1 to K.
- Calculate $\mathbf{P}_{\mathrm{FD}}$ and construct $\mathbf{P}_{HY}$ from (3.28).
- Setup $\mathcal{S}_{N_T}(\mathbf{P}_{HY}, \alpha)$ using (3.19).
- Feed α to RF chain number $K+1$ to drive $\mathcal{S}_{N_T}(\mathbf{P}_{HY}, \alpha)$

Comparison of the proposed SRCA techniques: The conventional hybrid architecture was particularly designed for beamforming and cannot generally achieve the performance of FD systems. However, the proposed SRCA techniques are purposefully designed to generate any desired signal by the means of the large scale antenna array. Therefore, the applications are not limited to beamforming and other wireless transmission and signal processing techniques can be implemented in hybrid architecture with reduced number of

RF chains, including FD channel estimation (CE), and space-time coding (STC). Table 3.1 compares the schemes proposed in this subsection, the FD system and the existing hybrid structures in terms of number of RF chain, phase-shifter update rate, and possible applications. While the proposed schemes achieve the performance of FD systems, the required number of RF chains is smaller than existing hybrid structures.

3.5.4 HBF Design for Single-User and Multi-User

In this section, we illustrate the application of the proposed SRCA schemes by considering specific design approaches for the FDP in both single-user and multi-user scenarios. Using these schemes, however, it is possible to realize any given FDP regardless of the particular design criteria being used for obtaining the corresponding precoder matrix.

Single-User: In what follows, we use spectral efficiency as a criterion to design the precoder. In order to maximize the spectral efficiency of a point-to-point MIMO link, the optimal FD precoder and combiner matrices can be obtained from,

$$\max_{\mathbf{W},\mathbf{P}_{su}} \quad \log_2(|\mathbf{I}_K| + \frac{\rho}{K}\mathbf{R}_n^{-1}\mathbf{W}^H\mathbf{H}_{su}\mathbf{P}_{su}\mathbf{P}_{su}^H\mathbf{H}^H\mathbf{W}) \tag{3.29a}$$

$$\text{subject to } Tr(\mathbf{P}_{su}\mathbf{P}_{su}^H) \leq P_T, \tag{3.29b}$$

where P_T is the power budget at the transmitter, $\mathbf{H}_{su}$ is the mmWave MIMO channel matrix, $\mathbf{P}_{su}$ and $\mathbf{W}$ are precoder and combiner matrices and $\mathbf{R}_n = \mathbf{W}^H\mathbf{W}$. Then, the optimal solution can be calculated [10] using the singular value decomposition of $\mathbf{H}_{su}$, i.e.,

$$\mathbf{H}_{su} = \mathbf{U}_{su}\mathbf{\Sigma}_{su}\mathbf{V}_{su}{}^H, \tag{3.30}$$

as,

$$\mathbf{P}_{su} = \mathbf{V}_{su}\mathbf{W}, \tag{3.31}$$

where the matrix $\mathbf{W}$ is obtained using water-filling technique [10].

Table 3.1: Comparison of different structures

-	N_{RF}	PS update	Applications
Phase-shifter bank	1	T_p	BF, STC, CE
Analog constellation	1	T_c	BF
SRCA with multiplier	2 to $K+1$	T_c	BF,STC,CE
Fully digital	N_T	-	BF,STC,CE
Existing hybrid designs	K to N_T	T_c	BF

Multi-User: In the case of single antenna MU, the optimal precoder is calculated using the ergodic sum-rate,

$$\max_{\mathbf{P}_{mu}} \sum_{k=1}^{K} \log_2(1+\mathrm{SINR}_k) \tag{3.32a}$$

$$\text{subject to } \mathrm{SINR}_k = \frac{|\mathbf{h}_k \mathbf{P}_{mu_k}|^2}{\sum_{m\neq k} |\mathbf{h}_k \mathbf{P}_{mu_k}|^2 + \frac{1}{\rho}}, \tag{3.32b}$$

where $\mathbf{P}_{mu_k}$ is the beamforming vector of the kth user and consequently, kth column of beamforming matrix $\mathbf{P}_{mu}$. Moreover, $\mathbf{h}_k$ is the channel vector between the kth user and the BS; thus, the channel matrix seen by the BS can be written as,

$$\mathbf{H}_{\mathrm{mu}} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_K]. \tag{3.33}$$

The zero-force (ZF) beamforming with optimal power allocation is given by

$$\mathbf{P}_{mu} = \mathbf{H}_{mu}(\mathbf{H}_{mu}\mathbf{H}_{mu}^H)^{-1}\mathbf{\Gamma}^{\frac{1}{2}}, \tag{3.34}$$

where Γ is the diagonal weight matrix obtained by water-filling.

Conventional HBF and SRCA HBF: In conventional hybrid structures, the following constraints are added to all of the above optimization problems:

$$\mathbf{P} = \mathbf{P}_{\mathrm{A}}\mathbf{P}_{\mathrm{D}} \tag{3.35a}$$

$$\mathbf{P}_{\mathrm{A}} \in \mathbb{U}^{N_T \times N_{RF}}, \tag{3.35b}$$

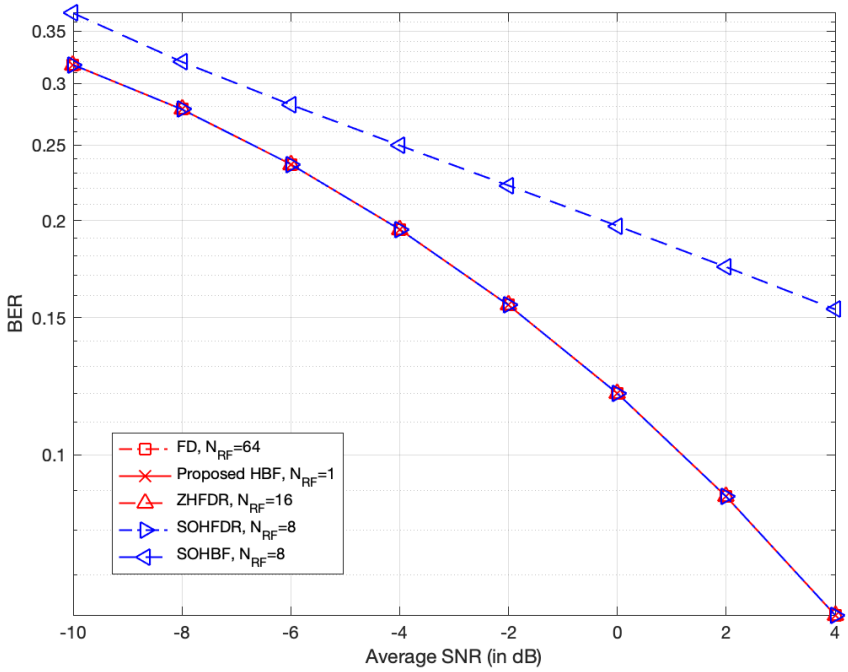


Fig. 3.11: BER versus SNR for different methods.

which makes the ensuing problems non-convex and therefore very difficult to solve. The proposed schemes, on the other hand, are all capable of realizing any given FD precoding scheme. For the optimal precoding, we first need to design the optimal precoders using (3.31) and (3.34) for SU and MU, respectively. Then, as discussed earlier, the proposed schemes can be accordingly configured to achieve the performance of FD systems. In the next section, we present computer simulations to illustrate the performance of the proposed schemes compared to recent HBF designs.

3.6 Simulation Results

In this section, simulation results are presented first for the SRCA developed in Section 3.4, and subsequently for the three SRCA-based schemes introduced in Section 3.5.

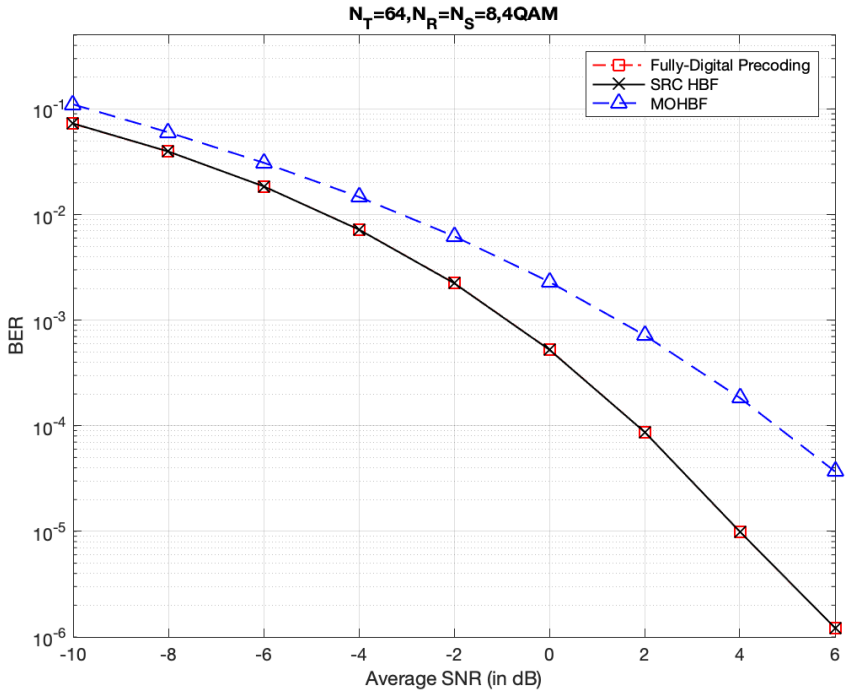


Fig. 3.12: BER versus SNR for FD precoding and our design in a 64×8 massive-MIMO system.

Single RF chain HBF: Since the proposed HBF design can realize any FDP, it exhibits a similar performance as other methods for realizing FDP in HBF, i.e., [16, 22, 23] but with only a smaller required number of RF chains. For example, while our design requires a single RF chain, the multi-stream extension of ZHFDR [16] needs $2N_s$ RF chains, and the hybrid designs in [22, 23] require N_s RF chains. Here, we present simulation results to compare the performance of the proposed HBF realization against the optimal baseband FDP, HBF realizations of this FDP based on [22] called SOFDR, and the direct HBF design in [22] called SOHBF. For the benchmark baseband FDP, the standard waterfilling based eigenmode transmission is used which achieves the optimal spectral efficiency. The direct HBF design in [22], however, separates the digital and analog precoder design in order to maximize the spectral efficiency, and utilizes an iterative optimization algorithm to find the optimal analog precoder.

In our simulations, we consider a similar set up as [22], that is: 64×8 massive-MIMO sys-

tem with an $L = 15$ paths directional vector channel model. At the receiver, fully digital decoding (combining) is used for all the methods. Fig. 3.11 shows the bit error rate (BER) performance versus signal-to-noise ratio (SNR) of the different methods for $N_s = 8$ symbols and 16-QAM modulation. It is observed that the optimal baseband FDP and consequently, the corresponding HBF realizations (including the proposed design) outperform the direct HBF design in [22]. Fig. 3.11 also illustrates that the proposed HBF design perfectly matches the performance of the other HBF realizations of FDP as well as the baseband FDP. While FDP implementation needs 64 RF chains, our design require 1 RF chain, FDP realization in [16] and [22] need 16 and 8 RF chains, respectively, and direct HBF design in [22] utilizes $N_{RF} = 8$ RF chains. The optimal FDP is chosen as in [22] and exhibit a considerable gain over the hybrid design from the same reference. As expected, all the hybrid FDP realization techniques including the proposed design has exactly identical performance to FDP. Finally, in order to compare the complexity of different methods, runtime of simulated algorithms with the same computer are presented in Table 3.2.

Hybrid Schemes Based on SRCA: Here, simulation results are presented for the three SRCA-based hybrid designs in Section 3.5. We perform simulations for a massive-MIMO BS equipped with $N_T = 64$ antennas with uniform linear configuration, located at the centre of a single-cell wireless system for both SU and MU cases.

For the SU scenario, 4-QAM constellation is used along with the channel model for mmWave massive-MIMO with sparse scattering environments as in [22,26,74]. For UE with $N_T = 8$ transmit antennas, $K = 8$ symbols per transmission, Fig. 3.12 depicts the BER performance versus SNR (SNR=P_T) for FD scheme described in section 3.5.4, our three proposed SRCA schemes in Section 3.5, and the hybrid robust design in [26] called MOHBF. Our schemes match the performance of FD beamforming, and outperform MOHBF by more

	T_c-Calculations	T_s-Calculations
Direct design [22] SOHBF	0.6911	1.8221e-05
FDP realization [22] SOFDR	0.0011	1.5301e-05
Proposed FDP realization	8.8858e-04	1.4921e-04

Table 3.2: Runtime of T_c and T_s-calculations in seconds.

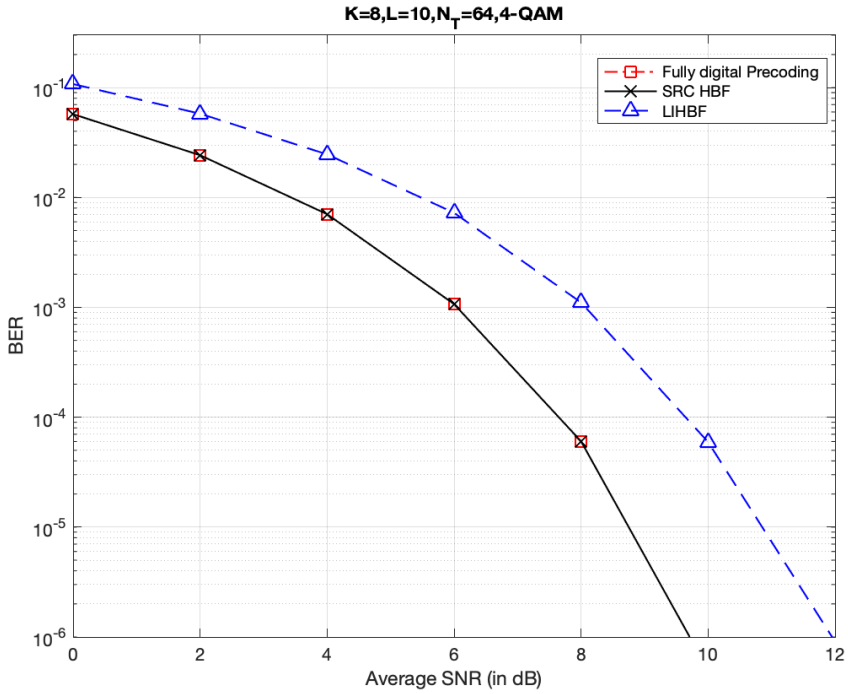


Fig. 3.13: BER versus SNR for FD precoder and our design in a MU setup with a 64 massive-MIMO BS.

than 3 dB for instance at BER 10^{-4}.

For the multi-user case, 4-QAM constellation along with the independent multipath channel model in [27] is used. We compared our proposed scheme with the hybrid design in [27], called LIHBF, as well as the FD beamforming scheme presented in Section 3.5.4. For $K = 8$ single-antenna users, and independent mmWave channels with $L_k = 10$, Figs. 3.13 illustrates the BER performance versus SNR. It can be observed that our scheme has a margin of more than 2 dB to the hybrid design in [27] while achieving the same performance as FD precoding.

3.7 Conclusion

We provided a solution for an efficient implementation of HBF for massive-MIMO systems. The proposed design requires a minimum number of RF chains while matching the performance of optimal baseband FDP. First, we presented the generalized system formulation of

the hybrid architecture. Then, we minimized the number of RF chains such that any given FDP can be realized in the extended HBF architecture. Based on the obtained optimum solutions, we then presented our hybrid design for realizing any FDP in the HBF architecture with a single RF chain, referred to as SRCA. We then proposed three novel SRCA-based beamformer schemes to overcome certain limitations of the original SRCA while achieving the performance of FD systems. Moreover, the applications of these systems for optimal precoding in both SU and MU cases were studied. The presented simulation results verify that the proposed single RF chain hybrid design outperforms a direct HBF design scheme and can actually achieve the same performance as the optimal baseband FDP. In the next chapter, we extend our design to wide-band systems.

Chapter 4

Realizing Wide-Band FD Precoding with Hybrid A/D Architecture

In this chapter[1], we explore the minimum required number of RF chains for realizing FD MIMO-OFDM beamforming with hybrid architecture. After presenting the MIMO-OFDM system model for FD and HBF, we focus on extending our proposed HBF design form the previous chapter to wide-band systems, as the adaptation of these designs to wide-band is not trivial and there is also a trade-off between the system bandwidth and the number of RF chains. Finally, simulation results are presented to compare the waveform and BER performance of FD beamforming and the proposed HBF.

4.1 Introduction

In this chapter, we develop a hybrid design for realizing FD precoding with minimum number of RF chains in case of wide-band wireless channels. Frequency selectivity is the main issue in this case as it can cause severe inter-symbol interference (ISI). Interference cancellation techniques have been studied for many years; however, in addition to the very

[1]Parts of the work presented in this chapter have been patented and presented in [72, 78]:

- A. Morsali, A. Haghighat, and B. Champagne, "Efficient Implementation of Hybrid Beamforming," *U.S. Patent,* 10,790,889, issued September 29, 2020.
- A. Morsali and B. Champagne, "Achieving Fully-Digital Performance by Hybrid Analog/Digital Beamforming in Wide-Band Massive-MIMO Systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, 2020, pp. 5125-5129.

complex calculations, the performance of these techniques is not acceptable for modern systems. Alternatively, it was shown in [79] that by increasing the symbol duration, ISI can be mitigated. This idea is elegantly used in OFDM by dividing the wide-band channel into many narrow-band sub-channels while increasing the symbol period to maintain the data rate [80–82]. OFDM is a time-limited version of multi carrier (MC) modulation which was first introduced in [81], but MC can further be traced back to the 1960s [83,84]. OFDM is widely used in communication systems, from digital subscriber lines [85] to space-time-frequency coding [86].

Since the advent of OFDM, many other MC modulation techniques have been proposed. There are a few versions of OFDM such as filtered-OFDM and windowed-OFDM. Besides, filter bank MC (FBMC) systems have been conceived based on polyphase filter bank analysis [87]. Circular shifted filters are used in generalized frequency devision multiplexing (GFDM) to reduce the out-of-block leakage [88]. In [89], the concept of universal filtered MC (UFMC) was introduced which performs OFDM on a smaller number of frequency tones and then combines them by means of shifted prototype filters. Recently, non-orthogonal multiple access (NOMA) has also attracted much attention for 5G, but this scheme falls the scope of our research [90, 91]. Although in this chapter we focus on OFDM, the same approach and guidelines can be used for extending the proposed hybrid design to other MC techniques.

In the OFDM-based hybrid A/D structure, the subcarrier can use different digital precoders but must share a common RF precoder. Consequently, designing hybrid precoders is more complicated than in the narrow-band scenario. Compared to narrow-band HBF, only a limited number of studies have been published on MIMO-OFDM HBF in the literature, some of which are reviewed in what follows.

A wide-band FDP realization is presented in [23] where the number of RF chains is equal to the rank of the combined digital precoder matrices. In [24], the hybrid precoder design is formulated as a matrix factorization problem, and an alternating minimization algorithm is presented for fully-connected and partially-connected RF precoder structures. For a given RF codebook, the optimal hybrid precoding is derived in [60] which is used as a benchmark for other heuristic algorithms. Subsequently, codebooks are developed for spatial multiplexing in wide-band mmWave systems. In [64], based on a closed-form solution for the fully-connected OFDM-based hybrid analog/digital precoder, a technique that dynamically constructs the hybrid subarrays using the long-term channel characteristics is

presented.

In this chapter, we focus on realizing any given FDP with the HBF architecture in wide-band massive-MIMO-OFDM systems. We first introduce the system formulation and channel model, and then, we present time-domain OFDM signal reconstruction via HSP. We next present the required conditions for realizing any FDP with HBF architecture in a general case, before focusing on a specific case and deriving a systematic design for arbitrary number of RF chains. The trade-off between the number of RF chains and bandwidth of the output is investigated. After developing a design for two RF chains, we finally present and discuss the simulation results.

4.2 Wide-Band massive-MIMO System Model

Let us consider OFDM based mmWave massive-MIMO transmitter with N_T antennas and N_{RF} RF chains, respectively. Each OFDM symbol has N_c sub-carriers with cyclic prefix of length N_{cp} where at each sub-carrier N_s are transmitted via multiple antennas. Let $\mathbf{s}[k] = [s_{k,1}, s_{k,2}, ..., s_{k,N_s}]^T$ be the symbol vector of the kth sub-carrier, where $s_{k,i}$ is taken from a discrete constellation $\mathcal{A}$ (such as M-QAM or M-PSK).

4.2.1 mmWave Wide-Band Channel Model

There are two main categories for wireless channel models, namely *parametric* and *non-parametric*. While the latter is mostly used for theoretical analysis and stochastic evolution of the channel matrix (e.g., the Kronecker model [92, 93]), parametric models which are based on the geometry of the propagation environment, better represent the performance of the system in realistic scenarios. The Saleh-Valenzuela model [46] which was first introduced for indoor mmWave communications is also adapted and widely used for cellular mmWave systems as it can adequately capture the propagation characteristics of the mmWave signals [41, 60]. We use the channel model in [64] because of its accuracy and compactness of presentation. The channel matrix at kth sub-carrier can be modeled as,

$$\mathbf{H}[k] = \sum_{c=1}^{N_{sc}} \sum_{r=1}^{N_r} \alpha_{c,r} w_{\tau_{c,r}}[k] \mathbf{a}_R(\phi_{c,r}^R, \theta_{c,r}^R) \mathbf{a}_T(\phi_{c,r}^T, \theta_{c,r}^T)^H, \tag{4.1}$$

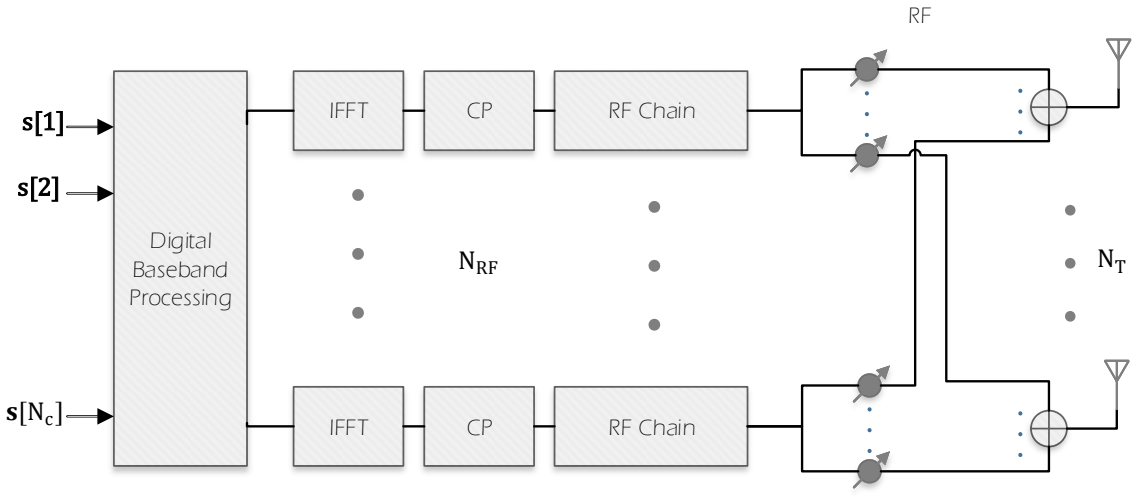


Fig. 4.1: MIMO-OFDM transmitter with HBF.

where

$$w_{\tau_l}[k] = \sum_{d=0}^{D-1} p(dT_s - \tau_l) e^{-\frac{j2\pi kd}{N_c}}. \tag{4.2}$$

where N_{sc} is the number of scattering clusters with N_r scatterers (subrays) in each cluster. Moreover, α is the complex path gain and τ_l is the path delay. The azimuth angles of departure and arrival are denoted by ϕ^T and ϕ^R, respectively, also, θ^T and θ^R are elevation angles of departure and arrival. The overall impulse response of the combination of all the filters (pulse shaping, matched filter, etc.) is represented by $p(\tau)$.

4.2.2 FDP and HBF Architectures

In FDP systems, precoding can be performed at each sub-carrier as

$$\mathbf{x}_{\mathrm{T}}[k] = \mathbf{P}_{\mathrm{FD}}[k]\mathbf{s}[k], \tag{4.3}$$

where $\mathbf{x}_{\mathrm{T}}[k]$ and $\mathbf{P}_{\mathrm{FD}}[k]$ are transmitted signal and FDP matrix at kth sub-carrier. In this scenario, $\mathbf{P}_{\mathrm{FD}}[k]$'s are $N_T \times N_s$ matrices which requires N_T RF chains. Since it is desirable to reduce the number of RF chains, HBF architecture can play a significant

role. However, there are more restrictions compared to narrow-band case. Since the RF processing is performed in time domain, the phase-shifters in RF precoder are assumed to have a flat frequency response which means the same analog precoding is performed on all the sub-carriers.

Most prior works in the literature consider frequency signal reconstruction where the transmitted signal is written as

$$\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}[k] = \mathbf{P}_{\mathrm{A}}\mathbf{P}_{\mathrm{D}}[k]\mathbf{s}[k], \tag{4.4}$$

where $\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}[k]$ and $\mathbf{P}_{\mathrm{D}}[k]$ are the transmitted signal and digital precoder of kth sub-carrier, respectively, and $\mathbf{P}_{\mathrm{A}}$ is the analog precoder for all sub-carriers. Using the same notation as Section 3, the baseband signal after digital precoding can be also written as,

$$\mathbf{x}_{\mathrm{BB}}^{\mathrm{HY}}[k] = \mathbf{P}_{\mathrm{D}}[k]\mathbf{s}[k]. \tag{4.5}$$

Fig. 4.1. depicts the transmitter of a conventional wide-band HBF [24, 60–65]. Symbols are first precoded at each sub-carrier, then N_{RF} OFDM symbols are constructed by performing IFFT and appending cyclic prefixes. After being converted to RF signals, the OFDM symbols are precoded in time domain by RF precoder.

4.3 Realizing FDP in OFDM HBF Architecture

In most prior works [24, 60–65], frequency domain analysis is used for designing hybrid precoders. We, on the other hand, focus on time domain HBF design.

4.3.1 Problem Formulation

The OFDM signal for each antenna i can be calculated in frequency domain according to the chosen FDP using (4.3). Thus, we have the desired time domain transmitted signal of each antenna denoted by $\mathbf{x}_i^{\mathrm{FD}}(t)$. Let $x_i^{\mathrm{TI}}(t)$ denote the transmitted signal in HBF from ith antenna during $t \in [0, NT_s]$ where $T_s = 1/F_s$, and F_s being the sampling frequency and $N = N_c + N_{cp}$. We can thus write,

$$\mathbf{x}_i^{\mathrm{TI}}[n] \triangleq x_i^{\mathrm{TI}}(nT_s), \tag{4.6}$$

which for $n = 0, 2, ..., N_c - 1$, we have,

$$\mathbf{x}_i^{\mathrm{TI}}[n] = \sum_{k=0}^{N_c} \mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}[k, i] e^{-j2\pi n/N_c}, \tag{4.7}$$

where $\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}[k, i]$ denotes the ith entry of $\mathbf{x}_{\mathrm{T}}^{\mathrm{HY}}[k]$. In practice, $\mathbf{x}_i^{\mathrm{TI}}$ is passed through DAC to obtain $x_i^{\mathrm{TI}}(t)$. However, since we want to reconstruct $x_i^{\mathrm{TI}}(t)$, it is helpful to have the reverse relationships. Now assuming the same hardware constrictions as Chapter 3, we can derive conditions for realization of any FDP with HBF.

Proposition 4.1. *Any FDP can be realized in hybrid architecture with N_{RF} RF chains if for $i = 1, 2, ..., N_T$, $m = 0, 1, 2, ..., M - 1$ and $p = 0, 1, 2, ..., N_{RF} - 1$, there exist: $e^{j\phi_{m,p}^i}$, and $x_{m,p}(t)$, where for each p and $m \neq \hat{m}$,*

$$\mathrm{supp}(x_{m,p}(t)) \neq \mathrm{supp}(x_{\hat{m},p}(t)), \tag{4.8}$$

such that for all i's and $t \in [0, NT_s]$

$$\mathbf{x}_i^{FD}(t) = x_i^{TI}(t), \tag{4.9}$$

with

$$x_i^{TI}(t) = \sum_{p=0}^{N_{RF}-1} \sum_{m=0}^{M-1} e^{j\phi_{m,p}^i}\ x_{m,p}(t). \tag{4.10}$$

Proof. Since (4.9) ensures the realization of FDP, we must show that a hybrid system with N_{RF} RF chains can have (4.10) as its output. If the condition in (4.8) is satisfied, then all $x_{m,p}(t)$s for $m = 1, 2, ..., M$ can be generated by the pth RF chain:

$$x_p(t) = \sum_{m=0}^{M-1} x_{m,p}(t). \tag{4.11}$$

Assuming the phase-shifters can be updated M times during $t \in [0, NT_s]$ period (which can be also implemented by the system designs in 3.5), the transmitted signal of ith antenna is given by (4.8). □

Proposition 4.1 shows that, in order to design an HBF which can realize any FDP, a set of MN_TN_{RF} angels ($\phi_{m,p}^i$ for $i = 1, 2, ..., N_T$, $m = 1, 2, ..., M$ and $p = 1, 2, ..., N_{RF}$)

are required as well as MN_{RF} functions $x_{m,p}(t)$. In the next section, we present a class of solutions with perfect reconstruction at sampling points (PRSP) by restricting $x_i^{\mathrm{TI}}(t)$s to periodic functions.

4.3.2 Minimum Number of RF Chains

In this section, we present PRSP HBF design for FDP realization. We first discuss one fundamental trade-off for time-domain signal reconstruction of PRSP. Then, we present the general conditions for realizing FDP with HBF with arbitrary number of RF chains. Finally, we present an HBF design using raised cosine pulse with two RF chains.

Perfect Reconstruction at Sampling Point

Although the transmitted and received OFDM signals in RF domain are continuous time functions, the information is encoded at the sampling points. Therefore, if we have prefect synchronization, it is not necessary to have the same continuous time function as long as the sample values are the same. In other words, instead of the strict condition (4.9), only the sample value must be equal:

$$\mathbf{x}_i^{\mathrm{FD}}(nT_s) = x_i^{\mathrm{TI}}(nT_s). \tag{4.12}$$

Indeed, according to Nyquist–Shannon sampling theorem, (4.12) and (4.9) are equivalent if the bandwidths of $\mathbf{x}_i^{\mathrm{FD}}(t)$ and $x_i^{\mathrm{TI}}(t)$ are the same. However, since we have do not have this restriction in general, it is possible to have (4.12) whilst (4.9) being violated. It is clear that (4.9) is the ideal case but harder to attain. Consequently, under perfect time offset synchronization the performance is the same and the design is simplified.

In (4.11), we defined the generated signal from pth RF chain as summation of M non-overlapping functions $x_{1,p}(t), x_{2,p}(t), ..., x_{M,p}(t)$. On way of reducing the design parameters is to use shifted version of a windowed function to build $x_{m,p}(t)$'s. Let us consider an energy signal $p_p(t)$ such that $p_p(t) = 0$ for $t \notin [0, \tau]$. Thus, if $\frac{M}{N}$ is an integer and we have,

$$\tau = \frac{N}{M} T_s, \tag{4.13}$$

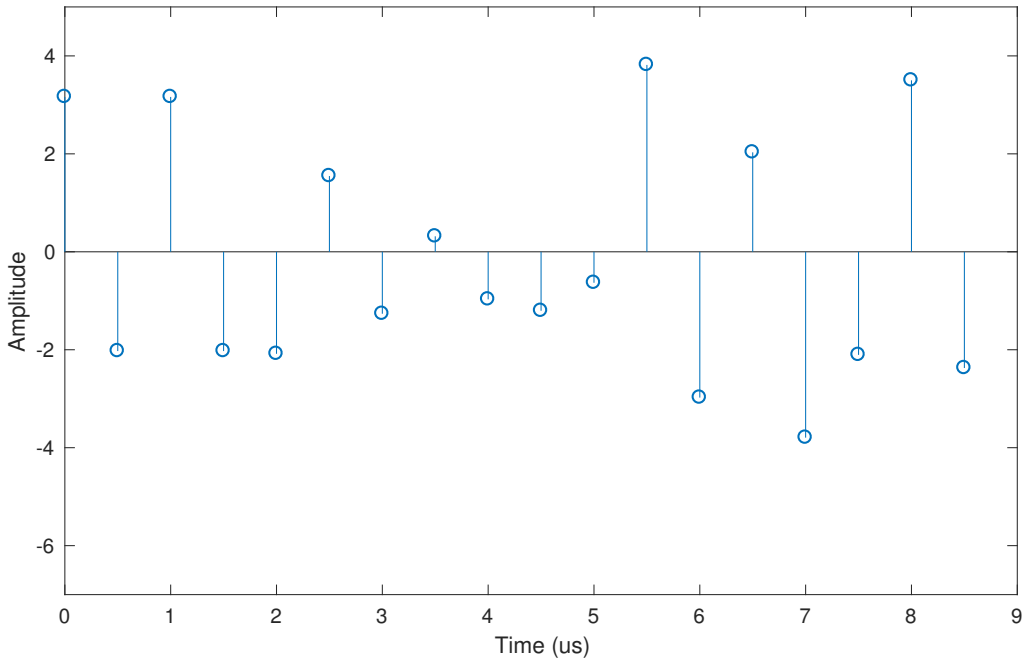


Fig. 4.2: Discrete-time OFDM frame.

where the generated signal of pth RF chain(4.11) can be written as,

$$x_p(t) = \sum_{m=0}^{M-1} p_p(t - m\tau). \tag{4.14}$$

Note that $p_p(t)$ must be an energy signal so that it can be generated by an RF chain. Now, we can present the PRSP HBF realization of any given FDP.

Theorem 4.1. *Any FDP can be perfectly reconstructed at sampling points by a hybrid architecture with $N_{RF} = N/M$ RF chains and $2N_TN_{RF}$ phase shifters.*

Proof. From Section. 3.4, we know it is possible to have two phase shifters per each antenna

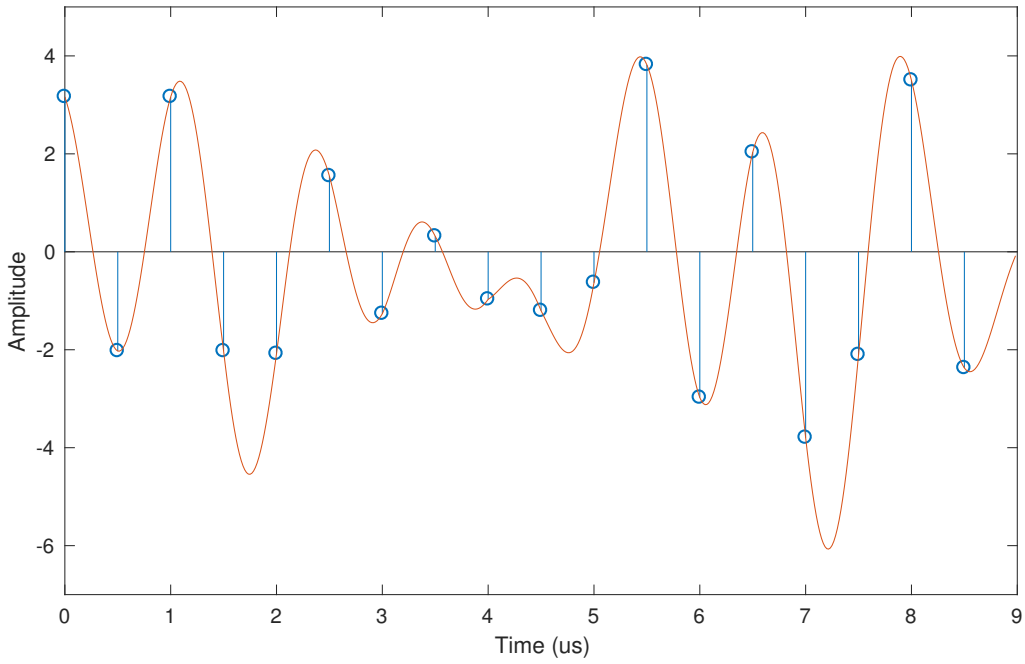


Fig. 4.3: Analog OFDM waveform.

for each RF chain, therefore, (4.10) can be further extended to,

$$x_i^{\mathrm{TI}}(t) = \sum_{p=0}^{N_{RF}-1} \sum_{m=0}^{M-1} (e^{j\phi^i_{m,p,1}} + e^{j\phi^i_{m,p,2}})\ x_{m,p}(t). \tag{4.15}$$

Using (4.14), the above equation can be written as,

$$x_i^{\mathrm{TI}}(t) = \sum_{p=0}^{N_{RF}-1} \sum_{m=0}^{M-1} (e^{j\phi^i_{m,p,1}} + e^{j\phi^i_{m,p,2}})\ p_p(t - m\tau). \tag{4.16}$$

Now, let us define $p_p(t)$'s as,

$$p_p(t) = p(t - pT_s), \tag{4.17}$$

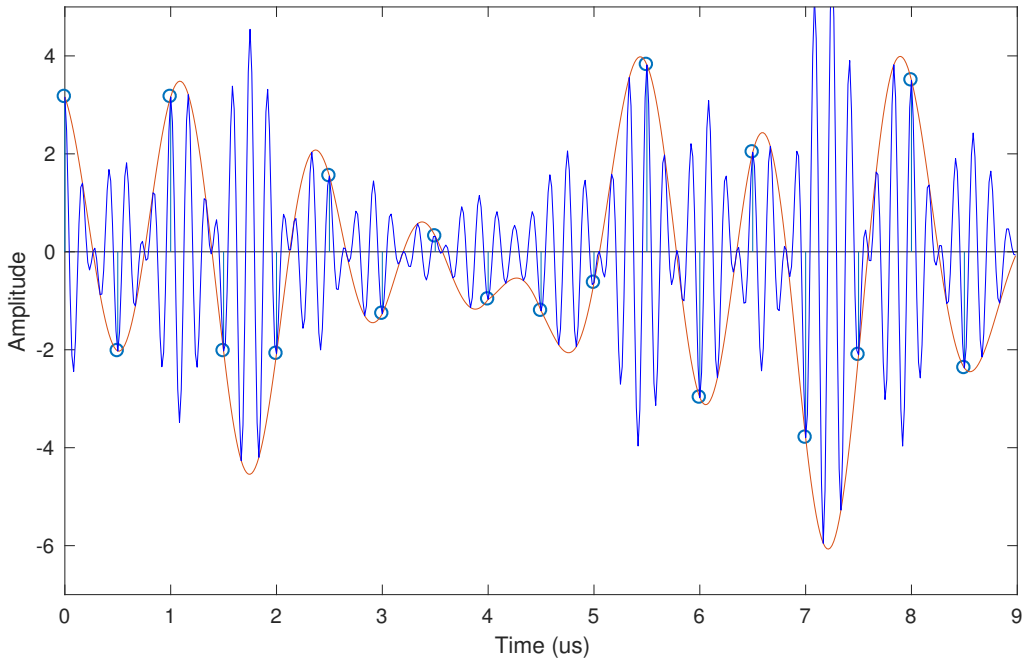


Fig. 4.4: RF Modulated OFDM waveform.

where $p(t)$ satisfies the Nyquist filters criterion, i.e.,

$$p(nT_s) = \begin{cases} 1; & n = 0 \\ 0; & n \neq 0 \end{cases}. \tag{4.18}$$

Hence, we can write

$$x_i^{\mathrm{TI}}(t) = \sum_{p=0}^{N_{RF}-1} \sum_{m=0}^{M-1} (e^{j\phi_{m,p,1}^i} + e^{j\phi_{m,p,2}^i})\; p(t - pT_s - m\tau). \tag{4.19}$$

Now, from (4.19), (4.12) and (4.13), and using the definition in (4.6), for PRSP realization it is sufficient to have,

$$\mathbf{x}_i^{\mathrm{FD}}[n] = \sum_{p=0}^{N_{RF}-1} \sum_{m=0}^{M-1} (e^{j\phi_{m,p,1}^i} + e^{j\phi_{m,p,2}^i})\; p[n - p - mN_{RF}]. \tag{4.20}$$

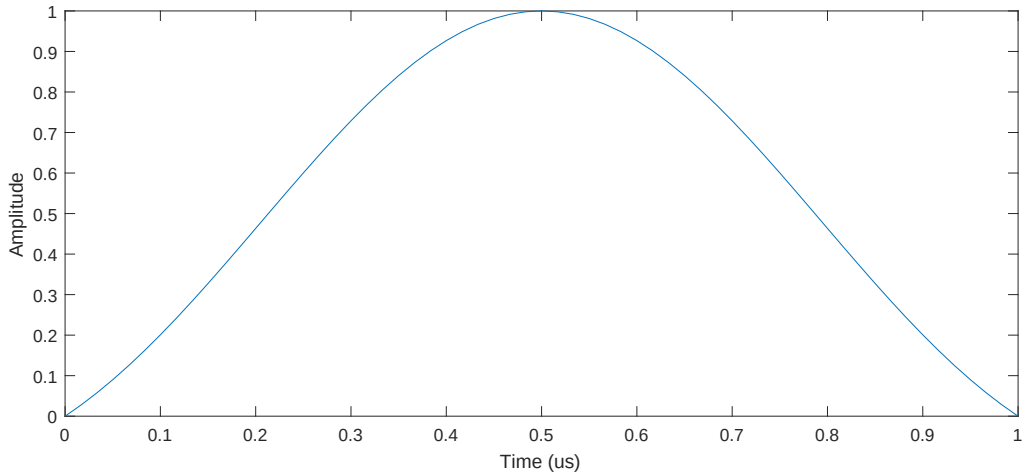


Fig. 4.5: Raised cosine filter for $\beta = 0.5$.

Without loss of generality, let us assume $\mathbf{x}_i^{\mathrm{FD}}[n]$'s for all i's and n's are less than 2. From (4.18) and Theorem 3.2 it is guaranteed that $\phi^i_{m,p,1}$ and $\phi^i_{m,p,2}$ exist such that (4.20) is satisfied.

□

Remark 4.1. *The minimum number of required RF chains for realizing any given FDP is $N_{RF} = 1$. However, reducing the number of RF chains comes with the price of extra bandwidth. On the other hand, for $N_{RF} = N$ RF chains, sinc function can be generated by $p(t)$ which satisfies* (4.9) *and results in the most favourable spectral characteristics.*

Proposition 4.2. *There is a trade-off between number of RF chains and bandwidth of the output signal, where decreasing N_{RF} results in an increase in bandwidth.*

Proof. Based on the OFDM parameters, N is given, and from (4.13) we have,

$$\tau = N_{RF} T_s. \tag{4.21}$$

Therefore, decreasing N_{RF} also decreases τ which is the length of $p(t)$. Reducing τ will increase the bandwidth of $p(t)$ which in turn increases the bandwidth of the output signal.

□

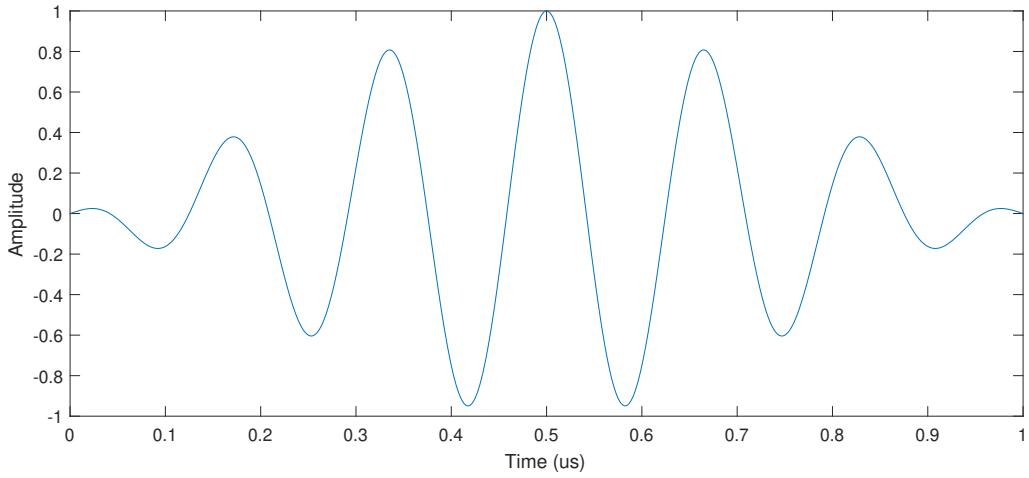


Fig. 4.6: RF modulated raised cosine filter with $\beta = 0.5$.

FDP realization with two RF chains: In order to have a proper frequency characteristics (as discussed in Proposition. 4.2) we present an example for FDP realization with two RF chains. The OFDM signal $x_i^{\mathrm{TI}}(t)$ is generally a complex function in time. In communication systems however, $cos(\omega_c t)$ and $sin(\omega_c t)$ are used as carriers for the real and imaginary parts of $x_i^{\mathrm{TI}}(t)$, respectively. Let $\mathcal{R}_i(t)$ and $\mathcal{I}_i(t)$ represent the in-phase and quadrature RF component of the HBF output. Thus, we have

$$\mathcal{R}_i(t) = \sum_{p=1}^{N_{RF}} \sum_{m=1}^{M} x_{m,p}^{I}(t) cos(\omega_c t + \phi_{m,p}^{i}) \tag{4.22}$$

$$\mathcal{I}_i(t) = \sum_{p=1}^{N_{RF}} \sum_{m=1}^{M} x_{m,p}^{Q}(t) sin(\omega_c t + \theta_{m,p}^{i}), \tag{4.23}$$

where $x_{m,p}^{I}(t)$ and $x_{m,p}^{Q}$ are real values functions in time. The main objective is then to find $x_{m,p}^{I}(t)$,$x_{m,p}^{Q}(t)$, $x_{m,p}^{Q}$,$\phi_{m,p}^{i}$ and $\theta_{m,p}^{i}$ such that

$$Real(\mathbf{x}_i^{\mathrm{FD}}[n]) = \mathcal{R}_i(nT_s), \tag{4.24}$$

$$Imag(\mathbf{x}_i^{\mathrm{FD}}[n]) = \mathcal{I}_i(nT_s). \tag{4.25}$$

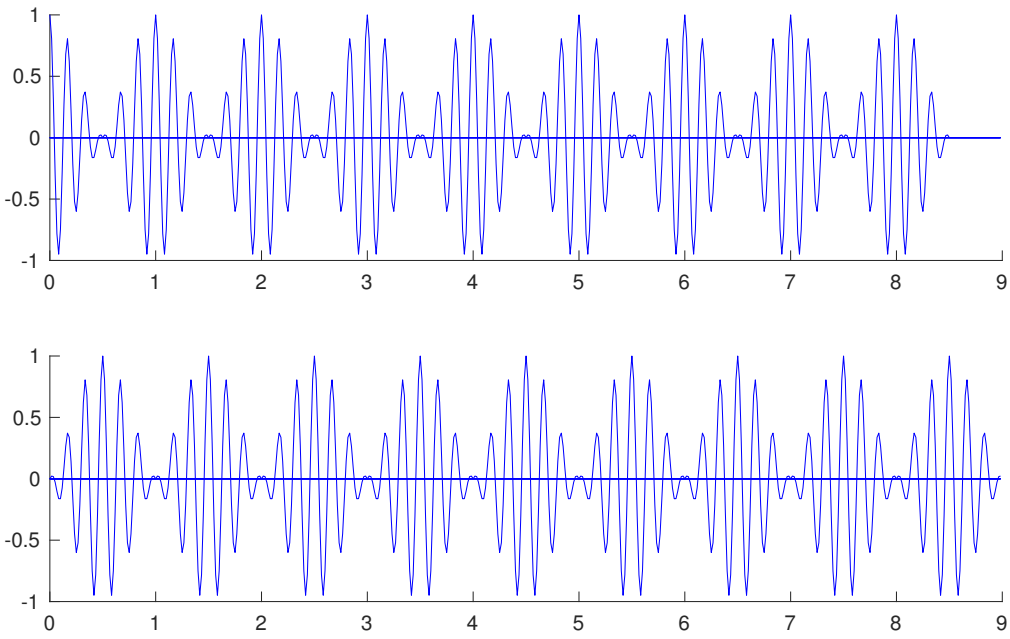


Fig. 4.7: Output signal of RF chains.

Here, we present an FDP realization with two RF chains as an example. First, let us take $p(t)$ as the windowed raised cosine filter with roll off factor 0.5 with duration of $2T_s$ which satisfies (4.14) and (4.18). Thereby, considering the polar representation $\mathbf{x}_i^{\mathrm{FD}}[n] = |\mathbf{x}_i^{\mathrm{FD}}[n]| \exp(j\angle \mathbf{x}_i^{\mathrm{FD}}[n])$, for $n = p + mN_{RF}$, the phase-shifter parameters are then calculated as,

$$\phi_{m,p,1}^{i} = \angle \mathbf{x}_i^{\mathrm{FD}}[n] + \cos^{-1}\left(\frac{|\mathbf{x}_i^{\mathrm{FD}}[n]|}{2c_0}\right), \tag{4.26a}$$

$$\phi_{m,p,2}^{i} = \angle \mathbf{x}_i^{\mathrm{FD}}[n] - \cos^{-1}\left(\frac{|\mathbf{x}_i^{\mathrm{FD}}[n]|}{2c_0}\right). \tag{4.26b}$$

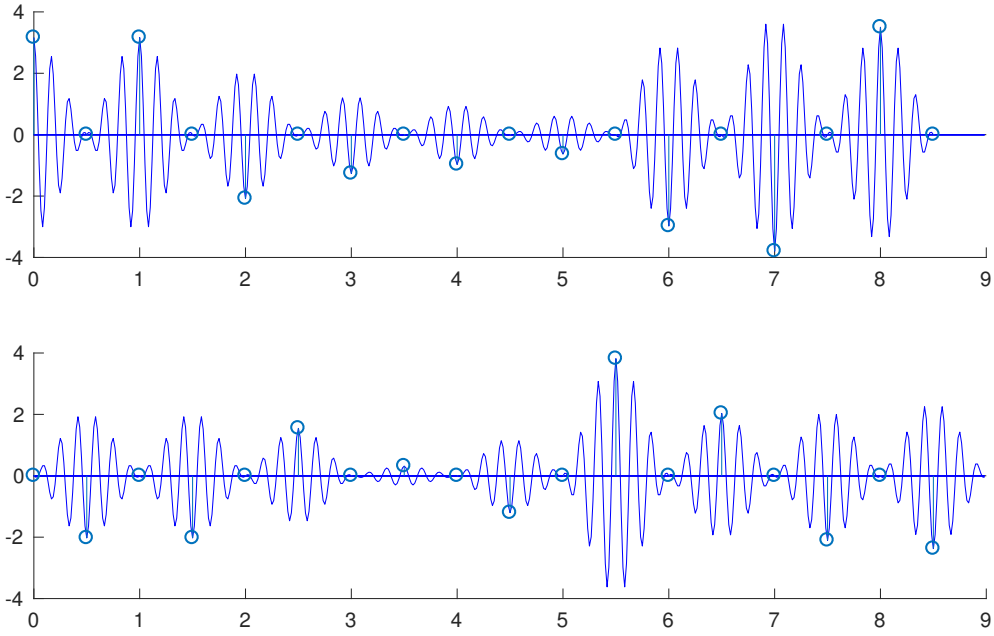


Fig. 4.8: Output signal of phase-shifters for one of the antenna elements.

4.4 Simulation Results

In this section, first waveforms of the proposed HBF and FDP are compared and then, we present simulation results for FDP system our proposed HBF and the HBF design in recently published work [65] called SOFDBF.

4.4.1 FDP and HBF Waveforms

To clarify the concept of PRSP FDP realization, an example for the the real part of the signal (In-phase component) is presented below. We consider an OFDM system with $N_c = 16$ subcarrier, and bandwidth of $BW = 2MHz$ with carrier frequencies of $F_c = 6GHz$. For 16-QAM constellation is used and the number of cyclic prefixes is set to $N_{cp} = 2$. Simulations are performed with sampling frequency of $10F_c$. Without loss of generality, we only present the signal of one arbitrary antenna element. Fig. 4.2. depicts the digital OFDM signal whereas the analog baseband and analog RF signals are illustrated in Fig. 4.3. and Fig. 4.4., respectively. We used raised cosine filter with roll of factor 0.5 for $p(t)$ which

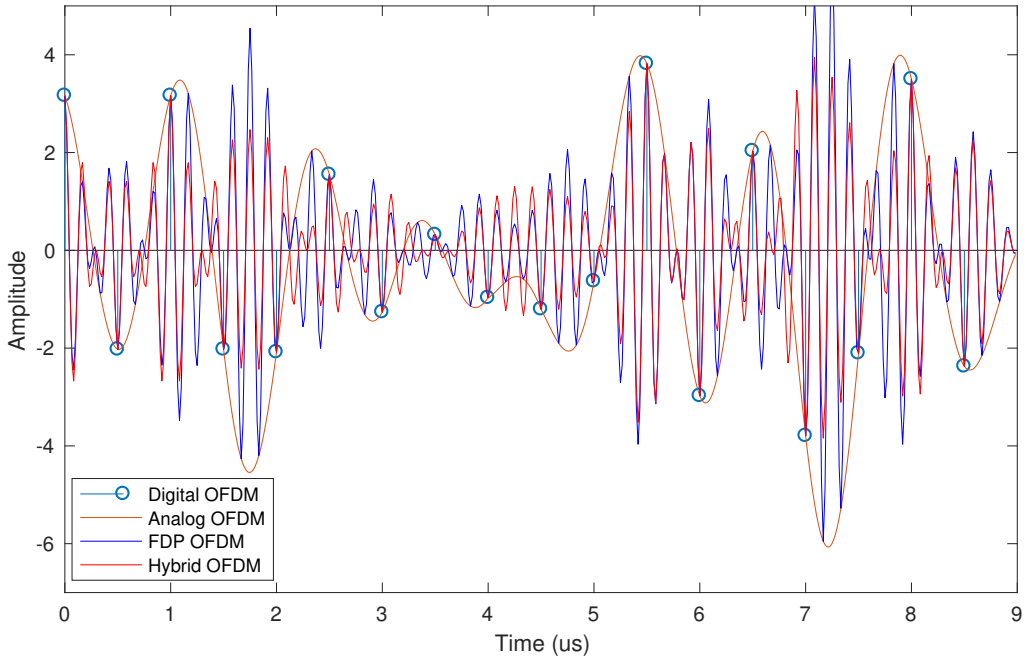


Fig. 4.9: Comparison of FDP and HBF waveforms.

is depicted in baseband and RF in Fig. 4.5 and Fig. 4.4, respectively. Consequently, for $N_{RF} = 2$, we have $M = 2$ and thus, the RF signals generated by two RF chains which are illustrated in Fig. 4.7 can be written as:

$$x^1_{RF}(t) = \sum_{m=1}^{9} x^I_{m,1}(t) cos(\omega_c t), \tag{4.27a}$$

$$x^2_{RF}(t) = \sum_{m=1}^{9} x^I_{m,2}(t) cos(\omega_c t). \tag{4.27b}$$

After, passing $x^1_{RF}(t)$ and $x^2_{RF}(t)$ through RF precoder, the result is shown in Fig. 4.8.

In Fig. 4.9. RF FDP and HBF signals are presented as well as the digital and baseband analog OFDM signals. It can be observed that in the sampling points all the signals have the same value as we also showed mathematically. The HBF signal, however, has higher frequencies due to the choose of $p(t)$ which can be filtered or altered by available degrees of freedom in the system.

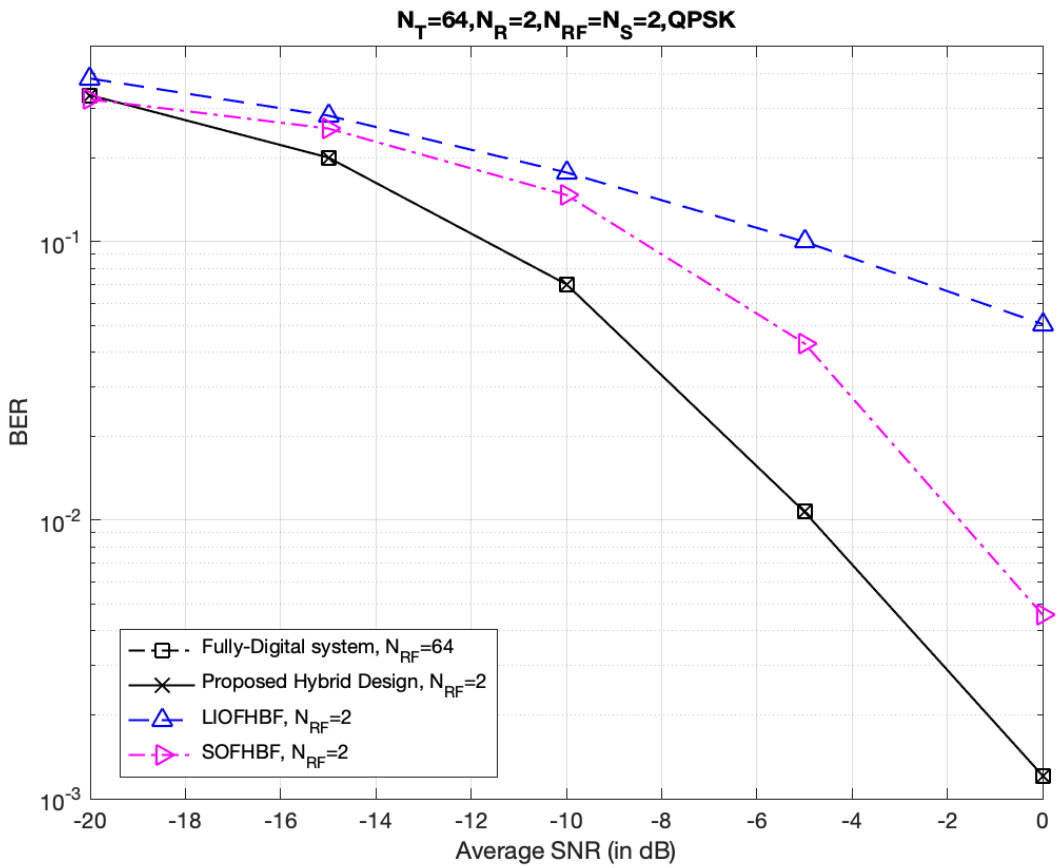


Fig. 4.10: BER versus SNR for different methods.

4.4.2 BER Performance

Simulations are performed for massive-MIMO transmitter with $N_T = 64$ transmit antennas and a user with $N_R = 16$ receive antennas. OFDM modulation with bandwidth of $BW = 2MHz$ and carrier frequency of $F_c = 6GHz$ is considered where the number of subcarrier and cyclic prefixes are set to $N_c = 128$ and $N_{cp} = 16$, respectively. Moreover, Q-PSK modulation modulation is used for $N_s = 8$ symbols per subcarrier. The receiver performs FD combining and optimal precoding is used for FDP and the proposed HBF at the transmitter.

The mmWave channel is modeled with $N_{sc} = 6$ clusters and $N_r = 2$ rays per cluster. Channel gain $\alpha_{c,r}$ is assumed to be an i.i.d complex Gaussian with variance $N_T N_R / N_{sc} N_r$. Azimuth angles of departure and arrival $\phi^T_{c,r}$ and $\phi^R_{c,r}$ are uniformly distributed in $[0, \pi]$.

Moreover, τ_l is uniformly distributed in $[0, N_{cp}T_s]$ and τ_l/T_s are integers. For simplicity, we further assume $p(\tau) = \delta(\tau)$.

Fig. 4.10. shows the BER performance versus SNR of the FD system, HBF in [65] and [75],(i.e., SOFHBF, and LIOFHBF) and our proposed HBF design. It can be seen that while our design matches the FDP, it could outperform the design in [65] by more than 5 dB.

4.5 Conclusion

In this chapter, we studied wide-band massive-MIMO mmWave communications with hybrid architecture. After reviewing the conventional frequency domain analysis for HBF, we introduced the notion of time domain signal reconstruction of OFDM signals. Then, we formulated the problem and presented the conditions for the realization of any given FDP in HBF architecture. After introducing the PRSP realization, we discussed the general solution for an arbitrary number of RF chains and as an example, presented a design for two RF chains. Finally, simulations results were presented to confirm the merits of our proposed method over recently published works. In the next two chapters, which comprise the second part of the thesis, we explore the design of hybrid A/D combining.

Part II

Hybrid A/D Combining

Chapter 5

Massive-MIMO Receiver with Hybrid A/D Architecture

In this chapter[1], we present a novel HBF structure for massive-MIMO which can be used to realize any given FD combiner. Considering the unit-modulus element constraint of analog beamforming matrices, we exploit the fact that any given complex matrix can be written as a scaled sum of two matrices with unit-modulus entries. By taking advantage of this decomposition, we propose a novel hybrid structure for realizing any FD combiner with the minimum number of radio-frequency (RF) chains required by an HBF structure to match the performance of an FD beamformer. We then focus on a more practical scenario where phase-shifters can realize a finite number of phase angles. Accordingly, we propose a modified hybrid structure by introducing an additional degree of freedom, i.e., the phase-offset between the finite-resolution phase-shifts and optimize this parameter via close-form approximations which are shown to result in the same solution as an exhaustive search. To illustrate that the proposed hybrid structure can be used for different applications, we develop a novel pilot-based channel estimation technique that can achieve the same performance as FD optimum linear estimation. Simulation results confirm that all our proposed HBF schemes can achieve the same performance as their FD counterparts, and therefore outperform existing hybrid designs with the same number of RF chains by a large

[1]Parts this chapter have been submitted to USPTO:

- A. Morsali, A. Haghighat, and B. Champagne,"Efficient Receiver Combining for Hybrid Analog/Digital Beamforming," *U.S. Patent,* PCT/US2020/024318, March 23, 2020.

margin.

5.1 Introduction

MIMO has revolutionized modern wireless communications by unveiling the potential increase in system capacity resulting from the deployment of multiple antennas at the transmitter and/or receiver side(s) of the link [6]. In recent years, asymptotic analysis of large-scale antenna arrays has revealed that massive-MIMO systems exhibit a linear increase in capacity with the minimum number of antennas employed at either the transmitter or receiver regardless of the channel characteristics [7,8]. Moreover, mmWave communications has been proposed for indoor and fixed outdoor communications, with the aim to enable gigabit-per-second data rates [10, 11].

However, mmWave signals experience severe path loss, penetration loss and atmospheric absorption compared with microwave signals, which until recently has precluded their use in wireless cellular and local area networks. Recent advances in mmWave hardware and, more importantly, newly discovered capabilities of massive-MIMO have revived mmWave communications. Especially, the highly selective beam steering capabilities provided by large-scale antenna array and sophisticated beamforming algorithms can mitigate the intrinsic limitations of mmWave communications [12]. In practice, beamforming can be employed at both the transmitter and the receiver ends of a wireless link, where it is referred to as precoding and combining, respectively. In this chapter, we focus on the mmWave massive-MIMO receiver and specifically explore the combiner design.

In the FD implementation of MIMO systems, each antenna element is connected to a dedicated RF chain. While this approach is suitable for commonly used small scale MIMO systems, it is not applicable to mmWave massive-MIMO systems with large number of antenna elements (e.g., on the order of 100 or more) due to the high production costs and power consumption of the associated RF circuitry. Therefore, although mmWave massive-MIMO is the prime candidate for B5G wireless networks, the implementation of such systems still faces many technical challenges, and to date remains a topic of ongoing research [16, 17, 94].

HBF is an ingenious and effective approach to facilitate the implementation of mmWave massive-MIMO transceivers [22, 23]. In an HBF receiver structure, the high dimensional received RF signal vector from the antenna array is first combined into a lower dimensional

signal by means of an analog beamformer, comprised of basic RF components such as phase-shifters, adders and splitters. The resulting analog signal vector is then down-converted to a digital signal vector by means of RF chains and fed to a digital beamformer for final combining. Consequently, the HBF requires a smaller number of RF chains compared to the FD structure. However, a unit-modulus constraint is imposed on the entries of the analog beamformer due to the use of phase-shifters. In the existing HBF literature, due to this constraint, the ensuing optimization problem for designing the analog and digital beamformer matrices is non-convex [17,22]. In this chapter, using the same RF components, we propose an alternative hybrid structure which lifts the non-convex constraint and allows the realization of any FD beamformer.

Related Works

Realization of FD beamfoming in hybrid structure has been studied in [16, 22, 23, 73, 74, 78, 94]. In [16], it has been shown that for a single stream of data, two RF chains are required to achieve the performance of a FD combiner. More generally, for multiple data streams (layers), the number of RF chains must be at least twice the number of the data streams [22]. The technique in [16] can be modified to design hybrid precoders which only require the same number of RF chains as the number of data streams [22,23,94]. In [72,74], we proposed a single RF chain scheme for realizing any given FD precoding based on the use of fast adjustable phase-shifters.

An alternative approach is to design the analog and digital beamformers directly by optimizing different cost functions. However, due to the constant modulus constraint on the analog beamformer entries, the ensuing optimization problems are non-convex and only sub-optimal solutions can be obtained. Minimizing the Euclidean distance between the hybrid and FD beamformers was considered as the objective function in [21,24,29,60]. Particularly, in [21, 60], compressed sensing techniques were used to exploit sparse characteristics of the mmWave channels while in [24, 29], a manifold optimization algorithm and a simultaneous matrix diagonalization technique were introduced, respectively.

Sparsity of the channel was also considered in [22], [65] where iterative orthogonalization algorithms were proposed for designing spectrally efficient HBF transceivers. Gram–Schmidt orthogonalization was used in [27] to design a robust hybrid combiner with low complexity for an uplink MU scenario. Considering that closed-form expressions with fixed amount

calculation are often more attractive in applications, non-iterative HBF design algorithms exploiting this type of solutions were proposed in [25,26]. The mean square error (MSE) was considered as the performance metric in [75], where an alternating minimization technique is used to design the HBF matrices. The authors in [28] implemented a low-complexity HBF based on orthogonal beamforming codebooks and a local search scheme.

HBF fundamentally changes the structure of the MIMO communication system and due to limited number of RF chains, the main reception and transmission functionalities must be redesigned accordingly. For instance, estimating the CSI must be revisited in the context of hybrid structure to account for these structural constraints. Adaptive techniques were proposed in [66, 67], where both transmitter and receiver scan for the dominant paths within the angular domain and narrow the search to obtain a quantized solution. Inspired by the random convolutional measurement process, a structured random sensing codebook was proposed in [95] to estimate mmWave channels by exploiting the sparse nature of mmWave channels. Similar ideas for wideband MIMO systems were presented in [71, 96]. A general framework for channel estimation with hybrid structure was presented in [97] where algorithmic solutions are proposed for different scenarios.

Contributions

In this chapter, our goal is to design HBFs that lift the unit modulus constraint and, as a result, can achieve the performance of FD systems at the receiver with the minimum number of RF chains. We also consider a more practical scenario where only finite-resolution phase-shifters are used for implementation of the analog beamformer and optimize the proposed hybrid design to take advantage of available degrees of freedom. The proposed HBF realizations are based on simple decomposition properties of analog beamformer matrices and use the same hardware elements as existing HBF solutions. Specifically, our main contributions can be summarized as follows:

- Considering the structure of phase-only analog beamformers and the limited number of allowed RF chains, we propose a technique for decomposing any given complex matrix into a scaled sum of two matrices with unit-modulus entries.
- This decomposition is exploited to develop new and flexible HBF structures for FD realization with minimum number of required RF chains. The new structure can be

used for different reception functionalities (such as beamforming and channel estimation) and under various scenarios and conditions (such as multi/single user and perfect/imperfect CSI).

- A modified hybrid structure is presented to improve the performance of the hybrid beamforming when finite resolution phase-shifters are used. By taking advantage of available degrees of freedom in our design, we minimize the maximum error between the RF beamformer and its FD counterpart.
- We also present a novel channel estimation method based on the proposed hybrid structure which achieves the performance of optimal linear FD channel estimation.
- Simulation results are presented for the proposed beamformer designs as well as the channel estimation method all of which demonstrate the same performance as their FD counterpart, and therefore achieve significant improvement over existing HBF and hybrid channel estimation methods.

5.2 System Model

We consider the uplink connection of a massive-MIMO system where the BS is equipped with N_R receive antennas and N_{RF} RF chains, $N_{RF} \ll N_R$. A conventional mmWave massive-MIMO receiver with hybrid analog/digital structure is depicted in Fig. 5.1. In order to develop a solution for realizing any given digital combiner, we consider the two main reception modes for uplink connection, i.e., SU and MU. In what follows the system models for both cases are presented.

Single-User: For a UE with N_T antennas, the received signal at the BS is given by,

$$\mathbf{y}_{\text{su}} = \sqrt{\rho}\mathbf{H}_{\text{su}}\mathbf{P}\mathbf{s} + \mathbf{n}, \tag{5.1}$$

where $\mathbf{H}_{\text{su}} \in \mathbb{C}^{N_R \times N_T}$ is the point-to-point mmWave MIMO channel matrix, $\mathbf{P} \in \mathbb{C}^{N_T \times K}$ and $\mathbf{s} \in \mathcal{A}^K$ are the precoder matrix and information symbol vector, respectively, where $\mathcal{A}$ is the selected constellation such as PSK or QAM and K is the number transmitted symbols. Moreover, ρ is the average transmit power and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_R})$ is the additive white

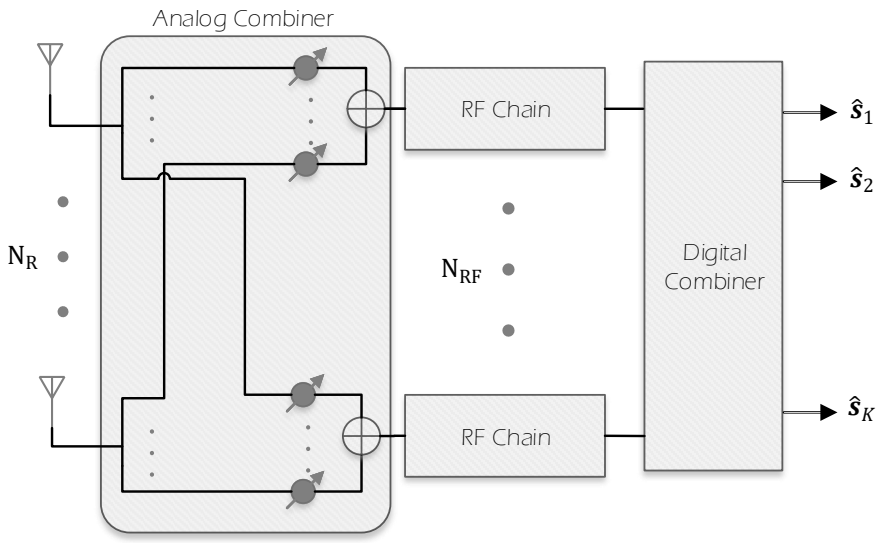


Fig. 5.1: Conventional hybrid combiner structure.

Gaussian noise (AWGN) vector. We consider the following channel model for mmWave massive-MIMO with sparse scattering environments [22, 26, 74]:

$$\mathbf{H}_{\text{su}} = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^{L} g^l \mathbf{a}_r(\phi_r^l)\mathbf{a}_t(\phi_t^l)^H, \tag{5.2}$$

where, $g^l \sim \mathcal{CN}(0,1)$ is the complex gain of the l^{th} path, $\mathbf{a}_r$ and $\mathbf{a}_t$ are the antenna array responses of receiver and transmitter, respectively. ϕ_r^l and ϕ_t^l are the arrival and departure angles and have uniform distribution over $[0, 2\pi)$. The array response for the widely-used uniform linear configuration is given by,

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{N_R}}[1, e^{jkd\sin(\phi)}, ..., e^{jkd(N_R-1)\sin(\phi)}], \tag{5.3}$$

where $k = 2\pi/\lambda$, and for wavelength of λ we let $d = \lambda/2$.

Multi-User: For K single antenna users transmitting simultaneously, the received signal at the BS can be written as,

$$\mathbf{y}_{\text{mu}} = \mathbf{H}_{\text{mu}}\mathbf{G}\mathbf{s} + \mathbf{n}, \tag{5.4}$$

where $\mathbf{G} = diag\{\sqrt{\rho_1}, ..., \sqrt{\rho_K}\}$ and ρ_k is the average transmit power of the k^{th} user. Under statistical channel inversion power control scheme [27], we have $\rho = \rho_k$ for $k = 1, 2, ..., K$. The channel matrix $\mathbf{H}_{\text{mu}} \in \mathbb{C}^{N_R \times K}$ can be expressed as,

$$\mathbf{H}_{\text{mu}} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_K], \tag{5.5}$$

where $\mathbf{h}_k$ is the uplink fading channel between the k^{th} user and BS. Subsequently, $\mathbf{s} = [s_1, s_2, ..., s_K]^T$ is the symbol vector where s_k denotes the transmitted symbol of the k^{th} user.

The mmWave channel vector of the k^{th} user can be modeled as,

$$\mathbf{h}_k = \sqrt{\frac{N_R}{L_K}} \sum_{l=1}^{L_k} g^{k,l} \mathbf{a}_r(\phi^l), \tag{5.6}$$

where $g^{k,l} \sim \mathcal{CN}(0, p_{k,l})$ is the complex gain of the l^{th} path and,

$$\frac{1}{L_k} \sum_{l=1}^{L_K} p_{k,l} = 1, \tag{5.7}$$

for normalization purposes.

Hybrid Analog/Digital Beamforming: For a transmitter with limited number of RF chains, i.e., $N_{RF} \ll N_T$, the received signal after hybrid precoding is given by (5.1) for,

$$\mathbf{P} = \mathbf{P}_{\text{A}}\mathbf{P}_{\text{D}}, \tag{5.8}$$

where $\mathbf{P}_{\text{D}} \in \mathbb{C}^{N_{RF} \times K}$ is the digital precoder and $\mathbf{P}_{\text{A}} \in \mathbb{U}^{N_R \times N_{RF}}$ is the analog precoder with,

$$\mathbb{U} = \{z \in \mathbb{C} \colon |z| = 1\}. \tag{5.9}$$

To formulate the combining scheme for both SU and MU cases, let us define the total

equivalent channel at the BS for $\mu = \{\text{su}, \text{mu}\}$, as,

$$\mathbf{H}_{\mu}^{eq} \triangleq \begin{cases} \mathbf{H}_{\text{su}}\mathbf{P} & \mu = \text{su} \\ \mathbf{H}_{\text{mu}} & \mu = \text{mu} \end{cases}. \tag{5.10}$$

The received signal at the BS can be written as

$$\mathbf{y}_{\mu} = \sqrt{\rho}\mathbf{H}_{\mu}^{eq}\mathbf{s} + \mathbf{n}. \tag{5.11}$$

In hybrid combining structure, to reduce the dimension of the received signal to N_{RF}, the signal first goes through the analog combiner $\mathbf{W}_{\text{A}} \in \mathbb{U}^{N_R \times N_{RF}}$ as shown in Fig. 5.1:

$$\mathbf{x}_{\mu} = \mathbf{W}_{\text{A}}^{H}\mathbf{y}_{\mu}. \tag{5.12}$$

The RF chains then convert the analog vector $\mathbf{x}_{\mu}$ to digital which in turn is used to estimate the transmitted symbols by

$$\hat{\mathbf{s}}_{\text{HQ}} = \mathbf{W}_{\text{D}}^{H}\mathbf{x}_{\mu}, \tag{5.13}$$

where $\mathbf{W}_{\text{D}} \in \mathbb{C}^{N_{RF} \times K}$ is the digital combiner.

In FD systems, since there are no limitations on the number of RF chains, the transmitted symbols can be directly estimated by the digital combiner $\mathbf{W}_{\text{FD}} \in \mathbb{C}^{N_R \times K}$ as,

$$\hat{\mathbf{s}}_{\text{FD}} = \mathbf{W}_{\text{FD}}^{H}\mathbf{y}_{\mu}. \tag{5.14}$$

5.3 Achieving FD Performance in Hybrid Structure

In this section, we first review the motivation for designing HBF structures that match the performance of arbitrary FD combiners. We then present a new hybrid structure for realizing any given FD combiner. Let us start by revisiting two well-known facts:

- Any hybrid combining scheme can be realized by FD combining because the hybrid combiner comprises of analog and digital combiners $\mathbf{W}_{\text{A}}$ and $\mathbf{W}_{\text{D}}$ which can be modeled by the digital combiner $\mathbf{W}_{\text{FD}} = \mathbf{W}_{\text{A}}\mathbf{W}_{\text{D}}$. In a more general case, any analog operation (even a non-linear one) can be simulated by the digital processor and thus,

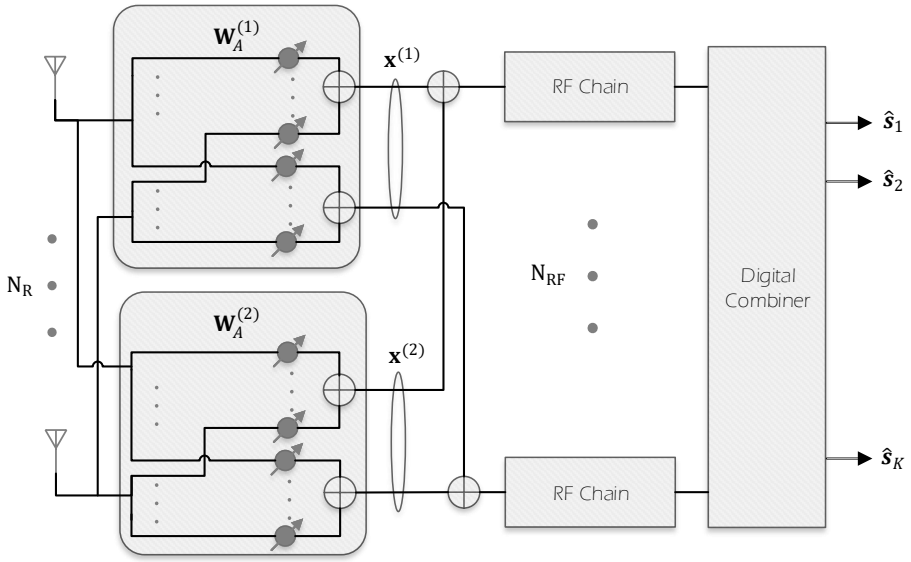


Fig. 5.2: Proposed hybrid structure.

there does not exist a hybrid combining scheme that can outperform the FD combining. Consequently, *the best possible design for hybrid combiner is to match the performance of FD one, i.e., realizing any given FD combiner.*

- In FD systems we have $\text{rank}(\mathbf{W}_{\text{FD}}) \leq K$. For hybrid combiners, however, we can write $\text{rank}(\mathbf{W}_{\text{A}}\mathbf{W}_{\text{D}}) \leq N_{RF}$. Assuming that the digital combiner has the maximum rank, i.e., $\text{rank}(\mathbf{W}_{\text{FD}}) = K$. Hence, *the minimum number of required RF chains for a hybrid combiner to match the performance of the FD combiner is* $N_{RF} = K$.

In the rest of this chapter, we let $N_{RF} = K$ to ensure that the minimum required hardware is utilized. The following theorem paves the way for hybrid realization of digital combiners.

Theorem 5.1. *Any given complex matrix* $\mathbf{A} \in \mathbb{C}^{N_R \times K}$ *can be written as a scaled sum of two matrices* $\mathbf{R}^{(1)}, \mathbf{R}^{(2)} \in \mathbb{U}^{N_R \times K}$*, with unit modulus entries, i.e.,*

$$\mathbf{A} = c(\mathbf{R}^{(1)} + \mathbf{R}^{(2)}), \tag{5.15}$$

for some $c \geq \frac{1}{2}|vec(\mathbf{A})|_\infty$.

Proof. It is sufficient to show that for all $p = 1, 2, ..., N_R$ and $q = 1, 2, ..., K$, we have,

$$A(p,q) = c\big(R^{(1)}(p,q) + R^{(2)}(p,q)\big). \tag{5.16}$$

Since $\mathbf{R}^{(1)}, \mathbf{R}^{(2)} \in \mathbb{U}^{N_R \times K}$, the above equation can be further written as,

$$A(p,q) = c\big(e^{j\theta^1_{p,q}} + e^{j\theta^2_{p,q}}\big), \tag{5.17}$$

where $e^{\theta^1_{p,q}}$ and $e^{\theta^2_{p,q}}$ represent the entry on the p^{th} row and q^{th} column of $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(2)}$, respectively. Since $2c$ is greater than the absolute value of all the entries of $\mathbf{A}$, from Lemma 3.1, there exist non-unique $\theta^1_{p,q}$ and $\theta^1_{p,q}$ such that (5.17) holds which proves the theorem. □

In order to achieve the same performance of FD system in hybrid structure, the following must hold

$$\hat{\mathbf{s}}_{\mathrm{HY}} = \hat{\mathbf{s}}_{\mathrm{FD}} = \mathbf{W}^H_{\mathrm{FD}}\mathbf{y}_\mu. \tag{5.18}$$

The following theorem provides a new simple hybrid structure for realizing any given FD combiner.

Theorem 5.2. *Any given FD combiner* $\mathbf{W}_{FD} \in \mathbb{C}^{N_R \times K}$ *can be realized in HBF structure shown in Fig. 5.2 such that* $\mathbf{W}_{FD} = (\mathbf{W}^{(1)}_A + \mathbf{W}^{(2)}_A)\mathbf{W}_D$ *where* $\mathbf{W}_D \in \mathbb{C}^{N_{RF} \times K}$ *is a digital combiner and* $\mathbf{W}^{(1)}_A$,$\mathbf{W}^{(2)}_A \in \mathbb{U}^{N_R \times N_{RF}}$ *are analog combiners.*

Proof. The estimated signal in hybrid structure shown in Fig. 5.2 can be written as,

$$\hat{\mathbf{s}}_{\mathrm{HY}} = \mathbf{W}^H_{\mathrm{D}}(\mathbf{x}^{(1)}_\mu + \mathbf{x}^{(2)}_\mu), \tag{5.19}$$

where $\mathbf{x}^{(1)}_\mu = {\mathbf{W}^{(1)}_{\mathrm{A}}}^H \mathbf{y}_\mu$ and $\mathbf{x}^{(2)}_\mu = {\mathbf{W}^{(2)}_{\mathrm{A}}}^H \mathbf{y}_\mu$, thus,

$$\hat{\mathbf{s}}_{\mathrm{HY}} = \mathbf{W}^H_{\mathrm{D}}({\mathbf{W}^{(1)}_{\mathrm{A}}}^H \mathbf{y}_\mu + {\mathbf{W}^{(2)}_{\mathrm{A}}}^H \mathbf{y}_\mu), \tag{5.20}$$

which can be further simplified to,

$$\hat{\mathbf{s}}_{\mathrm{HY}} = \mathbf{W}^H_{\mathrm{D}}(\mathbf{W}^{(1)}_{\mathrm{A}} + \mathbf{W}^{(2)}_{\mathrm{A}})^H \mathbf{y}_\mu. \tag{5.21}$$

Algorithm 1 Fully-Digital Realization in Hybrid Structure

Given: $\mathbf{H}_{\mu}^{eq}$
1. Calculate the optimal FD combiner $\mathbf{W}_{\mathrm{FD}}$.
2. $c = \frac{1}{2}\max_{i,j}|w_{p,q}|$.
3. Construct the baseband combiner as $\mathbf{W}_{\mathrm{D}} = c\mathbf{I}_K$.
4. Calculate $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$ from (5.26).

Outputs: $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$, $\mathbf{W}_{\mathrm{D}}$.

Using Theorem 5.1, there exist analog matrices $\mathbf{W}_{\mathrm{A}}^{(1)}$ and $\mathbf{W}_{\mathrm{A}}^{(2)}$ such that $\mathbf{W}_{\mathrm{FD}} = c(\mathbf{W}_{\mathrm{A}}^{(1)} + \mathbf{W}_{\mathrm{A}}^{(2)})$. By setting $\mathbf{W}_{\mathrm{D}} = c\mathbf{I}_K$, the FD combiner can be expressed as $\mathbf{W}_{\mathrm{FD}} = \tilde{\mathbf{W}}_{\mathrm{A}}\mathbf{W}_{\mathrm{D}}$, where $\tilde{\mathbf{W}}_{\mathrm{A}} = (\mathbf{W}_{\mathrm{A}}^{(1)} + \mathbf{W}_{\mathrm{A}}^{(2)})$, which brings us to

$$\mathbf{W}_{\mathrm{FD}} = (\mathbf{W}_{\mathrm{A}}^{(1)} + \mathbf{W}_{\mathrm{A}}^{(2)})\mathbf{W}_{\mathrm{D}}. \tag{5.22}$$

Consequently, by substituting the above equation in (5.21), we have

$$\hat{\mathbf{s}}_{\mathrm{HY}} = \mathbf{W}_{\mathrm{FD}}^{H}\mathbf{y}_{\mu}, \tag{5.23}$$

which results (5.18) and proves the theorem. □

In the proposed HBF structure, the constant modulus constraint is lifted which makes it possible to achieve the same performance as FD systems for various reception modes and scenarios. In the conventional hybrid structure, the following constraints are added to the optimization procedures:

$$\mathbf{W}_i = \mathbf{W}_{\mathrm{A}}\mathbf{W}_{\mathrm{D}} \tag{5.24a}$$

$$\mathbf{W}_{\mathrm{A}} \in \mathbb{U}^{N_R \times N_{RF}}, \tag{5.24b}$$

for $i = \{\mathrm{su}, \mathrm{mu}\}$ which make the beamformer design optimizations non-convex and therefore, extremely difficult so solve. Many sub-optimal solutions are studied in the literature [21,24–29,60,65,75]. Nevertheless, simulations results in [21,24,28,29,60] confirm that FD realization achieves the best performance as it is also discussed in Section 6.3.

FD Combiner Realization with HBF: According to Theorem. 5.2, for any given FD combiner $\mathbf{W}_{\mathrm{FD}} = \mathbf{W}_i \in \mathbb{C}^{N_R \times K}$, there exist matrices $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$ and $\mathbf{W}_{\mathrm{D}}$ that match the performance of $\mathbf{W}_{\mathrm{FD}}$.

Note that for a given FD matrix $\mathbf{W}_{\mathrm{FD}}$ there are non-unique matrices $\mathbf{W}_{\mathrm{A}}^{(1)}$,$\mathbf{W}_{\mathrm{A}}^{(2)}$ and $\mathbf{W}_{\mathrm{D}}$ such that $\mathbf{W}_{\mathrm{FD}} = (\mathbf{W}_{\mathrm{A}}^{(1)} + \mathbf{W}_{\mathrm{A}}^{(2)})\mathbf{W}_{\mathrm{D}}$ holds. Here, we present an easy way to obtain $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$ and $\mathbf{W}_{\mathrm{D}}$. By denoting the polar representation of $\mathbf{W}_{\mathrm{FD}}$'s elements as $\mathbf{W}_{\mathrm{FD}}(p,q) = |w_{p,q}| \exp(j\vartheta_{p,q})$, the digital combiner is given by $\mathbf{W}_{\mathrm{D}} = c\mathbf{I}_K$ where,

$$c = \frac{1}{2} \max_{i,j} |w_{p,q}|. \tag{5.25}$$

The elements of analog combiners $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$ are calculated by,

$$\mathbf{W}_{\mathrm{A}}^{(1)}(p,q) = e^{j\left(\vartheta_{p,q} + \cos^{-1}\left(\frac{|w_{p,q}|}{2c}\right)\right)} \tag{5.26a}$$

$$\mathbf{W}_{\mathrm{A}}^{(2)}(p,q) = e^{j\left(\vartheta_{p,q} - \cos^{-1}\left(\frac{|w_{p,q}|}{2c}\right)\right)}. \tag{5.26b}$$

By simple mathematical manipulation one can easily check the validity of presented solutions in (5.25) and (5.26) (for more information see [22]). The proposed hybrid realization of FD combiners is summarized in Algorithm 1.

5.4 Finite Resolution Phase-Shifters

In this section, we consider the HBF design with finite-resolution phase-shifters for massive-MIMO receivers. In previous sections, we assumed that infinite resolution phase-shifters are used in the RF beamformer, i.e., the elements of RF beamformers can have any arbitrary phase angles. Since such components are expensive, and due to the fact that the number of phase-shifters in hybrid structure is proportional to the number of antennas, finite-resolution phase-shifters are preferred in many scenarios. In this section, we consider RF beamforming with finite resolution phase-shifters, i.e., $\mathbf{W}_{\mathrm{A}} \in \mathbb{U}_M^{N_R \times N_{RF}}$ with

$$\mathbb{U}_M = \{e^{j2\pi m/M} : m = 0, 1, \ldots, M-1\}, \tag{5.27}$$

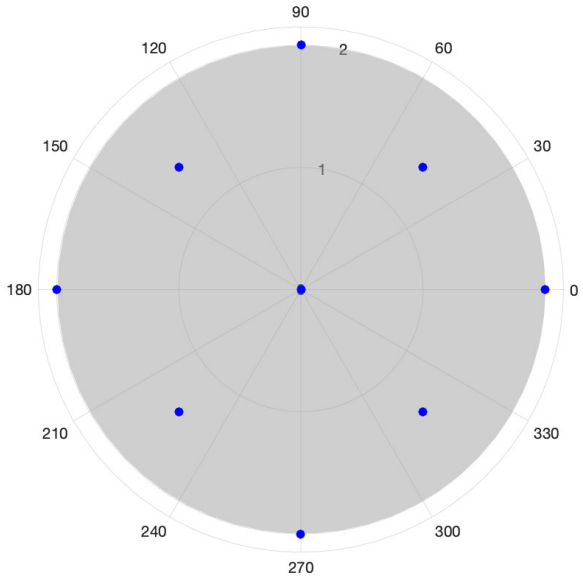


Fig. 5.3: The shaded area represents the disk $\mathbb{D}$ and the dots show the polar grid $\mathbb{G}_4$.

and M is the number of realizable phase angles where M is usually a power of 2. However, in this study, our only assumption is that $M \geq 2$.

From Theorem 5.1, any given matrix $\mathbf{A} \in \mathbb{D}^{N_R \times K}$ for $\mathbb{D} = \{z \in \mathbb{C} \colon |z| \leq 1\}$ can be written as,

$$\mathbf{A} = \frac{1}{2}(\mathbf{R}^{(1)} + \mathbf{R}^{(2)}), \tag{5.28}$$

where $\mathbf{R}^{(1)}$ and $\mathbf{R}^{(1)}$ are implemented with infinite resolution phase-shifters, i.e., $\mathbf{R}^{(1)}, \mathbf{R}^{(2)} \in \mathbb{U}^{N_R \times K}$. However, if phase-shifters with resolution M are used, i.e., $\mathbf{R}_M^{(1)}, \mathbf{R}_M^{(2)} \in \mathbb{U}_M^{N_R \times K}$, the equality in (5.28) cannot be guaranteed. In this case, the RF beamformers can be designed by solving

$$\min_{\mathbf{R}_M^{(1)}, \mathbf{R}_M^{(2)}} \|\mathbf{A} - \frac{1}{2}(\mathbf{R}_M^{(1)} + \mathbf{R}_M^{(2)})\|_F^2, \tag{5.29}$$

which requires finding the closest point in the polar grid $\mathbb{G}_M = \{e^{j2\pi m_1/M} + e^{j2\pi m_2/M} \colon m_1, m_2 = 0, 1, \ldots, M-1\}$, to a given point in $\mathbb{D}$. In Fig. 5.3, the disk $\mathbb{D}$ and the polar gird $\mathbb{G}_4$ are illustrated. In this case, each phase-shifter can take 4 possible values and thus, there are 16 possible combination of the two phase-shifters corresponding to $e^{j2\pi m_1/4}$ and $e^{j2\pi m_2/4}$. However, it can be seen form Fig. 5.3, that $\mathbb{G}_4$ contains only 10 points which shows some

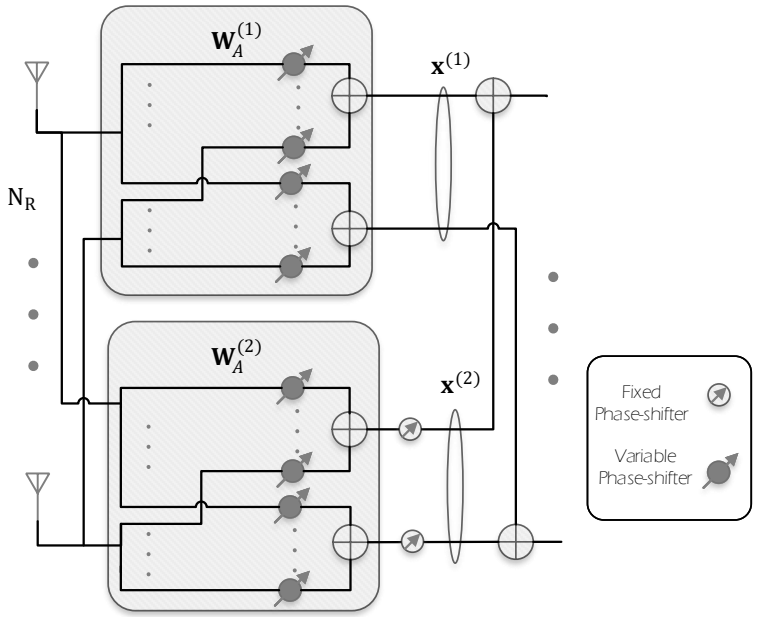


Fig. 5.4: Proposed hybrid structure with finite-resolution phase-shifters.

of the points are overlapping. Thus, 4 more points can be included which improves the performance of the hybrid system. Note that for $M > 4$ the number of overlapping point increase. Moreover, we do not have control over the distribution of points in $\mathbb{G}_M$ and it would be of interest to be able to optimize the distribution of the points to improve system performance. This motivates us to investigate whether it is possible to have alternative polar grids for designing the RF beamformers and if so what is the best configuration for points in such polar grid.

By taking advantage of an additional degree of freedom, we can solve the above issue with finite resolution phase-shifters. The main practical constraint for finite resolution phase-shifters is that the phase-shifters can realize only M different phase angles. However, it is possible to have fixed phase offset between the phase-shifters.

While, in principle, it is possible to consider various phase-offsets for all of the phase-shifters, for the sake of simplicity and practicality, we only introduce a phase offset between $\mathbf{R}_M^{(1)}$ and $\mathbf{R}_M^{(2)}$. This can be implemented by either using phase-shifters which are manufac-

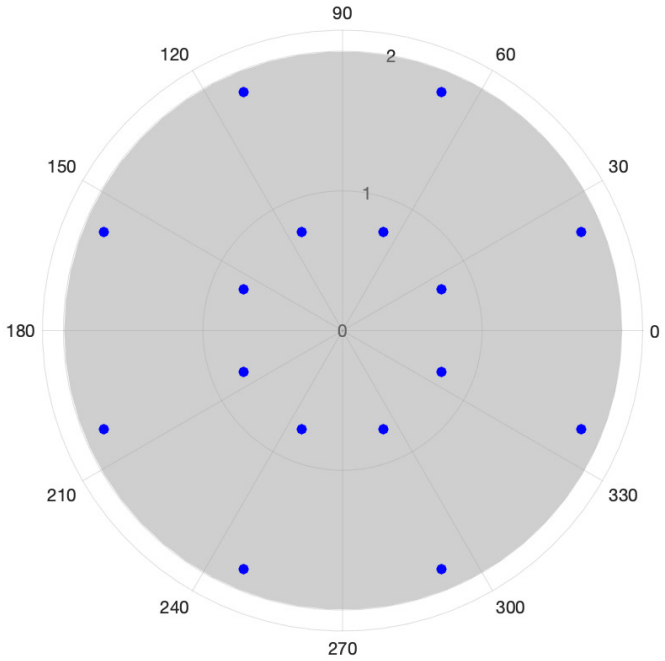


Fig. 5.5: The shaded area represents the disk $\mathbb{D}$ and the dots show the polar grid $\mathbb{G}_4^{\pi/4}$.

tured with a fixed phase difference or placing N_{RF} fixed-valued phase-shifter after the RF bemformer $\mathbf{R}_M^{(2)}$ as shown in Fig. 5.4.

As mentioned earlier, we wish to change the distribution of points in the polar gird $\mathbb{G}_M$ to improve the system performance. Since, the phase offset is selected and hard coded to the system alongside the resolution of the phase-shifters and cannot be changed, we focus on finding the optimal phase-offset β for any given resolution M such that the maximum error between a given matrix $\mathbf{A} \in \mathbb{D}^{N_R \times K}$ and its hybrid realization $\mathbf{R}_M^{(1)}, \mathbf{R}_M^{(2)}$ is minimized. Thus, the phase offset can be obtained by,

$$\beta^* = \arg\min_{\beta} \max_{\mathbf{A}} \min_{\mathbf{R}_M^{(1)}, \mathbf{R}_M^{(2)}} \quad \|2\mathbf{A} - (\mathbf{R}_M^{(1)} + e^{j\beta}\mathbf{R}_M^{(2)})\|_F^2 \tag{5.30a}$$

$$\text{subject to} \quad \mathbf{R}_M^{(1)}, \mathbf{R}_M^{(2)} \in \mathbb{U}_M^{N_R \times K} \tag{5.30b}$$

$$\mathbf{A} \in \mathbb{D}^{N_R \times K}, \tag{5.30c}$$

which minimizes the maximum error between the given FD beamformer and the effective

hybrid beamformer, i.e., $(\mathbf{R}_M^{(1)} + e^{j\beta}\mathbf{R}_M^{(2)})$.

Lemma 5.1. *The following optimization is equivalent to* (5.30)*,*

$$\beta^* = \arg\min_{\beta} \quad \max_{a} \quad \min_{m_1,m_2} \quad |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2 \tag{5.31a}$$

$$\text{subject to} \qquad \beta \in [0, 2\pi] \tag{5.31b}$$

$$a \in \{z \in \mathbb{C} \colon |z| \leq 1\} \tag{5.31c}$$

$$m_1, m_2 \in \{1, 2, \ldots, M\}. \tag{5.31d}$$

Proof. By expanding (5.30a) as, $\sum_{i=1}^{N_R}\sum_{j=1}^{K}|2A(i,j) - (R_M^{(1)}(i,j) + \beta R_M^{(2)}(i,j))|^2$, we can see the optimization parameters are independent and can be decoupled by rewriting (5.30) as,

$$\begin{aligned}\beta^* = \arg\min_{\beta} \quad \sum_{i=1}^{N_R}\sum_{j=1}^{K} \max_{A(i,j)} \quad \min_{R_M^{(1)}(i,j), R_M^{(2)}(i,j)} \\ |2A(i,j) - (R_M^{(1)}(i,j) + \beta R_M^{(2)}(i,j))|^2,\end{aligned} \tag{5.32}$$

subject to the constraints (5.39b) to (5.39d). Since the max-min optimization has the same form for all i, j and admits the same function for β, we can write the above problem as (5.31). □

The optimization (5.31) can be solved numerically, but since it is a mixed-integer problem, the optimization procedure is cumbersome and inefficient. Therefore, in what follows, we obtain a closed form solution based on a tight upper bound of (5.32). The closed-form solution is equal to the result of the numerical optimizations for every value of M we tested, i.e., $M = 2^k$ for $k = 1, 2, \ldots, 7$.

By defining,

$$f(\beta) = \max_{a} \min_{m_1,m_2} |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2 \tag{5.33a}$$

$$\hat{f}(\beta) = \min_{m_1,m_2} \max_{a} |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2, \tag{5.33b}$$

from the max-min inequality, we can write $f(\beta) \leq \hat{f}(\beta)$. Consequently, we have,

$$f(\beta^*) \leq \hat{f}(\hat{\beta}^*) = \min_{\beta} \min_{m_1,m_2} \max_{a} |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2. \tag{5.34}$$

Lemma 5.2. *The following optimization has at least one solution in set*

$$\{e^k | k = \frac{j}{2}\tau(m_1 + m_2) + \beta\}$$

where $\tau = 2\pi/M$,

$$\max_{a} \quad |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2, \tag{5.35a}$$

$$\text{subject to} \quad a \in \{z \in \mathbb{C} \colon |z| \leq 1\}. \tag{5.35b}$$

Proof. Since the objective function is convex with respect to a and the feasible set of a is convex and compact, the maximum of the function occurs on the boundary of the feasible set. Consequently, we can rewrite the above problem as,

$$\max_{a} \quad |2e^{j\alpha} - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2, \tag{5.36a}$$

$$\text{subject to} \quad \alpha \in [0, 2\pi]. \tag{5.36b}$$

The objective function of (5.36) can be expanded as,

$$\begin{aligned}\Delta(\alpha) = (2\cos(\alpha) - \cos(\tau m_1) - \cos(\tau m_2 + \beta))^2 \times \\ (2\sin(\alpha) - \sin(\tau m_1) - \sin(\tau m_2 + \beta))^2.\end{aligned} \tag{5.37}$$

To find the optima, by taking the derivitive of $\Delta(\alpha)$ and setting it to zero, we arrive at,

$$\begin{aligned}\frac{d\Delta(\alpha)}{d\alpha} &= \cos(\alpha)\big(\sin(\tau m_1) + \sin(\tau m_2 + \beta)\big) \\ &\quad - \cos(\alpha)\big(\sin(\tau m_1) + \sin(\tau m_2 + \beta)\big) = 0,\end{aligned} \tag{5.38}$$

which has a solution for $\alpha = \frac{\tau(m_1+m_2)+\beta}{2}$: □

Theorem 5.3. *The optimal solution for the following optimization problem is* $\hat{\beta}^* = \pi/M$.

$$\hat{\beta}^* = \arg\min_{\beta} \min_{m_1,m_2} \max_{a} |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2, \tag{5.39a}$$

$$\text{subject to} \quad \beta \in [0, 2\pi] \tag{5.39b}$$

$$a \in \{z \in \mathbb{C} \colon |z| \leq 1\} \tag{5.39c}$$

$$m_1, m_2 \in \{1, 2, \ldots, M\}. \tag{5.39d}$$

Proof. From Lemma 5.2, the objective function (5.39a) is maximized by at least one point in set $\mathcal{E} = \{\mathrm{e}^k | k = \frac{j}{2}2\pi/M(m_1 + m_2) + \beta\}$. Moreover, since $f(\beta, \alpha) = |2a - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2$ is periodic for β and α with period π/M, without loss of generality, the feasible domain of β and a can be reduced to,

$$\beta \in [0, 2\pi/M] \tag{5.40a}$$

$$a \in \{e^{j\alpha} \colon \alpha \in [0, 2\pi/M]\}. \tag{5.40b}$$

Subsequently, the points in $\mathcal{E}$ that satisfy (5.40b), are $a^* = \exp(\frac{j}{2}\tau(m_1 + m_2) + \beta)$ where $m_1 + m_2 = 1$ and $m_1 + m_2 = 0$ for $\beta \leq \pi/M$, and $\pi/M \leq \beta \leq 2\pi/M$, respectively. Thus, we have to consider each case separately.

- For $\pi/M \leq \beta \leq 2\pi/M$, we have $m_1 = m_2 = 0$, therefore, $a^* = \exp(\frac{j\beta}{2})$ which brings us to the following problem:

$$\hat{\beta}^* = \arg\min_{\beta} \quad |2e^{j\beta/2} - (1 + e^{j\beta})|^2, \tag{5.41a}$$

$$\text{subject to} \quad \beta \in [\pi/M, 2\pi/M]. \tag{5.41b}$$

  By mathematical manipulation the objective function can be written as $6 - 8\cos(\beta/2) + 2\cos(\beta)$, which is non-decreasing function for $\beta \in [\pi/M, 2\pi/M]$[2], hence, $\hat{\beta}^* = \pi/M$.

- For $0 \leq \beta \leq \pi/M$, we have $m_1 + m_2 = 1$, therefore, $a^* = \exp(\frac{j(\beta+\pi/M)}{2})$. We fist must

[2]Note that $M \geq 2$ as the resolution of phase-shifter is at least 2.

Algorithm 2 Proposed Finite Resolution Hybrid Design

Given: $\mathbf{W}_{\mathrm{FD}}$
1. Calculate the analog beamformers $\mathbf{W}_{\mathrm{A}}^{(1)}$ and $\mathbf{W}_{\mathrm{A}}^{(2)}$ by solving: $\min_{\mathbf{W}_{\mathrm{A}}^{(1)},\mathbf{W}_{\mathrm{A}}^{(2)}} \| \frac{\mathbf{W}_{\mathrm{FD}}}{|\mathrm{vec}(\mathbf{W}_{\mathrm{FD}})|_\infty} - 1/2(\mathbf{W}_{\mathrm{A}}^{(1)} + e^{j\beta}\mathbf{W}_{\mathrm{A}}^{(2)}\|_F^2$.
2. Update the phase-shifters according to $\mathbf{W}_{\mathrm{A}}^{(1)}$ and $\mathbf{W}_{\mathrm{A}}^{(2)}$.
3. Update the baseband beamformer using (5.44).

Outputs: $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$, $\mathbf{W}_{\mathrm{D}}$.

find m_1 and m_2 from the following problem:

$$\min_{m_1,m_2} \quad | \; 2e^{\frac{j}{2}(\beta+\pi/M)} - (e^{j\frac{2\pi}{M}m_1} + e^{j\beta}e^{j\frac{2\pi}{M}m_2})|^2, \tag{5.42a}$$

$$\text{subject to} \quad m_1, m_2 = 1 \tag{5.42b}$$

$$\beta \in [\pi/M, 2\pi/M]. \tag{5.42c}$$

By comparing the only two possible feasible points: $m_1 = 0, m_2 = 1$ and $m_1 = 1, m_2 = 0$, it can be checked that the latter is the solution, i.e., $m_1 = 1, m_2 = 0$. Now, β can be obtained from,

$$\hat{\beta}^* = \arg\min_{\beta} \quad |2e^{\frac{j}{2}(\beta+\pi/M)} - (e^{j\frac{2\pi}{M}} + e^{j\beta})|^2, \tag{5.43a}$$

$$\text{subject to} \quad \beta \in [0, \pi/M]. \tag{5.43b}$$

By mathematical manipulations, the objective function can be written as $6-8\cos\big((\beta-2\pi/M)/2\big) + 2\cos\big((\beta-2\pi/M)/2\big)$, which is non-decreasing for $\beta \in [0, \pi/M]$, thus, $\hat{\beta}^* = \pi/M$.

□

In our simulations, we considered $R = 4$ where the optimal phase-difference is calculated as $\beta^* = \pi/4$. It is worth noting that solving the optimization (5.31) with numerical techniques also results in $\beta^* = \pi/4$. In Fig. 5.5, possible values of the analog beamformer entries is shown. It can be seen that as expected 16 points corresponding to the each combination of the two phase-shifters corresponding to $e^{j2\pi m_1/4}$ and $e^{j2\pi m_2/4}$ are scattered such that the minimum distance between two adjacent points is maximized.

Subsequently, for a given β, the equivalent analog bemformer is written as, $\mathbf{A}_{\mathbb{G}} = \mathbf{R}_M^{(1)} + \mathbf{R}_M^{(2)}$, therefore, in the proposed structure, the received signal is first combined by $\mathbf{A}_{\mathbb{G}} \in {\mathbb{G}_M^{\beta}}^{N_R \times K}$ where $\mathbb{G}_M^{\beta} = \{e^{j2\pi m_1/M} + e^{j(\beta+2\pi m_2/M)} \colon m_1, m_2 = 0, 1, \ldots, M-1\}$. The baseband beamforming is then performed which is obtained by,

$$\mathbf{W}_{\mathrm{D}} = (\mathbf{A}_{\mathbb{G}}^H(\mathbf{H}\mathbf{P}_{\mathrm{FD}}{\mathbf{P}_{\mathrm{FD}}}^H\mathbf{H}^H + \sigma\mathbf{I}_{N_R})\mathbf{A}_{\mathbb{G}})^{-1}\mathbf{A}_{\mathbb{G}}^H\mathbf{H}\mathbf{P}_{\mathrm{FD}}. \tag{5.44}$$

The proposed finite resolution hybrid structure is summarized in Algorithm 3.

5.5 Application to Channel Estimation

In this section, we propose a novel hybrid channel estimation technique based on the hybrid structure in Section 6.3. We first discuss the problem formulation and then present a non-parametric hybrid channel estimation that achieves the same performance as of the FD MMSE channel estimation.

5.5.1 Problem Formulation

We consider channel estimation in a single-cell wireless system where a massive-MIMO BS serves K single-antenna users, i.e., mobile stations (MSs). The proposed scheme can be easily adapted to multiple antenna SU channel estimation by letting $N_T = K$ and considering each column of the channel matrix as the channel of a virtual single antenna user. The BS is equipped with N_R and N_{RF} receive antennas and RF chains, respectively.

In time division duplexing, the channel between the BS and users can be obtained by the uplink channel estimation and put into use in downlink as a result of the reciprocity. In this case, the transmitted training signals also known as pilots must be orthogonal to each other to avoid pilot contamination. Therefore, the duration of pilot transmission needs to be at least K time slots.

Let us consider the same assumptions as [97] where one MS sends pilots in each time slot and the remaining MSs stay silent. This technique exhibits similar performance to other orthogonal pilots in block-fading channels [97]. Consequently, it is sufficient to present channel estimation procedure for one user within the proposed hybrid structure.

Without loss of generality, we consider user number one and for simplicity, we drop the subscript 1 from $\mathbf{h}_1$. For each user, there are T measurements available at the BS for

Algorithm 3 Proposed Hybrid Channel Estimation

Given: $\mathbf{R}$,$\mathbf{r}_c$
1. Calculate the eigenvalue decomposition $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$.
2. Calculate $\mathbf{W}_c^{\mathrm{Opt}}$ from (5.49).
3. Construct $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$ from (5.26) for $\mathbf{W}_{\mathrm{FD}} = \mathbf{W}_c^{\mathrm{Opt}}$.
4. $c = \frac{1}{2}\max_{i,j}|w_{p,q}|$.
5. Calculate the baseband combiner as $\mathbf{W}_{\mathrm{D}} = c\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}$.
6. Estimated channel is given by (5.55)

Outputs: $\mathbf{W}_{\mathrm{A}}^{(1)}$, $\mathbf{W}_{\mathrm{A}}^{(2)}$, $\mathbf{W}_{\mathrm{D}}$.

T transmitted pilot signals. The t^{th} baseband measurement within the proposed hybrid structure can be written as,

$$\mathbf{r}_t = \varphi^* \mathbf{W}_{A_t}^{\mathrm{H}}(\sqrt{\rho}\mathbf{h}\varphi + \tilde{\mathbf{n}}_t), \tag{5.45}$$

where φ is the training symbol such that $\varphi\varphi^* = 1$ which is assumed to be the same signal for all T pilot transmissions and $\tilde{\mathbf{n}}_t$ is the AWGN vector of the t^{th} measurement for $t = \{1, 2, \ldots, T\}$. Letting $\mathbf{n}_t = \varphi^*\tilde{\mathbf{n}}_t$ which retains the distribution of $\tilde{\mathbf{n}}_t$ and concatenating all T received signals as $\mathbf{r}_c = [\mathbf{r}_1^{\mathrm{H}}, \mathbf{r}_2^{\mathrm{H}}, ..., \mathbf{r}_T^{\mathrm{H}}]^{\mathrm{H}}$, we can write,

$$\mathbf{r}_c = \sqrt{\rho}\mathbf{W}_c^{\mathrm{H}}\mathbf{h} + \tilde{\mathbf{W}}_c^{\mathrm{H}}\mathbf{n}_c, \tag{5.46}$$

where $\mathbf{W}_c = [\mathbf{W}_{A_1}, ..., \mathbf{W}_{A_T}]$, $\tilde{\mathbf{W}}_c = \mathrm{diag}\{\mathbf{W}_{A_1}, ..., \mathbf{W}_{A_T}\}$ and the concatenated noise vector is written as, $\mathbf{n}_c = [\mathbf{n}_1^{\mathrm{T}}, ..., \mathbf{n}_T^{\mathrm{T}}]^{\mathrm{T}}$.

5.5.2 Channel Estimation With Proposed Hybrid Structure

Similar to hybrid beamforming techniques, due to non-convexity of the analog combiner, hybrid channel estimation techniques also suffer from high complexity and poor performance. We make the same assumptions as [97] in which the authors presented a linear channel estimation based on MMSE. The spatially correlated MIMO channel model is considered, which is the typical model in MIMO scenarios specifically as we are dealing with mmWave massive-MIMO systems and the spatial channel covariance matrix is given by $\mathbf{R} = \mathbb{E}\{\mathbf{h}\mathbf{h}^H\}$. While in [97] it is assumed that $\mathbf{R}$ is full-rank our solution is not limited to

such channels and we let $\text{rank}(\mathbf{R}) = r$.

Using the Karhunen–Loeve representation, the channel vector can be expressed as,

$$\mathbf{h} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\eta, \tag{5.47}$$

for some $\eta \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_r)$ where $\mathbf{U} \in \mathbb{C}^{N_R \times r}$ and $\mathbf{\Lambda} \in \mathbb{C}^{r \times r}$ are obtained from eigenvalue decomposition $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$. We fix $T = \lfloor r/N_{RF} \rfloor$ and by substituting (5.47) in (5.46), we can estimate η by solving the following optimization problem:

$$\begin{aligned} \min_{\mathbf{W}_c} \quad & \mathbb{E}\big\{\|\hat{\eta} - \eta\|_2^2\big\} && \text{(5.48a)} \\ \text{subject to} \quad & \hat{\eta} = \sqrt{\rho}\mathbf{W}_c^{\mathrm{H}}\mathbf{h} + \tilde{\mathbf{W}}_c^{\mathrm{H}}\mathbf{n}_c && \text{(5.48b)} \\ & \mathbf{h} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\eta. && \text{(5.48c)} \end{aligned}$$

The above optimization does not directly admit the conventional optimal solution of a linear MMSE problem because a rearranged version of the optimization parameter, i.e., $\tilde{\mathbf{W}}_c$, is multiplied to the noise vector. The optimal combiner $\mathbf{W}_c^{\text{Opt}}$ is given by,

$$\mathbf{W}_c^{\text{Opt}} = \sqrt{\rho}\big(\frac{\rho}{\sigma^2}(\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}})^{\mathrm{H}}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}} + \mathbf{I}\big)^{-1}(\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}})^{\mathrm{H}}, \tag{5.49}$$

as shown bellow.

The optimization problem (5.48) has the well-known form of linear MMSE except for the fact that the noise vector $\mathbf{n}_c$ and estimation parameter vector η have different sizes and the corresponding matrices multiplied to these vectors are also different. By expanding (5.48a), we have,

$$\mathbb{E}\big\{\|\hat{\eta} - \eta\|_2^2\big\} = (\sqrt{\rho}\mathbf{W}_c^{\mathrm{H}}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\eta + \tilde{\mathbf{W}}_c^{\mathrm{H}}\mathbf{n}_c - \eta)^H(\sqrt{\rho}\mathbf{W}_c^{\mathrm{H}}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\eta + \tilde{\mathbf{W}}_c^{\mathrm{H}}\mathbf{n}_c - \eta). \tag{5.50}$$

To find the optimal solution, we can write the gradient of the above equation as

$$\frac{\partial}{\partial \mathbf{W}_c}\mathbb{E}\big\{\|\hat{\eta} - \eta\|_2^2\big\} = 2\rho\mathbf{W}_c^{\mathrm{H}}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}})^{\mathrm{H}} - 2\sqrt{\rho}(\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}})^{\mathrm{H}} + \sigma^2\frac{\partial}{\partial \mathbf{W}_c}\text{Tr}\big(\tilde{\mathbf{W}}_c\tilde{\mathbf{W}}_c^{\mathrm{H}}\big), \tag{5.51}$$

by invoking the fact that η and $\mathbf{n}_c$ are independent random vectors. To solve this equation, we must first write (5.51) in terms of only $\mathbf{W}_c$. By simple matrix manipulations, one can

verify that,

$$\mathrm{Tr}\big(\tilde{\mathbf{W}}_c\tilde{\mathbf{W}}_c^{\mathrm{H}}\big) = \mathrm{Tr}\big(\mathbf{W}_c\mathbf{W}_c^{\mathrm{H}}\big). \tag{5.52}$$

Therefore, by substituting (5.52) in (5.51) and solving $\mathbb{E}\big\{\|\hat{\eta}-\eta\|_2^2\big\} = 0$, the optimal solution is given by,

$$\mathbf{W}_c^{\mathrm{Opt}} = \sqrt{\rho}(\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}})^{\mathrm{H}}\big(\rho\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}(\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}})^{\mathrm{H}} + \sigma^2\mathbf{I}\big)^{-1}, \tag{5.53}$$

which by using the matrix inversion lemma [98] can be written as (5.49).

Implementation can be done using the same approach as Section. 5.3 and by setting $\mathbf{W}_{\mathrm{FD}} = \mathbf{W}_c^{\mathrm{Opt}}$, analog matrices $\mathbf{W}_{\mathrm{A}}^{(1)}$ and $\mathbf{W}_{\mathrm{A}}^{(2)}$ are calculated as (5.26). Finally, to estimate the channel, by setting $c = \frac{1}{2}\max_{i,j}|w_{p,q}|$, the digital combiner is designed as,

$$\mathbf{W}_{\mathrm{D}} = c\mathbf{U}\boldsymbol{\Lambda}^{\frac{1}{2}}. \tag{5.54}$$

The estimated channel is then obtained by,

$$\hat{\mathbf{h}} = \mathbf{W}_{\mathrm{D}}\mathbf{r}_c. \tag{5.55}$$

In Algorithm 2, the proposed channel hybrid estimation is summarized. For the multi-antenna SU, this technique can be similarly implemented to estimate each column of channel matrix. Each antenna of the multi-antenna user can be viewed as a separate user and therefore the MU channel estimation technique can be used for multi-antenna user by estimating the channel between each antenna of the user and BS.

5.6 Simulation Results

In this section, we present simulation results for different scenarios and compare the FD system with our proposed hybrid structure and recent hybrid designs the literature.

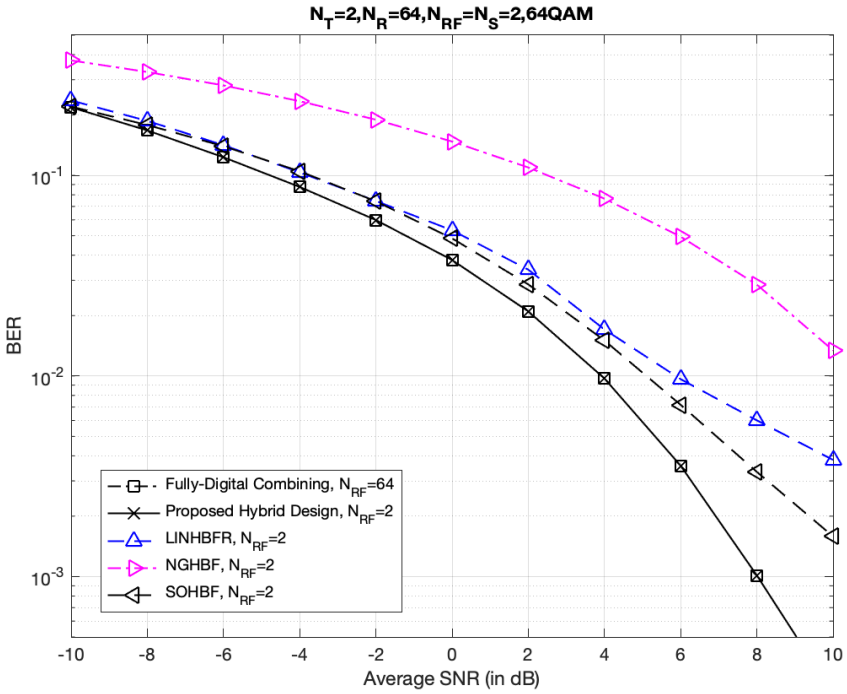


Fig. 5.6: BER versus SNR of different methods for a massive-MIMO BS with $N_R = 64$ antennas.

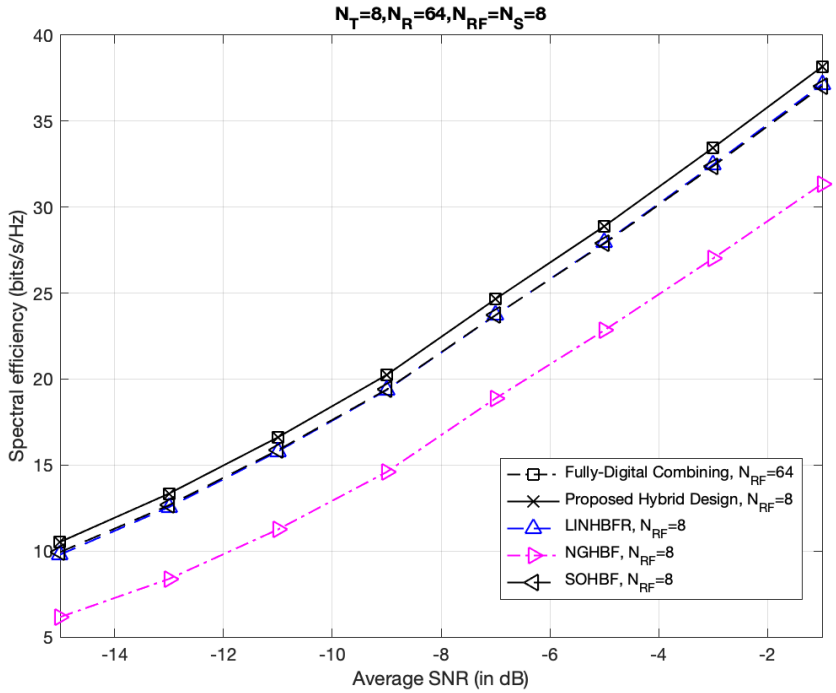


Fig. 5.8: Spectral efficiency versus SNR of different methods for a 8×64 massive-MIMO system.

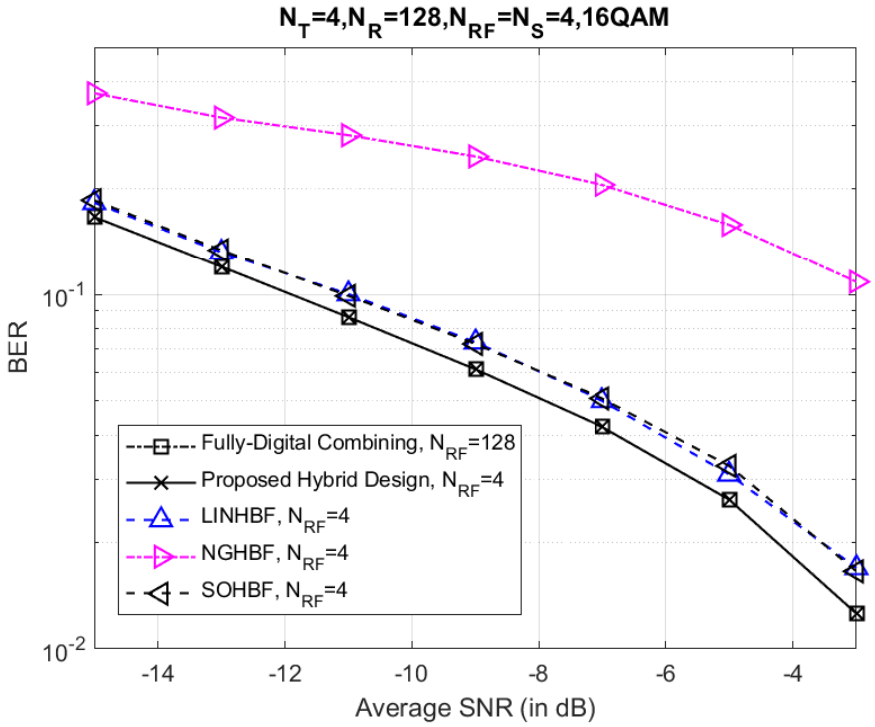


Fig. 5.7: BER versus SNR of different methods for massive-MIMO BS with $N_R = 128$ antennas.

First, we present simulation results for SU and MU uplink combining under perfect CSI where infinte resolution phase-shifters are used in the hybrid structure. We consider the mmWave channel models introduced in Section II-A with $L = 15$ for SU scenario where a uniform linear array configuration is used. The BER performance versus signal-to-noise ratio (SNR) (SNR=P_T/σ^2) for 2×64 and 4×128 point-to-point MIMO setups are depicted in Fig. 5.6 and 5.7, respectively. FD combining, our proposed hybrid realization as well as hybrid designs in [22, 75, 99], i.e., SOHBF, LINHBFR, and NGHBF, respectively, are compared by performing computer simulations.

Fig. 5.6 depicts simulation results for a single UE with $N_T = 2$ antennas where 64-QAM constellation is used with $K = 2$ symbols per transmission to the massive-MIMO BS with $N_R = 64$ antennas and $N_{RF} = 2$ RF chains. For a UE with $N_T = K = 4$ antennas,

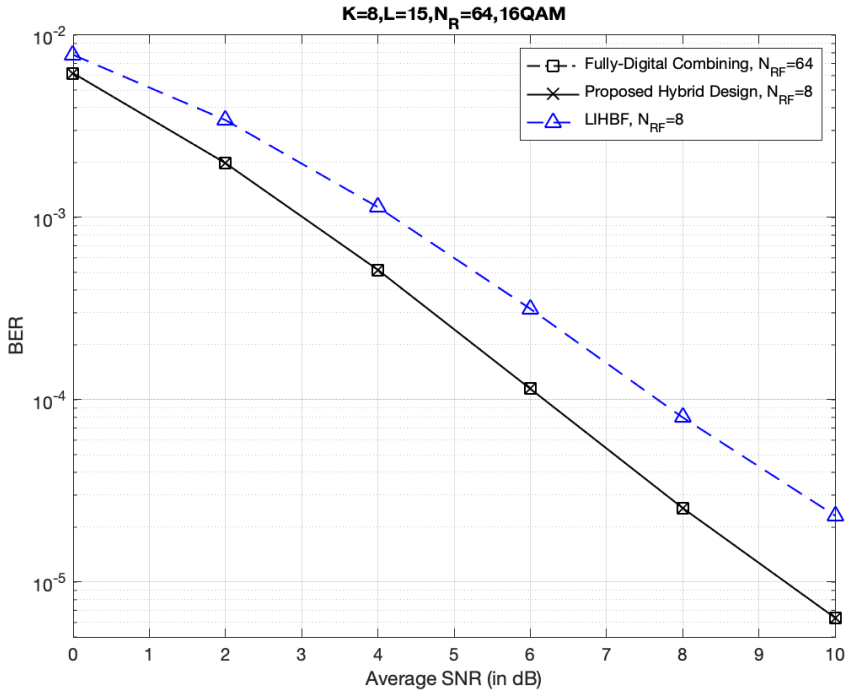


Fig. 5.10: BER versus SNR of FD combining and our design.

16 QAM constellation and BS with $N_R = 128$ receive antenna where $N_{RF} = 4$, BER performance is illustrated in Fig. 5.7. It can be seen that in both cases the proposed hybrid realization matches the performance of the FD systems while outperforming the existing hybrid designs. The FD systems require $N_{RF} = 64$ and 128 RF chains whereas the proposed design achieves the same performance with $N_{RF} = 2$ and 4 RF chains, respectively.

The spectral efficiency of our proposed hybrid realization is compared to that of the FD combiner and HBF designs in [22, 75, 99] (i.e., SOHBF, LINHBFR, and NGHBF, respectively) for $N_{RF} = N_T = K = 8$, $N_R = 64$ and $N_{RF} = N_T = K = 4$, $N_R = 128$ in Fig. 5.8 and Fig. 5.9, respectively. As expected, the proposed realization achieves the same rate as FD systems and has higher rate than existing designs.

For multi-user case, 16-QAM constellation and independent multipath channel model in Section II-B is used with $L_k = 15$. We compare FD, our hybrid realization, and hybrid design in [27] i.e., LIHBF. For $K = 8$ single antenna users, Fig. 5.10 illustrates the BER performance versus SNR. It is observed that our design achieves the same performance as the FD combining and outperforms the robust design in [27], i.e., LIHBF by more than 2

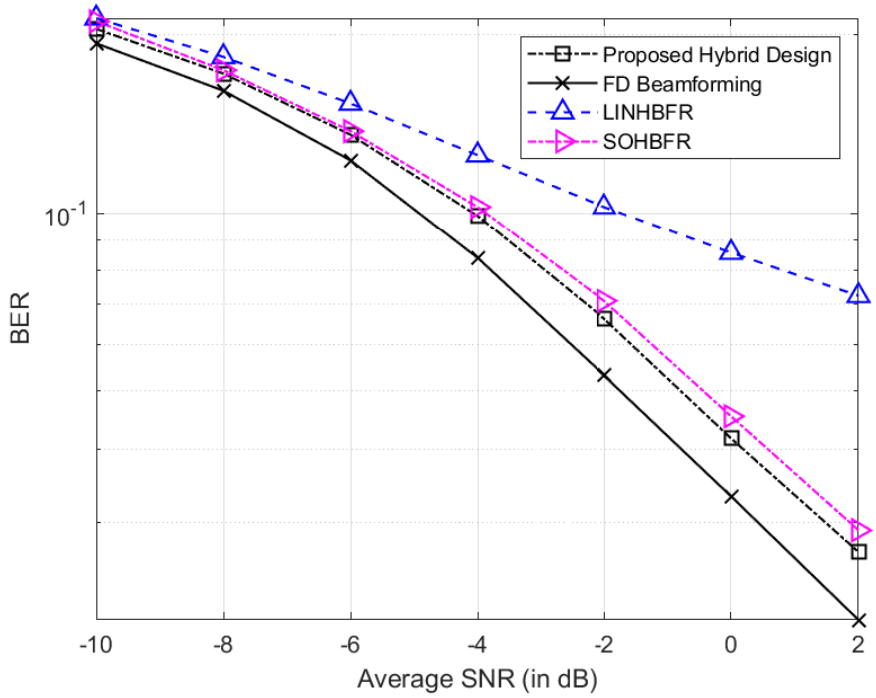


Fig. 5.11: BER versus SNR of finite-resolution hybrid designs for a 4×32 massive-MIMO system.

dB.

We also study the effect of finite resolution phase-shifters on the performance of the hybrid beamforming in uplink direction of a massive-MIMO system. We compare our approach to hybrid designs for finite resolution phase-shifters in [22, 75], i.e., SOHBFR and LINHBFR, respectively. It is assumed that low resolution phase-shifters, i.e., $M = 4$ are available for our design. Phase-shifters with twice of this resolution, i.e., $M = 8$ is considered for hybrid structures in [22,75], in order to present a fair comparison. Fig. 5.11 depicts simulation results for a single UE with $N_T = 4$ antennas where 16-QAM constellation is used with $K = 4$ symbols per transmission to the massive-MIMO BS with $N_R = 32$ antennas and $N_{RF} = 4$ RF chains. For a UE with $N_T = K = 8$ antennas, 4-QAM constellation and BS with $N_R = 64$ receive antenna where $N_{RF} = 8$, BER performance is illustrated in Fig. 5.12. It can be seen that in both cases the proposed hybrid realization outperforms the existing hybrid designs [22, 75] (i.e., SOHBFR and LINHBFR, respectively).

In order to evaluate the performance of the proposed channel estimation technique, we

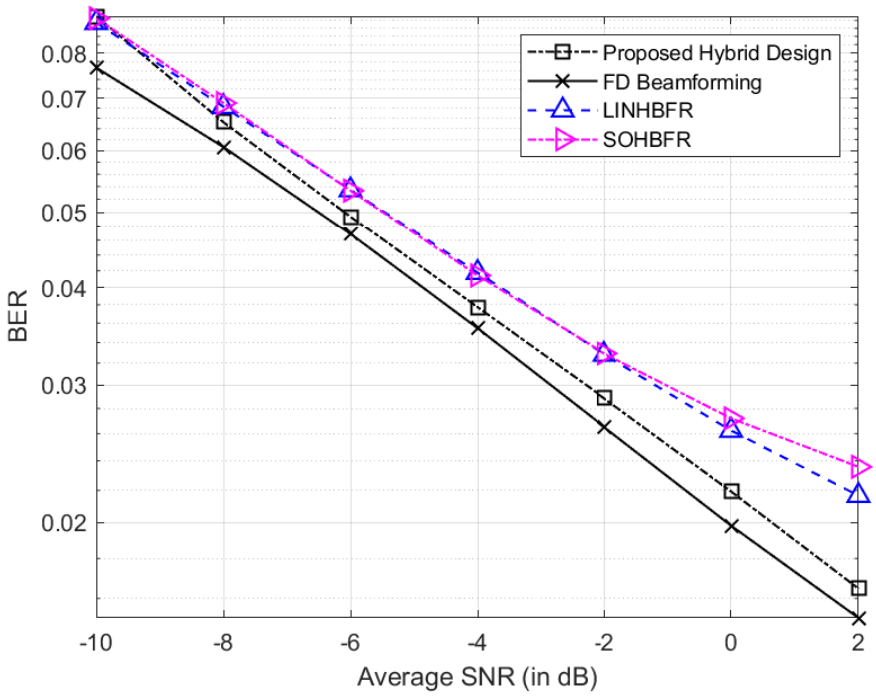


Fig. 5.12: BER versus SNR of finite-resolution hybrid designs for 8×64 massive-MIMO system.

consider $K = 10$ single antenna users and independent mmWave channels with $L_k = 8$ paths. At each SNR, 100 random cluster positions are generated and for each covariance matrix, 1000 channel realizations for all the users are simulated. We compare our design with hybrid design in [97], i.e., PANCH as well as FD channel estimation as the benchmark. The normalized MSE between the actual and estimated MU channel matrix versus SNR is plotted in Fig. 5.13 for all the channel estimation techniques. It can be seen that our design achieves the same performance as the FD channel estimation, while up to 1dB performance gain is obtained compared to hybrid design in [97].

5.7 Conclusion

In this chapter, we proposed a novel HBF structure for the receiver side of massive-MIMO communication systems that matches the performance of any given FD combiner. We took advantage of a matrix decomposition technique that allowed us to realize the FD combiner

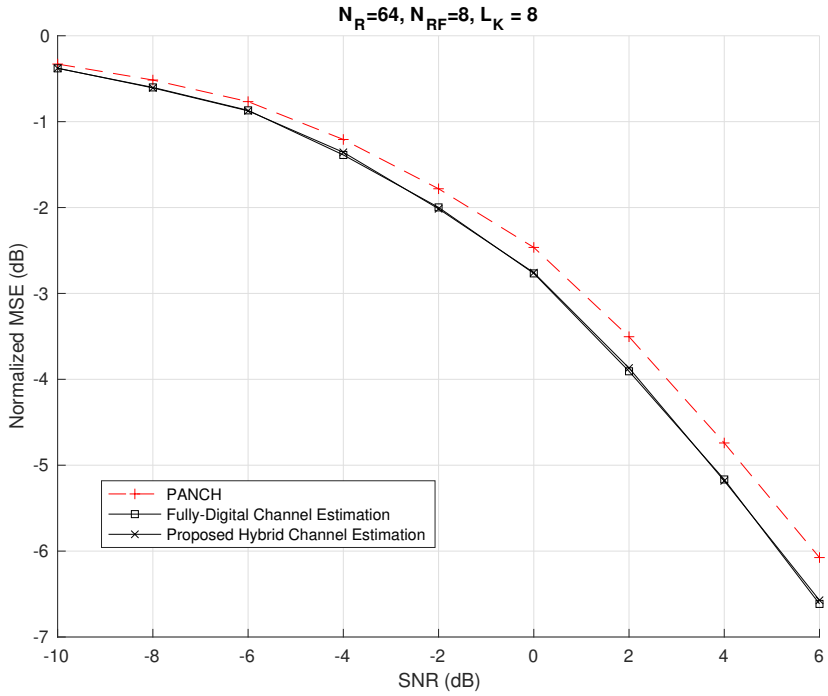


Fig. 5.13: Normalized MSE versus SNR of FD channel estimation and our design for a massive-MIMO BS with $N_R = 64$ antennas.

by HBF. We also focused on finite-resolution phase-shifters and presented a modified hybrid structure with phase-offset as an additional degree of freedom in order to improve the system performance. We then found a closed-form solution for the phase-offset which minimizes the error between FD combiner and its analog realization. A novel hybrid channel estimation was also presented based on the proposed hybrid structure that achieves FD performance. Finally, simulation results were presented to illustrate the advantages of the proposed hybrid designs to recently published works. In the next chapter, we study robust hybrid combiner design with imperfect CSI.

Chapter 6

Robust HBF for Uplink Massive-MIMO with Imperfect CSI

In this chapter[1], we study the design of hybrid A/D beamformers for uplink connection in massive-MIMO systems under imperfect-CSI. The norm-bounded channel error model is used to capture characteristics of imperfect CSI in practical systems. The objective function for the design is formulated based on MMSE worst-case robustness. We consider both SU and MU reception modes of a mmWave massive-MIMO BS. For the SU scenario, we study hierarchical beamformer optimization as well as joint precoder/combiner optimization for users with limited and extended computational capabilities, respectively. These optimization techniques are subsequently extended to the MU case where a new hybrid robust combiner design is proposed. Simulation results are presented confirming the superiority of our designs when compared to recent robust hybrid designs in the literature.

6.1 Introduction

Deployment of multiple antennas at both transmitter and receiver, i.e., MIMO was undoubtedly a huge step up for wireless communication systems. While bandwidth is limited

[1]Parts of this chapter have been presented at [100]:

- A. Morsali and B. Champagne, ""Robust Hybrid Analog/Digital Beamforming for Uplink Massive-MIMO with Imperfect CSI," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC),* Marrakesh, Morocco, 2019.

and increasing SNR increases the capacity only logarithmically, it has been shown that MIMO channel capacity increases linearly with the number of antennas [6]. Generally, this is only true if the channel matrix is full-rank which is not always the case, especially in mmWave systems [9, 11]. However, since in massive-MIMO the number of antennas can be very large, asymptotic limits of random matrix theory do apply and from an information theoretic point of view, it follows that regardless of the channel characteristics, capacity increases linearly with the minimum number of antennas employed at either the transmitter or the receiver [7, 8].

Nevertheless, the practical implementation of mmWave massive-MIMO systems faces many technical difficulties, and to this day remains very challenging and costly. In particular, since each antenna element must be driven by a RF chain, the conventional implementation of massive-MIMO requires as many RF chains as the number of antenna elements. Even if the design and implementation of a MIMO system with such large number of RF chains was possible and worth the cost, the power consumption of such a large number of RF components would seriously limit its potential for application [17].

One the most effective solutions to this problem is HSP [16, 17]. In conventional FD systems, signal processing is performed in the digital domain by means of dedicated processors and/or digital circuitry, which requires each antenna to be connected to a dedicated RF chain. Consequently, the received signal of each antenna is available in the digital domain [20, 101, 102]. In HSP, another layer of signal processing in the analog domain is incorporated to the system and by doing so, it becomes possible to reduce the number of RF chains [22].

In [16], the first attempt was made to realize any FD beamforming by the HBF structure with two RF chains for the case of a single data stream, i.e., one symbol per transmission. In [22, 23] multi-stream hybrid beamforming designs were presented in which the number of RF chains must be set equal to the number of symbols per transmission. We also proposed a single RF chain scheme for realizing any given FD precoding in [74]. The HBF has an intricate structure where the entries of the RF precoder matrix satisfy unit modulus constraint (i.e., phase-shifters). Since the ensuing precoder optimization is non-convex, many works have alternatively focused on designing HBF directly using heuristic iterative algorithms or reconstruction algorithms [21, 22]. Specifically, in [26] and [27], robust hybrid combiner designs for SU and MU are presented, respectively.

In this chapter, we investigate the design of hybrid analog/digital beamformers for

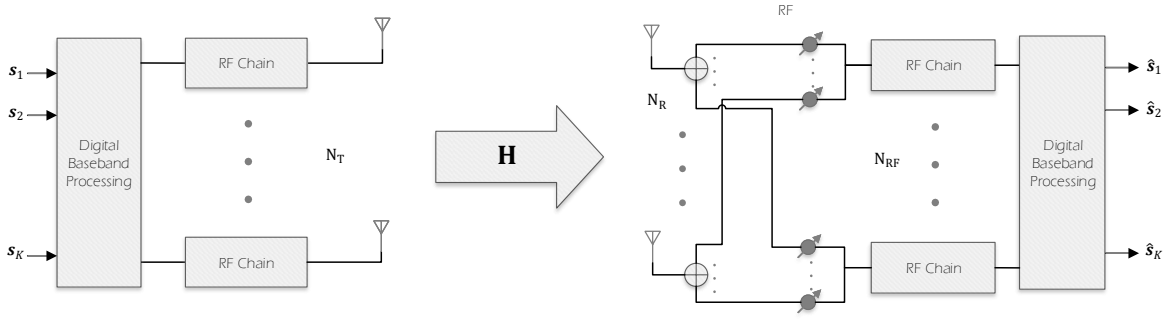


Fig. 6.1: Single-user massive-MIMO system.

uplink connection in massive-MIMO systems under imperfect-CSI for both SU and MU scenarios. The norm-bounded channel error model is used to represent the imperfect CSI aspects of practical systems while worst-case robustness MMSE is selected as the criterion to formulate the optimization problem. In the SU scenario, we consider the computational capabilities of the user for designing the beamformer. For users with limited computational resources, hierarchical optimization is presented which only puts the burden of robust calculations on the BS. For users with extended computational capabilities, a more flexible joint precoder/combiner design is proposed. We then extend these optimization techniques to propose a new robust hybrid combiner design for MU. Simulation results are presented confirming the superiority of our design compared to recently published robust hybrid designs in the literature.

6.2 System Model with Imperfect CSI

We consider the uplink connection of a single-cell wireless system where the massive-MIMO BS has N_R receive antennas and N_{RF} RF chains with $N_{RF} \ll N_R$. In what follows, the system formulations of both SU and MU for uplink connection are presented.

SU System Model: For the SU scenario, we consider the general point-to-point MIMO system where the transmitter is considered to be a multiple-antenna UE. As illustrated in Fig. 7.1, the UE is equipped with N_T antennas and the same number RF chains. The

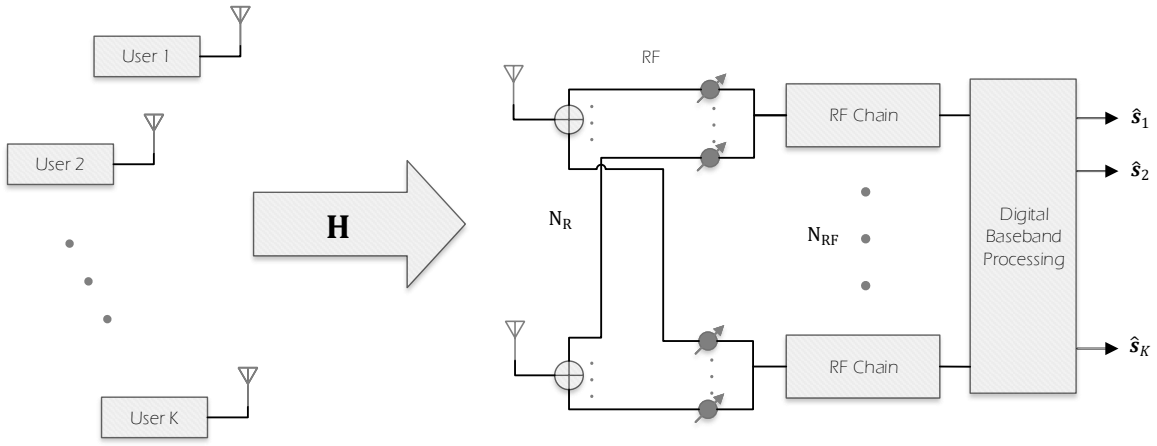


Fig. 6.2: Multi-user massive-MIMO system.

received signal at the BS can be written as,

$$\mathbf{y}_{\text{su}} = \mathbf{H}_{\text{su}}\mathbf{P}\mathbf{s} + \mathbf{n}, \tag{6.1}$$

where $\mathbf{H}_{\text{su}} \in \mathbb{C}^{N_R \times N_T}$ is the point-to-point mmWave MIMO channel matrix, $\mathbf{P} \in \mathbb{C}^{N_T \times K}$ and $\mathbf{s} \in \mathcal{A}^K$ are the precoder matrix and information symbol vector, respectively, where $\mathcal{A}$ is the selected constellation such as PSK or QAM and K is the number transmitted symbols. Moreover, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_R})$ is the AWGN vector.

MU System Model: Without loss of generality, we assume K single antenna users are served by the BS as depicted in Fig. 6.2. The combined received signal from all K users with statistical channel inversion power control scheme [27] can be written as,

$$\mathbf{y}_{\text{mu}} = \mathbf{H}_{\text{mu}}\mathbf{s} + \mathbf{n}, \tag{6.2}$$

where $\mathbf{H}_{\text{mu}} \in \mathbb{C}^{N_R \times K}$ can be expressed as,

$$\mathbf{H}_{\text{mu}} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_K], \tag{6.3}$$

with $\mathbf{h}_k$ being the the uplink fading channel between kth user and BS. Moreover, $\mathbf{s} = [s_1, s_2, ..., s_K]^T$ is the symbol vector where $\mathbf{s}_i$ denotes the transmitted symbol of ith user.

Hybrid Decoding: For $\mathbf{H}_i$ where $i = \{\text{su}, \text{mu}\}$, the total equivalent channel at the BS can be defined as,

$$\mathbf{H}_i^{eq} \triangleq \begin{cases} \mathbf{H}_{\text{su}}\mathbf{P} & i = \text{su} \\ \mathbf{H}_{\text{mu}} & i = \text{mu} \end{cases}. \tag{6.4}$$

We can thus formulate the decoded symbol vector for both cases after hybrid processing as,

$$\hat{\mathbf{s}} = (\mathbf{D}_{\text{A}}\mathbf{D}_{\text{D}})^H \mathbf{H}_i \mathbf{s} + \mathbf{n_e}, \tag{6.5}$$

where, $\mathbf{n_e} = \mathbf{D}_{\text{D}}^H \mathbf{D}_{\text{A}}^H \mathbf{n}$ is the effective noise vector after linear combining. Matrices $\mathbf{D}_{\text{D}} \in \mathbb{C}^{N_{RF} \times K}$ and $\mathbf{D}_{\text{A}} \in \mathbb{U}^{N_R \times N_{RF}}$ are digital and analog combiners where

$$\mathbb{U} = \{z \in \mathbb{C} \colon |z| = 1\}. \tag{6.6}$$

Analog combiner is implemented by RF phase shifters; therefore, all the entries of matrix $\mathbf{D}_{\text{A}}$ are constrained to have the same magnitude. For convenience, let us assume all analog matrices and vectors have unit magnitude entries, further practical discussions can be found in [74].

Finally, assuming $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$, MSE between transmitted and decoded signals is given by,

$$\text{MSE} \triangleq E\{\|\hat{\mathbf{s}} - \mathbf{s}\|^2\} = \|(\mathbf{D_A}\mathbf{D_D})^H \mathbf{H}_i^{eq} - \mathbf{I}\|_F^2 + \sigma^2 \|(\mathbf{D_A}\mathbf{D_D})^H\|_F^2. \tag{6.7}$$

Imperfect CSI: To represent the imperfect CSI characteristics of a practical system, the actual channel $\mathbf{H}_i$ can be expressed as,

$$\mathbf{H}_i = \hat{\mathbf{H}}_i + \boldsymbol{\Delta}, \tag{6.8}$$

where $\hat{\mathbf{H}}_i$ is the nominal channel known by the transceiver and $\boldsymbol{\Delta}$ is the error between the actual channel and the nominal channel which lies in an uncertainty region of $\mathcal{E}$:

$$\mathcal{E} \triangleq \{\boldsymbol{\Delta} : \|\boldsymbol{\Delta}\|_F \leq \varepsilon\}. \tag{6.9}$$

6.3 SU Hybrid Combiner Design

We formulate the beamforming design problem based on the philosophy of worst-case robustness. The robust beamformer for SU can be obtained by solving the following optimization problem in (6.10) where P_T is the total power budget at the transmitter:

$$\min_{\mathbf{D}_\mathrm{A},\mathbf{D}_\mathrm{D},\mathbf{P}} \quad \max_{\mathbf{\Delta}\in\mathcal{E}} \quad \|(\mathbf{D_A}\mathbf{D_D})^H(\hat{\mathbf{H}}_{su}+\mathbf{\Delta})\mathbf{P}-\mathbf{I}\|_F^2+\sigma^2\|(\mathbf{D_A}\mathbf{D_D})^H\|_F^2 \tag{6.10a}$$

$$\text{subject to} \quad \mathrm{Tr}(\mathbf{P}\mathbf{P}^H)\leq P_T, \tag{6.10b}$$

$$\mathbf{D}_\mathrm{A}\in\mathbb{U}^{N_R\times N_{RF}}. \tag{6.10c}$$

For designing the analog combiner, there are limitations in adjusting the magnitude of the combiner rows which is crucial in designing the robust combiner. Therefore, we design the analog combiner for the nominal channel and rely on the digital processing for robustness. Before designing the analog combiner, let us fix $N_{RF}=K$, as it is the minimum possible number of RF chains for decoding the received signal using linear operations [22]. Having the singular value decomposition of the nominal channel as,

$$\hat{\mathbf{H}}_{su}=\mathbf{U}_{su}\mathbf{\Sigma}_{su}\mathbf{V}_{su}^H. \tag{6.11}$$

The optimal FD combiner for eigenbeam transmission is given by [103, 104] as,

$$\mathbf{D}=\mathbf{U}_{su}^a, \tag{6.12}$$

where $\mathbf{U}_{su}=[\mathbf{U}_{su}^a,\mathbf{U}_{su}^b]$ and $\mathbf{U}_{su}^a$ contains the first K columns of $\mathbf{U}_{su}$. We now formulate the analog decoder design as,

$$\min_{\mathbf{D}_\mathrm{A}} \quad \|\mathbf{D}_\mathrm{A}-\mathbf{D}\|_F^2 \tag{6.13a}$$

$$\text{subject to} \quad \mathbf{D}_\mathrm{A}\in\mathbb{U}^{N_R\times N_{RF}}, \tag{6.13b}$$

$$\mathbf{D}_\mathrm{A}^H\mathbf{D}_\mathrm{A}=N_R\mathbf{I}_{N_{RF}}. \tag{6.13c}$$

Our intentions behind introducing the constraint (6.13c) are threefold: firstly, to avoid coloring the noise; secondly, maintaining the noise power, and thirdly, to avoid amplifying

the channel error power after analog decoding. Since this problem is non-convex, we first relax the unit modulus constraint:

$$\hat{\mathbf{D}}_A^* = \min_{\mathbf{D}_\mathrm{A}} \quad \|\hat{\mathbf{D}}_A - \mathbf{D}\|_F^2 \tag{6.14a}$$

$$\text{subject to} \quad \|\hat{\mathbf{D}}_A\|_F^2 \leq N_{RF} N_R, \tag{6.14b}$$

$$\hat{\mathbf{D}}_A^H \hat{\mathbf{D}}_A = \mathbf{I}_{N_{RF}}. \tag{6.14c}$$

Fortunately, the optimization problem (6.14) admits a closed form solution as,

$$\hat{\mathbf{D}}_A^* = \mathbf{D}(\mathbf{D}^H \mathbf{D})^{(1/2)}. \tag{6.15}$$

Using (6.15), we arrive at a solution for analog decoder by projecting $\hat{\mathbf{D}}_A^*$ on the ring $\mathbb{U}$:

$$\mathbf{D}_A = \mathrm{proj}_{\mathbb{U}}(\hat{\mathbf{D}}_A^*), \tag{6.16}$$

where $\mathrm{proj}_{\mathbb{U}}$ is the following element-wise operation on each entry of the given matrix:

$$e^{j\theta} = \mathrm{proj}_{\mathbb{U}}(\alpha e^{j\theta}). \tag{6.17}$$

For simplicity, let us define the equivalent channel after analog combining as,

$$\hat{\mathbf{H}}_{su}^e \triangleq \mathbf{D}_\mathrm{A}^H \hat{\mathbf{H}}_{su}. \tag{6.18}$$

Having designed the analog combiner, we can omit the constraint (6.10c) from the optimization problem (6.10). This problem however is required to be solved by both BS and UE to jointly optimize the precoder and combiner for worst-case robustness. Nevertheless, it is possible that the UE does not have enough computational power to perform the extra required computations. Therefore, we propose beamformer designs for both cases where the UE can and cannot perform the robustness calculations.

6.3.1 Hierarchical Robust Hybrid Beamformer Design

Here, we explore the case that the UE does not have enough computational resources to perform joint robust optimization. Thus, hierarchical optimization is performed to design

beamformers. To do so, we first design the optimal precoding at UE for the equivalent nominal channel after analog combining i.e., $\hat{\mathbf{H}}_{su}^{e}$. Under the assumption that the receiver is capable of optimal decoding, we can decouple the transmitter and the receiver designs. Therefore, the precoder can be obtained by solving the following optimization problem [22]:

$$\max_{\mathbf{P}} \quad \log_2\left(\det(\mathbf{I}_K + \frac{1}{\sigma^2}\mathbf{R}_n^{-1}\hat{\mathbf{H}}_{su}^{e}\mathbf{P}\mathbf{P}^H\hat{\mathbf{H}}_{su}^{e})\right) \tag{6.19a}$$

$$\text{subject to} \quad \mathrm{Tr}(\mathbf{P}\mathbf{P}^H) \leq P_T, \tag{6.19b}$$

$$\mathbf{R}_n = \mathbf{D}_{\mathrm{A}}^H\mathbf{D}_{\mathrm{A}}. \tag{6.19c}$$

Since we have already designed analog combiner with the constraint (6.13c), for large N_R, we have,

$$\mathbf{D}_{\mathrm{A}}^H\mathbf{D}_{\mathrm{A}} \approx N_R\mathbf{I}_K. \tag{6.20}$$

Then, the optimal solution of (6.19) can be analytically calculated [10] using the singular value decomposition of $\hat{\mathbf{H}}_{su}^{e}$:

$$\hat{\mathbf{H}}_{su}^{e} = \mathbf{U}_{su}^{e}\mathbf{\Sigma}_{su}^{e}{\mathbf{V}_{su}^{e}}^H. \tag{6.21}$$

The optimal non-robust precoder is given by,

$$\mathbf{P}_{nr} = \mathbf{V}_{su}^{e}\mathbf{W}, \tag{6.22}$$

where the diagonal weight matrix $\mathbf{W}$ is calculated via water filling [10]. Now, we can rewrite the worst-case optimization problem as,

$$\min_{\mathbf{D}_{\mathrm{D}}} \quad \max_{\mathbf{\Delta}\in\mathcal{E}} \quad \|{\mathbf{D}_{\mathbf{D}}}^H(\hat{\mathbf{H}}_{su}^{e} + \mathbf{\Delta})\mathbf{P} - \mathbf{I}\|_F^2 + \sigma^2\|{\mathbf{D}_{\mathbf{D}}}^H\|_F^2 \tag{6.23a}$$

$$\text{subject to} \quad \mathcal{E} = \{\mathbf{\Delta} : \|\mathbf{\Delta}\|_F \leq \varepsilon\|\mathbf{D}_{\mathbf{A}}\|_F\}. \tag{6.23b}$$

Using Theorem 1 in [103], for the chosen precoder, the optimal solution of the above problem is given by,

$$\mathbf{D}_{\mathrm{D}}^H = \mathbf{\Sigma}_D{\mathbf{U}_{su}^{e}}^H, \tag{6.24}$$

with $\mathbf{\Sigma}_D = \text{diag}\big([d_1, d_2, ..., d_K]\big)$ and d_i's are roots of the following equation:

$$\phi(d_i) = -\frac{2\mu w_i(d_i\gamma_i w_i - 1)(d_i w_i - \mu\gamma_i)}{(\mu - d_i^2 w_i^2)^2} + 2\sigma^2 d_i = 0, \tag{6.25}$$

where μ is an auxiliary variable, w_i's and γ_i's are diagonal entries of $\mathbf{W}$ and $\mathbf{\Sigma}_{su}^e$, respectively. Simple algorithmic solutions of (6.25) are given in [103].

6.3.2 Joint Robust Hybrid Precoder/Combiner Design

If the SU has enough computational power to perform the robust calculations, the worst-case beamformer design for the selected hybrid combiner (6.16) can be written as,

$$\min_{\mathbf{D}_\text{D},\mathbf{P}} \quad \max_{\mathbf{\Delta}\in\mathcal{E}} \quad \|\mathbf{D_D}^H(\hat{\mathbf{H}}_{su}^e + \mathbf{\Delta})\mathbf{P} - \mathbf{I}\|_F^2 + \sigma^2\|\mathbf{D_D}^H\|_F^2 \tag{6.26a}$$

$$\text{subject to} \quad \text{Tr}(\mathbf{P}\mathbf{P}^H) \leq P_T. \tag{6.26b}$$

This problem can be efficiently solved using alternating optimization techniques [17, 22, 103]. By fixing $\mathbf{P}$ in (6.26), we arrive at (6.23) which we have solved in previous subsection and by fixing $\mathbf{D}_\text{D}$, we can write,

$$\min_{\mathbf{P}} \quad \max_{\mathbf{\Delta}\in\mathcal{E}} \quad \|\mathbf{D_D}^H(\hat{\mathbf{H}}_{su}^e + \mathbf{\Delta})\mathbf{P} - \mathbf{I}\|_F^2 + \sigma^2\|\mathbf{D_D}^H\|_F^2 \tag{6.27a}$$

$$\text{subject to} \quad \text{Tr}(\mathbf{P}\mathbf{P}^H) \leq P_T. \tag{6.27b}$$

The optimal solution of (6.27) can be written as [104],

$$\mathbf{P}_{rob} = \mathbf{V}_{su}^e \text{diag}(\mathbf{w}), \tag{6.28}$$

where $\mathbf{w} = [w_1, \ldots, w_K]^T$ and w_i's is the allocated power of ith symbol. Here, we present a suboptimal solution with low computational complexity which can be used in alternating joint optimization. Pursuing the same guidelines we used to solve (6.23), we can first calculate $\hat{\mathbf{w}} = [\hat{w_1}, \ldots, \hat{w_K}]^T$ for the unconstrained optimization (6.27a), where $\hat{w_1}$'s are the roots of,

$$\psi(\hat{w_i}) = -\frac{2\mu\hat{w_i}(d_i\gamma_i\hat{w_i} - 1)(d_i\hat{w_i} - \mu\gamma_i)}{(\mu - d_i^2\hat{w_i}^2)^2} + 2\sigma^2 d_i = 0. \tag{6.29}$$

Note that (6.29) can be solved in a similar fashion as (6.25). Then, to satisfy (6.27b), by

Algorithm 1 Joint Robust Precoder/Combiner Design

Given: $\mathbf{U}_{su}^{e}$, $\mathbf{V}_{su}^{e\,H}$, P_T, $\varepsilon\|\mathbf{D_A}\|_F$
Initializing: $\mathbf{P} = \mathbf{V}_{su}^{e}$, $\mathbf{D}_{\mathrm{D}}^{H} = \mathbf{U}_{su}^{e\,H}$
1. $\mathbf{\Sigma}_D = \mathrm{diag}([d_1, d_2, ..., d_K])$ where d_i's are roots of (6.25).
2. $\mathbf{D}_{\mathrm{D}}^{H} = \mathbf{\Sigma}_D \mathbf{U}_{su}^{e\,H}$.
3. Calculate $\mathbf{w}$ from (6.30).
4. $\mathbf{P} = \mathbf{V}_{su}^{e}\mathrm{diag}(\mathbf{w})$.
5. Iterate until convergence .
Outputs: $\mathbf{P}_{rob} = \mathbf{P}$, $\mathbf{D}_{\mathrm{D}}^{H}$.

defining $\rho = \sqrt{P_T}/\|\hat{\mathbf{w}}\|$, we have,

$$\mathbf{w} = \rho\hat{\mathbf{w}}. \tag{6.30}$$

We summarized our proposed joint robust transceiver design based on alternating optimization in algorithm 1. Note that, (6.30) is a simple low-complexity suboptimal power allocation. In [104], an algorithmic solution is presented to obtain optimal w_i's. Therefore, if the system can handle the required complicated computations, the optimal robust precoder/combiner design can obtained by changing line 3 of algorithm 1 to: "*3. Calculate* $\mathbf{w}$ *from Algorithm 1 in [104]* ".

6.4 MU Hybrid Combiner Design

In this section, we explore robust combiner design for MU scenario by formulating the worst-case optimization as,

$$\min_{\mathbf{D}_{\mathrm{A}},\mathbf{D}_{\mathrm{D}}} \max_{\mathbf{\Delta}\in\mathcal{E}} \|(\mathbf{D_A D_D})^H(\hat{\mathbf{H}}_{mu} + \mathbf{\Delta}) - \mathbf{I}\|_F^2 + \sigma^2\|(\mathbf{D_A D_D})^H\|_F^2 \tag{6.31a}$$

$$\text{subject to} \quad \mathbf{D}_{\mathrm{A}} \in \mathbb{U}^{N_R \times N_{RF}}. \tag{6.31b}$$

Comparing the above optimization with (6.10), one can observe that if we set $\mathbf{P} = \mathbf{I}_K$ in (6.10) we arrive at (6.31). Therefore, In order to solve the above optimization problem we can use the proposed solution in Section 6.3 by assuming $\mathbf{P} = \mathbf{I}_K$.

As a result, considering the singular value decomposition of the nominal MU channel:

$$\hat{\mathbf{H}}_{mu} = \mathbf{U}_{mu}\mathbf{\Sigma}_{mu}\mathbf{V}_{,u}^{H}, \tag{6.32}$$

the analog combiner is given as,

$$\mathbf{D}_A = \mathrm{proj}_{\mathbb{U}}\big(\mathbf{D}(\mathbf{D}^H\mathbf{D})^{(1/2)}\big), \tag{6.33}$$

with,

$$\mathbf{D} = \mathbf{U}_{mu}^{a}, \tag{6.34}$$

where $\mathbf{U}_{mu} = [\mathbf{U}_{mu}^{a}, \mathbf{U}_{mu}^{b}]$ i.e., $\mathbf{U}_{mu}^{a}$ contains the first K columns of $\mathbf{U}_{mu}$. Then, defining the equivalent channel after analog combining as,

$$\hat{\mathbf{H}}_{mu}^{e} \triangleq \mathbf{D}_{\mathrm{A}}^{H}\hat{\mathbf{H}}_{mu}, \tag{6.35}$$

and having its singular value decomposition as,

$$\hat{\mathbf{H}}_{mu}^{e} = \mathbf{U}_{mu}^{e}\mathbf{\Sigma}_{mu}^{e}{\mathbf{V}_{mu}^{e}}^{H}. \tag{6.36}$$

The digital combiner is then obtained from,

$$\mathbf{D}_{\mathrm{D}}^{H} = \mathbf{\Sigma}_{D}{\mathbf{U}_{mu}^{e}}^{H}, \tag{6.37}$$

with $\mathbf{\Sigma}_D = \mathrm{diag}([d_1, d_2, ..., d_K])$ and d_i's are roots of the following equation:

$$\phi(d_i) = -\frac{2\mu(d_i\gamma_i - 1)(d_i - \mu\gamma_i)}{(\mu - d_i^2)^2} + 2\sigma^2 d_i = 0, \tag{6.38}$$

where μ is an auxiliary variable, and γ_i's are diagonal entries of $\mathbf{\Sigma}_{su}^{e}$. The roots of (6.38) can be obtained in a similar fashion as (6.25).

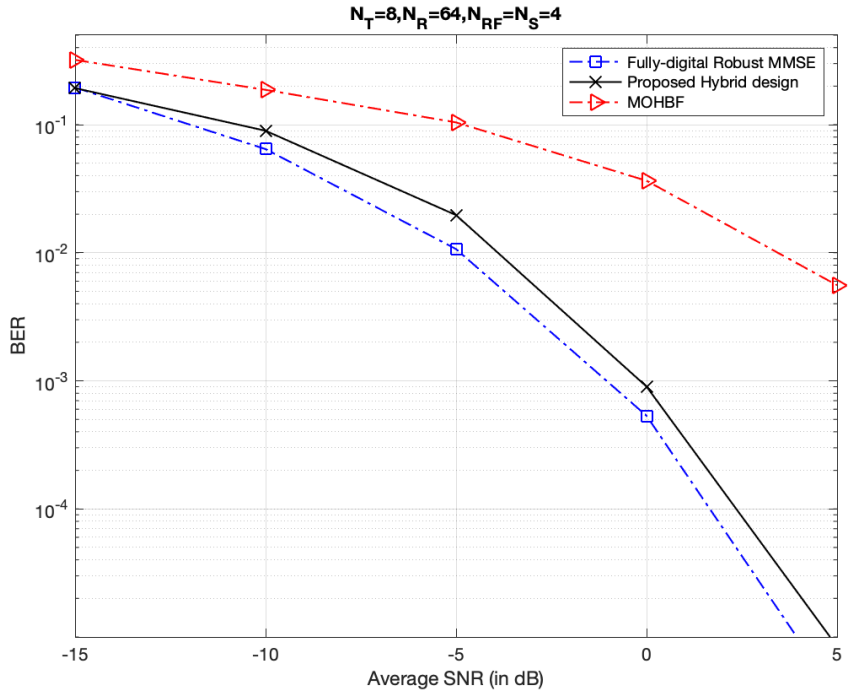


Fig. 6.3: BER versus SNR of robust FD combining and our design in a 8×64 massive-MIMO system.

6.5 Simulation Results

In this section, we present simulation results for both SU and MU cases. To obtain realistic results, we do not perform simulations only for the worst possible channel in the uncertainty region, instead the elements of the channel error $\mathbf{\Delta}$ are randomly generated according to zero-mean, i.i.d. Gaussian distributions such that $\|\mathbf{\Delta}\|_F = \varepsilon$. The radius of uncertainty region ε is calculated by $\varepsilon^2 = s\|\hat{\mathbf{H}}\|_F^2$ where $s \in [0, 1)$. For all the simulations, number of receive antennas at the BS is set to $N_R = 64$ and the uniform linear configuration is used. Thus, the array response is given by,

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{N_R}}[1, e^{jkd\sin(\phi)}, ..., e^{jkd(N_R-1)\sin(\phi)}], \tag{6.39}$$

where $k = 2\pi/\lambda$, and for wavelength of λ we have $d = \lambda/2$.

For SU scenario, 8-PSK constellation is used and we consider the following channel

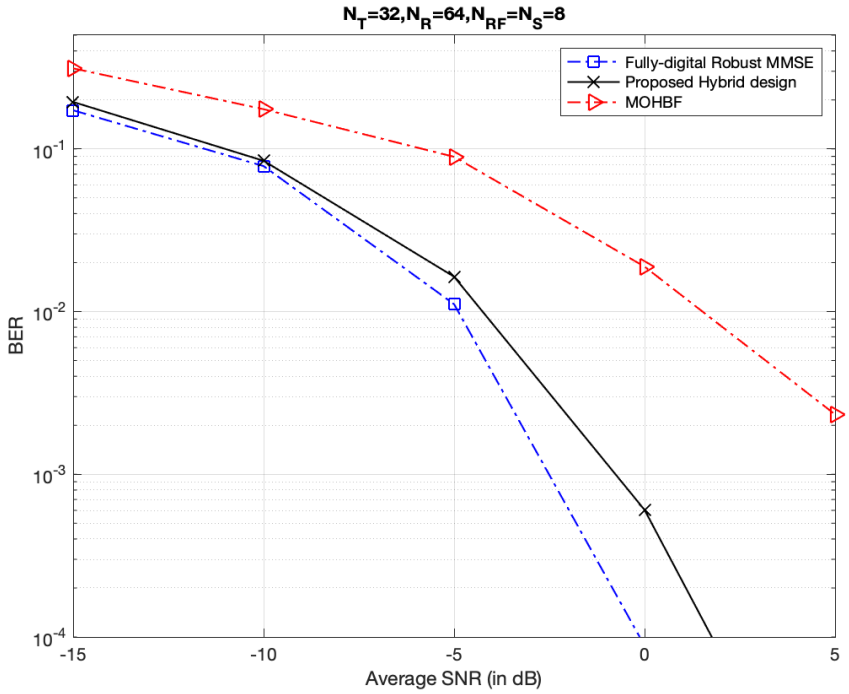


Fig. 6.4: BER versus SNR of robust FD combining and our design for a 32×64 massive-MIMO system.

model for mmWave massive-MIMO with sparse scattering environments [22, 26, 74]:

$$\mathbf{H} = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^{L} \alpha^l \mathbf{a}_r(\phi_r^l)\mathbf{a}_t(\phi_t^l)^H, \tag{6.40}$$

where, $\alpha^l \sim CN(0,1)$ is the complex gain of l^{th} path, $\mathbf{a}_r$ and $\mathbf{a}_t$ are the antenna array responses of receiver and transmitter, respectively. ϕ_r^l and ϕ_t^l are arrival and departure angles and have uniform distribution over $[0, 2\pi)$.

For UE with $N_T = 8$ transmit antennas, $N_s = 4$ symbols per transmission, $s = 0.01$ and mmWave channel with $L = 10$ paths, Fig. 6.3 depicts the BER performance versus SNR (SNR$=P_T/\sigma^2$) for fully digital beamforming, our proposed hierarchical design in Section 6.3.1, and hybrid robust design in [26], i.e., MOHBF. While there is around less than 1 dB gap between our design and fully digital beamforming, our proposed design outperforms [26] by more than 6 dB. Simulation results for the case where both BS and UE

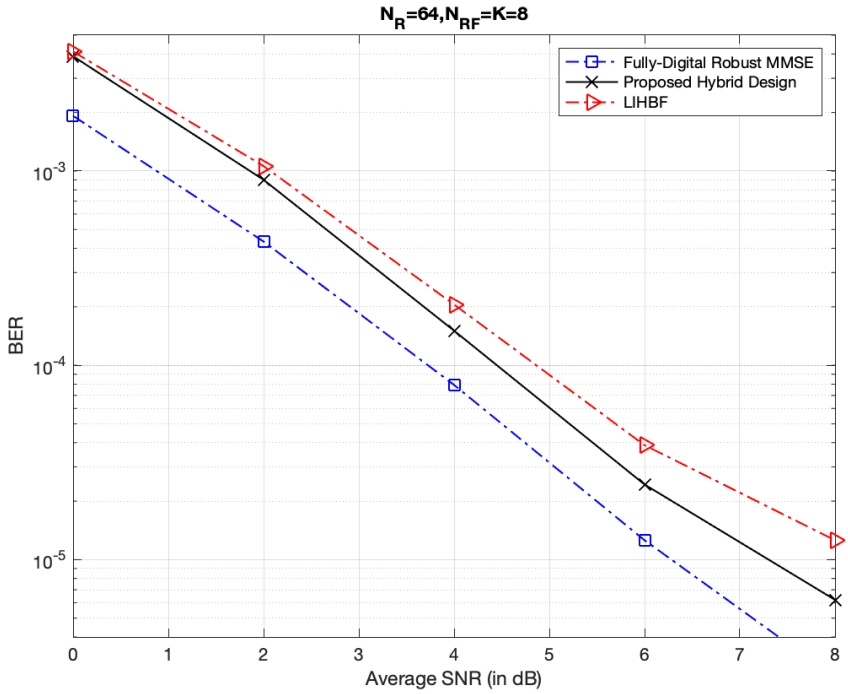


Fig. 6.5: BER versus SNR of robust FD combining and our design for a 64 massive-MIMO BS and $K = 8$ users.

perform joint robust optimization are illustrated in Fig. 6.4. The UE is equipped with $N_T = 32$ transmit antennas, and the reset of parameters are set as follows: $N_s = 8$ symbols per transmission, $s = 0.03$ and mmWave channel with $L = 15$. The results expectedly confirm the superiority of our design to the recently published robust design [26].

For multi-user case, 16-QAM constellation and independent multipath channel model [27] is used. The mmWave channel vector of k^{th} user can be modelled as,

$$\mathbf{h}_k = \sqrt{\frac{N_R}{L_K}} \sum_{l=1}^{L_k} \alpha^{k,l} \mathbf{a}_r(\phi^l), \tag{6.41}$$

where $\alpha^{k,l} \sim CN(0, p_{k,l})$ is the complex gain of l^{th} path and,

$$\frac{1}{L_k} \sum_{l=1}^{k} p_{k,l} = 1, \tag{6.42}$$

for normalization purposes.

We compared our design with robust hybrid design in [27], i.e., LIHBF, as well as FD decoder as the benchmark. For $K = 8$ single antenna users, and independent mmWave channels with $L_k = 15$ with $s = 0.001$, Figs. 6.5 illustrates the BER performance versus SNR. It cn be observed that our design has a margin of 1 dB to the fully digital combining and outperforms the robust design in [27] by more than 1 dB.

6.6 Conclusion

In this chapter, we proposed robust HBF designs for uplink connection in massive-MIMO communication systems. Specifically, both SU and MU modes of reception in the uplink direction of a single-cell configuration with massive-MIMO BS were considered. The norm-bounded channel error was used to capture the imperfect CSI conditions while the objective function was formulated based on the worst-case robustness MMSE. For the SU uplink scenario, we presented hierarchical optimization as well as joint optimization depending on the UE capabilities to perform extra calculations required for robust transceiver design. We then proposed robust hybrid combiner design for the MU uplink connection. Finally, simulation results were presented to demonstrate the superiority of our design over recently published hybrid designs. In the next chapter, we present a generalized framework for analog/digital signal processing.

Part III

Generalized Hybrid A/D Signal Processing

Chapter 7

Generalized Framework for Hybrid A/D Signal Processing

In this chapter[1], we explore the hybrid A/D structure as a general framework for signal processing in massive and ultra-massive-MIMO systems. To exploit the full potential of the analog domain, we first focus on the analog signal processing (ASP) network. We investigate a mathematical representation suitable for any arbitrarily connected feed-forward ASP network comprised of the common RF hardware elements in the context of hybrid A/D systems, i.e., phase-shifter and power-divider/combiner.

A novel ASP structure is then proposed which is not bound to the unit modulus constraint, thereby facilitating the design of hybrid A/D systems. We then study MIMO transmitter and receiver designs to exploit the full potential of digital processing as well. It is shown that replacing the linear transformation in the digital domain with a generic mapping can improve the system performance. In some cases, the performance of optimal FD MIMO systems can be achieved without extra calculations compared to sub-optimal hybrid A/D techniques. An optimization model based on the proposed structure is presented that can be used for hybrid A/D system design. Specifically, precoding and combining

[1]Parts of this chapter have been submitted to USPTO and published in [105]:

- A. Morsali, A. Haghighat and B. Champagne, "Generalized Framework for Hybrid Analog/Digital Signal Processing in Massive and Ultra-Massive-MIMO Systems," *IEEE Access*, vol. 8, pp. 100262–100279, 2020.
- A. Morsali, A. Haghighat, and B. Champagne," Efficient Receiver Combining for Hybrid Analog/Digital Beamforming," *U.S. Patent,* PCT/US2020/024318, March 23, 2020.

designs under different conditions are discussed as examples. Finally, simulation results are presented which illustrate the superiority of the proposed architecture compared to the conventional hybrid designs for massive-MIMO systems.

7.1 Introduction

Massive-MIMO and ultra-massive (UM)-MIMO systems operating in mmW/Terahertz (THz) bands are the prime candidates for B5G cellular networks [7, 8, 11, 13]. In fact, BSs with 64 antennas have been recently deployed for commercial use in some countries [14]. Moreover, an extensive theory for massive MIMO has been developed in recent years, including capacity and spectral efficiency analysis, system design for high energy efficiency, pilot contamination, etc. However, implementation of such systems faces many technical difficulties, and to this day remains very challenging and costly [9, 15]. In conventional FD MIMO systems, each antenna element requires a dedicated RF chain. The direct FD implementation for massive-MIMO/UM-MIMO systems, however, is not practical and efficient due to the ensuing high production costs and more importantly, huge power consumption.

HSP is an effective approach to overcome this problem by cascading an ASP network to the baseband digital signal processor [16,17]. While in conventional FD MIMO transmitters [18–20], each antenna element is directly controlled by the digital processor, in an HSP-based transmitter, the digital processor generates a low-dimensional RF signal vector, whose size is then increased by analog circuitry for driving the large-scale antenna array. Similarly, in an HSP-based receiver, the size of the large-dimensional vector of antenna signals is reduced by an ASP network, whose outputs are then converted to the digital domain by means of RF chains for baseband processing.

There are practical constraints in the implementation and design of ASP networks and only a few types of RF components are commonly used in practice. Specifically, the power-divider (splitter), power-combiner (adder), and phase-shifter are the key analog components of the ASP design [21–29, 99]. In the existing hybrid beamforming structures, due to the particular configuration of the aforementioned analog components, a constant modulus constraint is imposed on the analog beamformer weights which generally turn beamforming design into intractable non-convex optimization problems [21, 22].

Related Works

In one of the earliest works in this field [16], it is shown that for a single data stream, two RF chains are required to achieve the performance of a FD combiner. This technique was extended to multiple stream beamforming (i.e., precoding/combining) where the required number of RF chains must be twice the number of the data streams [22,23]. In Chapter 3, we proposed a single RF chain FD precoding realization. Many researchers, however, focused on developing the hybrid beamformers directly by solving non-convex design optimization problems [21–29].

In [21], the beamformer design was formulated as the minimization of the Euclidean distance between the hybrid beamformer and the FD one. Then, by taking into account the sparse characteristics of the mmWave channels, compressed sensing (CS) techniques were presented to solve the underlying optimization problems. The same authors, extended their results to wide-band systems in [60]. This approach was later used in [29] and [24] where in the latter, manifold optimization algorithms as well as other low-complexity algorithms were used for hybrid beamformer design. Directly tackling the non-convex design optimization problems was attempted in [22] where the authors took advantage of orthogonalization techniques and exploited the sparsity of the channel for designing the hybrid beamformers. These results were then extended to wide-band systems in [65]. In [75], the Gram-Schmidt method was used specifically in uplink MU scenario for designing robust low-complexity beamformers. Robust beamformers for SU were studied in [26] by minimizing the sum-power of interfering signals. In [25], a simple non-iterative algorithm was proposed for hybrid regularized channel diagolnalization and in [75] the MSE was chosen as the cost function for designing the hybrid beamformers.

The majority of the above works consider a fully-connected architecture, i.e., each RF chain is connected to all of the antenna elements. Alternatively, in a sub-connected architecture, only a subset of RF chains are connected to each antenna [24,106,107]. Recently, a dynamic sub-connected hybrid architecture has been proposed in [107] for multi-user equalization in wideband millimeter-wave massive MIMO systems, based on the minimization of the sum of MSE over multiple subcarriers. Although sub-connected designs require less RF components, fully-connected ones can achieve a superior performance in theory. Hence, in this study, we investigate properties of fully-connected ASP networks.

Contributions and Organization

In this chapter, our goal is to investigate and exploit the full potential of HSP in massive-MIMO systems. Aiming at this challenge, we can summarize our contributions as follows:

- We first explore the degrees of freedom in the analog domain by developing a compact mathematical representation for any given feed-forward ASP network with arbitrary connections of any number of RF components, i.e., phase-shifters, power dividers and power combiners.
- Based on the above generalization, a simple and novel ASP architecture is conceived out of the above RF components, which is not bound to the constant modulus constraint. Removing this constraint facilitates system design as non-convex optimizations are difficult to solve and global optimality of the solutions cannot usually be guaranteed.
- The transmitter and receiver sides are then studied separately by exploiting the newly proposed ASP architecture and generalizing generalizing the digital processing. Specifically, the optimization problem for the HSP beamformer is reformulated within the new representation framework, which facilitates its solution under a variety of constraints and requirements for the massive MIMO system.
- The realization of optimal FD by HSP and the problem of RF chain minimization are presented as guideline examples to illustrate potential applications of the proposed theoretical framework.
- Simulation results of optimal beamformer designs with the proposed architecture are finally presented. The results demonstrate that the new designs can achieve the same performance as the corresponding optimal FD system and hence, outperform recently published hybrid beamformer designs.

7.2 System Model

We consider a generic point-to-point massive-MIMO system where the transmitter and receiver are equipped with N_T and N_R antennas as well as M_T and M_R RF chains, re-

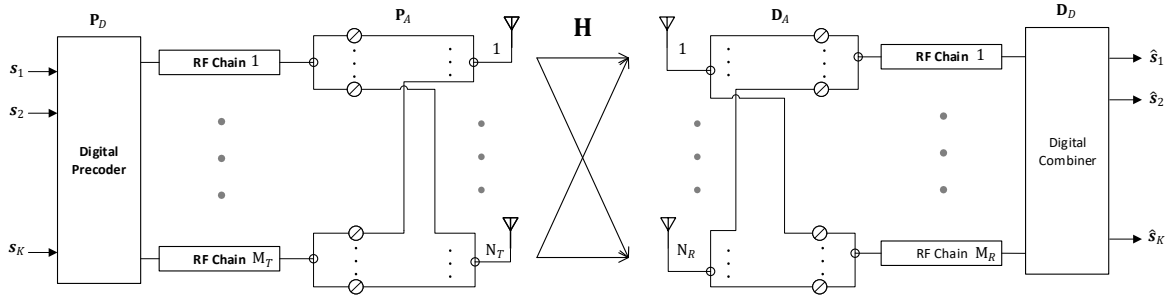


Fig. 7.1: Conventional HSP architecture for single-user massive-MIMO System.

spectively. In the context of HSP, due to practical constraints, it is further assumed that $M_T \ll N_T$ and $M_R \ll N_R$.

7.2.1 Conventional hybrid beamforming

Fig. 7.1 illustrates a point-to-point massive-MIMO system with conventional hybrid beamforming implemented at both ends. The transmitted signal over one symbol duration T_s can be formulated as,

$$\mathbf{x} = \sqrt{\rho}\,\mathbf{P}_A\mathbf{P}_D\mathbf{s}, \tag{7.1}$$

where $\mathbf{s} = [s_1, s_2, ..., s_K]^T$ is the symbol vector with zero-mean random information symbols s_k's taken from a discrete constellation $\mathcal{A}$ (such as M-QAM or M-PSK), normalized such that $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_K$ and, ρ is the average transmit power. Matrices $\mathbf{P}_D \in \mathbb{C}^{M_T \times K}$ and $\mathbf{P}_A \in \mathbb{U}^{N_T \times M_T}$ are the digital and analog precoders, respectively, where $\mathbb{U} = \{z \in \mathbb{C}\colon |z| = 1\}$ and for normalization purposes, it is further assumed that $\|\mathbf{P}_A\mathbf{P}_D\|_F^2 = 1$.

The received signal can then be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{7.2}$$

where $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the MIMO flat fading channel matrix such that $\mathbb{E}\{\|\mathbf{H}\|_F^2\} = N_T N_R$ and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_R})$ is an AWGN vector. The decoded symbols after hybrid processing can be expressed as

$$\hat{\mathbf{s}} = \mathbf{D_D}\mathbf{D_A}\mathbf{y}, \tag{7.3}$$

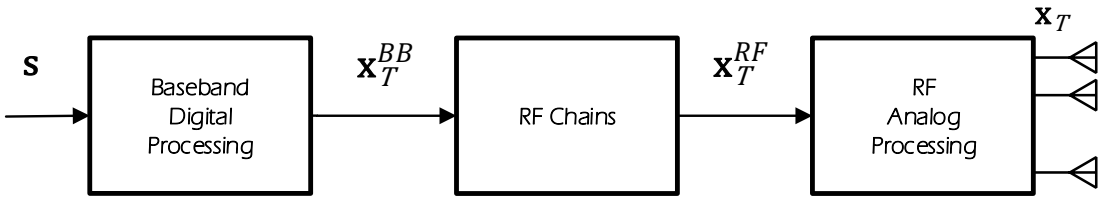


Fig. 7.2: HSP-based massive-MIMO transmitter modules.

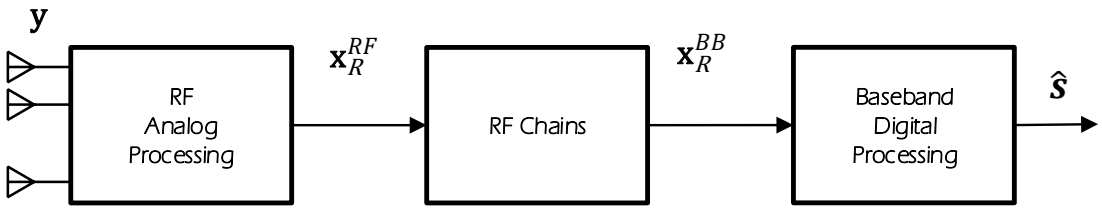


Fig. 7.3: HSP-based massive-MIMO receiver modules.

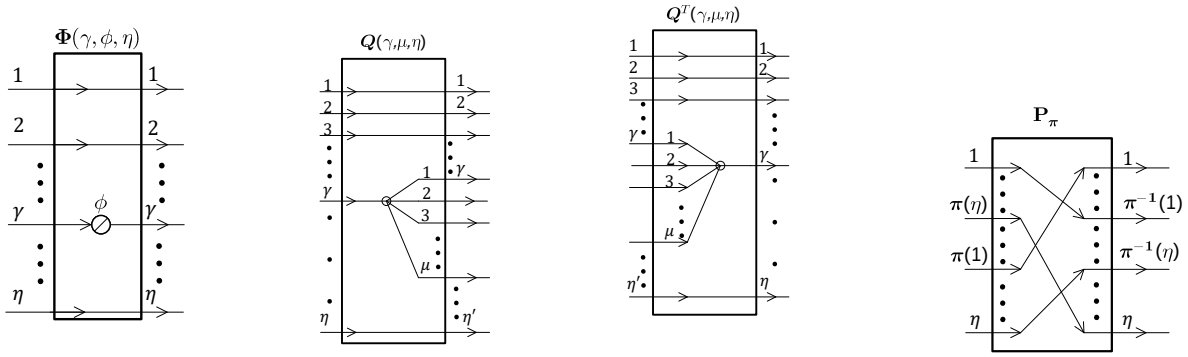
where $\mathbf{D_D} \in \mathbb{C}^{K \times M_R}$ and $\mathbf{D_A} \in \mathbb{U}^{M_R \times N_R}$ are the digital and analog combiners, respectively.

7.2.2 Generalized HSP system formulation

In this work, we consider a more general formulation for HSP that extends the cascaded structure of analog and digital linear transformations presented in Subsection 7.2.1. We will see that this formulation can in fact bring simplifications to the conventional linear MIMO precoding/combining techniques.

In the generalized HSP-based massive-MIMO transmitter, as shown in Fig. 7.2, the symbol vector $\mathbf{s}$ is first applied as input to the digital signal processor, whose output is a baseband signal vector expressed as,

$$\mathbf{x}_T^{BB} = \mathcal{F}_T(\mathbf{s}) \in \mathbb{C}^{M_T}, \tag{7.4}$$



(a) Single phase-shifter. (b) Single power divider. (c) Single power combiner. (d) Permutation matrix.

Fig. 7.4: Block diagram of ASP components.

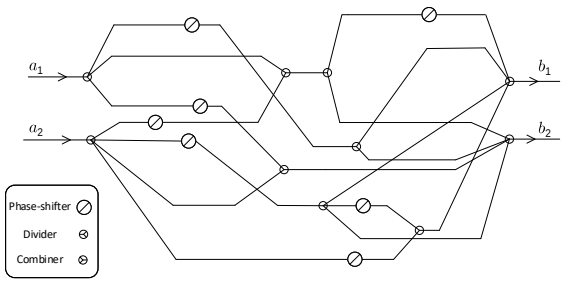


Fig. 7.5: An example of an arbitrary ASP network.

where $\mathcal{F}_T : \mathcal{A}^K \to \mathbb{C}^{M_T}$ is the corresponding mapping from $\mathcal{A}^K$ to $\mathbb{C}^{M_T}$. Then, M_T parallel RF chains convert the baseband signal vector $\mathbf{x}_T^{BB}$ into a bandpass modulated RF signal vector $\mathbf{x}_T^{RF}$. The latter is next input to the ASP network whose output is the transmit signal vector, which can be expressed as

$$\mathbf{x}_T = \sqrt{\rho}\, \mathcal{G}_T(\mathbf{x}_T^{RF}) \in \mathbb{C}^{N_T}, \tag{7.5}$$

where $\mathcal{G}_T : \mathbb{C}^{M_T} \to \mathbb{C}^{N_T}$ is the corresponding mapping.

As shown in Fig. 7.3, the received RF signal $\mathbf{y}$ following from the noisy MIMO transmission as in (7.2) is first applied as input to the ASP network, yielding,

$$\mathbf{x}_R^{RF} = \mathcal{G}_R(\mathbf{y}) \in \mathbb{C}^{M_R}, \tag{7.6}$$

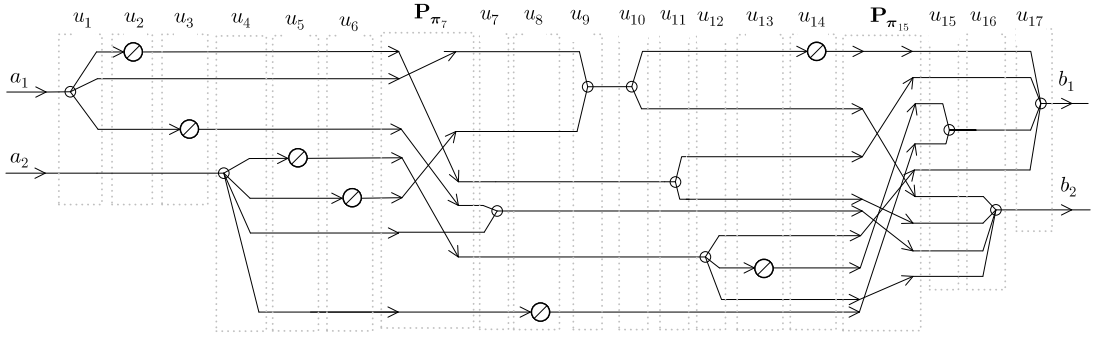


Fig. 7.6: Using Proposition 7.1 for rearranging the ASP network in Fig. 7.5.

where $\mathcal{G}_R : \mathbb{C}^{N_R} \to \mathbb{C}^{M_R}$. The RF signal vector $\mathbf{x}_R^{RF}$ is next converted to baseband vector $\mathbf{x}_R^{BB}$ by M_R RF chains. Finally, $\mathbf{x}_R^{BB}$ is processed in digital domain to obtain the decoded symbols,

$$\hat{\mathbf{s}} = \mathcal{F}_R(\mathbf{x}_R^{BB}). \tag{7.7}$$

where $\mathcal{F}_R : \mathbb{C}^{M_R} \to \mathcal{A}^K$.

While only a power constraint is imposed on the baseband mappings $\mathcal{F}_R$ and $\mathcal{F}_T$, the RF mappings $\mathcal{G}_R$ and $\mathcal{G}_T$ must be implemented by RF analog components which constrain these transformations as discussed in the following section.

7.3 Analog Signal Processing Network

In this section, aiming at exploiting the full potential of the analog domain, we develop a mathematical formulation for the ASP network represented by the RF mappings $\mathcal{G}_T$ and $\mathcal{G}_R$ in the previous section. Specifically, instead of focusing on the conventional analog beamformer structure used in the recent literature [21–29], we consider an arbitrarily connected network of phase-shifters, power dividers and power combiners. In our developments, signal-flow graph concepts are used which provide valuable insights for analysis of linear networks [108, 109].

Let us start by formally introducing the individual RF components comprising the ASP networks. The input-output (I/O) relationship of a phase-shifter is given by $b = e^{j\theta}a$ where $a, b \in \mathbb{C}$ are the input and output, respectively, and $\theta \in [0, 2\pi]$ controls the phase difference

between them. In this work, in order to explore the performance limits of ASP networks and find a compact representation for any arbitrarily connected ASP with common RF components, we consider infinite resolution phase shifters. The passive power combiner and power divider are implemented by the same RF multi-port network but their port configuration is different. For instance, the ideal μ-way Wilkinson power divider is an $\mu+1$ port RF network which can act as an equi-power divider if the input signal is applied to its port 1 and the outputs are taken from ports 2 to $\mu+1$ [110]. Conversely, it acts as a combiner if the inputs are applied to port 2 to $\mu+1$ and the output is taken from port 1.

To obtain a unified model for any possible ASP network with M input ports and N output ports using primary modules (i.e., phase-shifter, power divider and power combiner), we first present a convenient multi-port matrix representation of each component. We also include a permutation operation which does not require additional hardware and is used mainly for the sake of mathematical simplification. The I/O relationship of the components are defined below in terms of their input and outputs represented by vector $\mathbf{a}$ and $\mathbf{b}$, respectively.

- *Single phase-shifter*: As illustrated in Fig. 7.4a, for vector $\mathbf{a}, \mathbf{b} \in \mathbb{C}^{\eta}$, the corresponding $\eta \times \eta$ matrix only changes the phase of the γ^{th} element of the RF input signal $\mathbf{a}$, which can be expressed as,

$$\mathbf{b} = \mathbf{\Phi}(\gamma, \phi, \eta)\mathbf{a}, \tag{7.8a}$$

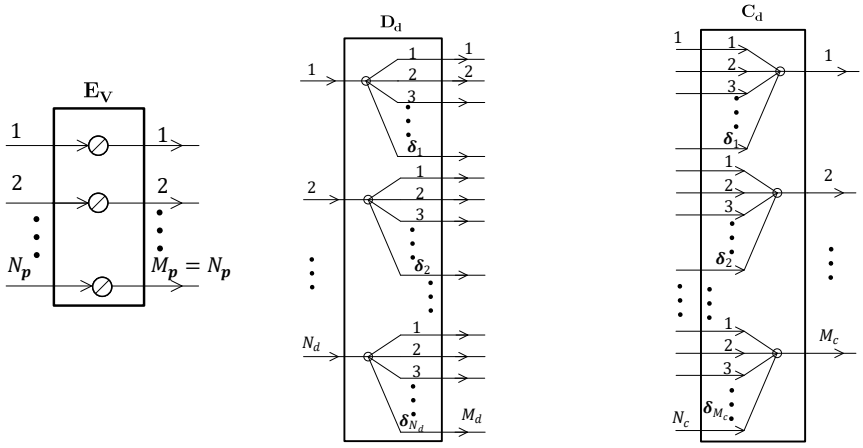
$$\mathbf{\Phi}(\gamma, \phi, \eta) = \mathrm{bd}(\mathbf{I}_{\gamma-1}, e^{j\phi}, \mathbf{I}_{\eta-\gamma}) \in \mathbb{C}^{\eta\times\eta}. \tag{7.8b}$$

- *Single power divider*: For input vector $\mathbf{a} \in \mathbb{C}^{\eta}$ and output vector $\mathbf{b} \in \mathbb{C}^{\eta'}$, the corresponding $\eta' \times \eta$ matrix divides the γ^{th} element of the input RF signal into μ equi-power signals and the remaining RF branches are not altered, and hence, $\eta' = \eta + \mu - 1$. As illustrated in Fig. 7.4b, this operation can be described by a block diagonal matrix:

$$\mathbf{b} = \mathbf{Q}(\gamma, \mu, \eta)\mathbf{a}, \tag{7.9a}$$

$$\mathbf{Q}(\gamma, \mu, \eta) = \mathrm{bd}(\mathbf{I}_{\gamma-1}, \frac{1}{\sqrt{\mu}}\mathbf{1}_{\mu}, \mathbf{I}_{\eta-\gamma}). \tag{7.9b}$$

- *Single power combiner*: This transformation can be represented by the transpose of



(a) Phase-shifter network. (b) Power divider network. (c) Power combiner network.

Fig. 7.7: ASP sub-networks

the single power divider matrix $\mathbf{Q}(\gamma, \mu, \eta)$. Consequently, for input vector $\mathbf{a} \in \mathbb{C}^{\eta'}$ and output vector $\mathbf{b} \in \mathbb{C}^{\eta}$ the corresponding matrix combines μ adjacent RF signals into the γ^{th} output signal and the rest of the RF branches are not altered. As seen from Fig. 7.4c, we can write,

$$\mathbf{b} = \mathbf{Q}^t(\gamma, \mu, \eta)\mathbf{a}. \tag{7.10}$$

- *Permutation matrix*: This transformation shown in Fig. 7.4d corresponds to rearrangement of the elements of vector $\mathbf{a} \in \mathbb{C}^{\eta}$ into vector $\mathbf{b} \in \mathbb{C}^{\eta}$ according to a permutation $\pi : \{1, ..., \mu\} \rightarrow \{1, ..., \mu\}$. This can be expressed as,

$$\mathbf{b} = \mathbf{P}_{\pi}\mathbf{a}, \tag{7.11}$$

where $\mathbf{P}_{\pi} = [\mathbf{e}_{\pi_1}, ..., \mathbf{e}_{\pi_M}]^t$, and $\mathbf{e}_i$ denotes a column vector of zeros except for its i^{th} element which is one.

Having introduced a matrix representation of the RF components, we can now seek the mathematical formulation for any given ASP in terms of these matrices.

Proposition 7.1. *Any given RF network, with N input and M output ports, implemented by arbitrary feed-forward connections of T RF components (i.e., phase-shifters, power com-*

biners and power dividers) can be modeled as follows,

$$\begin{aligned}\mathbf{b} &= \mathbf{A}_{u_T}(\theta_T)\mathbf{P}_{\pi_T}\ldots\mathbf{A}_{u_2}(\theta_2)\mathbf{P}_{\pi_2}\mathbf{A}_{u_1}(\theta_1)\mathbf{P}_{\pi_1}\mathbf{a} \\ &\triangleq \prod_{i=1}^{T}\mathbf{A}_{u_i}(\theta_i)\mathbf{P}_{\pi_i}\mathbf{a},\end{aligned} \tag{7.12}$$

$$\mathbf{A}_{u_i}(\theta_i) = \begin{cases}\mathbf{\Phi}(\gamma,\phi,\eta), & \text{where } \theta_i \equiv (\gamma,\phi,\eta) \quad \text{if } u_i = 1 \\ \mathbf{Q}(\gamma,\mu,\eta), & \text{where } \theta_i \equiv (\gamma,\mu,\eta) \quad \text{if } u_i = 2 \\ \mathbf{Q}^t(\gamma,\mu,\eta), & \text{where } \theta_i \equiv (\gamma,\mu,\eta) \quad \text{if } u_i = 3\end{cases},$$

where $\mathbf{a} \in \mathbb{C}^N$ *and* $\mathbf{b} \in \mathbb{C}^M$ *are the input and output RF signals, respectively, and* θ_i *is a* 3*-tuple containing the parameters of the* i^{th} *RF component.*

Proof. The matrix representation of the RF components in (7.8)-(7.10) are introduced such that the input and output signals can be of any size and thus can include RF branches that are not affected by the RF component. Consequently, we can sort the RF components such that the input of each RF component is the output of another RF component except for the first component. Let us denote the input and output of the i^{th} RF component as $\mathbf{a}_i$ and $\mathbf{b}_i$, respectively. Consequently, we have $\mathbf{b}_{i-1} = \mathbf{a}_i$, $\mathbf{a}_1 = \mathbf{a}$ and $\mathbf{b} = \mathbf{b}_T$. To be more precise the following algorithm is used to assign the index i for $i = 1, 2, ..., T$ to each RF element:

for $i \leftarrow 1$ **to** T **do**

1. *Find an RF component whose input is* $\mathbf{a}_i$;
2. *Assign index* i *to that RF component*;
3. *Denote the output as* $\mathbf{b}_i$;
4. $\mathbf{a}_{i+1} = \mathbf{b}_i$;

end

Note that step 1 has always an answer because of how $\mathbf{a}_i$ and $\mathbf{b}_i$ are defined. Moreover, it is possible that more than one RF component satisfy the condition in step 1. In these cases, the components are parallel, i.e, the signals are simultaneously entering them and any ordering of these components is acceptable. Now, for $i = 1, 2, ..., T$ we can write $\mathbf{b}_i$ in terms of $\mathbf{a}_i$. If the i^{th} RF component is a phase-shifter, a power divider, or a power combiner, then we have $\mathbf{b}_i = \mathbf{\Phi}(\gamma,\phi,\eta)\mathbf{P}_\pi\mathbf{a}_i$, $\mathbf{b}_i = \mathbf{Q}(\gamma,\mu,\eta)\mathbf{P}_\pi\mathbf{a}_i$, or $\mathbf{b}_i = \mathbf{Q}^t(\gamma,\mu,\eta)^t\mathbf{P}_\pi\mathbf{a}_i$, respectively. Note that, if the order of the signals is not changed before the i^{th} component, we have

$\mathbf{P}_\pi = \mathbf{I}$. Hence, the given ASP can be expressed as in (7.12). □

To illustrate the application of this result, consider the ASP network example in Fig. 7.5. By using the indexing scheme introduced in the Proof of Proposition 7.1, this network can be reorganized as a product of basic RF transformations as shown in Fig. 7.6. Note that in the latter figure, permutation matrices only appear before the 7^{th} and 15^{th} RF components; for the remaining components the permutation is an identity matrix (not shown for simplicity). It is worth mentioning that the indexing is not unique and parallel components can be swapped, for instance, the order of u_2, u_3 and u_4 does not affect the I/O relationship of the ASP network.

In the following theorem, we present five commutative properties of matrices $\mathbf{\Phi}(\gamma, \phi, \eta)$, $\mathbf{P}_\pi$, $\mathbf{Q}(\gamma, \mu, \eta)$, and $\mathbf{Q}^t(\gamma, \mu, \eta)$ which later will be used to rearrange the RF components for further simplifications.

Theorem 7.1. *For each one of the following products of two basic RF component matrices on the left, there exists an equivalent matrix factorization as given on the right of the equality sign:*

$$\mathbf{Q}(\gamma, \mu, \eta)\mathbf{P}_\pi = \mathbf{P}_{\pi'}\mathbf{Q}(\pi(\gamma), \mu, \eta), \tag{7.13}$$

$$\mathbf{P}_\pi\mathbf{Q}^t(\gamma, \mu, \eta) = \mathbf{Q}^t(\pi(\gamma), \mu, \eta)\mathbf{P}_{\pi'}, \tag{7.14}$$

$$\begin{aligned}&\mathbf{Q}(\gamma_1, \mu_1, \eta_1)\mathbf{Q}^t(\gamma_2, \mu_2, \eta_2)\\ &= \prod_{j=1}^{J'} \mathbf{Q}^t(\gamma'_j, \mu'_j, \eta'_j)\mathbf{P}_{\pi'} \prod_{j'=1}^{J''} \mathbf{Q}(\gamma''_{j'}, \mu''_{j'}, \eta''_{j'}),\end{aligned} \tag{7.15}$$

$$\mathbf{Q}(\gamma_1, \mu, \eta_1)\mathbf{\Phi}(\gamma_2, \phi, \eta_2) = \prod_{j=1}^{J} \mathbf{\Phi}(\gamma'_j, \phi'_j, \eta'_j)\mathbf{Q}(\gamma', \mu', \eta'), \tag{7.16}$$

$$\mathbf{\Phi}(\gamma_1, \phi, \eta_1)\mathbf{Q}^t(\gamma_2, \mu, \eta_2) = \mathbf{Q}^t(\gamma', \mu', \eta') \prod_{j=1}^{J} \mathbf{\Phi}(\gamma'_j, \phi'_j, \eta'_j). \tag{7.17}$$

The definitions of the parameters appearing on the right hand side of these identities are given in the proof.

Proof. See Appendix A. □

Next, we introduce three ASP sub-networks and their compact equivalent representations; these will play a key role in establishing our main results in Theorems 7.2 and 7.3.

- *Phase-shifter network*: This sub-network is obtained by cascading J basic phase shifter matrices (with accompanying permutations) of common size $N_{\mathbf{p}}$, i.e.,

$$\prod_{j=1}^{J} \mathbf{\Phi}(\gamma_j, \phi_j, N_{\mathbf{p}}) \mathbf{P}_{\pi_j} = \mathbf{E}_{\mathbf{v}} \mathbf{P}_{\pi}, \tag{7.18}$$

  where, as illustrated in Fig. 7.7a:

$$\mathbf{E}_{\mathbf{v}} = \mathrm{diag}(\mathbf{v}), \tag{7.19}$$

  with $\mathbf{v} = [e^{j\phi_1}, e^{j\phi_2}, ..., e^{j\phi_{N_{\mathbf{P}}}}]^t \in \mathbb{U}^{N_{\mathbf{P}}}$.

- *Power divider network*: By cascading J power divider matrices of compatible sizes, we obtain,

$$\prod_{j=1}^{J} \mathbf{Q}(\gamma_j, \mu_j, \eta_j) = \mathbf{P}_{\pi} \mathbf{D}_{\mathbf{d}} \mathbf{P}_{\pi'}, \tag{7.20}$$

  where, as illustrated in Fig. 7.7b,

$$\mathbf{D}_{\mathbf{d}} = \mathrm{bd}(\frac{1}{\sqrt{\delta_1}} \mathbf{1}_{\delta_1}, \frac{1}{\sqrt{\delta_2}} \mathbf{1}_{\delta_2}, \ldots, \frac{1}{\sqrt{\delta_{N_d}}} \mathbf{1}_{\delta_{N_d}}, \mathbf{I}), \tag{7.21}$$

  with $\mathbf{d} = [\delta_1, \delta_2, \ldots, \delta_{N_d}]^t$, and $\sum_{i=1}^{N_d} \delta_i = M_d$, which is equivalent to an RF network that divides N_d RF signals into a total of M_d signals. The presence of the identity matrix in (7.21) accounts for branches that are not divided.

- *Power combiner network*: By cascading J power combiner matrices, we obtain,

$$\prod_{j=1}^{J} \mathbf{Q}^t(\gamma_j, \mu_j, \eta_j) \mathbf{P}_{\pi_j} = \mathbf{P}_{\pi} \mathbf{C}_{\mathbf{d}} \mathbf{P}_{\pi'}, \tag{7.22}$$

  where, as illustrated in Fig. 7.7c,

$$\mathbf{C}_{\mathbf{d}} = \mathrm{bd}(\frac{1}{\sqrt{\delta_1}} \mathbf{1}_{\delta_1}^t, \frac{1}{\sqrt{\delta_2}} \mathbf{1}_{\delta_2}^t, \ldots, \frac{1}{\sqrt{\delta_{M_c}}} \mathbf{1}_{\delta_{M_c}}^t, \mathbf{I}), \tag{7.23}$$

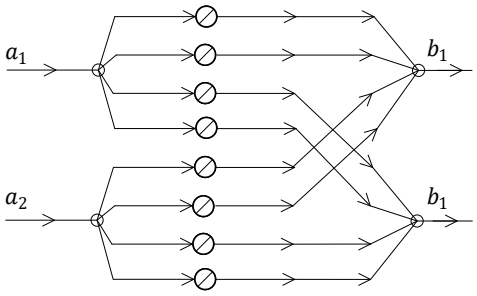


Fig. 7.8: ASP network equivalent to the ones in Figs. 7.5 and 7.6.

with $\mathbf{d} = [\delta_1, \delta_2, \ldots, \delta_{M_c}]$ and $\sum_{i=1}^{M_c} \delta_i = N_c$, which is equivalent to an RF network that combines N_c RF signals into M_c signals.

The validity of the identities in (7.18), (7.20) and (7.22) is demonstrated in Appendix B.1. We can now derive a mathematical expression for the representation of any given ASP network.

Theorem 7.2. *Any arbitrarily connected feed-forward ASP network with M inputs and N outputs, implemented by a total number of T phase-shifters, power dividers, and power combiners can be modeled as,*

$$\mathbf{b} = \frac{1}{\sqrt{MN}}\mathbf{A}\mathbf{a}, \tag{7.24a}$$

$$\mathbf{A} \in \breve{\mathbb{U}}^{N\times M}, \tag{7.24b}$$

where $\mathbf{a} \in \mathbb{C}^M$ and $\mathbf{b} \in \mathbb{C}^N$ are the input and output signals, respectively, and $\breve{\mathbb{U}} = \{z \in \mathbb{C}\colon |z| \leq 1\}$. That is, all the entries of matrix $\mathbf{A}$ have magnitude less or equal to 1.

Proof. See Appendix C □

Going back to our previous example in Fig. 7.5 and Fig. 7.6, the ASP network in the latter figure can be transformed into that of Fig. 7.8, for which the 2×2 transformation

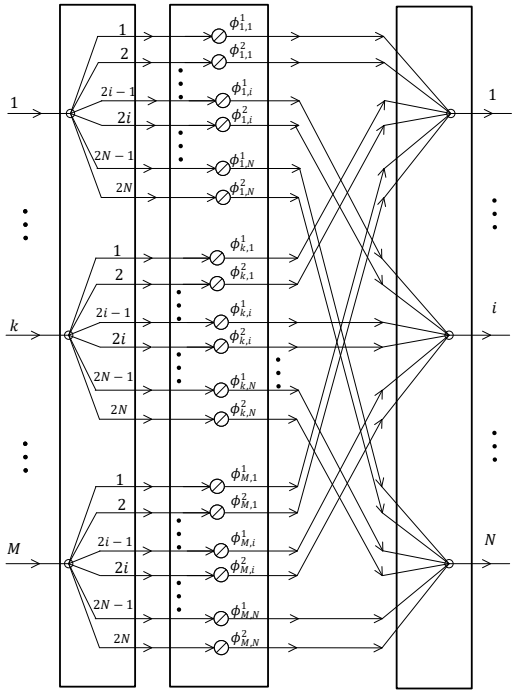


Fig. 7.9: Proposed ASP architecture.

matrix $\mathbf{A}$ satisfies the condition of the theorem. Now, we investigate whether any matrix in the convex set $\breve{\mathbb{U}}^{N\times M}$ can be realized by an ASP.

Theorem 7.3. *Any given matrix* $\mathbf{A} \in \breve{\mathbb{U}}^{N\times M}$ *can be realized by an ASP network with a total number of* $T = 2MN + M + N$ *RF components, i.e.,* N *dividers,* M *combiners, and* $2NM$ *(unit-modulus) phase shifters, as shown in Fig. 7.9.*

Proof. The output of the ASP in Fig. 7.9, corresponding to the input vector $\mathbf{a}$, can be expressed as,

$$b_i = \frac{1}{\sqrt{2M}} \sum_{k=1}^{M} \Big(\frac{a_k}{\sqrt{2N}} e^{j\phi^1_{k,i}} + \frac{a_k}{\sqrt{2N}} e^{j\phi^2_{k,i}} \Big) \tag{7.25a}$$

$$= \frac{1}{\sqrt{MN}} \sum_{k=1}^{M} a_k \Big(\frac{1}{2} \sum_{l=1}^{2} e^{j\phi^l_{k,i}} \Big). \tag{7.25b}$$

In (7.25a), since b_i is the output of a $2M$-way combiner, the normalization factor $\frac{1}{\sqrt{2M}}$

appears from (7.9). Similarly, the k^{th} input, i.e., a_k is divided into $2N$ branches which according to (7.9) introduces a normalization factor of $\frac{1}{\sqrt{2N}}$. Subsequently, for a given $\mathbf{A} \in \breve{\mathbb{U}}^{N\times M}$, we have $A_{ki} \leq 1$, and by invoking Lemma 7.1 (presented after the proof), there exist angles $\phi_{k,i}^{l}$'s such $A_{ki} = \frac{1}{L}\sum_{l=1}^{L} e^{j\phi_{k,i}^{l}}$ where the minimum possible value of L is two, i.e., $A_{ki} = \frac{1}{2}(e^{j\phi_{k,i}^{1}} + e^{j\phi_{k,i}^{2}})$. Therefore, we have,

$$b_i = \frac{1}{\sqrt{MN}} \sum_{k=1}^{M} A_{ki} a_k, \tag{7.26}$$

where $\mathbf{A} \in \breve{\mathbb{U}}^{N\times M}$. Moreover, $2M$ phase-shifters are required for each element of $\mathbf{b}$ and consequently, a minimum of $2MN$ phase-shifters are needed. □

Lemma 7.1. *Any complex number z where $0 \leq |z| \leq L$ for $L \geq 2$ can be written as: $z = \sum_{l=1}^{\bar{L}} e^{j\theta_l}$ where $\bar{L} = L + (L \bmod 2)$, $\theta_l \in [0, 2\pi]$ and θ_l's may be non-unique.*

Proof. The proof for $L = 2$ is presented in [23], thus, for a given $0 \leq |z'| \leq 2$ we have $z' = e^{j\theta_1} + e^{j\theta_2}$. Thus, it is sufficient to provide the proof for $L = 2L'$ and then $L = 2L' + 1$ when $L \geq 2$.

For $L = 2L'$: We have $\bar{L} = 2L'$, thus, we can write $z = L'z'$ where $0 \leq |z'| \leq 2$. Therefore, we can similarly write $z = L'(e^{j\theta_1} + e^{j\theta_2}) = L'e^{j\theta_1} + L'e^{j\theta_2}$. Then by writing $L'e^{j\theta_1} = \sum_{l=1}^{L'} e^{j\theta_1}$ and $L'e^{j\theta_2} = \sum_{l=1}^{L'} e^{j\theta_2}$ we can write $z = \sum_{l=1}^{L'} e^{j\theta_1} + \sum_{l=1}^{L'} e^{j\theta_2}$ which follows $z = \sum_{l=1}^{2L'} e^{j\theta'_l}$ where $\theta'_l = \theta_1$ for $l = 1, 2, ..., L'$ and $\theta'_l = \theta_2$ for $l = L' + 1, ..., L$.

For $L = 2L' + 1$: We have $\bar{L} = 2(L' + 1)$, thus from the above case there exist θ'_l's such that $z = \sum_{l=1}^{2(L'+1)} e^{j\theta'_l}$.

□

Remark 7.1. *The significant result of Theorem 7.3, is that any $\mathbf{A} \in \breve{\mathbb{U}}^{N\times M}$ can be implemented with an ASP structure using conventional RF components, i.e. combiners, dividers and phase shifters, whose input-output relationship is not bound to the unit modulus constraint. That is, while the individual phase-shifter components satisfy this constraint, the overall transformation matrix implemented by the proposed structure in Fig. 7.9 is no longer restricted to the unit modulus constraint. Thus, the troubling non-convexity constraint found in the literature on hybrid beamforming literature can be lifted from the design optimization problems.*

Remark 7.2. *According to the above proof, non-unique solutions for phase-shifter may exist. This additional degree of freedom can be considered when designing the ASP network based on the requirements and constraints of the analog system. By writing* $A_{p,q} = |A_{p,q}| \exp(j\angle A_{p,q})$*, one possible solution for* $\phi^1_{k,i}$ *and* $\phi^2_{k,i}$ *is given by,*

$$\phi^1_{k,i} = \angle A_{p,q} + \cos^{-1}(|A_{p,q}|) \tag{7.27a}$$

$$\phi^2_{k,i} = \angle A_{p,q} - \cos^{-1}(|A_{p,q}|). \tag{7.27b}$$

In the conventional hybrid structure $T = MN + M + N$ *RF components are required [21–29]. In contrast, the proposed ASP structure requires MN additional phase-shifters, for a total of* $T = 2MN+M+N$ *RF components. These additional components, when employed as in Fig. 7.9, allow to lift the constant modulus constraint for the overall transformation.*

Remark 7.3. *It is worth mentioning that since for wide-band systems it is desirable to have a common ASP network for the entire band [60, 65, 75] the proposed structure can be used for MIMO-OFDM systems. Particularly, since the proposed ASP structure is not bound to constant modulus constraint, it simplifies the design of hybrid MIMO-OFDM beamformers.*

7.4 Transmitter and Receiver Design with Generalized HSP

While the previous section focused on the realization of the RF mappings $\mathcal{G}_R$ and $\mathcal{G}_T$, as defined in (7.5) and (7.6), using basic RF components, in this section we turn our attention to the baseband mappings $\mathcal{F}_R$ and $\mathcal{F}_T$ as defined in (7.7) and (7.4), respectively. To this end, we consider the ASP network in Fig. 7.9 for $\mathcal{G}_T$ and $\mathcal{G}_R$ and consequently, (7.5) and (7.6) are replaced by,

$$\mathcal{G}_T(\mathbf{x}_T^{RF}) = \mathbf{A}_T \mathbf{x}_T^{RF}, \tag{7.28a}$$

$$\mathcal{G}_R(\mathbf{y}) = \mathbf{A}_R \mathbf{y}, \tag{7.28b}$$

where $\mathbf{A}_T \in \breve{\mathbb{U}}^{N_T \times M_T}$ and $\mathbf{A}_R \in \breve{\mathbb{U}}^{M_R \times N_R}$. We first focus on the transmitter and then on the receiver design.

7.4.1 HSP Design at the Transmitter

Considering (7.4), (7.5), and (7.28a), the transmitted signal of the generalized HSP can be written as follows,

$$\mathbf{x}_T = \sqrt{\rho}\mathbf{A}_T\mathcal{F}_T(\mathbf{s}). \tag{7.29}$$

In the literature on hybrid beamforming, $\mathcal{F}_T$ is usually a linear transformation, i.e., $\mathbf{x}_T = \sqrt{\rho}\mathbf{A}_T\mathbf{P}\mathbf{s}$, where $\mathbf{P} \in \mathbb{C}^{M_T \times K}$ is the precoding matrix. We first explore the properties and implementation of $\mathcal{F}_T$, and then discuss the design of $\mathcal{F}_T$ and $\mathbf{A}_T$ at the HSP-based transmitter.

Let $\mathcal{D}_T(\mathbf{s})$ denote the transformation that generates the desired transmitted signal from the given vector symbol $\mathbf{s}$. In effect, this function can represent a generic communication technique at the transmitter side. For instance, the optimal eigen-mode precoding is obtained by solving the following problem:

$$\max_{\mathbf{P}} \quad \log_2 \det(\mathbf{I}_{N_R} + \mathbf{H}\mathbf{P}\mathbf{P}^H\mathbf{H}^H), \tag{7.30a}$$

$$\text{subject to} \quad \mathrm{Tr}(\mathbf{P}\mathbf{P}^H) \leq P_T. \tag{7.30b}$$

The solution is given by

$$\mathbf{P} = \mathbf{V}\boldsymbol{\Upsilon}, \tag{7.31}$$

where the diagonal weight matrix $\boldsymbol{\Upsilon}$ is calculated via water filling [10] and $\mathbf{V}$ is a unitary matrix obtained from singular value decomposition of the channel matrix, i.e.,

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H. \tag{7.32}$$

Consequently, for this particular precoding scheme, we have,

$$\mathcal{D}_T(\mathbf{s}) = \mathbf{V}\boldsymbol{\Upsilon}\mathbf{s}. \tag{7.33}$$

Note that nonlinear beaforming, channel estimation, space-time coding and many other techniques can also be represented by $\mathcal{D}_T(\mathbf{s})$.

From (7.29), in order to generate the same transmit signal as a given $\mathcal{D}_T(\mathbf{s})$ via an

HSP-based transmitter, we need to find $\mathbf{A}_T$ and $\mathcal{F}_T(.)$ such that,

$$\mathbf{A}_T \mathcal{F}_T(\mathbf{s}) = \mathcal{D}_T(\mathbf{s}), \tag{7.34}$$

holds for all symbol vectors $\mathbf{s}$. Hence, since $\mathcal{D}_T(\mathbf{s})$ is given, $\mathcal{F}_T(\mathbf{s})$ can be defined as the following set, or multi-valued function:

$$\mathcal{F}_T(\mathbf{s}) \triangleq \{\mathbf{x} \in \mathbb{C}^K : \mathbf{A}_T \mathbf{x} = \mathcal{D}_T(\mathbf{s})\}. \tag{7.35}$$

Note that while it might be very difficult to explicitly construct the mapping $\mathcal{F}_T(.)$, obtaining its output, i.e., $\mathcal{F}_T(\mathbf{s})$ is simple because the value of $\mathcal{D}_T(\mathbf{s})$ is available. In other words, since the output of the HSP-based transmitter is given, i.e. $\mathcal{D}_T(\mathbf{s})$, it is sufficient to calculate the desired output of $\mathcal{F}_T(.)$ rather than implementing the mapping itself.

From (7.4) and (7.35), we can rewrite (7.34) as,

$$\mathbf{A}_T \mathbf{x}_T^{BB} = \mathcal{D}_T(\mathbf{s}), \tag{7.36}$$

which means that in general the HSP objective is *to find* $\mathbf{A}_T$ *and* $\mathbf{x}_T^{BB}$ *such that* (7.36) *is satisfied for the given* $\mathcal{D}_T(\mathbf{s})$. This objective guarantees that the HSP-based system achieves the same performance as the FD one, i.e., $\mathcal{D}_T(\mathbf{s})$. However, many variations can be derived according to the conditions and constraints of the system, which opens new avenues for investigation in this area.

In practice, depending on the system constraints, one may wish to design $\mathbf{A}_T$, $\mathbf{x}_T^{BB}$ and possibly some other system parameters represented by vector $\mathbf{p}$ on the basis of some optimization criterion. For instance, the following generic optimization problem can be used for obtaining the HSP parameters,

$$\min_{\mathbf{A}_T, \mathbf{x}_T^{BB}, \mathbf{p}} \quad \mathbb{E}\{\|\mathbf{A}_T \mathbf{x}_T^{BB} - \mathcal{D}_T(\mathbf{s})\|^2\}, \tag{7.37a}$$

$$\text{subject to} \quad C(\mathbf{A}_T, \mathbf{x}_T^{BB}, \mathbf{p}), \tag{7.37b}$$

where $C(\mathbf{A}_T, \mathbf{x}_T^{BB}, \mathbf{p})$ represents the system constraints. Alternatively, this could be for-

mulated as

$$\underset{\mathbf{A}_T, \mathbf{x}_T^{BB}, \mathbf{p}}{\text{optimize}} \quad f(\mathbf{A}_T, \mathbf{x}_T^{BB}, \mathbf{p}), \tag{7.38a}$$

$$\text{subject to} \quad \mathbf{A}_T \mathbf{x}_T^{BB} = \mathcal{D}_T(\mathbf{s}). \tag{7.38b}$$

where $f(.)$ is the chosen cost function based on the objectives of the system. Note that the power constraint is not necessary as it can be taken into account when designing $\mathcal{D}_T(\mathbf{s})$. One obvious choice is $f(\mathbf{A}_T, \mathbf{x}_T^{BB}, \mathbf{p}) = 1$, in which case $\mathbf{A}_T$ must be designed such that for some set $\mathbb{S} \subset \mathcal{A}^K$, we have $\mathcal{D}_T(\mathbf{s}) \in \text{span}(\mathbf{A}_T), \forall \mathbf{s} \in \mathbb{S}$, where $\text{span}(\mathbf{A}_T)$ denotes the span of $\mathbf{A}_T$. Consequently, the baseband signal is obtained from $\mathbf{x}_T^{BB} = \mathbf{A}_T^{\dagger} \mathcal{D}_T(\mathbf{s})$ where $\mathbf{A}_T^{\dagger}$ is the Moore Penrose inverse of matrix $\mathbf{A}_T$. In what follows, we present different cost functions for designing precoding matrices with HSP.

Unconstrained FD Precoding for $M_T \geq K$

For $M_T \geq K$ it is possible to realize any given FD precoder. As an example, we explore optimal eigen-mode precoding, although any other precoding matrix can be obtained in the same fashion. We first consider the case $M_T = K$ and subsequently discuss the modifications needed for $M_T \geq K$.

From (7.33) and (7.36), both $\mathbf{A}_T$ and $\mathbf{x}_T^{BB}$ must be designed such that,

$$\mathbf{A}_T \mathbf{x}_T^{BB} = \mathbf{V} \boldsymbol{\Upsilon} \mathbf{s}. \tag{7.39}$$

Since $\mathbf{A}_T$ is of size $N_T \times M_T$, this problem, for $M_T = K$, has the following simple solution:

$$\mathbf{A}_T = \frac{1}{p_0} \mathbf{V} \boldsymbol{\Upsilon}, \tag{7.40a}$$

$$\mathbf{x}_T^{BB} = p_0 \mathbf{s}, \tag{7.40b}$$

where $p_0 = \|\text{vec}(\mathbf{V}\boldsymbol{\Upsilon})\|_{\infty}$.

In the case $M_T > K$, one possible solution that achieves the same performance as the FD precoding is to append $M_T - K$ zeros to the solution $\mathbf{x}_T^{BB}$ in (7.40b) and set the corresponding columns of $\mathbf{A}_T$ in (7.40a) to zero.

Note that no constraint is enforced on the system and similar to existing hybrid solutions in the literature, $\mathbf{A}_T$ must be updated according to the channel coherence time, denoted as T_c in the sequel. Since $\mathbf{s}$ changes after every symbol duration T_s, $\mathbf{x}_T^{BB}$ is also updated every T_s.

Unconstrained FD Precoding for $M_T < K$

In this case, from using either (7.38) or (7.38), it is possible to obtain various hybrid beamformer designs depending on the system requirements. Here, we aim at minimizing the Euclidean distance between the eigen-mode FD precoder in (7.31) and the hybrid beamforming matrix $\mathbf{A}_T$. However, since the former has size $N_T \times M_T$ while the latter has size $N_T \times K$, we first find a beamforming matrix $\hat{\mathbf{A}}_T$ of size $N_T \times M_T$ subject to a rank M_T constraint, i.e.,

$$\min_{\hat{\mathbf{A}}_T} \quad \|\hat{\mathbf{A}}_T - \mathbf{V}\boldsymbol{\Upsilon}\|^2 \tag{7.41a}$$

$$\text{subject to} \quad \text{rank}(\hat{\mathbf{A}}_T) = M_T. \tag{7.41b}$$

Since here $\Upsilon = [\text{diag}(\upsilon_1, \upsilon_2, ..., \upsilon_K), \mathbf{0}_{K\times(N_T-K)}]^t$, we can write the solution for the above problem as,

$$\hat{\mathbf{A}}_T = \mathbf{V} \begin{bmatrix} \text{diag}(\upsilon_1, \upsilon_2, ..., \upsilon_{M_T}, 0, ..., 0) \\ \mathbf{0}_{(N_T-K)\times K} \end{bmatrix}. \tag{7.42}$$

Now by defining,

$$\mathbf{A}_T = \mathbf{V} \begin{bmatrix} \text{diag}(\upsilon_1, \upsilon_2, ..., \upsilon_{M_T}) \\ \mathbf{0}_{(N_T-M_T)\times M_T} \end{bmatrix}, \tag{7.43}$$

we can obtain $\mathbf{x}_T^{BB}$ by solving,

$$\min_{\mathbf{x}_T^{BB}} \quad \|\mathbf{A}_T\mathbf{x}_T^{BB} - \mathbf{V}\boldsymbol{\Upsilon}\mathbf{s}\|^2, \tag{7.44}$$

which yields,

$$\mathbf{x}_T^{BB} = [\mathbf{I}_{M_T}, \mathbf{0}_{M_T\times(K-M_T)}]\mathbf{s}. \tag{7.45}$$

Minimum number of RF chains with fast phase-shifters

If we do not have a constraint on the update rate of the analog components, we can reduce the number of RF chains by solving the following problem:

$$\min_{\mathbf{A}_T, \mathbf{x}_T^{BB}} \quad M_T, \tag{7.46a}$$

$$\text{subject to} \quad \mathbf{A}_T \mathbf{x}_T^{BB} = \mathcal{D}_T(\mathbf{s}). \tag{7.46b}$$

This problem is shown to have non-unique solution for $M_T = 1$ where $\mathcal{D}_T(\mathbf{s}) = \mathbf{V}\boldsymbol{\Upsilon}\mathbf{s}$ in [74] but essentially the same solution is valid for any other transmit function $\mathcal{D}_T(\mathbf{s})$. Note that in this case the ASP must be updated after every symbol duration T_s.

7.4.2 HSP Design at the Receiver

Similar to the previous subsection, let us assume that the ideal FD decoder that maps the received RF signal $\mathbf{y}$ into the detected symbols $\hat{\mathbf{s}}$, represented by the mapping $\mathcal{D}_R(\mathbf{y})$, is known. Since in massive-MIMO systems beamforming and multiplexing are key techniques, linear detection is of great interest due to its simplicity. In this case, which is considered in our discussion, $\mathcal{D}_R(\mathbf{y}) = \mathbf{Z}\mathbf{y}$ where $\mathbf{Z} \in \mathbb{C}^{K \times N_R}$ is the FD combiner matrix. However, at the price of increased computational complexity, $\mathcal{D}_R(\mathbf{y})$ can be extended to more sophisticated detectors such as maximum likelihood or sphere decoding.

By substituting (7.6) and (7.28b) in (7.7), the estimated signal at the receiver is written as,

$$\hat{\mathbf{s}} = \mathcal{F}_R\big(\mathbf{A}_R(\mathbf{H}\mathbf{x}_T + \mathbf{n})\big). \tag{7.47}$$

Clearly, the same approach used in Subsection 7.4.1 for realizing the transformation $\mathcal{F}_T(.)$ cannot be applied here because the desired output of $\mathcal{F}_R(.)$ is unknown, i.e., we need this mapping to implement the decoding function. Ideally, we want to find a mapping $\mathcal{F}_R(\cdot)$ and $\mathbf{A}_R$ such that,

$$\mathcal{F}_R\big(\mathbf{A}_R\mathbf{y}\big) = \mathcal{D}_R(\mathbf{y}), \tag{7.48}$$

or all $\mathbf{y}$. Similar to the HSP literature [21–29], we consider linear transformation for the baseband processing, i.e., $\mathcal{F}_R(\mathbf{x}_R^{BB}) = \mathbf{W}\mathbf{x}_R^{BB}$ where $\mathbf{W} \in \mathbb{C}^{K \times M_R}$ is the corresponding transformation matrix; however, extension to types of transformations is straightforward by using (7.48). Consequently, the following generic optimization problem can be considered

for obtaining the HSP parameters:

$$\min_{\mathbf{A}_R,\mathbf{W},\mathbf{p}} \quad \mathbb{E}\{\|\mathbf{W}\mathbf{A}_R - \mathbf{Z}\|_F^2\}, \tag{7.49a}$$

$$\text{subject to} \quad C(\mathbf{A}_R, \mathbf{W}, \mathbf{p}), \tag{7.49b}$$

where $C(\mathbf{A}_R, \mathbf{W}, \mathbf{p})$ represents the system constraints. Alternatively, this could be formulated as,

$$\underset{\mathbf{A}_R,\mathbf{W},\mathbf{p}}{\text{optimize}} \quad f(\mathbf{A}_R, \mathbf{W}, \mathbf{p}), \tag{7.50a}$$

$$\text{subject to} \quad \mathbf{W}\mathbf{A}_R = \mathbf{Z}, \tag{7.50b}$$

where $f(.)$ is a cost function designed to satisfy the requirements of the system. In what follows, FD combining for point-to-point MIMO is presented as an example.

Unconstrained FD Combining for $M_R \geq K$

We first consider the case where $M_R = K$ and subsequently discuss the case $M_R > K$. The optimal FD combiner for a point-to-point MIMO can be obtained from,

$$\max_{\mathbf{Z}} \ \log_2 \det\left(\mathbf{I}_K + \rho(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Z}\mathbf{H}\mathbf{H}^H\mathbf{Z}^H\right). \tag{7.51}$$

From (7.32), the solution is given by,

$$\mathbf{Z}^H = \mathbf{U}^a, \tag{7.52}$$

where $\mathbf{U} = [\mathbf{U}^a, \mathbf{U}^b]$ and $\mathbf{U}^a$ contains the first K columns of $\mathbf{U}$, corresponding to the K dominant singular values of the channel matrix $\mathbf{H}$. Thus, $\mathbf{A}_R$ and $\mathbf{W}$ must be jointly designed such that,

$$\mathbf{W}\mathbf{A}_R = \mathbf{Z}, \tag{7.53}$$

where $\mathbf{A}_R \in \breve{\mathbb{U}}^{M_R \times N_R}$. Note that if $M_T = K$, for any FD combiner $\mathbf{Z} \in \mathbb{C}^{K \times N_R}$, this problem has the following solution:

$$\mathbf{A}_R = \frac{1}{p_1}\mathbf{Z}, \tag{7.54a}$$

$$\mathbf{W} = p_1 \mathbf{I}_K, \tag{7.54b}$$

where $p_1 = \|\mathrm{vec}(\mathbf{Z})\|_\infty$.

The above design can be extended to the case $M_R > K$, although here including more RF chains adds to the cost and complexity of the system while no improvement is gained. One trivial solution that guarantees the same performance as the FD solution is to set the additional $M_R - K$ columns of $\mathbf{W}$ to 0, i.e., using $\mathbf{W} = p_1[\mathbf{I}_K, \mathbf{0}_{K\times(M_R-K)}]$.

The FD realization for the multi-user case can be similarly obtained. First (7.51) must be replaced by the desired optimization problem for finding the FD combiner. Analog and digital combiners are then calculated by (7.54).

FD Combining for $M_T < K$

In the case of linear decoding, there must be at least K independent equations to recover K transmitted symbols. Hence, the minimum number of required RF chains is $M_R = K$. Consequently, combiner design for $M_R < K$ is not practical in this case.

Minimum number of RF chains with fast phase-shifters

Even with the same assumption as in Subsection IV-A.3, i.e. the phase-shifters can be updated every T_s, at least K RF chains are required. Since only the channel matrix is known at the receiver which changes each T_c, a faster update rate of the phase-shifters does not provide any extra degrees of freedom and hence does not help in reducing the number of RF chains at the receiver. Consequently, the minimum number of possible RF chains for digital linear combining is $M = K$.

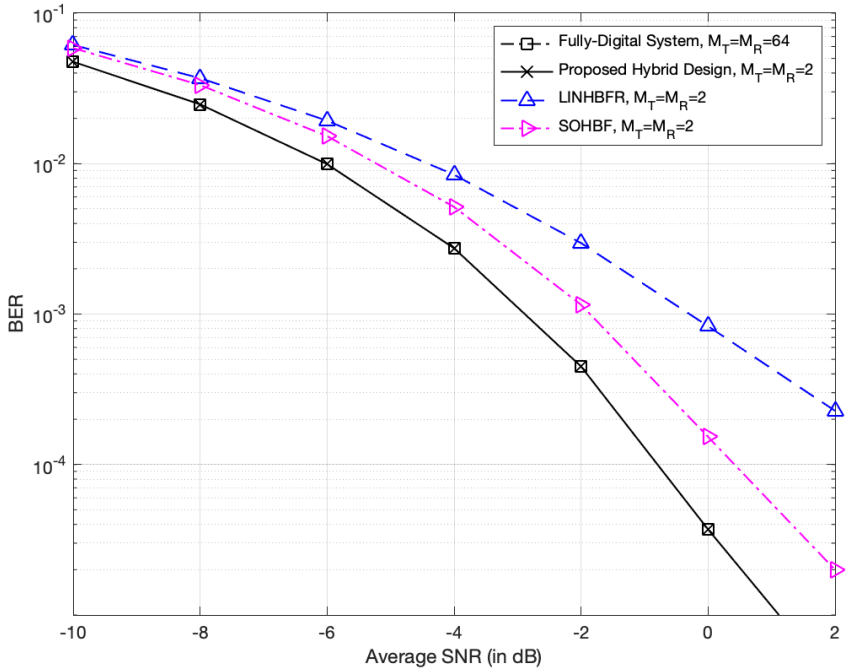


Fig. 7.10: BER versus SNR of different methods for a point-to-point massive-MIMO system with $N_T = N_R = 64$.

7.5 Simulation Results

In this section, we present simulation results for different scenarios and compare the FD system with our proposed hybrid architecture as well as existing hybrid designs in the literature.

The following channel models is used for all the simulations:

$$\mathbf{H} = \sqrt{\frac{N_T N_R}{N_{\mathrm{c}} N_{\mathrm{ray}}}} \sum_{i=1}^{N_{\mathrm{c}}} \sum_{j=1}^{N_{\mathrm{ray}}} \alpha_{ij} \mathbf{a}_{\mathrm{r}}(\theta_{ij}^{\mathrm{r}}) \mathbf{a}_{\mathrm{t}}(\theta_{ij}^{\mathrm{t}})^H, \tag{7.55}$$

where $N_{\mathrm{c}} = 5$ is the number of clusters, and $N_{\mathrm{ray}} = 10$ is the number of rays in each cluster. Similar to [22, 75], the path gains are independently generated as $\alpha_{ij} \sim \mathcal{CN}(0, 1)$. The transmit and receive antenna responses are denoted by $\mathbf{a}_{\mathrm{r}}(\theta_{ij}^{\mathrm{r}})$ and $\mathbf{a}_{\mathrm{t}}(\theta_{ij}^{\mathrm{t}})$ respectively,

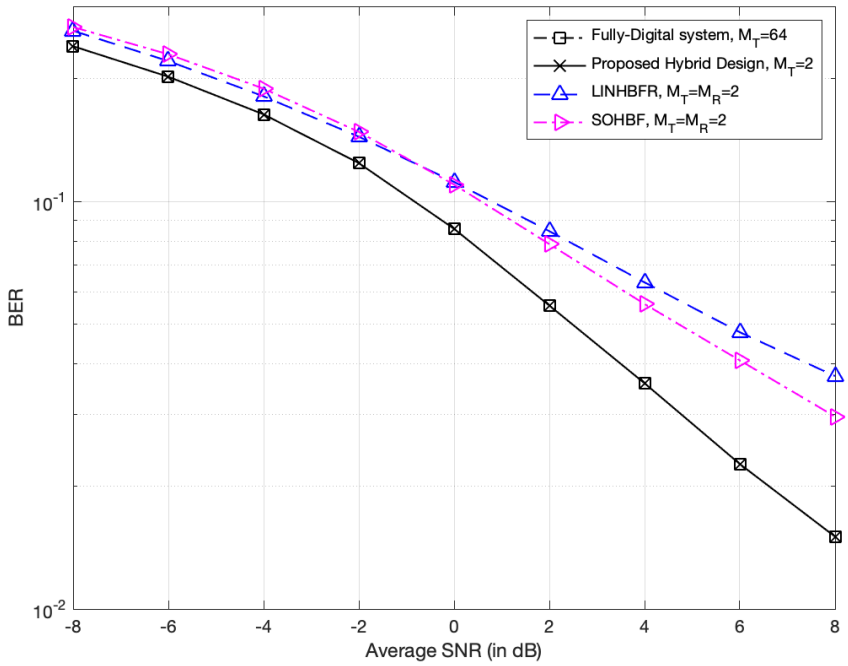


Fig. 7.11: BER versus SNR of different methods for a downlink connection with $N_T = 64$ antenna massive-MIMO BS and a single user with $N_R = 2$ antennas.

where,

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{N}}[1, e^{j\pi \sin(\phi)}, \ldots, e^{j(N-1)\pi \sin(\phi)}], \tag{7.56}$$

for uniform linear arrays of size N. The angles of arrival θ_{ij}^{r} and departure θ_{ij}^{t} are independently generated according to the Laplacian distribution with the mean cluster angles $\bar{\theta}_{ij}^{\mathrm{r}}$ and $\bar{\theta}_{ij}^{\mathrm{t}}$, uniformly distributed in $[0, 2\pi]$. The angular spread is 10 degrees within each cluster. We further assume that the channel estimation and system synchronization are perfect.

For the massive MIMO simulations, the number of antennas N_T and N_R vary between 2 and 64; larger values are used for UM-MIMO, larger number of antennas are used, as later indicated. For the proposed and existing hybrid designs, the number of RF chains and transmitted symbols are set as $M_T = M_R = K$. Unless otherwise indicated, we set $K = 2$ and 64-QAM modulation is used for all the simulations. For the FD systems, we set $M_T = N_T$ and $M_R = N_R$, while the same value as for the hybrid designs is used.

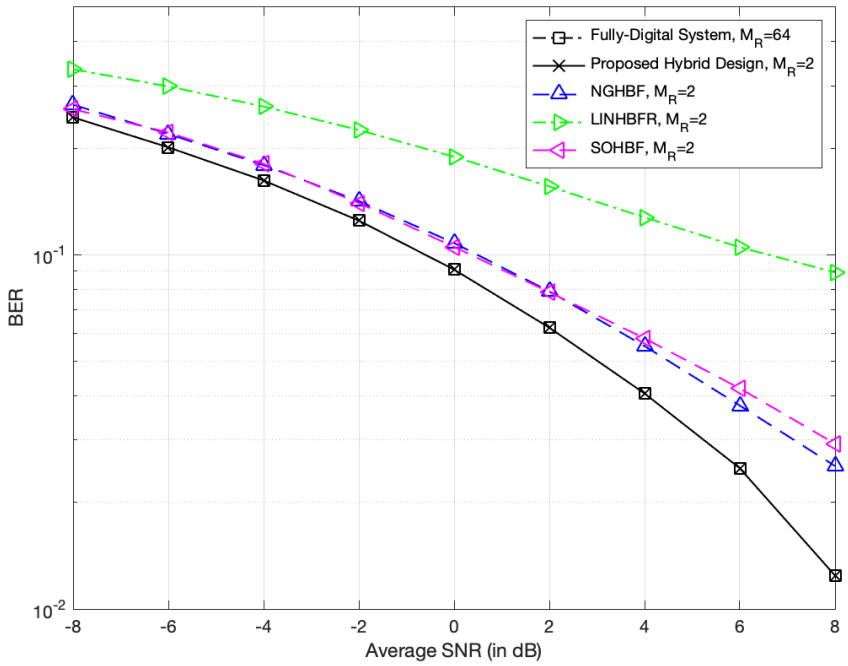


Fig. 7.12: BER versus SNR of different methods for an uplink connection with $N_R = 64$ antenna massive-MIMO BS and a single-user with $N_T = 2$.

Simulation results are presented for the optimal FD precoder and combiner, our proposed hybrid precoder and combiner realization of FD in Subsections IV-A.1 and IV-B.1, as well as selected hybrid designs from [22,75]. For M RF chains and N antennas, the proposed and the conventional structures require $T = 2MN + M + N$ and $T = MN + M + N$ RF components, respectively.

7.5.1 BER performance

BER performance versus SNR (SNR=ρ/σ^2) for three different setups is shown in Fig. 7.10 to 7.12. Fig. 7.10 presents the results for a massive-MIMO system with $N_T = N_R = 64$ antennas (and $M_T = M_R = 2$ RF chains). The downlink BER performance of a massive-MIMO BS with $N_T = 64$ antennas transmitting to a single user with $N_R = 2$ antenna is shown in Fig. 7.11, while the uplink BER performance for the system is shown in Fig. 7.12. It can be seen that in all the simulated scenarios the proposed hybrid realization matches the performance of the FD systems while outperforming the existing hybrid designs. The

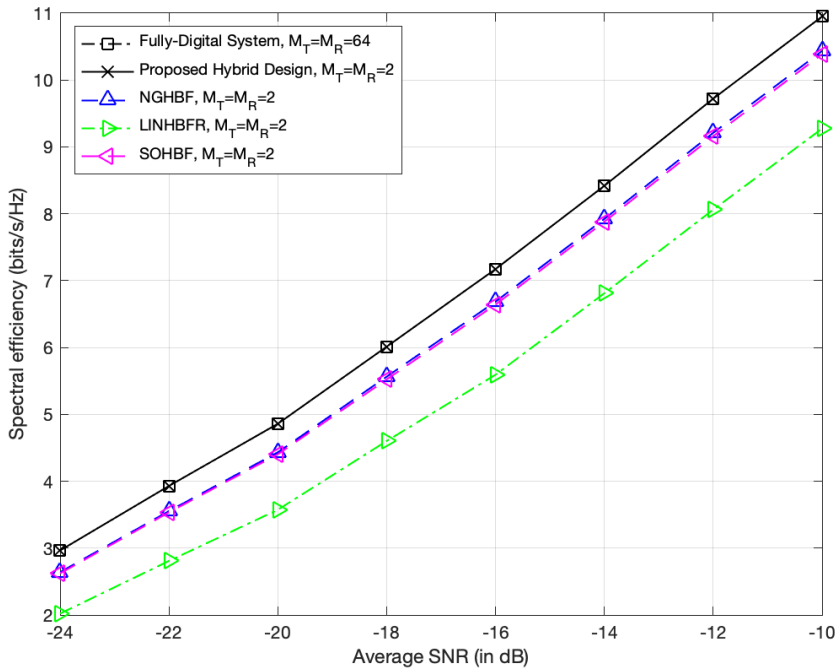


Fig. 7.13: Spectral efficiency versus SNR of different methods for a point-to-point massive-MIMO system with $N_T = N_R = 64$.

FD systems require $M_T = M_R = 64$ RF chains whereas the proposed design achieves the same performance with only 2 RF chains. Consequently, the proposed design outperforms the existing hybrid designs with the same number of RF chains. As discussed in Subsections IV-A.1 and IV-B.1, the proposed hybrid design generates the same signals as the FD system with limited number of RF chains by employing the proposed ASP network. In particular, since the RF output of the proposed structure is identical to that of the desired FD system, the same performance as the optimal FD beamforming can be achieved.

7.5.2 Spectral efficiency

The spectral efficiency (in bits/s/Hz) of optimal FD beamforming, proposed hybrid realizations of FD as well as the hybrid designs from [22, 75, 99] for massive-MIMO system with $N_T = N_R = 64$ antennas is shown in Fig. 7.13. The spectral efficiency of an uplink

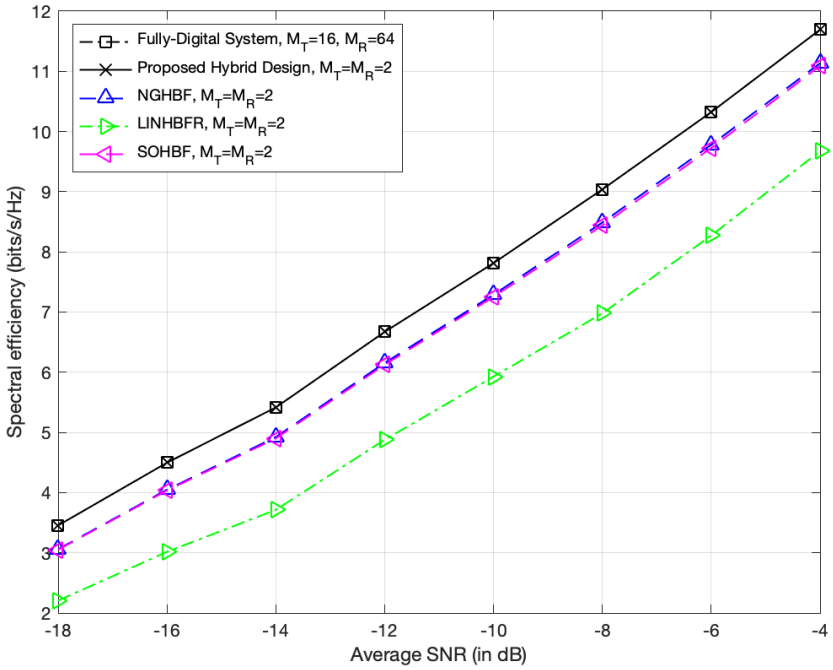


Fig. 7.14: Spectral efficiency versus SNR of different methods for an uplink connection with $N_R = 64$ antenna massive-MIMO BS and a single-user with $N_T = 16$ antennas.

connection for a single user with $N_T = 16$ antennas and a massive-MIMO BS with $N_R = 64$ antennas is presented in Fig. 7.14. Furthermore, Fig. 7.15 shows the spectral efficiency of a downlink connection for a massive-MIMO BS with$N_T = 64$ antennas and a single user with $N_R = 4$ antennas. As expected, the proposed ASP-based realizations achieve the same rate as their FD counterparts and and therefore outperform existing hybrid designs. In order to evaluate the performance of the proposed ASP structure when number of antennas grows larger, simulations are performed for ultra-massive MIMO system configurations. Spectral efficiency versus number of transmitter antennas N_T is plotted in Fig. 7.16 for different number of receive antennas. For the FD system the number of RF chains is equal to the number of transmitter antennas, i.e., $M_T = N_T$ whereas for the proposed hybrid structure the number of antennas is kept equal to the number of transmitted symbols, i.e., $M_T = K$. It can be seen that in all cases, the hybrid design with the proposed ASP architecture achieves the same performance as the corresponding FD system. For instance, for an ultra-

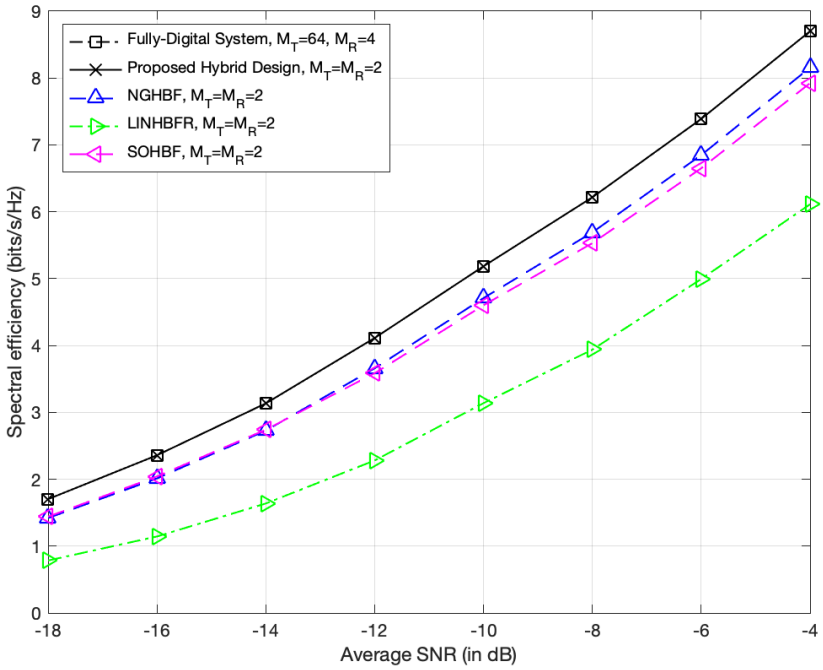


Fig. 7.15: Spectral efficiency versus SNR of different methods for an downlink connection with $N_T = 64$ antenna massive-MIMO BS and a single-user with $N_R = 4$.

massive MIMO transmitter with $N_T = 1024$ antennas and receiver with $M_T = 2$ antennas, the FD structure requires $N_R = 1024$ RF chains while the proposed structure guarantees the same performance with $M_T = 2$ RF chains.

7.5.3 Computational Complexity

The proposed ASP architecture is implemented with the same RF components as the conventional hybrid structures [21–29]. Moreover, since the constant unit modulus is not imposed on the entries for the resulting analog transformation matrix with our approach, the computational complexity of designing the analog and digital beamformers can be reduced. Compared to the FD system design, the additional computations required for the proposed ASP approach lie in the calculation of the phase-shifter parameters as given in (7.27). In the case of an eigen-mode FD beamformer for instance, the calculations in (7.27), in terms of complexity order, are dominated by the SVD and water filling algorithm needed for FD design, as represented by (7.31). Moreover, existing hybrid designs use sophisticated

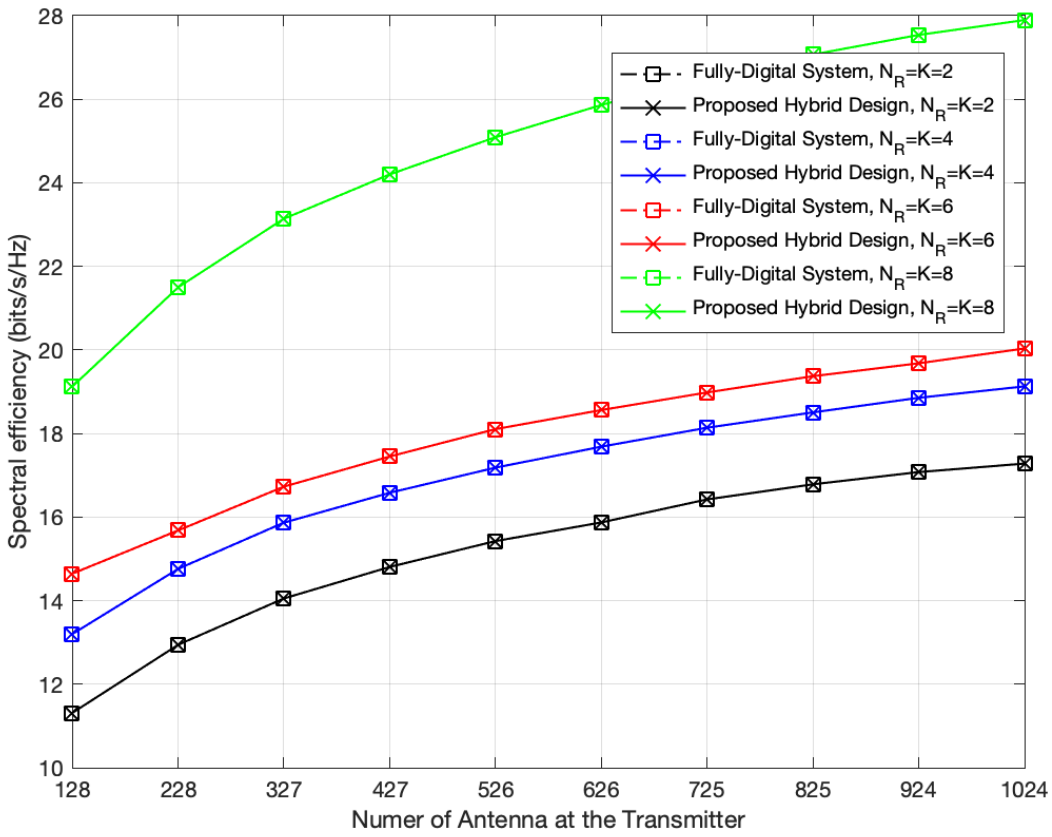


Fig. 7.16: Spectral efficiency versus number of transmitter antennas for proposed and FD beamforming, with different numbers of receive antennas.

optimization or reconstruction techniques to handle the constant modulus constraint. For instance, the iterative algorithms in [22] and [75] require matrix inversion in each iteration. Consequently, the computational complexity of the proposed FD realizations with ASP is less than each iteration in these hybrid designs.

7.6 Conclusion

In this chapter, we investigated the hybrid A/D structure as a general framework for signal processing in massive and ultra-massive-MIMO systems. We first explored the ASP network in details by developing a mathematical representation for any arbitrarily connected

feed-forward ASP network comprised of phase-shifters, power-dividers and power combiners. Then, a novel ASP structure was proposed which is not bound to the unit modulus constraint. Subsequently, we focused on the transmitter and receiver sides by exploiting the newly proposed ASP architecture and generalizing generalizing the digital processing. Specifically, the optimization problem for the HSP beamformer was reformulated within the new representation framework, which facilitates its solution under a variety of constraints and requirements for the massive MIMO system. Finally simulation results were presented illustrating the superiority of the proposed architecture to the conventional hybrid designs for massive-MIMO systems. Next chapter, summarizes this thesis and provides some topics for future research.

Chapter 8

Conclusion

In this chapter, the main contributions of this thesis are summarized and some possible avenues for future works are discussed.

8.1 Summary

In this thesis, we studied hybrid analog/digital signal processing. Our main focus was on achieving performance of FD systems with limited number of RF chains. After presenting the literature survey and system model, we focused on the HBF design for the transmitter in part I and then studied the receiver HBF in part II. Finally, in part III, a generalized hybrid structure for hybrid signal processing was presented.

In the first part of this thesis, we considered HBF at the transmitter side and study the minimum number of required RF chains for realizing any MIMO FDP with the HSP architecture. We first presented a solution for an efficient implementation of HSP for massive-MIMO systems in Chapter 3. The proposed design requires a minimum number of RF chains while matching the performance of optimal baseband FDP. We introduced the generalized system formulation of the hybrid architecture. Then, we minimized the number of RF chains such that any given FDP can be realized in the extended HSP architecture. Based on the obtained optimum solutions, we then presented our hybrid design for realizing any FDP in the HSP architecture with a single RF chain. Then, we presented three novel beamformer schemes which eliminate the defects of SRCA, and can achieve the performance of FD beamforming schemes. Moreover, the applications of these systems for optimal precoding in both SU and MU cases was studied.

We then extended our design to wideband massive-MIMO mmWave communications in Chapter 4. The notion of time domain signal reconstruction of OFDM signals was introduced. Then, we formulated the problem and presented the conditions for realization of any given FDP in HSP systems. Next, after introducing the PRSP realization, we discussed the general solution for any arbitrary number of RF chains and as an example, presented a design for two RF chains.

In the second part of this thesis, HSP at the receiver side was studied. In Chapter 5, we presented a novel HBF structure for massive-MIMO communication systems that matches the performance of any given FD combiner. We took advantage of a decomposition technique which allowed us to realize the FD combiner by HBF. Finite-resolution phase-shifters were considered to design a modified hybrid structure with phase-offset as an additional degree of freedom in order to improve the system performance. A closed-form solution for the phase-offset which minimizes the error between FD combiner and its analog realization was given. A novel hybrid channel estimation was also presented based on the proposed hybrid structure that achieves FD performance.

Subsequently, in Chapter 6, we studied robust hybrid analog/digital bemformer designs for uplink in massive-MIMO communication systems. Both SU and MU reception modes in uplink direction of a single-cell configuration with massive-MIMO BS were considered. The norm-bounded channel error was used to capture the imperfect CSI conditions and the objective function is formulated based on the worst-case robustness MMSE. For SU scenario, a hierarchical optimization as well as joint optimization based on UE capabilities were presented. We also proposed robust hybrid combiner design for MU uplink connection.

Finally, in the third part of this thesis, the hybrid A/D structure as a general framework for signal processing in massive and ultra-massive-MIMO systems was studied. In Chapter 7, the ASP network was studied in details by developing a mathematical representation for any arbitrarily connected feed-forward ASP network comprised of phase-shifters, power-dividers and power combiners. We then presented a novel ASP structure which is not bound to the unit modulus constraint. Subsequently, the transmitter and receiver sides were considered. Specifically, the optimization problem for the HSP beamformer was reformulated within the new representation framework, which facilitates its solution under a variety of constraints and requirements for the massive MIMO system. Finally, precoding and combining designs under different conditions are discussed as examples.

8.2 Future Works

We considered wide-band hybrid beamforming at the transmitter in Chapter 4, however, similar technique can be developed for hybrid combiner design at the receiver which remains a topic of future work. Moreover, the robust combiner design in Chapter 6 can be extended to robust precoder design under imperfect CSI.

The proposed ASP architecture in Chapter 7 can still be applied with finite-precision phase-shifters, but some of the results derived in the here will become approximative. This raises interesting questions from a theoretical perspective, e.g.: how to characterize the effect of phase-shifter quantization on the performance of the overall ASP network, and how to select the quantization levels of the phase-shifters to ensure a given accuracy in the final ASP design? These questions remain of interest for future research.

In light of the huge success of machine learning (ML) and particularly deep learning (DL) in various fields of engineering, recently, deep neural networks (DNNs) have attracted considerable attention among researchers for designing HSP systems.

Moreover, since in conventional HSP techniques, the focus was on designing the analog and digital beamformers, the DNN-based HSP techniques also followed the same guideline. However, this approach imposes a fundamental limit on the minimum number of required RF chain. Specifically, the minimum number of required RF chain must be greater or equal to the number of transmitted/received symbols or the rank of FD matrix. Due to the so-called expressive power of DNNs, it is theoretically possible for DNN-based HSP to achieve FD performance with smaller number of RF chines.

One of the possible directions for the future work is to develop analog deep neural networks using conventional RF components in HSP systems. A deep learning framework for HSP in mmWave massive-MIMO systems can be developed to enable HSP-based systems to achieve the same performance as the FD system with much smaller number of RF chains.

Appendix A

Proof of Theorem 7.1

By transposing (7.14) and (7.17), and the fact that the transpose of a permutation matrix is also a permutation matrix, we arrive at (7.13) and (7.16). The proofs for the remaining properties are presented below.

A.1 Proof of (7.13)

We show that for any vector $\mathbf{x}$ there exist π' such that, if we have $\hat{\mathbf{x}}^{(1)} = \mathbf{Q}(\gamma, \mu, \eta)\mathbf{P}_{\pi}\mathbf{x}$ and $\hat{\mathbf{x}}^{(2)} = \mathbf{P}_{\pi'}\mathbf{Q}(\pi(\gamma), \mu, \eta)\mathbf{x}$ then $\hat{\mathbf{x}}^{(1)} = \hat{\mathbf{x}}^{(2)}$. By denoting $\mathbf{x}^{(\pi)} = \mathbf{P}_{\pi}\mathbf{x}$, we can write,

$$\hat{\mathbf{x}}^{(1)} \overset{\pi}{=} [{\mathbf{x}^{(\pi)}}^t, x_{\gamma}^{(\pi)}\mathbf{1}_{\mu-1}^t]^t, \tag{A.1}$$

since $x_{\gamma}^{(\pi)} = x_{\pi(\gamma)}$ we can further write,

$$\hat{\mathbf{x}}^{(1)} \overset{\pi}{=} [\mathbf{x}^t, x_{\pi(\gamma)}\mathbf{1}_{\mu-1}^t]^t. \tag{A.2}$$

On the other hand we can write,

$$\mathbf{Q}\big(\pi(\gamma), \mu, \eta\big)\mathbf{x} \overset{\pi}{=} [\mathbf{x}^t, x_{\pi(\gamma)}\mathbf{1}_{\mu-1}^t]^t, \tag{A.3}$$

thus, we can conclude $\hat{\mathbf{x}}^{(1)} \overset{\pi}{=} \mathbf{Q}\big(\pi(\gamma), \mu, \eta\big)\mathbf{x}$; therefore, there exist $\mathbf{P}_{\pi'}$ such that $\hat{\mathbf{x}}^{(1)} = \hat{\mathbf{x}}^{(2)}$. ■

A.2 Proof of (7.15)

For $\gamma_1 < \gamma_2$ and $\gamma_1 > \gamma_2$, we show that $J = J' = 1$ and for $\gamma_1 = \gamma_2$ it will be shown that $J = \mu_1$ and $J' = \mu_2$. First, considering $\gamma_1 > \gamma_2$, we can write,

$$\begin{aligned}&\mathbf{Q}(\gamma_1, \mu_1, \eta)\mathbf{Q}^t(\gamma_2, \mu_2, \eta) = \\ &\mathrm{bd}(\mathbf{I}_{\gamma_1-1}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta-\gamma_1})\mathrm{bd}(\mathbf{I}_{\gamma_2-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta-\gamma_2}).\end{aligned} \tag{A.4}$$

With simple matrix manipulation, we have,

$$\begin{aligned}&\mathrm{bd}(\mathbf{I}_{\gamma_1-1}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta-\gamma_1})\mathrm{bd}(\mathbf{I}_{\gamma_2-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta-\gamma_2}) = \\ &\mathrm{bd}(\mathbf{I}_{\gamma_2-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta-\gamma_2+\mu_1-1})\mathrm{bd}(\mathbf{I}_{\gamma_1-2+\mu_2}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta-\gamma_1}),\end{aligned} \tag{A.5}$$

we can further write,

$$\begin{aligned}&\mathrm{bd}(\mathbf{I}_{\gamma_2-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta-\gamma_2+\mu_1-1})\mathrm{bd}(\mathbf{I}_{\gamma_1-2+\mu_2}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta-\gamma_1}) = \\ &\mathbf{Q}^t(\gamma_2, \mu_2, \eta + \mu_1 - 1) \\ &\mathbf{Q}(\gamma_1 + \mu_2 - 1, \mu_1, \eta + \mu_2 - 1).\end{aligned} \tag{A.6}$$

In case of $\gamma_1 < \gamma_2$, we can accordingly write,

$$\begin{aligned}&\mathbf{Q}(\gamma_1, \mu_1, \eta)\mathbf{Q}^t(\gamma_2, \mu_2, \eta) = \\ &\mathrm{bd}(\mathbf{I}_{\gamma_1-1}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta-\gamma_1})\mathrm{bd}(\mathbf{I}_{\gamma_2-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta-\gamma_2}).\end{aligned} \tag{A.7}$$

Similar to (A.5), we have,

$$\begin{aligned}&\mathrm{bd}(\mathbf{I}_{\gamma_1-1}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta-\gamma_1})\mathrm{bd}(\mathbf{I}_{\gamma_2-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta-\gamma_2}) = \\ &\mathrm{bd}(\mathbf{I}_{\gamma_2+\mu_1-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta+\mu_1-1})\mathrm{bd}(\mathbf{I}_{\gamma_1-1}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta+\mu_2-1}).\end{aligned} \tag{A.8}$$

We can then write,

$$\begin{aligned}&\mathrm{bd}(\mathbf{I}_{\gamma_2+\mu_1-1}, \frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}, \mathbf{I}_{\eta+\mu_1-1})\mathrm{bd}(\mathbf{I}_{\gamma_1-1}, \frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I}_{\eta+\mu_2-1}) = \\ &\mathbf{Q}^t(\gamma_2+\mu_1-1, \mu_2, \eta+\mu_1-1)\mathbf{Q}(\gamma_1, \mu_1, \eta+\mu_2-1).\end{aligned} \tag{A.9}$$

For $\gamma_1 = \gamma_2$, without loss of generality, we only provide the proof for $\gamma_1 = \gamma_2 = 1$ and $\eta = 1$, extending to other values of γ_1,γ_2 and η straightforward but tedious. Hence, we show,

$$\begin{aligned}&\mathbf{Q}(1, \mu_1, 1)\mathbf{Q}^t(1, \mu_2, 1) = \\ &\prod_{j=1}^{\mu_1}\mathbf{Q}^t(\gamma'_j, \mu'_j, \eta'_j)\mathbf{P}_\pi\prod_{j'=1}^{\mu_2}\mathbf{Q}(\gamma''_{j'}, \mu''_{j'}, \eta''_{j'}).\end{aligned} \tag{A.10}$$

The left hand side can be written as,

$$\mathbf{Q}(1, \mu_1, 1)\mathbf{Q}^t(1, \mu_2, 1) = \frac{1}{\sqrt{\mu_1\mu_2}}\mathbf{1}_{\mu_1\times\mu_2}. \tag{A.11}$$

Now, by letting $\gamma''_{j'} = (j'-1)\mu_1 + 1$, $\mu''_{j'} = \mu_1$ and $\eta''_{j'} = (j'-1)\mu_1 + \mu_2 - j' + 1$, we have,

$$\begin{aligned}&\prod_{j'=1}^{\mu_2}\mathbf{Q}\big((j'-1)\mu_1+1, \mu_1, (j'-1)\mu_1+\mu_2-j'+1\big) = \\ &\mathbf{I}_{\mu_2}\otimes\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}.\end{aligned} \tag{A.12}$$

We can then form the permutation matrix $\mathbf{P}_\pi$ with,

$$\pi_i = \lceil i/\mu_1\rceil + \mu_2\big((i-1) \bmod (\mu_1)\big), \tag{A.13}$$

which results in,

$$\mathbf{P}_{\pi_i}\big(\mathbf{I}_{\mu_2}\otimes\mathbf{1}_{\mu_1}\big) = \mathbf{1}_{\mu_1}\otimes\mathbf{I}_{\mu_2}. \tag{A.14}$$

Then by letting $\gamma'_j = j$, $\mu'_j = \mu_1$ and $\eta'_j = (\mu_2 - j)\mu_1 + j$, we arrive at,

$$\prod_{j=1}^{\mu_1}\mathbf{Q}\big(j, \mu_2, (\mu_1-j)\mu_2+j\big) = \mathbf{I}_{\mu_1}\otimes\frac{1}{\sqrt{\mu_2}}\mathbf{1}^t_{\mu_2}. \tag{A.15}$$

From the above equation, (A.12) and (A.14), we can write,

$$\prod_{j=1}^{\mu_1} \mathbf{Q}^t(\gamma'_j, \mu'_j, \eta'_j)\mathbf{P}_\pi \prod_{j'=1}^{\mu_2} \mathbf{Q}(\gamma''_{j'}, \mu''_{j'}, \eta''_{j'}) = \frac{1}{\sqrt{\mu_1\mu_2}}\big(\mathbf{I}_{\mu_1} \otimes \mathbf{1}^t_{\mu_2}\big)\big(\mathbf{1}_{\mu_1} \otimes \mathbf{I}_{\mu_2}\big). \tag{A.16}$$

The right hand side of the above equation can be further simplified as,

$$\big(\mathbf{I}_{\mu_1} \otimes \mathbf{1}^t_{\mu_2}\big)\big(\mathbf{1}_{\mu_1} \otimes \mathbf{I}_{\mu_2}\big) = \mathbf{1}_{\mu_1\times\mu_2}, \tag{A.17}$$

which can be easily verified by invoking the mixed-product property[1]. Consequently, using the above equation,(A.16) and (A.11), we showed (A.10) is held which concludes the proof. ■

A.3 Proof of (7.16)

For $\gamma_1 > \gamma_2$, and $\gamma_1 < \gamma_2$,

$$\prod_{j=1}^{J} \mathbf{\Phi}(\gamma'_j, \phi'_j, \eta'_j)\mathbf{Q}(\gamma_1, \mu, \eta) = \mathbf{Q}(\gamma_1, \mu, \eta)\mathbf{\Phi}(\gamma_2, \phi, \eta), \tag{A.18}$$

is actually simplified to the case where J is equal to one. Thus, for $\gamma_1 > \gamma_2$, one can easily check,

$$\mathbf{\Phi}(\gamma_2, \phi, \eta + \mu - 1)\mathbf{Q}(\gamma_1, \mu, \eta) = \mathbf{Q}(\gamma_1, \mu, \eta)\mathbf{\Phi}(\gamma_2, \phi, \eta). \tag{A.19}$$

Similarly, for $\gamma_1 < \gamma_2$, we can write,

$$\mathbf{\Phi}(\gamma_2, \phi, \eta + \mu - 1)\mathbf{Q}(\gamma_1, \mu, \eta) = \mathbf{Q}(\gamma_1, \mu, \eta)\mathbf{\Phi}(\gamma_2, \phi, \eta). \tag{A.20}$$

[1] If $\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$ and $\mathbf{D}$ are matrices of appropriate sizes, then $\big(\mathbf{A} \otimes \mathbf{b}\big)\big(\mathbf{C} \otimes \mathbf{D}\big) = \big(\mathbf{AC}\big) \otimes \big(\mathbf{BD}\big)$.

In case of $\gamma_1 = \gamma_2$, $J = \mu$ and we can write,

$$\prod_{j=1}^{\mu} \mathbf{\Phi}(\gamma_1 + j - 1, \phi, \eta + \mu - 1)\mathbf{Q}(\gamma_1, \mu, \eta) = \mathbf{Q}(\gamma_1, \mu, \eta)\mathbf{\Phi}(\gamma_1, \phi, \eta). \tag{A.21}$$

which can be easily verified by matrix manipulations. ■

Appendix B

Proof of (7.18), (7.20) and (7.22)

B.1 Proof of (7.18)

To show that $\mathbf{E_v}$ is a diagonal matrix, we can use induction and the fact that for a diagonal matrix $\mathbf{D}$ and permutation matrix $\mathbf{P}_\pi$, the matrix $\hat{\mathbf{D}} = \mathbf{P}_\pi \mathbf{D} {\mathbf{P}_\pi}^t$ is also a diagonal matrix. For $J = 1$ the statement is true, and we must prove for $J = K + 1$ we have,

$$\hat{\mathbf{E}}_\mathbf{v} \hat{\mathbf{P}}_\pi = \prod_{j=1}^{K+1} \mathbf{\Phi}(\gamma_j, \phi_j, N_\mathbf{p}) \mathbf{P}_{\pi_j}. \tag{B.1}$$

By assuming for $J = K$, matrix $\mathbf{E_v}$ is diagonal and $\mathbf{P}_\pi$ is a permutation matrix, we can rewrite the above equation as,

$$\hat{\mathbf{E}}_\mathbf{v} \hat{\mathbf{P}}_\pi = \mathbf{E_v} \mathbf{P}_\pi \mathbf{\Phi}(\gamma_{K+1}, \phi_{K+1}, N_\mathbf{p}) \mathbf{P}_{\pi_{K+1}}. \tag{B.2}$$

We can therefore write,

$$\hat{\mathbf{E}}_\mathbf{v} \hat{\mathbf{P}}_\pi = \mathbf{E_v} \mathbf{P}_\pi \mathbf{\Phi}(\gamma_{K+1}, \phi_{K+1}, N_\mathbf{p}) \mathbf{P}_\pi^t \mathbf{P}_\pi \mathbf{P}_{\pi_{K+1}}. \tag{B.3}$$

Using the aforementioned property of permutation matrices, we know,

$$\hat{\mathbf{\Phi}} = \mathbf{P}_\pi \mathbf{\Phi}(\gamma_{K+1}, \phi_{K+1}, N_\mathbf{p}) \mathbf{P}_\pi^t,$$

is a diagonal matrix. We further know that $\hat{\mathbf{P}}_\pi = \mathbf{P}_\pi \mathbf{P}_{\pi_{K+1}}$ is a permutation matrix thus we can write,

$$\hat{\mathbf{E}}_\mathbf{v} \hat{\mathbf{P}}_\pi = \mathbf{E}_\mathbf{v} \hat{\mathbf{\Phi}} \hat{\mathbf{P}}_\pi. \tag{B.4}$$

Since $\mathbf{E}_\mathbf{v}$ and $\hat{\mathbf{\Phi}}$ are both diagonal so is $\hat{\mathbf{E}}_\mathbf{v}$. Furthermore, since all the diagonal entries are unit modulo complex numbers their products are also on the unit circle and thus $\mathbf{v} = [e^{j\phi_1}, e^{j\phi_2}, ..., e^{j\phi_{N_\mathbf{P}}}]^t \in \mathbb{U}^{N_\phi}$. ■

B.2 Proof of (7.20)

Induction can be used to prove this statement. For $J = 1$, one can easily find $\mathbf{P}_{\pi_1'}$ such that,

$$\mathbf{P}_{\pi_1'}^t \mathrm{bd}(\mathbf{I}_{\gamma-1}, \frac{1}{\sqrt{\mu}}\mathbf{1}_\mu, \mathbf{I}_{\eta-\gamma}) = \begin{bmatrix} 0 & \frac{1}{\sqrt{\mu}}\mathbf{1}_\mu & 0 \\ \mathbf{I}_{\gamma-1} & 0 & 0 \\ 0 & 0 & \mathbf{I}_{\eta-\gamma} \end{bmatrix}, \tag{B.5}$$

accordingly there exist $\mathbf{P}_{\pi_2'}$, such that,

$$\mathbf{P}_{\pi_1'}^t \mathrm{bd}(\mathbf{I}_{\gamma-1}, \frac{1}{\sqrt{\mu}}\mathbf{1}_\mu, \mathbf{I}_{\eta-\gamma}) \mathbf{P}_{\pi_2'}^t = \begin{bmatrix} \frac{1}{\sqrt{\mu}}\mathbf{1}_\mu & 0 & 0 \\ 0 & \mathbf{I}_{\gamma-1} & 0 \\ 0 & 0 & \mathbf{I}_{\eta-\gamma} \end{bmatrix}, \tag{B.6}$$

using the fact that permutation matrices are orthogonal, we can write,

$$\mathbf{Q}(1, \mu, \eta) = \mathbf{P}_{\pi_1'} \mathrm{bd}(\frac{1}{\sqrt{\mu}}\mathbf{1}_\mu, \mathbf{I}_\eta) \mathbf{P}_{\pi_2'}.$$

Now, let us assume for $J = K$, we have,

$$\prod_{j=1}^{K} \mathbf{Q}(\gamma_j, \mu_j, \eta_j) = \mathbf{P}_\pi \mathbf{D}_\mathbf{d} \mathbf{P}_{\pi'}, \tag{B.7}$$

We can thus write the following for $J = K + 1$:

$$\prod_{j=1}^{K+1} \mathbf{Q}(\gamma_j, \mu_j, \eta_j) = \mathbf{P}_\pi \mathbf{D}_\mathbf{d} \mathbf{P}_{\pi'} \mathbf{Q}(\gamma_{K+1}, \mu_{K+1}, \eta_{K+1}). \tag{B.8}$$

According to the $J = 1$ case, there exist $\mathbf{P}_{\pi'_3}$ and $\mathbf{P}_{\pi'_4}$ such that,

$$\mathbf{Q}(\gamma_{K+1}, \mu_{K+1}, \eta_{K+1}) = \mathbf{P}_{\pi'_3} \mathrm{bd}(\frac{1}{\sqrt{\mu_{K+1}}} \mathbf{1}_{\mu_{K+1}}, \mathbf{I}) \mathbf{P}_{\pi'_4}, \tag{B.9}$$

thus,

$$\prod_{j=1}^{K+1} \mathbf{Q}(\gamma_j, \mu_j, \eta_j) = \mathbf{P}_{\pi} \mathbf{D}_{\mathbf{d}} \mathbf{P}_{\pi'} \mathbf{P}_{\pi'_3} \mathrm{bd}(\frac{1}{\sqrt{\mu_{K+1}}} \mathbf{1}_{\mu_{K+1}}, \mathbf{I}) \mathbf{P}_{\pi'_4}. \tag{B.10}$$

Let us first define $\mathbf{P}_{\pi'_5} = \mathbf{P}_{\pi'} \mathbf{P}_{\pi'_3}$, then by considering $\mathbf{D}_{\mathbf{d}} \mathbf{P}_{\pi'_5}$, the permutation matrix $\mathbf{P}_{\pi'_5}$ rearranges the columns of $\mathbf{D}_{\mathbf{d}}$. Therefore, there exist permutation matrix $\mathbf{P}_{\pi'_6}$ that rearranges the rows of $\mathbf{D}_{\mathbf{d}} \mathbf{P}_{\pi'_5}$ to make a block diagonal matrix,

$$\mathbf{P}^t_{\pi'_6} \mathbf{D}_{\mathbf{d}} \mathbf{P}_{\pi'_5} = \mathbf{D}_{\mathbf{d}'}, \tag{B.11}$$

where $\mathbf{D}_{\mathbf{d}'} = \mathrm{bd}(\frac{1}{\sqrt{\delta'_1}} \mathbf{1}_{\delta'_1}, \frac{1}{\sqrt{\delta'_2}} \mathbf{1}_{\delta'_2}, \ldots, \frac{1}{\sqrt{\delta'_{N_d}}} \mathbf{1}_{\delta'_{N_d}})$. It is possible that $\delta'_i = 1$ for individual i or some consecutive number indices which result it diagonal block of identity matrices $\mathbf{I}$. From (B.10) and (B.11), we arrive at,

$$\begin{aligned} &\mathbf{P}_{\pi} \mathbf{D}_{\mathbf{d}} \mathbf{P}_{\pi'} \mathbf{Q}_{(\gamma_{K+1}, \mu_{K+1}, \eta_{K+1})} = \\ &\mathbf{P}_{\pi} \mathbf{P}_{\pi'_6} \mathbf{D}_{\mathbf{d}'} \mathrm{bd}(\frac{1}{\sqrt{\mu_{K+1}}} \mathbf{1} \mu_{K+1}, \mathbf{I}) \mathbf{P}_{\pi'_4}. \end{aligned} \tag{B.12}$$

The above equation can be further simplified as, $\mathbf{D}_{\mathbf{d}'} \mathrm{bd}(\frac{1}{\sqrt{\mu_{K+1}}} \mathbf{1}_{\mu_{K+1}}, \mathbf{I}) = \mathbf{D}_{\mathbf{d}''}$, where $\mathbf{D}_{\mathbf{d}''} = \mathrm{bd}(\frac{1}{\sqrt{\delta'_1 + \mu_{K+1}}} \mathbf{1}_{\delta'_1 + \mu_{K+1}}, \frac{1}{\sqrt{\delta'_2}} \mathbf{1}_{\delta'_2}, \ldots, \frac{1}{\sqrt{\delta'_{N_d}}} \mathbf{1}_{\delta'_{N_d}})$. Therefore by defining $\mathbf{P}_{\pi'_7} = \mathbf{P}_{\pi} \mathbf{P}_{\pi'_6}$, and from (B.8) and (B.12), we have,

$$\prod_{j=1}^{K+1} \mathbf{Q}(\gamma_j, \mu_j, \eta_j) = \mathbf{P}_{\pi'_7} \mathbf{D}_{\mathbf{d}''} \mathbf{P}_{\pi'_4}, \tag{B.13}$$

which proves the statement. ■

B.3 Proof of (7.22)

For $J = 1$, we have to show that $\mathbf{C}_{\mathbf{d}}$ has the block diagonal structure of (7.23) in $\mathbf{P}_{\pi} \mathbf{C}_{\mathbf{d}} \mathbf{P}_{\pi'} = \mathbf{Q}^t(\gamma, \mu, \eta) \mathbf{P}_{\pi_1}$. According to (B.6), we can write $\mathbf{Q}^t(\gamma, \mu, \eta) \mathbf{P}_{\pi_1} = \mathbf{P}_{\pi} \mathrm{bd}(\frac{1}{\sqrt{\mu}} \mathbf{1}_{\mu}, \mathbf{I}) \mathbf{P}_{\pi'} \mathbf{P}_{\pi_1}$.

Since the product of two permutation matrices is also a permutation matrix, we have: $\mathbf{P}_{\pi'} = \mathbf{P}_{\pi'}\mathbf{P}_{\pi_1}$. To continue the proof with induction, we assume that for $J = K$ there exist $\mathbf{P}_{\pi}$,$\mathbf{P}_{\pi'}$ and $\mathbf{C_d}$ such that $\mathbf{P}_{\pi}\mathbf{C_d}\mathbf{P}_{\pi'} = \prod_{j=1}^{J} \mathbf{Q}^t(\gamma_j, \mu_j, \eta_j)\mathbf{P}_{\pi_j}$. Now, for $J = K + 1$ we can write,

$$\prod_{j=1}^{K+1} \mathbf{Q}^t(\gamma_j, \mu_j, \eta_j)\mathbf{P}_{\pi_j} = \mathbf{Q}^t(\gamma_1, \mu_1, \eta_1)\mathbf{P}_{\pi_1}\mathbf{P}_{\pi}\mathbf{C_d}\mathbf{P}_{\pi'}. \tag{B.14}$$

According to the $J = 1$ case, there exist $\mathbf{P}_{\pi'_1}$ and $\mathbf{P}_{\pi'_2}$, such that,

$$\mathbf{Q}^t(\gamma_1, \mu_1, \eta_1) = \mathbf{P}_{\pi'_1}\mathrm{bd}(\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I})\mathbf{P}_{\pi'_2}, \tag{B.15}$$

Be defining a new permutation matrix, $\mathbf{P}_{\pi'_3} = \mathbf{P}_{\hat{\pi_2}}\mathbf{P}_{\pi_1}\mathbf{P}_{\pi}$, we can write the left hand-side of (B.14) as,

$$\mathbf{Q}^t(\gamma_1, \mu_1, \eta_1)\mathbf{P}_{\pi_1}\mathbf{P}_{\pi}\mathbf{C_d}\mathbf{P}_{\pi'} = \mathbf{P}_{\pi'_1}\mathrm{bd}(\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I})\mathbf{P}_{\pi'_3}\mathbf{C_d}\mathbf{P}_{\pi'}. \tag{B.16}$$

Considering $\mathbf{P}_{\pi'_3}\mathbf{C_d}$, permutation matrix $\mathbf{P}_{\pi'_3}$ rearranges the rows of $\mathbf{C_d}$. Therefore, there exist permutation matrix $\mathbf{P}_{\pi'_4}$ that rearranges the columns of $\mathbf{P}_{\pi'_3}\mathbf{C_d}$ to make a block diagonal matrix,

$$\mathbf{P}_{\pi'_3}\mathbf{C_d}\mathbf{P}_{\pi'_4} = \mathbf{C_{d'}}, \tag{B.17}$$

where $\mathbf{C_{d'}} = \mathrm{bd}(\frac{1}{\sqrt{\delta'_1}}\mathbf{1}^t_{\delta'_1}, \frac{1}{\sqrt{\delta'_2}}\mathbf{1}^t_{\delta'_2}, \ldots, \frac{1}{\sqrt{\delta'_{M_c}}}\mathbf{1}^t_{\delta'_{M_c}})$. Note that it is possible that $\delta'_i = 1$ for individual i or some consecutive number indices which result it diagonal block of identity matrices $\mathbf{I}$. From (B.16), (B.17) and the fact that $\mathbf{P}_{\pi'_4}\mathbf{P}^t_{\pi'_4}$, we can write,

$$\mathbf{P}_{\pi'_1}\mathrm{bd}(\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I})\mathbf{P}_{\pi'_3}\mathbf{C_d}\mathbf{P}_{\pi'} = \mathbf{P}_{\pi'_1}\mathrm{bd}(\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I})\mathbf{C_{d'}}\mathbf{P}^t_{\pi'_4}\mathbf{P}_{\pi'}. \tag{B.18}$$

To further simplify the above equation, we can write, $\mathrm{bd}(\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I})\mathbf{C_{d'}} = \mathbf{C_{d''}}$, where,

$$\delta''_i = \begin{cases} \sum_{j=1}^{\mu_1} \delta'_j & i = 1 \\ \delta'_{(i-1+\mu_1)} & \text{otherwise} \end{cases}. \tag{B.19}$$

To take the last step, there exist permutation matrices $\mathbf{P}_{\pi'_5}$ and $\mathbf{P}_{\pi'_6}$, such that, $\mathbf{P}^t_{\pi'_5}\mathbf{C}_{\mathbf{d}''}\mathbf{P}^t_{\pi'_6} = \mathbf{C}_{\mathbf{d}'''}$, where for some L, we can have,

$$\mathbf{C}_{\mathbf{d}'''} = \mathrm{bd}(\frac{1}{\sqrt{\delta_1}}\mathbf{1}^t_{\delta_1}, \frac{1}{\sqrt{\delta_2}}\mathbf{1}^t_{\delta_2}, \ldots, \frac{1}{\sqrt{\delta_L}}\mathbf{1}^t_{\delta_L}, \mathbf{I}). \tag{B.20}$$

From the above equation and (B.18), we have,

$$\mathbf{P}_{\pi'_1}\mathrm{bd}(\frac{1}{\sqrt{\mu_1}}\mathbf{1}_{\mu_1}, \mathbf{I})\mathbf{P}_{\pi'_3}\mathbf{C}_{\mathbf{d}}\mathbf{P}_{\pi'} = \mathbf{P}_{\pi'_1}\mathbf{P}_{\pi'_5}\mathbf{C}_{\mathbf{d}'''}\mathbf{P}_{\pi'_6}\mathbf{P}^t_{\pi'_4}\mathbf{P}_{\pi'}. \tag{B.21}$$

Now by defining permutation matrices $\mathbf{P}_{\pi'_7} = \mathbf{P}_{\pi'_1}\mathbf{P}_{\pi'_5}$, $\mathbf{P}_{\pi'_8} = \mathbf{P}_{\pi'_6}\mathbf{P}^t_{\pi'_4}\mathbf{P}_{\pi'}$ and from (B.14) to (B.21), we arrive at,

$$\prod_{j=1}^{K+1}\mathbf{Q}^t(\gamma_j, \mu_j, \eta_j)\mathbf{P}_{\pi_j} = \mathbf{P}_{\pi'_7}\mathbf{C}_{\mathbf{d}'''}\mathbf{P}_{\pi'_8}, \tag{B.22}$$

which proves the statement. ■

Appendix C

Proof of Theorem 7.2

Without loss of generality, let us assume there are a total of T RF components, i.e., P combiners, R dividers repectively, and Q phase-shifters, so that $T = P+Q+R$. According to properties (7.14), (7.15) and (7.17) in Theorem 7.1, we can rewrite (7.12) as follows by commuting the combiner matrices to the left hand side, i.e,

$$\prod_{i=1}^{T} \mathbf{A}_{u_i}(\theta_i) = \prod_{j=1}^{P} \mathbf{Q}^t(\theta_j) \prod_{i=1}^{T'} \mathbf{A}_{u_i}(\theta_i), \tag{C.1}$$

where $T' = T - P$. Similarly, the divider matrices can be moved to the right hand side using properties (7.13), (7.15) and (7.17), thus,

$$\prod_{i=1}^{T} \mathbf{A}_{u_i}(\theta_i) = \prod_{j=1}^{P} \mathbf{Q}^t(\theta_j) \prod_{i=1}^{T''} \mathbf{A}_{u_i}(\theta_i) \prod_{k=1}^{R} \mathbf{Q}(\theta_k), \tag{C.2}$$

where $T'' = T-P-R$. In (C.2) only the permutation and single phase-shifter matrices are in the middle of the expression. Therefore, without loss of generality and due to the fact that permutation and single phase-shifter matrices can be identity matrices we can write: $\prod_{i=1}^{T''} \mathbf{A}_{u_i}(\theta_i) = \prod_{i=1}^{J} \mathbf{\Phi}(\gamma_i, \phi_i, \eta_i)\mathbf{P}_{\pi_i}$, hence,

$$\prod_{i=1}^{T} \mathbf{A}_{u_i}(\theta_i) = \prod_{j=1}^{P} \mathbf{Q}^t(\theta_j) \prod_{i=1}^{Q} \mathbf{\Phi}(\gamma_i, \phi_i, \eta_i)\mathbf{P}_{\pi_i} \prod_{j=k}^{R} \mathbf{Q}(\theta_k). \tag{C.3}$$

Now, using (7.18), (7.20), (7.22), and the fact that product of permutation matrices is another permutation matrix, we have,

$$\prod_{i=1}^{T} \mathbf{A}_{u_i}(\theta_i) = \mathbf{P}_{\pi_1}\mathbf{C}_{\mathbf{d}}\mathbf{P}_{\pi_2}\mathbf{E}_{\mathbf{v}}\mathbf{P}_{\pi_3}\mathbf{D}_{\hat{\mathbf{d}}}\mathbf{P}_{\pi_4}, \tag{C.4}$$

which follows,

$$\mathbf{b} = \mathbf{P}_{\pi_1}\mathbf{C}_{\mathbf{d}}\mathbf{P}_{\pi_2}\mathbf{E}_{\mathbf{v}}\mathbf{P}_{\pi_3}\mathbf{D}_{\hat{\mathbf{d}}}\mathbf{P}_{\pi_4}\mathbf{a}. \tag{C.5}$$

By defining,

$$\begin{aligned} \mathbf{b} &= \mathbf{P}_{\pi_1}\mathbf{h}^{(1)}, \quad & \mathbf{h}^{(1)} &= \mathbf{C}_{\mathbf{d}}\mathbf{h}^{(2)}, \quad & \mathbf{h}^{(2)} &= \mathbf{P}_{\pi_2}\mathbf{h}^{(3)}, \\ \mathbf{h}^{(3)} &= \mathbf{E}_{\mathbf{v}}\mathbf{h}^{(4)}, \quad & \mathbf{h}^{(4)} &= \mathbf{P}_{\pi_3}\mathbf{h}^{(5)}, \quad & \mathbf{h}^{(5)} &= \mathbf{D}_{\hat{\mathbf{v}}}\mathbf{h}^{(6)}, \\ \mathbf{h}^{(6)} &= \mathbf{P}_{\pi_4}\mathbf{a}, \end{aligned} \tag{C.6}$$

we can further have,

$$\begin{aligned} b_i &= h^{(1)}_{\pi_1^{-1}(i)}, \quad & h^{(1)}_i &= \frac{1}{\sqrt{\delta_i}} \sum_{j=\psi_\delta(i-1)+1}^{\psi_\delta(i)} h^{(2)}_j, \\ h^{(2)}_i &= h^{(3)}_{\pi_2^{-1}(i)}, \quad & h^{(3)}_i &= e^{j\phi_i} h^{(4)}_i \\ h^{(4)}_i &= h^{(5)}_{\pi_3{}^{-1}(i)}, \quad & h^{(5)}_i &= \frac{1}{\sqrt{\hat{\delta}_{\varphi_{\hat{\delta}}(i)}}} h^{(6)}_{\varphi_{\hat{\delta}}(i)}, \\ h^{(6)}_i &= a_{\pi_4{}^{-1}(i)}, \end{aligned} \tag{C.7}$$

where $\psi_\delta(i) = \sum_{k=1}^{i} \delta_k$ and $\varphi_{\hat{\delta}}(i) = \{j \in \mathbb{N} | \psi_{\hat{\delta}}(j-1) \leq i \leq \psi_{\hat{\delta}}(j)\}$. Consequently, we arrive at,

$$b_i = \delta^{-1/2}_{\pi_1^{-1}(i)} \sum_{j=\psi_\delta(\pi_1^{-1}(i)-1)+1}^{\psi_\delta(\pi_1^{-1}(i))} \hat{\delta}^{-1/2}_{\varphi_{\hat{\delta}}(\pi_3^{-1}(\pi_2^{-1}(j)))} \exp\{j\phi_{\pi_2^{-1}(j)}\} a_{\pi_4^{-1}(\varphi_{\hat{\delta}}(\pi_3^{-1}(\pi_2^{-1}(j))))}, \tag{C.8}$$

where by defining,

$$u_{i,j} = \pi_2^{-1}\Big(\psi_\delta\big(\pi_1^{-1}(i)-1\big)+j\Big), \tag{C.9}$$

$$\dot{v}(n) = \varphi_{\hat{\delta}}\big(\pi_3^{-1}(n)\big), \tag{C.10}$$

$$\ddot{v}(n) = \pi_4^{-1}(\dot{v}(n)), \tag{C.11}$$

we can write,

$$b_i = \frac{1}{\sqrt{\delta_{\pi_1^{-1}(i)}}} \sum_{j=1}^{\delta_{\pi_1^{-1}(i)}} \frac{a_{\ddot{v}(\dot{v}(u_{(i,j)}))}}{\sqrt{\hat{\delta}_{\dot{v}(u_{(i,j)})}}} \exp\{j\phi_{u_{(i,j)}}\}. \tag{C.12}$$

Without loss of generality, we can further have,

$$b_i = \frac{1}{\sqrt{\delta_{\pi_1^{-1}(i)}}} \sum_{k=1}^{M} a_k \Big(\frac{1}{\sqrt{\hat{\delta}_{k'}}}\big(\sum_{<\dot{v}(u_{(i,j)})=k'>} e^{j\phi_{u_{(i,j)}}}\big)\Big), \tag{C.13}$$

for $\ddot{v}(k') = k$. By defining $\hat{\delta'}_k = \hat{\delta}_{k'}$ and $\phi'_{k,l} = \phi_{u_{(i,j)}}$, for $\ddot{v}(\dot{v}(u_{(i,j)})) = k$, we can write,

$$b_i = \frac{1}{\sqrt{\delta_{\pi_1^{-1}(i)}}} \sum_{k=1}^{M} \frac{1}{\sqrt{\hat{\delta'}_k}} a_k \big(\sum_{l=1}^{L_k} e^{j\phi'_{k,l}}\big). \tag{C.14}$$

Since in $\sum_{l=1}^{L_k} e^{j\phi'_{k,l}}$ by adding even number of phase-shifters which can cancel each other ($e^{j2\pi}$ and $e^{j\pi}$) the sum remains the same, we can have $L_k = L$. Furthermore, we can write,

$$\begin{aligned} \delta_{\pi_1^{-1}(i)} &= LM, \\ \hat{\delta'}_k &= LN, \end{aligned} \tag{C.15}$$

hence,

$$b_i = \frac{1}{\sqrt{MN}} \sum_{k=1}^{M} a_k \big(\sum_{l=1}^{L} \frac{e^{j\phi'_{k,l}}}{L}\big). \tag{C.16}$$

From Lemma C.1 (presented after the proof) we can write,

$$b_i = \frac{1}{\sqrt{MN}} \sum_{k=1}^{M} z_{k,i} a_k. \tag{C.17}$$

Thus, we arrive at,

$$\mathbf{b} = \frac{1}{\sqrt{MN}}\mathbf{A}\mathbf{a}. \tag{C.18}$$

■

Lemma C.1. *For arbitrary $\theta_l \in [0, 2\pi]$ where $l = 1, 2, ..., L$, if we have $z = \sum_{l=1}^{L} \frac{e^{j\theta_l}}{L}$, then $0 \leq |z| \leq 1$.*

Proof. The proof follows from successive applications of the triangle inequality in complex plane as $0 \leq |\sum_{l=1}^{L} e^{j\theta_l}| \leq L$. □

References

[1] I.-R. M.2370-0, "IMT traffic estimates for the years 2020 to 2030," July 2015.

[2] T. S. Rappaport, R. W. Heath, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Prentice Hall, 2015.

[3] D. J. J. Lee, G. Liu, "5G: Vision and requirements for mobile communication system towards year 2020," *Chinese Journal of Eng.*, 2016.

[4] T. Chen, M. Matinmikko, X. Chen, X. Zhou, and P. Ahokangas, "Software defined mobile networks: Concept, survey, and research directions," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 126–133, Nov. 2015.

[5] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Trans. Veh. Technol.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.

[6] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, no. 6, pp. 585–595, 1999.

[7] 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 Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[8] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[9] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai, and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Commun. Surveys Tutorials*, vol. 20, no. 2, pp. 836–869, 2018.

[10] E. Torkildson, U. Madhow, and M. Rodwell, "Indoor millimeter wave MIMO: Feasibility and performance," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4150–4160, Dec. 2011.

[11] R. C. Daniels and R. W. Heath, "60 GHz wireless communications: Emerging requirements and design recommendations," *IEEE Trans. Veh. Technol.*, vol. 2, no. 3, pp. 41–50, Sept. 2007.

[12] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[13] I. F. Akyildiz, C. Han, and S. Nie, "Combating the distance problem in the millimeter wave and terahertz frequency bands," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 102–108, June 2018.

[14] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. Marzetta, "Massive MIMO is a reality—What is next?" *Digit. Signal Process.*, vol. 94, pp. 3–20, 2019.

[15] S. A. Busari, K. M. S. Huq, S. Mumtaz, J. Rodriguez, Y. Fang, D. C. Sicker, S. Al-Rubaye, and A. Tsourdos, "Generalized hybrid beamforming for vehicular connectivity using THz massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8372–8383, June 2019.

[16] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.

[17] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.

[18] A. Morsali, S. S. Hosseini, B. Champagne, and X. Chang, "Design criteria for omnidirectional STBC in massive-MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1435–1439, June 2019.

[19] F. Koroupi, A. Morsali, V. Niktab, M. Shahabinejad, and S. Talebi, "Quasi-orthogonal space-frequency and space-time-frequency block codes with modified performance and simplified decoder," *IET Commun.*, vol. 11, no. 11, pp. 1655–1661, Sept. 2017.

[20] M. Samavat, A. Morsali, and S. Talebi, "Delay interleaved cooperative relay networks," *IEEE Commun. Lett.*, vol. 18, no. 12, pp. 2137–2140, Dec. 2014.

[21] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[22] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[23] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[24] 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 Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[25] F. Khalid, "Hybrid beamforming for millimeter wave massive multiuser MIMO systems using regularized channel diagonalization," *IEEE Commun. Lett.*, vol. 8, no. 3, pp. 705–708, June 2019.

[26] M. M. Molu, P. Xiao, M. Khalily, K. Cumanan, L. Zhang, and R. Tafazolli, "Low-complexity and robust hybrid beamforming design for multi-antenna communication systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1445–1459, Mar. 2018.

[27] J. Li, L. Xiao, X. Xu, and S. Zhou, "Robust and low complexity hybrid beamforming for uplink multiuser mmwave MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1140–1143, June 2016.

[28] C. Ho, H. Cheng, and Y. Huang, "Hybrid precoding processor for millimeter wave MIMO communications," *IEEE Trans. Circuits Syst. II*, vol. 66, no. 12, pp. 1992–1996, 2019.

[29] R. Mai, D. H. N. Nguyen, and T. Le-Ngoc, "MMSE hybrid precoder design for millimeter-wave massive MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, April 2016, pp. 1–6.

[30] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice Hall PTR, 2009.

[31] 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 J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[32] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 70–77, Sep. 2014.

[33] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[34] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, 2014.

[35] "IEEE standard for information technology– local and metropolitan area networks– Specific requirements– Part 15.3: Amendment 2: Millimeter-wave-based alternative physical layer extension," *IEEE Std 802.15.3c-2009 (Amendment to IEEE Std 802.15.3-2003)*, pp. 1–200, Oct. 2009.

[36] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, 2013.

[37] G. P. M. Gerla, E. Lee and U. Lee, "Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds," in *IEEE forum on internet of things (WF-IoT)*, Mar. 2014, pp. 241–246.

[38] P. Kumari, N. Gonzalez-Prelcic, and R. W. Heath, "Investigating the IEEE 802.11ad standard for millimeter wave automotive radar," in *IEEE 82nd Veh. Technol. Conf.*, Sep. 2015, pp. 1–5.

[39] K. Venugopal, M. C. Valenti, and R. W. Heath, "Device-to-device millimeter wave communications: Interference, coverage, rate, and finite topologies," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6175–6188, 2016.

[40] D. L. B. H. S. Chen, H. Xu and H. Wang, "A vision of IoT: Applications, challenges, and opportunities with china perspective," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 349–359, 2014.

[41] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1924–1934, july 2007.

[42] ECMA-387, "High rate 60 GHz PHY, MAC and HDMI PAL, 1st ed." Dec. 2008.

[43] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 1, pp. 167–178, First Quarter 2013.

[44] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[45] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 106–112, Apr. 2014.

[46] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Select. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[47] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, June 2014.

[48] 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. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sept. 2015.

[49] A. Ghazal, Y. Yuan, C. X. Wang, Y. Zhang, Q. Yao, H. Zhou, and W. Duan, "A non-stationary IMT-Advanced MIMO channel model for high-mobility wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2057–2068, Apr. 2017.

[50] F. Giannetti, M. Luise, and R. Reggiannini, "Mobile and personal communications in the 60 GHz band: A survey," *Wireless Pers. Commun.*, vol. 10, no. 2, pp. 207–243, july 1999.

[51] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[52] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.

[53] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[54] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan. 2018.

[55] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, Third Quarter 2016.

[56] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[57] A. Garcia-Rodriguez, V. Venkateswaran, P. Rulikowski, and C. Masouros, "Hybrid analog-digital precoding revisited under realistic RF modeling," *IEEE Commun. Lett.*, vol. 5, no. 5, pp. 528–531, Oct. 2016.

[58] I.-R. M.1457-10, "Detailed specifications of the terrestrial radio interfaces of international mobile telecommunications-2000," June 2011.

[59] "IEEE standard for telecommunications and information exchange between systems - LAN/MAN Specific requirements - Part 11: Wireless medium access control (MAC) and physical layer (PHY) specifications: High speed physical layer in the 5 GHz band," *IEEE Std 802.11a-1999*, pp. 1–102, Dec. 1999.

[60] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[61] C. Y. Chi, C. H. Peng, K. C. Huang, T. H. Tsai, and W. K. Ma, "A block-by-block blind post-FFT multistage beamforming algorithm for multiuser OFDM systems based on subcarrier averaging," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3238–3251, Aug. 2008.

[62] Y. Zhu, Q. Zhang, and T. Yang, "Low complexity hybrid precoding with dynamic beam assignment in mmWave OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3685–3689, 2018.

[63] D. Zhu, B. Li, and P. Liang, "A novel hybrid beamforming algorithm with unified analog beamforming by subspace construction based on partial CSI for massive MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 594–607, 2017.

[64] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmwave MIMO systems," vol. 16, no. 5, pp. 2907–2920, 2017.

[65] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmwave OFDM large-scale antenna arrays," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, july 2017.

[66] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[67] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Select. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sept. 2017.

[68] X. Lin, S. Wu, L. Kuang, Z. Ni, X. Meng, and C. Jiang, "Estimation of sparse massive MIMO-OFDM channels with approximately common support," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1179–1182, May 2017.

[69] Y. Zhang, D. Wang, J. Wang, and X. You, "Channel estimation for massive MIMO-OFDM systems by tracking the joint angle-delay subspace," *IEEE Access*, vol. 4, pp. 10 166–10 179, 2016.

[70] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, June 2016.

[71] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, "Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, july 2017.

[72] A. Morsali, B. J. Champagne, and A. Haghighat, "Efficient implementation of hybrid beamforming," *U.S. Patent*, September 29, 2020.

[73] A. Morsali, S. Norouzi, and B. Champagne, "Single RF chain hybrid analog/digital beamforming for mmWave massive-MIMO," in *IEEE Global Confer. Signal Inf. Process. (GlobalSIP)*, 2019.

[74] A. Morsali, A. Haghighat, and B. Champagne, "Realizing fully digital precoders in hybrid A/D architecture with minimum number of RF chains," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2310–2313, Oct. 2017.

[75] T. Lin, J. Cong, Y. Zhu, J. Zhang, and K. B. Letaief, "Hybrid beamforming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3693–3708, 2019.

[76] Y. S. Yeh, B. Walker, E. Balboni, and B. Floyd, "A 28-GHz phased-array receiver front end with dual-vector distributed beamforming," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1230–1244, 2017.

[77] K. Greene, A. Sarkar, and B. Floyd, "A 60-GHz dual-vector doherty beamformer," *IEEE J. Solid-State Circuits*, vol. 52, no. 5, pp. 1373–1387, 2017.

[78] A. Morsali and B. Champagne, "Achieving fully-digital performance by hybrid analog/digital beamforming in wide-band massive-MIMO systems," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, 2020, pp. 5125–5129.

[79] J. Chuang, "The effects of time delay spread on portable radio communications channels with digital modulation," *IEEE J. Select. Areas Commun.*, vol. 5, no. 5, pp. 879–889, June 1987.

[80] L. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, no. 7, pp. 665–675, july 1985.

[81] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Trans. Commun. Technol.*, vol. 19, no. 5, pp. 628–634, Oct. 1971.

[82] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[83] M. L. Doelz, E. T. Heald, and D. L. Martin, "Binary data transmission techniques for linear systems," *Proc. of the IRE*, vol. 45, no. 5, pp. 656–661, May 1957.

[84] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J.*, vol. 45, no. 10, pp. 1775–1796, Dec. 1966.

[85] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 773–775, 1995.

[86] A. Morsali and S. Talebi, "On permutation of space-time-frequency block codings," *IET Commun.*, vol. 8, no. 3, pp. 315–323, 2014.

[87] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 92–112, May 2011.

[88] N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, "Generalized frequency division multiplexing for 5th generation cellular networks," *IEEE Trans. Wireless Commun.*, vol. 62, no. 9, pp. 3045–3061, Sept. 2014.

[89] V. Vakilian, T. Wild, F. Schaich, S. ten Brink, and J. F. Frigon, "Universal-filtered multi-carrier technique for wireless systems beyond LTE," in *Proc. IEEE Global Telecommun. Conf (GLOBECOM)*, Dec. 2013, pp. 223–228.

[90] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[91] X. Chen, Z. Zhang, C. Zhong, R. Jia, and D. W. K. Ng, "Fully non-orthogonal communication for massive access," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2017.

[92] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[93] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[94] Y. Lu and W. Zhang, "Hybrid precoding design achieving fully digital performance for millimeter wave communications," *J. of Commun. Inf. Netw.*, vol. 3, pp. 74–84, 2018.

[95] C. Tsai and A. Wu, "Structured random compressed channel sensing for millimeter-wave large-scale antenna systems," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5096–5110, Oct 2018.

[96] Y. Zhang, D. Wang, J. Wang, and X. You, "Channel estimation for massive MIMO-OFDM systems by tracking the joint angle-delay subspace," *IEEE Access*, vol. 4, pp. 10 166–10 179, 2016.

[97] L. Pan, L. Liang, W. Xu, and X. Dong, "Framework of channel estimation for hybrid analog-and-digital processing enabled massive MIMO communications," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 3902–3915, 2018.

[98] G. Strang, *Introduction to Linear Algebra.* Wellesley - Cambridge Press, 2016.

[99] 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. 19 167–19 181, 2017.

[100] A. Morsali and B. Champagne, "Robust hybrid analog/digital beamforming for uplink Massive-MIMO with imperfect CSI," in *IEEE wireless Commun. Netw. Conf. (WCNC)*, 2019, pp. 1–6.

[101] S. Tofigh, H. M. Kermani, and A. Morsali, "A new design criterion for linear receiver stbcs based on full-rank spaces," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 207–210, Feb. 2015.

[102] A. Morsali and S. Talebi, "On permutation of space-time-frequency block codings," *IET Commun.*, vol. 8, no. 3, pp. 315–323, Feb. 2014.

[103] J. Wang and M. Bengtsson, "Joint optimization of the worst-case robust MMSE MIMO transceiver," *IEEE Signal Process. Lett.*, vol. 18, no. 5, pp. 295–298, May 2011.

[104] J. Wang and D. P. Palomar, "Robust MMSE precoding in mimo channels with prefixed receivers," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5802–5818, Nov. 2010.

[105] A. Morsali, A. Haghighat, and B. Champagne, "Generalized framework for hybrid analog/digital signal processing in massive and ultra-massive-mimo systems," *IEEE Access*, vol. 8, pp. 100 262–100 279, 2020.

[106] J. Du, W. Xu, H. Shen, X. Dong, and C. Zhao, "Hybrid precoding architecture for massive multiuser MIMO with dissipation: Sub-connected or fully connected structures?" *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5465–5479, June 2018.

[107] R. Magueta, D. Castanheira, A. Silva, R. Dinis, and A. Gameiro, "Hybrid multi-user equalizer for massive MIMO millimeter-wave dynamic subconnected architecture," *IEEE Access*, vol. 7, pp. 79 017–79 029, June 2019.

[108] C. E. Shannon, *The Theory and Design of Linear Differential Equation Machines*. Bell Telephone Laboratories, 1942.

[109] F. Bonchi, P. Sobociński, and F. Zanasi, "The calculus of signal flow diagrams I: Linear relations on streams," *Information and Computation*, vol. 252, pp. 2–29, Feb. 2017.

[110] D. M. Pozar, *Microwave Engineering*. John Wiley and Sons, 2005.