Efficient Pilot-Data Transmission and Channel Estimation in Next Generation Wireless Communication Systems

by

Leyuan Pan
B.Eng., Southeast University, China, 2010
M.Sc., Southeast University, China, 2013

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Electrical and Computer Engineering

Efficient Pilot-Data Transmission and Channel Estimation in Next Generation Wireless Communication Systems

by

Leyuan Pan
B.Eng., Southeast University, China, 2010
M.Sc., Southeast University, China, 2013

Supervisory Committee

Dr. Xiaodai Dong, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Hong-Chuan Yang, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Jianping Pan, Outside Member
(Department of Computer Science)

Supervisory Committee

Dr. Xiaodai Dong, Supervisor
(Department of Electrical and Computer Engineering)

Dr. Hong-Chuan Yang, Departmental Member
(Department of Electrical and Computer Engineering)

Dr. Jianping Pan, Outside Member
(Department of Computer Science)

ABSTRACT

To meet the urgent demand of high-speed data rate and to support large number of users, the massive multiple-input multiple-output (MIMO) technology is becoming one of the most promising candidates for the next generation wireless communications, namely the 5G. To realize the full potential of massive MIMO, it is necessary to have the channel state information (CSI) (partially) available at the transmitter. Hence, an efficient channel estimation is one of the key enablers and also critical challenges for 5G communications. Dealing with such problems, this dissertation investigates the design of efficient pilot-data transmission pattern and channel estimation in massive MIMO for both multipair relaying and peer-to-peer systems.

Firstly, this dissertation proposes a pilot-data transmission overlay scheme for multipair MIMO relaying systems employing either half- or full-duplex (HD or FD) communications at the relay station (RS). In the proposed scheme, pilots are transmitted in partial overlap with data to decrease the channel estimation overhead. The RS can detect the source data by exploiting the asymptotic orthogonality of massive MIMO channels. Due to the transmission overlay, the effective data period is extended, hence improving system throughput. Both theoretical and simulation results verify that the proposed pilot-data overlay scheme outperforms the conventional separate pilot-data design in the limited coherence interval scenario. Moreover, a power allocation problem is formulated to properly adjust the transmission power of source data transmission and relay data forwarding which further improves the system performance.

Additionally, this dissertation proposes and analyzes an efficient HD decode-and-

forward (DF) scheme, named sum decode-and-forward (SDF), with the physical layer network coding (PNC) in the multipair massive MIMO two-way relaying system. As comparison, a joint decode-and-forward (JDF) scheme applied to the multipair massive MIMO relaying is also proposed and investigated. In the SDF scheme, a half number of pilots are saved compared to the JDF scheme which in turn increases the spectral efficiency of the system. Both the theoretical analyses and numerical results verifies such superiority of the SDF scheme. Further, the power efficiency of the proposed schemes is also investigated. Simulation results show that the signal transmission power can be rapidly reduced if the massive antenna arrays are equipped on the RS and the required data transmission power can further decrease if the training power is fixed.

Finally, this dissertation investigates the general channel estimation problem in the massive MIMO system which employs the hybrid analog/digital precoding structure with limited radio-frequency (RF) chains. By properly designing RF combiners and performing multiple trainings, the performance of the proposed channel estimation can approach that of full-chain estimations depending on the degree of channel spatial correlation and the number of RF chains which is verified by simulation results in terms of both mean square error (MSE) and spectral efficiency. Moreover, a covariance matching method is proposed to obtain channel correlation in practice and the simulation verifies its effectiveness by evaluating the spectral efficiency performance in parametric channel models.

Contents

List of Figures

Abbreviations

3G	The Third Generation
3GPP	The 3rd Generation Partnership Project
4G	The Fourth Generation
5G	The Fifth Generation
AF	Amplify-and-Forward
AoA	Angle of Arrival
AoD	Angle of Departure
AWGN	Additive White Gaussian Noise
BC	Broadcasting
BS	Base Station
CDMA2000	Code Division Multiple Access 2000
CSI	Channel State Information
DAC	Digital/Analog Converter
DF	Decode-and-Forward
DPC	Dirty Paper Coding
FD	Full-Duplex
FDD	Frequency-Division Duplex
GSM	Global System for Mobile Communications
HD	Half-Duplex
IoT	Internet of Thing
IMT-Advanced	International Mobile Telecommunications-Advanced
IMT-MC	International Mobile Telecommunications Multi-Carrier
IPI	Inner-Pair Interference
ITU-R	International Telecommunication Union Radio Communication Sector
JDF	Joint Decode-and-Forward
LI	Loop Interference
LMMSE	Linear Minimum Mean Square Error
LTE	Long-Term Evolution
MA	Multiple Access

MI	Multipair Interference
MIMO	Multiple Input Multiple Output
MMSE	Minimum Mean Square Error
MRC	Maximum-Ratio Combining
MRT	Maximum-Ration Transmission
MS	Mobile Station
MSE	Mean Square Error
NC	Network Coding
P2P	Peer-to-Peer
QoS	Quality of Service
PNC	Physical Layer Network Coding
RF	Radio Frequency
RS	Relay Station
RSI	Residual Self-Interference
RV	Random Variable
SCA	Successive Convex Approximation
SDF	Sum Decode-and-Forward
SIC	Successive Interference Cancellation
SINR	Signal to Interference and Noise Ratio
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
TDD	Time-Division Duplex
TV	Television
TWR	Two-Way Relaying
TWRC	Two-Way Relaying Channel
UMTS	Universal Mobile Telecommunications System
ZF	Zero-Forcing
e2e	End-to-End
i.i.d.	Identically Independent Distribution
mmWave	Millimeter Wave

Notations

Unless stated otherwise, boldface upper-case and lower-case letters denote matrices and vectors respectively.

x^*	the conjugate of a complex scalar x
$\mathbf{X}^{\mathrm{T}}$	the transpose of matrix $\mathbf{X}$
$\mathbf{X}^{\mathrm{H}}$	the conjugate transpose (Hermitian) of matrix $\mathbf{X}$
$\mathbf{X}_{i,j}$	the (i,j)th element of $\mathbf{X}$
$\mathbf{X}_{[i:j]}$	the inclusive sub-matrix formed by the ith to jth columns of $\mathbf{X}$
$\mathrm{tr}\,(\mathbf{X})$	the trace of $\mathbf{X}$
$\mathbf{I}_M$	the identity matrix of dimension $M \times M$
$\mathbf{0}$	a zero vector or matrix
$\lvert x\rvert$	the absolute value of a real scalar x or magnitude of a complex scalar x
$\lVert\mathbf{x}\rVert$	the Euclidean norm of a vector $\mathbf{x}$
$\lVert\mathbf{X}\rVert_{\mathrm{F}}$	the Frobenius norm of a matrix $\mathbf{X}$
$\mathbb{E}\{\mathbf{X}\}$	the expectation of a random matrix $\mathbf{X}$
$\mathbb{V}\mathrm{ar}\{\mathbf{X}\}$	the variance of a random matrix $\mathbf{X}$
$\mathbb{R}$	the real number set
$\mathbb{C}$	the complex number set
$\mathbb{C}^m$	a set of complex column vectors with size m
$\mathbb{C}^{m\times n}$	a set of complex matrices with size $m \times n$
$\sim$	distributed according to
$\mathcal{CN}(\mathbf{m}, \mathbf{\Sigma})$	complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{\Sigma}$
$\overset{\text{a.s.}}{=}$	almost sure convergence
$\triangleq$	defined as
$\max$	maximize
$\min$	minimize
$\max\{x, y\}$	the maximum one of x and y
$\min\{x, y\}$	the minimum one of x and y
$\mathrm{diag}\{\mathbf{x}\}$	matrix with diagonal entries as the elements of $\mathbf{x}$
$\mathrm{blkdiag}\{\mathbf{A}_1, \cdots, \mathbf{A}_L\}$	matrix with diagonal entries as $\mathbf{A}_1, \cdots, \mathbf{A}_L$ in sequence

$\mathrm{colsp}(\mathbf{X})$	the space spanned by the column vectors of $\mathbf{X}$
$\mathbf{a} \succ \mathbf{b}$	$\mathbf{a}$ majorizes (or dominates) $\mathbf{b}$
$\mathbf{a} \prec \mathbf{b}$	$\mathbf{b}$ majorizes (or dominates) $\mathbf{a}$

ACKNOWLEDGEMENTS

First and foremost, I would like to attribute my greatest gratitude to my supervisor, Prof. *Xiaodai Dong*, and co-supervisor, Prof. *Tao Lu*, for their continuous support and supervision to my Ph.D study and research. I am grateful to them for their patient guidance, insightful comments, inspired instructions and dissertation revision every time, and providing me with excellent research atmospheres. I really appreciate all the efforts they have taken to help me complete my Ph.D study, efficiently and happily.

I am also indebted to my departmental committee member, Prof. *Hong-Chuan Yang*, for his meticulous guidance and insightful suggestions and advice. I also thank my outside committee member, Prof. *Jianping Pan*, for offering me valuable comments and suggestions. My sincere thanks also go to Prof. *Vijay Bhargava* for the insightful comments on my dissertation revision.

I thank my fellow team members and colleagues in the laboratories where I worked: *Zheng Xu*, *Ming Lei*, *Guang Zeng*, *Binyan Zhao*, *Lan Xu*, *Le Liang*, *Weiheng Ni*, *Ping Cheng*, *Yiming Huo*, *Farnoosh Talaei*, *Jun Zhou*, *Tianyang Li*, *Yuejiao Hui*, *Wanbo Li*, *Guowei Zhang*, *Tong Xue*, *Wenyan Yu*, *Xuan Du*, *Amin Cheraghi*, *Serge Vincent*, *Niloofar Sadeghi*, *Liao Zhang*, *Wen Zhou*, *Xiaoxuan Wang*, and *Fan Zhou*, for the research discussions and for all the fun we had in the last four years. I would like to thank my best friends: *Yunlong Shao*, *Fang Chen*, *Xiao Ma*, *Mengyue Cai*, *Xiao Feng*, *Po Zhang*, *Zhu Ye*, *Xiao Xie*, and *Feng Hu* for their supports and happy time they gave me in my life.

Last but not least, I express my endless gratitude to my wife, *Yongyu Dai*, for her unconditional supports, help, patience, sacrifices and love. She stands by me through the good times and the bad.

Finally, I would like to show my sincere appreciation to my parents and parents-in-law, who are always supporting me and encouraging me through my whole life.

Leyuan Pan

Burnaby, BC

March, 2017

DEDICATION

To
My parents
And
My wife
For everything

Chapter 1

Introduction

1.1 Overview

1.1.1 Evolution of Mobile Communications

In the past a few years, we are witnessing the explosive growing amount of data requirements in the mobile networks as the emerging of mobile intelligent devices, such as smart phones. The urgent demands of high data rate exchanging are spreading from conventional phone calls and text messages to all over the whole Internet, where the mobile services including the high-definition mobile video streaming and television (TV), online gaming, and real-time mobile conferences have been becoming the non-negligible life and business style all around our world. To support such high data rate demands, the *International Telecommunication Union Radio Communication Sector* (ITU-R) issued the International Mobile Telecommunications-Advanced (IMT-Advanced) standard in 2008, which is known as the fourth generation (4G) telecommunication technologies [1, 2]. The *3rd Generation Partnership Project* (3GPP) developed the Long-Term Evolution (LTE) standards which was first released in December 2008 by upgrading both Global System for Mobile Communications/Universal Mobile Telecommunications System (GSM/UMTS) and IMT Multi-Carrier (IMT-MC, also known as CDMA2000) networks to fulfill the requirements of 4G [3–7]. However, the technical specifications of LTE networks do not satisfy the IMT-Advanced requirements. Later, the 3GPP launched the research of LTE-Advanced as its successor and completed the standardization in March 2011 which can be treated as the "True 4G" [8]. Nevertheless, the development of mobile communications would not stop there.

As the rapid exploring of consumer requirements, we need the fifth generation (5G). The aspiration of 5G includes the huge increments of data rates, much less latencies and higher energy efficiency [9]. For the data rate, the 5G would be designed with roughly

1000x aggregate data rate increments compared to 4G, meeting 100 Mbps for 95% users while 1 Mbps for the edge rate to fulfill the seamless coverage, and achieving tens of Gbps for the peak rate. In addition, the energy efficiency of 5G would be increasing by about 100x than 4G. To support the requirements of 5G, more advanced technologies were proposed in recent years.

Among all the air interface technologies, multiple input multiple output (MIMO) is one of the most essential elements in the advanced wireless networks, which has been deployed in the current commercial LTE networks and also been employed in the LTE-Advanced standards. By exploiting the space-division multiplexing, MIMO can multiply the network capacity linearly as the number increment of antenna pairs. MIMO is often tracked back to the pioneering research done by Telatar [10], Foschini and Gans [11]. After their works, a tremendous amount of literature comes out involving this technology. Until 2010, almost all the works involving MIMO have focused on the limited number of antennas deployed on both transmitter and receiver (e.g., 2, 4 or at most 8 antennas), which is named as the small-scale MIMO system.

1.1.2 Massive MIMO and Millimeter Wave

In 2010, Marzetta published his seminal investigation of the system performance with the assumption that the number of antennas is tending to infinity, which is named as the large-scale MIMO system and is also known as the massive MIMO system [12]. The massive MIMO technology is one of the most promising solutions for the next-generation wireless communications to meet the urgent demands of both high-speed data transmissions and explosive growing numbers of user terminals[13–16]. Compared with conventional MIMO mechanisms, massive MIMO is capable to serve a large amount of mobile stations (MSs) simultaneously and achieve higher reliabilities, increased throughputs and improved energy efficiency by employing less complicated signal processing techniques, e.g., maximum-ratio combining/maximum-ratio transmission (MRC/MRT), with inexpensive and low-power components [9, 15]. Hence, a massive MIMO system can substantially reduce power consumption while improve the achievable rate performance.

However, the crowd frequency spectrum adopted by current mobile communications has been becoming one of the limitations to further improve the high-speed user experience with a yet unsustainable system cost. Therefore, it has captured the attention and imagination of researchers and engineers all around the world to seek wider bandwidth and more flexible spectrum for low cost but high-speed communications in recent years. Naturally, millimeter wave (mmWave) inevitably becomes one of the best candidates [17–19]. Different from the conventional radio frequency (RF) spectrum, i.e., usually below 3 GHz, the mmWave can provide a huge amount of available spectrum which sig-

nificantly benefits the massive MIMO technologies due to short wavelength leading to relatively small-size antenna arrays. With the increasing of antenna numbers, the huge amounts of RF chains are becoming more and more costly and the baseband signal processing is also becoming more and more complicated. To reduce such cost hardware and complexity of digital signal processing, an analog/digital hybrid precoding structure is proposed in literature [20].

Generally, one of the most critical and fundamental challenges in designing wireless communication systems is how to obtain the precise channel state information (CSI) by consuming limited resources, especially in the massive MIMO systems. In existing massive MIMO studies, time-division duplex (TDD) is the most widely considered implementation mode due to its more effective approaches of obtaining CSI than the frequency-division duplex (FDD) [21]. Thanks to the channel reciprocity, a TDD system exploits uplink pilot training to estimate channels which can be used in both uplink and downlink data transmissions within a coherence interval where the interval length is determined by the mobility of user equipment. Compared with the downlink one, the uplink pilot training saves a large amount of resources to estimate the channels of a large-scale antenna array because each pilot sequence can be used to estimate the channels between all base station antennas and a single-antenna user equipment. On the other hand, a pilot sequence can only be utilized to estimate the channels between one base station antenna and user equipments. It consumes a large amount of resource to estimate massive MIMO channels. Nevertheless, the uplink channel estimation also involves pilot contamination issues when the number of users is large. It has been reported in [13, 22] that pilot contamination reduces the system performance but cannot be suppressed by increasing the number of antennas. In general, the length of the pilot sequence in uplink channel estimations should be equal to or greater than the number of users to guarantee the orthogonality of pilot patterns among different users, which is to avoid the pilot contamination. In a massive MIMO system, due to the growing user number, it still requires a large amount of resources to transmit orthogonal pilot sequences. Hence, the overhead of channel estimation increases correspondingly, which degrades the effective system throughput.

Regarding the hybrid precoding structure of the massive MIMO and mmWave systems, it is very difficult to obtain the complete the CSI with the limited RF chains. Many compressed-sensing based channel estimation schemes are proposed in literature [18, 23]. However, such kind of methods are all with high complexities depending on the sparsity of channels. Hence, it is an urgent demand to develop an efficient channel estimation scheme in the limited-chain systems while the complexity is independent of the channel sparsities which is investigated in this dissertation.

1.1.3 Cooperative Wireless Communications by Relaying

To satisfy the seamless coverage requirements of the fifth-generation (5G) communications [9], the cooperative relaying system is a promising technique in the future wireless communication systems. The relaying technique is an emerging cooperative technology capable of scaling up the system performance by orders of magnitude to extend the coverage and reduce power consumption[24, 25]. Combining with massive MIMO technologies whereby the relay station (RS) is equipped with large scale antenna arrays, the performance of a relaying system can be dramatically improved [26–31]. Moreover, in spite of the conventional half-duplex (HD) system, the full-duplex (FD) relaying technique has attracted more interests recently due to the overlap of uplink and downlink data transmissions, whereby the overall system performance is further improved [30–32]. However, the full-duplex one-way relaying scheme suffers from the loop interference (LI) due to the signal leakage from the output to input antennas on the relay station (RS) while the HD two-way relaying (TWR) can increase the throughputs of multipair systems without importing LI [33–37]. In the HD TWR system, the inner-pair interference (IPI) is a considerable effect on improving the system performance. Nevertheless, the IPI of the HD TWR can be eliminated by applying the physical layer network coding (PNC) at the RS side [35]. The PNC technique was first proposed in [38], which is an apparatus similar to the conventional network coding (NC) while it adopts the proper modulation-and-demodulation technique at the RS to avoid the over-demodulation operations, and hence to reduce IPI. To further improve the performance of TWR, the adaptive channel-quantization PNC scheme with multiple antennas was proposed and analyzed by [39]. However, the scenario of only one pair of users are considered in the reported analyses.

Similar to the peer-to-peer (P2P) system, the multiuser relaying system also suffers from the critical channel estimation overhead within limited coherence time intervals, no matter of a HD or FD system. For a relaying system, it may be even worse as both source and destination users need to transmit pilots within the coherence interval, where the coherence interval determined by user pairs may be shorter than or at best equal to that by each user. Further, different from the P2P cellular system, the throughput of the whole relaying system is determined by the weaker one between the uplink (multi-access of the sources to RS) and downlink (broadcasting of RS to the destinations) connections. Thus, it is critical to co-consider both uplink and downlink transmissions to design pilot scheme for the relaying system, while previous work in the literature did not take it into account.

1.2 Summary of Contributions

In this dissertation, the main contributions are presented in Chapter 2, 3 and 4 which are summarized as follows.

Chapter 2 investigates and proposes an efficient pilot and data transmission scheme in multipair massive MIMO one-way relaying systems for both HD and FD communications. Due to the massive antennas equipped on RS, it is found that the source-relay and relay-destination channels are asymptotically orthogonal to each other, and thereby the transmission phase of pilots and data can be shifted to overlap each other to reduce the overhead of pilot transmission and accordingly to improve the system performance. Based on this consideration, the transmission schemes with pilot-data overlay in both HD and FD communications are proposed in Chapter 2. With the proposed schemes, the effective data transmission duration increases within a coherence interval and hence improves the system achievable rate performance. However, due to the overlapped pilot data transmission, pilot contamination and data interference emerges at the RS side. Nevertheless, by exploiting the asymptotic orthogonality of massive MIMO channels, Chapter 2 demonstrates that the received data and pilots can be well separated from each other with only residues of additive thermal noise by applying the MRC processing. For theoretical verifications, Chapter 2 derives *closed-form expressions* of the ergodic achievable rates of the considered relaying systems with the proposed scheme. Numerical and simulation results both support the superiority of the proposed scheme to the conventional ones. Further, Chapter 2 designs an optimal power allocation for the FD overlay scheme to minimize the interference between pilot and data transmissions by properly regulating their transmit power and proposes a successive convex approximation (SCA) approach to solve the non-convex optimization problem. Finally, the numerical result verifies the proposed power allocation algorithm and confirms the convergence of the SCA.

In Chapter 3, the HD multipair massive MIMO TWR system is considered. Due to the asymptotic orthogonal of massive MIMO channels in the TWR system, Chapter 3 reduces the pilot transmission overhead by estimating the sum of source and destination channels instead of estimating them respectively. In a multipair TWR system, the effective data transmission period increases as the decreasing of pilot overhead and hence improves the system achievable rate performance. The theoretical analyses show that the estimated sum channels can both be used to combine the received uplink signal and to precode the downlink broadcasting signal. In Chapter 3, we present the conventional joint decode-and-forward (JDF) scheme [40] for the multipair TWR and propose a novel sum decode-and-forward (SDF) scheme to reduce the number of pilots used for channel estimations. Other than the JDF scheme, SDF employs the PNC technique at the RS to improve the performance of the decode-and-forward scheme. By deploying the

massive antenna arrays on the RS, the the physical layer network coding residual self-interference (RSI) is rapidly reduced and the end-to-end (e2e) communication rates of the SDF scheme can achieve that of the JDF schemes. Furthermore, because of the decreased number of pilot symbols, the proposed SDF scheme outperforms the JDF scheme at high SNR regions. Additionally, to evaluate the spectral efficiency, we employ the statistical CSI to analyze the system performance as presented in [41], and compare it with the performance obtained by the instantaneous CSI. The numerical results reveal that the gap between these two methods is negligible. Finally, Chapter 3 investigates the power efficiency of relaying system by adopting massive antenna arrays on the RS. From the theoretical analyses and numerical evaluations, it is shown that the required transmitting power can be rapidly decreased while the quality of service (QoS) is guaranteed, e.g., the spectral efficiency of each communication pair achieves 1 bit/s/Hz.

Chapter 4 considers the uplink channel estimations of a massive MIMO system in a single cell where the hybrid RF-baseband processing structure is employed and propose an efficient channel estimation scheme. Note that the considered system structure can also be extended to the mmWave communications. The main task of the channel estimation in a hybrid precoding system is to recover the channel vector from the limited observations from the limited RF chains. To improve the estimation performance in the limited RF-chain scenario, multiple training phases are employed. Optimal design of the RF combiners for different training phases needs to be considered and properly designed to capture the channel energy and then recover the channel as accurately as possible with a small number of observations. In this dissertation, the RF combiners for the single training scenario is designed following the minimum mean square error (MMSE) criteria. The theoretical optimizer and the closed-form expression of the MSE is derived by relaxing the constant-magnitude constraint. Due to limited RF chains, the single training cannot achieve the full-chain performance in channel estimations. To compensate such performance loss, this dissertation proposes to estimate channels with multiple pilot trainings and the combiners for single-training scenario are extended to multiple trainings with the proposed Block Selection, Sequential Selection and Semi-joint Selection. To implement the proposed channel estimation scheme in the practical system, a covariance matching method is proposed to generate channel correlations. Final, the numerical and simulation results are presented to verify the performance of the proposed scheme by examine both the mean square error (MSE) of channel estimations and the spectral efficiency of the hybrid precoding scheme.

1.3 Organizations

The rest parts of this dissertation are organized as follows. In Chapter 2, this dissertation considers the pilot-data transmission scheme design in the half- and full-duplex massive MIMO multipair one-way relaying systems and proposes a pilot-data overlay transmission scheme where the performance analyses show the superiority of the proposed scheme over the conventional ones. To further improve the performance of the massive MIMO relaying system, Chapter 3 proposes two pilot transmission schemes in the two-way relaying system and analyzes the performance of the proposed schemes where only half-duplex mode is considered. The following Chapter 4 considers the channel estimations in the hybrid precoding massive MIMO systems with limited RF chains. Finally, Chapter 5 draws out the conclusions of this dissertation and the expectation of future works.

Chapter 2

Multipair Massive MIMO Relaying with Pilot-data Transmission Overlay

To satisfy the requirements of seamless coverage in 5G, the cooperative communications, such as multipair massive multiple-input multiple-output (MIMO) relaying, have attracted considerable research interests from both academia and industry. Further, the availability of channel state information (CSI) at the transmitter is one of the key enablers to realize the full potential of massive MIMO. However, the large amount of training symbols in multipair relaying systems increases the overhead of signal transmissions. Dealing with such problem, this chapter focuses on investigating and developing an efficient pilot-data transmission scheme in the multipair massive MIMO relaying system, and analyzing the performance of the proposed scheme in terms of both theoretical and numerical results.

2.1 Introduction

MASSIVE multiple-input multiple-output (MIMO) technology is becoming one of the most promising solutions for the next-generation wireless communication to meet the urgent demands of high-speed data transmissions and explosive growing numbers of user terminals, such as the traditional mobile equipments and the new Internet of Things (IoT) devices [12–16]. Compared with conventional MIMO mechanisms, massive MIMO is capable of achieving higher reliabilities, increased throughputs and improved energy efficiency by employing less complicated signal processing techniques, e.g., maximum-ratio combining/maximum-ratio transmission (MRC/MRT), with inexpensive and low-power components [15]. Despite of the advantages, massive MIMO is also facing significant challenges on the way towards practical applications. How to obtain precise channel state information (CSI) while consuming limited resources is most

critical and fundamental.

In existing massive MIMO studies, time-division duplex (TDD) is more widely considered than frequency-division duplex (FDD) because it is potentially easier and more feasible to obtain CSI [21]. Pioneering authors in the field of massive MIMO have recently raised this issue as a critical open question: can the massive MIMO work in FDD operation? [42] The answer to this question is still unclear now, which is mostly because of the prohibitively high complexity on CSI acquisition in FDD massive MIMO systems. In contrast, a TDD operation consumes much less resource on CSI acquisition. Thanks to the channel reciprocity, a TDD system exploits uplink pilot training to estimate channels which can be used in both uplink and downlink data transmissions within a coherence interval. Compared with the downlink one, the uplink pilot training saves a large amount of resources to estimate the channels of a large-scale antenna array because each pilot sequence can be used to estimate the channels between all base station antennas and a single-antenna user equipment. However, the uplink channel estimation has to deal with pilot contamination issues as the user number grows. It is reported in [13, 22] that pilot contamination reduces the system performance but cannot be suppressed by increasing the number of antennas. Generally, the length of pilot sequence should be equal to or greater than the number of users to guarantee the orthogonality of pilot patterns among different users, in order to avoid pilot contamination. In a massive MIMO system, due to the large user number, orthogonal pilot sequences become very long, causing significant overhead for channel estimation and thus degrading the effective system throughput. When the channel varies with time due to medium to high mobility, i.e., relatively short coherence interval, the pilot overhead issue gets more severe as channel estimation needs to be done frequently. There are some efforts in the literature to reduce pilot overhead in the massive MIMO cellular system serving a large number of users within a finitely long coherence interval. Zhang *et al.* proposed a semi-orthogonal pilot design in [43] and [44] to transmit both data and pilots simultaneously where a successive interference cancellation (SIC) method was employed to reduce the contaminations of interfering pilots. You *et al.* investigated the performance of a pilot reuse scheme in the single-cell scenario which distinguishes users by the angle of arrival and thereby reuses pilot patterns [45]. A time-shifted pilot based scheme was proposed in [46] and it was then extended to the finite antenna regime in [47] and [48] to cope with the multi-cell scenario. Nevertheless, all the research introduced above focused on point-to-point (P2P) communications. Few work has studied the pilot scheme design and optimization in massive MIMO relaying systems, especially for one-way multipair communications.

The relaying technique is an emerging cooperative technology capable of scaling up the system performance by orders of magnitude, extending the coverage and reducing power consumption[25]. Combining with massive MIMO technologies whereby the re-

lay station (RS) is equipped with large-scale antenna arrays, the performance of a relaying system can be dramatically improved [26–30, 49]. Moreover, in spite of the conventional half-duplex (HD) system, the full-duplex (FD) relaying technique has attracted more interests recently due to the simultaneous uplink (the sources transmit signals to the RS) and downlink (the RS broadcasting signals to the destinations) data transmissions, whereby the overall system performance is further improved [30, 32, 50, 51]. However, similar to the P2P system, the multiuser relaying system also suffers from the critical channel estimation overhead within limited coherence intervals. For a relaying system, it may be even worse as both source and destination users need to transmit pilots within the coherence interval, where the coherence interval determined by user pairs may be shorter than or at best equal to that by each user. Further, different from the P2P cellular system, the throughput of the whole relaying system is determined by the weaker one between the uplink and downlink connections. Thus, it is critical to co-consider both uplink and downlink transmissions when designing the pilot scheme for the relaying system, while previous work in the literature did not take this into consideration.

This chapter investigates the pilot and data transmission scheme in multipair massive MIMO one-way[1] decode-and-forward (DF) relaying systems for both HD and FD communications. Due to the massive antennas equipped on the RS, the source-relay and relay-destination channels are asymptotically orthogonal to each other, and thereby the transmission phase of pilots and data can be shifted to overlap each other to reduce the overhead of pilot transmission and accordingly to improve the system performance. Based on this consideration, a transmission scheme with pilot-data overlay in both HD and FD communications is proposed in this chapter. In practical pilot assisted transmissions, there is always a tradeoff between pilots and data transmissions [52]. With the proposed scheme, the system achievable rate increases due to the extension of the data transmission duration even though the pilot training is interfered. Apparently, the system rate rises linearly as the extension of the data transmission duration while only degrades in a logarithm tendency due to the enhancement of effective signal to interference and noise ratio (SINR) of the received signal. In fact, both theoretical analyses and simulation results in this chapter confirm that a tradeoff between the transmission of pilots and data is achieved with the proposed scheme which improves system throughputs. In details, the main contributions of this chapter are summarized as follows:

- **Pilot-data overlay transmission scheme design:** A transmission scheme with pilot-data overlay in both HD and FD multipair massive MIMO relaying systems is proposed and designed. In the HD overlay scheme, destination pilots are transmitted simultaneously with source data transmission, such that the effective data transmission

[1] It is notable that the proposed scheme in this chapter can be easily extended to the two-way relaying system with minor adjustments.

duration is increased. Moreover, both source and destination pilots are transmitted along with data transmission in the FD system and thus the effective data transmission duration can be further increased. However, pilot and data contaminate each other at the RS due to the simultaneous transmission. Nevertheless, by exploiting the asymptotic orthogonality of massive MIMO channels, this chapter demonstrates that the received data and pilots can be well separated from each other with only residues of additive thermal noise by applying the MRC processing. After all, the effective data transmission duration is extended within the limited coherence interval and therefore the overall system performance is improved.

- **Closed-form achievable rates and comparison with conventional schemes[2]:** This chapter derives *closed-form expressions* of the ergodic achievable rates of the considered relaying systems with the proposed scheme. The derived expression reveals that the loop interference (LI) in the FD overlay scheme can be effectively suppressed by the growing number of RS antennas and no error propagation exists with the proposed scheme, which is a critical issue in [43] where a semi-orthogonal pilot design is applied to the P2P system. Numerical results show that the superiority of the proposed scheme persists even with 30 dB residual LI power. For quantitative comparison between the proposed scheme and conventional ones, asymptotic achievable rates at ultra-low SNR are derived and the superiority of the proposed scheme is proved theoretically.

- **Power allocation design:** This chapter designs a near-optimal power allocation for the FD overlay scheme to minimize the interference between pilot and data transmissions by properly regulating the source and relay data transmission power for a fixed pilot power and proposes a successive convex approximation (SCA) approach to solve the non-convex optimization problem, where the SCA method is broadly employed for resource allocations in relaying systems[53–55]. Simulation results indicate that the proposed approach further improves the achievable rate compared with the equal power allocation. In addition, the proposed approach is computationally efficient and converges fast. With typical configurations (e.g., total data transmission energy at 20 dB), the simulation shows that the proposed approach converges to a relative error tolerance at $\epsilon = 10^{-5}$ after a few, say 4, iterations.

Organization: The rest of this chapter is organized as follows. The channel and signal models are presented in Section 2.2 within which the conventional and overlay scheme is presented and proposed, respectively. In the following Section 2.3, channel estimations

[2]In this chapter, the conventional schemes for both HD and FD systems represent the existing pilot-data transmission schemes in practical systems where the pilot training is separated from data transmissions.

of the proposed scheme applying to both HD and FD relaying systems are elaborated in details and the system achievable rates are derived theoretically in Section 2.4. Section 2.5 and 2.6 extend the analyses to the asymptotic scenario and power allocation consideration, respectively. The results presented in Section 2.7 reveal the performance comparisons numerically. Section 2.8 concludes the work of this chapter.

2.2 System Model

2.2.1 Signal and Channel Model

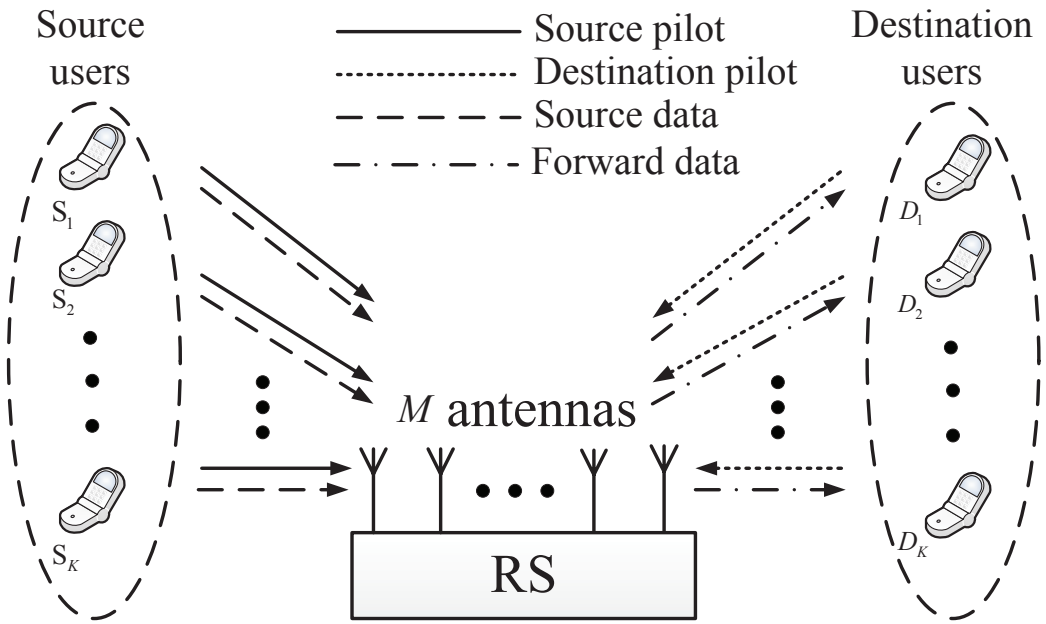
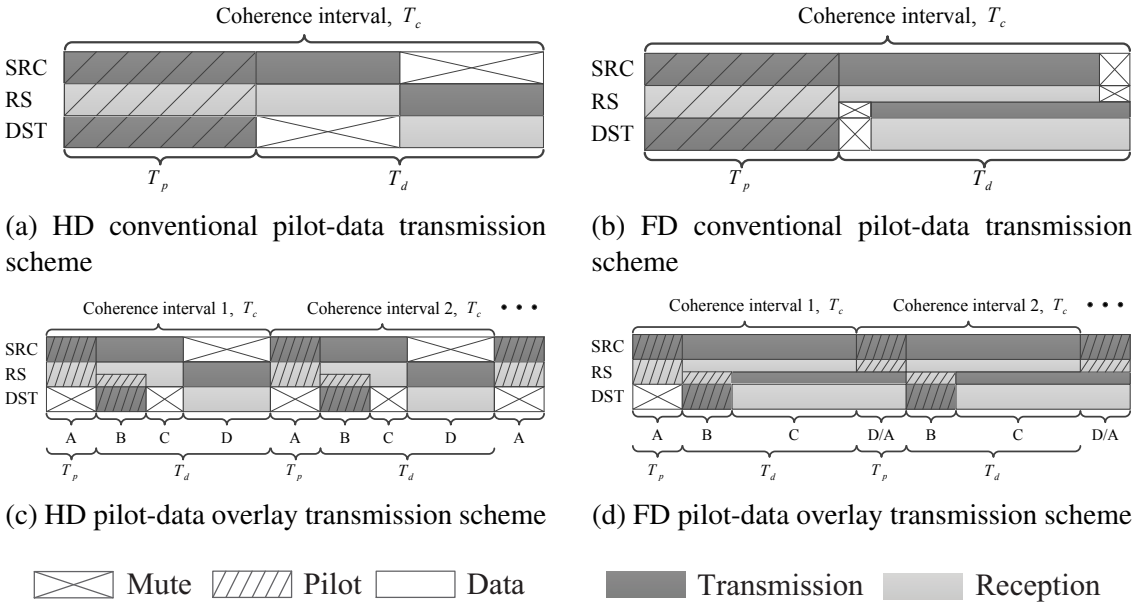


Figure 2.1: System diagram.

As depicted in Fig. 2.1, this chapter considers both HD and FD one-way relaying systems where K pairs of single-antenna source and destination users are served by the RS equipped with M ($M \gg K \gg 1$) antennas. In the following sections, the noise power is normalized to 1. Let ρ_{p}, ρ_{s} and ρ_{d} be the transmission power of pilots, source and forward data, respectively. The channel matrices from sources and destinations to the RS are denoted by $\mathbf{G}_{\mathrm{s}} \in \mathbb{C}^{M\times K}$ and $\mathbf{G}_{\mathrm{d}} \in \mathbb{C}^{M\times K}$, which are concisely named as source and destination channels, respectively, where the kth column of either matrix, $\mathbf{g}_{\mathrm{s}k}$ or $\mathbf{g}_{\mathrm{d}k}$, stands for the channel vector from the kth corresponding user to the RS. Both channel matrices are decomposed as $\mathbf{G}_{\mathrm{s}} = \mathbf{H}_{\mathrm{s}}\mathbf{D}_{\mathrm{s}}^{1/2}$ and $\mathbf{G}_{\mathrm{d}} = \mathbf{H}_{\mathrm{d}}\mathbf{D}_{\mathrm{d}}^{1/2}$, where the large-scale fading matrices $\mathbf{D}_{\mathrm{s}}$ and $\mathbf{D}_{\mathrm{d}}$ are both diagonal with the kth entries as $\beta_{\mathrm{s}k}$ and $\beta_{\mathrm{d}k}$, respec-

tively, and the small-scale fading matrices $\mathbf{H}_{\mathrm{s}}$ and $\mathbf{H}_{\mathrm{d}}$ are constructed by independent identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ random variables (RVs). Note that the large-scale fading factors $\beta_{\mathrm{s}k}$ and $\beta_{\mathrm{d}k}$ are slow-varying and can be treated as constants within the considered communication duration. The leakage channel of the FD communication is modeled as $\mathbf{G}_{\mathrm{LI}} \in \mathbb{C}^{M\times M}$ which represents the residual loop interference after imperfect hardware cancellation [56, 57]. In this chapter, the LI channel is modeled as the Rayleigh fading distribution following the common assumption existing in the literature [30] that the elements of $\mathbf{G}_{\mathrm{LI}}$ are modeled as i.i.d. $\mathcal{CN}(0,\beta_{\mathrm{LI}})$ RVs, where β_{LI} can be treated as the leakage power gain after hardware LI cancellation. Due to the fixation of RS antennas and hardware LI cancellation mechanism, the slow-varying component of the LI channel, β_{LI}, is also assumed to be fixed within the considered communication duration [30]. Further, it is assumed that the same frequency band is reused for both uplink and downlink transmissions, and they obey the reciprocity, i.e., the channel matrices are consistent within a coherence interval for both uplink and downlink communications. Finally, it is assumed that the RS can obtain long-term parameters, such as large-scale fading factors, user numbers, and pilot/data transmission power and inform all users this information via control channels, the design of which is out of the scope of this chapter.



(a) HD conventional pilot-data transmission scheme

(b) FD conventional pilot-data transmission scheme

(c) HD pilot-data overlay transmission scheme

(d) FD pilot-data overlay transmission scheme

Figure 2.2: Conventional and proposed pilot-data transmission diagrams in both HD and FD one-way relaying systems, where T_c, T_p and T_d denote the lengths of coherence, pilot and data transmission intervals, respectively. (SRC: source users, RS: relay station, DST: destination users.)

2.2.2 Conventional Pilot-data Transmission Scheme

This subsection considers the conventional pilot-data transmission scheme for both HD and FD communications.

With regard to the HD system shown by Fig. 2.2(a), TDD is selected as the working mode where the source and destination users firstly transmit training pilots to help the RS estimate CSI, following which the sources send uplink data to the RS, and then the data is forwarded to destination users [12, 46, 58]. Similarly, the FD system depicted in Fig. 2.2(b) shows that the CSI is also estimated first and the data is transmitted subsequently. Yet the FD RS forwards data to destinations simultaneously with source data transmission by some negligible processing time delay[56].

In conventional MIMO systems, the pilot scheme is always designed by utilizing orthogonal pilot sequences to prevent the inner-cell pilot contamination which requires the length of pilot sequences to be not smaller than the number of users. Therefore, $T_p \geq 2K$, where T_p is the length of orthogonal pilot sequences. The overhead of channel estimation is at least $2K/T_{\mathrm{c}}$ of each terminal, where T_{c} is the length of coherence interval in terms of the number of symbol duration [30]. During the rest of the duration of the coherence interval, denoted by T_d, the data transmission takes place. In the massive MIMO relaying system, this overhead is extremely large because the number of users becomes large with the increasing antenna number while the coherence interval is, to some extent, fixed mainly depending on the mobility of terminals. Therefore, especially for the massive MIMO with a relatively short T_{c}, most of the effective resource would be occupied by pilots, which makes the massive MIMO transmission inefficient. In order to reduce the pilot overhead and accordingly improve the data transmission efficiency, this chapter proposes a pilot-data transmission overlay scheme in the following subsection.

2.2.3 Pilot-data Overlay Transmission Scheme

This chapter proposes a pilot-data overlay transmission scheme for both HD and FD one-way relaying systems. The general design of the proposed scheme is explained in this subsection while the detailed signal transmissions will be described mathematically in Sections 2.3 and 2.4.

With respect to the HD relaying system, the pilot-data overlay is depicted in Fig. 2.2(c). To concisely describe signal transmissions, in this chapter, a coherence interval is separated into four phases denoted by A, B, C and D. During phases A and B, pilots are transmitted from users to the RS, while in C and D, the sources and the RS conduct data transmissions, respectively. Within phase A, all source users send piecewisely orthogonal pilot sequences to assist the RS in estimating channels while destination users keep mute, thus source channels can be estimated at the RS without being

contaminated. Subsequently, destination users start transmitting pilots in phase B, while sources can send uplink data to the RS simultaneously. The RS observes both source data and destination pilots in this phase. With the source channel estimated in phase A and the quasi-orthogonality of source and destination massive MIMO channels, the RS detects the source data and then is able to cancel it from the received signal, from which obtains the estimates of the destination channels. Thereafter, sources keep sending uplink data in phase C and the RS forwards downlink data to destination users in phase D. In the following coherence intervals, the HD relaying system repeats these communication procedures.

Regarding the FD pilot-data overlay scheme, the communication procedure is shown in Fig. 2.2(d). In the first coherence interval, the pilot transmissions during phases A and B are correspondingly the same as those of the HD scheme. In phase C, due to the ability of FD, the RS receives the source data as well as forwarding the downlink data to destinations. Different from the HD overlay scheme, the downlink data forwarding during phase D is exactly overlapped by the source pilot transmission in phase A of the subsequent coherence interval. The subsequent phases of the second coherence interval correspondingly repeats those of the first interval. As for the third and following coherence intervals, the communication procedures are identical to those of the second interval.

Remark 2.1. The pilot transmission overhead can be calculated by $\eta_p = T_p/T_c$, where $T_p^{\mathrm{C}} = 2K$ and $T_p^{\mathrm{P}} = K$ for the conventional and proposed schemes, respectively. Hence, it is straightforward to obtain the overheads of the conventional and proposed schemes to be $\eta_p^{\mathrm{C}} = 2K/T_c$ and $\eta_p^{\mathrm{P}} = K/T_c$, respectively. It is obvious that $\eta_p^{\mathrm{C}} = 2\eta_p^{\mathrm{P}}$ which explicitly indicates lower overheads of the proposed scheme in pilot transmissions.

Remark 2.2. For a non-buffered relaying system where the number of forwarding data exactly equals to that of the source data, the portion of data transmission within a coherence interval can be calculated as $\eta_d^{\mathrm{HD}} = T_d^{\mathrm{P}}/2T_c$ in the HD overlay system, where $T_d^{\mathrm{P}} = T_c - K$. Compared to that, it is further increased to $\eta_d^{\mathrm{FD}} = \mathcal{L}T_d^{\mathrm{P}}/(\mathcal{L}T_{\mathrm{c}} + T_p^{\mathrm{P}}) \approx (T_{\mathrm{c}} - K)/T_{\mathrm{c}}$ in the FD system, where $\mathcal{L}$ is the total number of successive coherence intervals used for communications and the approximation is taken when $\mathcal{L}$ is large. FD almost doubles the efficiency of HD data transmission due to the FD property and the proposed pilot-data overlay structure. Note that the source pilot transmission is not indented to overlap the downlink data transmission of the previous coherence interval in the HD mode. It is argued that a TDD RS can not receive source pilots when the transceiver is working in the transmission mode. Therefore, the FD overlay scheme economizes more resources for data transmissions than the HD one.

2.3 Channel Estimation

In this section, the mathematical formulations of channel estimations with the proposed scheme are presented for both duplex relaying systems.

2.3.1 Source Channel Estimation of the First Interval

In the proposed pilot-data overlay structure, the source channel estimations during the first coherence interval for both duplex systems are identical (see Figs. 2.2(c) and 2.2(d)).

In the first coherence interval, the RS receives source pilots without contamination during phase A while the transmitter of the RS and all destination users keep mute. Supposing source users send the pilot matrix $\mathbf{\Phi} \in \mathbb{C}^{K\times K}$ to the RS with power ρ_{p} per user, where the kth row of the matrix, $\boldsymbol{\phi}_k$, is the pilot sequence sent by the kth source user and $\mathbf{\Phi}\mathbf{\Phi}^{\mathrm{H}} = \mathbf{I}_K$ due to orthogonality, the received signal at the RS can be expressed as

$$\mathbf{R}^{\mathrm{A}}[1] = \sqrt{\rho_{\mathrm{p}} K}\mathbf{G}_{\mathrm{s}}[1]\mathbf{\Phi} + \mathbf{N}^{\mathrm{A}}[1], \tag{2.1}$$

where $\mathbf{N}^{\mathrm{A}}[1]$ is the additive white Gaussian noise (AWGN) matrix constructed by $\mathcal{CN}(0,1)$ RVs. By employing the MMSE criteria[59, 60], the estimate of the source channels can be obtained as

$$\hat{\mathbf{G}}_{\mathrm{s}}[1] = \frac{1}{\sqrt{\rho_{\mathrm{p}} K}}\mathbf{R}^{\mathrm{A}}[1]\mathbf{\Phi}^{\mathrm{H}}\tilde{\mathbf{D}}_{\mathrm{s}}[1], \tag{2.2}$$

where $\tilde{\mathbf{D}}_{\mathrm{s}}[1] \triangleq (\mathbf{I}_K + \frac{1}{\rho_{\mathrm{p}} K}\mathbf{D}_{\mathrm{s}}^{-1})^{-1}$ is the coefficient matrix of MMSE channel estimator. Due to the property of MMSE estimation, the channel matrix can be decomposed into two independent components as

$$\mathbf{G}_{\mathrm{s}}[1] = \hat{\mathbf{G}}_{\mathrm{s}}[1] + \mathcal{E}_{\mathrm{s}}[1], \tag{2.3}$$

where $\mathcal{E}_{\mathrm{s}}[1]$ is the error matrix constructed by columns mutually independent of the corresponding column entries of $\hat{\mathbf{G}}_{\mathrm{s}}[1]$, mathematically,

$$\mathbf{g}_{\mathrm{s}k}[1] = \hat{\mathbf{g}}_{\mathrm{s}k}[1] + \boldsymbol{\varepsilon}_{\mathrm{s}k}[1], \tag{2.4}$$

where $\hat{\mathbf{g}}_{\mathrm{s}k}[1]$ and $\boldsymbol{\varepsilon}_{\mathrm{s}k}[1]$ are the kth ($k = 1, 2, \cdots, K$) column vector of $\hat{\mathbf{G}}_{\mathrm{s}}[1]$ and $\mathcal{E}_{\mathrm{s}}[1]$, respectively, $\hat{\mathbf{g}}_{\mathrm{s}k}[1] \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{s}k}^2[1]\mathbf{I}_M)$, $\boldsymbol{\varepsilon}_{\mathrm{s}k}[1] \sim \mathcal{CN}(\mathbf{0}, \varepsilon_{\mathrm{s}k}^2[1]\mathbf{I}_M)$, $\sigma_{\mathrm{s}k}^2[1] = \rho_{\mathrm{p}} K\beta_{\mathrm{s}k}^2/(1+\rho_{\mathrm{p}} K\beta_{\mathrm{s}k})$, and $\varepsilon_{\mathrm{s}k}^2[1] \triangleq \beta_{\mathrm{s}k} - \sigma_{\mathrm{s}k}^2[1]$.

2.3.2 Source Channel Estimation of Subsequent Intervals

According to Remark 2.2 and Fig. 2.2(c), the scenario of HD source pilot transmission (phase A) is identical for all coherence intervals and hence the HD source channel estimation is fully addressed in Subsection 2.3.1. However, the story differs regarding the FD where the transmitter of the RS is working on forwarding the downlink data (phase D) while the receiver is receiving source pilots simultaneously (phase A) in the second coherence interval and after. Therefore, the RS receives both the source pilots and downlink data leakages, which means that the source pilot is contaminated by LI. This subsection considers the source channel estimation of the FD mode in the second and succeeding coherence intervals and characterizes the estimation errors introduced by both AWGN and LI.

Without loss of generality, take the ιth ($\iota > 1$) coherence interval for instance. The received source pilots at the RS can be expressed as

$$\mathbf{R}^{\mathrm{A}}[\iota] = \underbrace{\sqrt{\rho_{\mathrm{p}} K}\mathbf{G}_{\mathrm{s}}[\iota]\mathbf{\Phi}}_{\text{desired signal}} + \underbrace{\sqrt{\rho_{\mathrm{d}}}\alpha[\iota-1]\mathbf{G}_{\mathrm{LI}}\hat{\mathbf{G}}_{\mathrm{d}}[\iota-1]\mathbf{X}^{\mathrm{D}}[\iota-1]}_{\text{LI}} + \underbrace{\mathbf{N}^{\mathrm{A}}[\iota]}_{\text{AWGN}}, \tag{2.5}$$

where ρ_{p} and ρ_{d} represent the transmission power of source pilots and the RS forwarding data, respectively, $\hat{\mathbf{G}}_{\mathrm{d}}[\iota-1]$ (given by (2.10)) denotes the MRT precoding matrix of the forwarding data $\mathbf{X}^{\mathrm{D}}[\iota-1]$ with a power normalization factor $\alpha[\iota-1]$, and $\mathbf{N}^{\mathrm{A}}[\iota] \in \mathbb{C}^{M\times K}$ is the noise matrix constructed by $\mathcal{CN}(0,1)$ RVs. Refer to Sections 2.4.1 and 2.4.2 for detailed descriptions of the LI term and the expression of $\alpha[\iota-1]$, respectively. By applying MMSE channel estimation, the estimate of the source channels is obtained by

$$\hat{\mathbf{G}}_{\mathrm{s}}[\iota] = \frac{1}{\sqrt{\rho_{\mathrm{p}} K}}\mathbf{R}^{\mathrm{A}}[\iota]\mathbf{\Phi}^{\mathrm{H}}\tilde{\mathbf{D}}_{\mathrm{s}}[\iota], \tag{2.6}$$

where $\tilde{\mathbf{D}}_{\mathrm{s}}[\iota] \triangleq \left(\mathbf{I}_K + \frac{\rho_{\mathrm{d}}\beta_{\mathrm{LI}}+1}{\rho_{\mathrm{p}}K}\mathbf{D}_{\mathrm{s}}^{-1}\right)^{-1}$ denotes the coefficient matrix of MMSE channel estimator. Similarly, the channel matrix can also be decomposed into mutually independent two components as follows:

$$\mathbf{G}_{\mathrm{s}}[\iota] = \hat{\mathbf{G}}_{\mathrm{s}}[\iota] + \mathcal{E}_{\mathrm{s}}[\iota], \tag{2.7}$$

where $\mathcal{E}_{\mathrm{s}}[\iota]$ denotes the error matrix of estimations. With regard to the kth columns of matrices $\hat{\mathbf{G}}_{\mathrm{s}}[\iota]$ and $\mathcal{E}_{\mathrm{s}}[\iota]$, there exist $\hat{\mathbf{g}}_{\mathrm{s}k}[\iota] \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathrm{s}k}^2[\iota]\mathbf{I}_M\right)$ and $\boldsymbol{\varepsilon}_{\mathrm{s}k}[\iota] \sim \mathcal{CN}\left(\mathbf{0}, \varepsilon_{\mathrm{s}k}^2[\iota]\mathbf{I}_M\right)$, where $\sigma_{\mathrm{s}k}^2[\iota] \triangleq \rho_{\mathrm{p}}K\beta_{\mathrm{s}k}^2/(\rho_{\mathrm{d}}\beta_{\mathrm{LI}}+1+\rho_{\mathrm{p}}K\beta_{\mathrm{s}k})$ and $\varepsilon_{\mathrm{s}k}^2[\iota] \triangleq \beta_{\mathrm{s}k}-\sigma_{\mathrm{s}k}^2[\iota]$, respectively, for k from 1 to K.

2.3.3 Destination Channel Estimation

In this subsection, the communication in phase B is formulated during which the source data and destination pilots are transmitted simultaneously. To simplify the description, the source data is separated into two successive parts as $\mathbf{S} = [\mathbf{S}^{\mathrm{B}}\ \mathbf{S}^{\mathrm{C}}]$, where $\mathbf{S}^{\mathrm{B}} \in \mathbb{C}^{K\times K}$ is the source data transmitted within phase B and $\mathbf{S}^{\mathrm{C}} \in \mathbb{C}^{K\times(T_{\mathrm{d}}-K)}$ is within phase C. Here, T_{d} is the total length of the source data to transmit by each user within a coherence interval. Without loss of generality, it is assumed that $T_{\mathrm{d}} > K$. The source data $\mathbf{S}^{\mathrm{B}}$ is transmitted alongside destination pilots transmission. The following elaboration reveals that $\mathbf{S}^{\mathrm{B}}$ can be exactly detected from the received signal and the contamination to destination channel estimation can be suppressed by applying source data cancellation with a large number of RS antennas, due to the orthogonality between uplink and downlink channels.

At first, the RS detects the source data from the interfered received signal which can be expressed as

$$\mathbf{R}^{\mathrm{B}}[\iota] = \sqrt{\rho_{\mathrm{s}}}\mathbf{G}_{\mathrm{s}}[\iota]\mathbf{S}^{\mathrm{B}}[\iota] + \sqrt{\rho_{\mathrm{p}}K}\mathbf{G}_{\mathrm{d}}[\iota]\mathbf{\Psi} + \mathbf{N}^{\mathrm{B}}[\iota], \tag{2.8}$$

where $\mathbf{\Psi} \in \mathbb{C}^{K\times K}$ denotes the destination pilot matrix transmitted at power ρ_{p} per user and $\mathbf{N}^{\mathrm{B}}[\iota]$ is the AWGN matrix consisting of $\mathcal{CN}(0,1)$ RVs. The MRC[3] is applied to combine signals received by the RS antennas, where the combiner is $\hat{\mathbf{G}}_{\mathrm{s}}[\iota]$ given by (2.6). Hence, the combined signal is

$$\tilde{\mathbf{S}}^{\mathrm{B}}[\iota] = \hat{\mathbf{G}}_{\mathrm{s}}^{\mathrm{H}}[\iota]\mathbf{R}^{\mathrm{B}}[\iota]. \tag{2.9}$$

Finally, the source data detections are summarized by the following Proposition 2.1.

Proposition 2.1. *With MRC processing, the source data $\mathbf{S}^{\mathrm{B}}[\iota]$ can be exactly detected from $\tilde{\mathbf{S}}^{\mathrm{B}}[\iota]$, if a large number of reception antennas are equipped at the RS, i.e., $M \to \infty$.*

Proof: Using the same manipulation as presented in [43], the proposition can be directly obtained by employing the law of large numbers[61]. ■

Hereby, it is ready to estimate the destination channel by canceling the detected source data from the received signal to reduce pilot contaminations. By recalling the received signal from (2.8) and subtracting the product of the detected source signal and the hermitian of the estimated source channels, the MMSE estimation of the destination

[3] In this chapter, the MRC processing is employed for illustration purpose. The proposed scheme can also be applied to other combining methods, such as zero-forcing combining (ZFC), with minor adjustments.

channels can be obtained as

$$\hat{\mathbf{G}}_{\mathrm{d}}[\iota] = \frac{1}{\sqrt{\rho_{\mathrm{p}} K}} \left(\mathbf{R}^{\mathrm{B}}[\iota] - \sqrt{\rho_{\mathrm{s}}} \hat{\mathbf{G}}_{\mathrm{s}}[\iota] \mathbf{S}^{\mathrm{B}}[\iota] \right) \boldsymbol{\Psi}^{\mathrm{H}} \tilde{\mathbf{D}}_{\mathrm{d}}[\iota], \tag{2.10}$$

where $\tilde{\mathbf{D}}_{\mathrm{d}}[\iota] \triangleq \left(\mathbf{I}_K + \frac{\rho_{\mathrm{s}} \sum_{i=1}^{K} \varepsilon_{\mathrm{s}i}^2[\iota] + 1}{\rho_{\mathrm{p}} K} \mathbf{D}_{\mathrm{d}}^{-1} \right)^{-1}$. And the MMSE estimation follows

$$\mathbf{G}_{\mathrm{d}}[\iota] = \hat{\mathbf{G}}_{\mathrm{d}}[\iota] + \mathcal{E}_{\mathrm{d}}[\iota], \tag{2.11}$$

where $\hat{\mathbf{G}}_{\mathrm{d}}[\iota]$ and $\mathcal{E}_{\mathrm{d}}[\iota]$ are independent of each other. Particularly, the kth columns of both matrices, $\hat{\mathbf{g}}_{\mathrm{d}k}[\iota]$ and $\boldsymbol{\varepsilon}_{\mathrm{d}k}[\iota]$, are mutually independent random vectors, following distribution $\mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{d}k}^2[\iota]\mathbf{I}_M)$ and $\mathcal{CN}(\mathbf{0}, \varepsilon_{\mathrm{d}k}^2[\iota]\mathbf{I}_M)$, respectively, where $\sigma_{\mathrm{d}k}^2[\iota] \triangleq \rho_{\mathrm{p}} K \beta_{\mathrm{d}k}^2 / \left(\rho_{\mathrm{s}} \sum_{i=1}^{K} \varepsilon_{\mathrm{s}i}^2[\iota] + 1 + \rho_{\mathrm{p}} K \beta_{\mathrm{d}k} \right)$ and $\varepsilon_{\mathrm{d}k}^2[\iota] \triangleq \beta_{\mathrm{d}k} - \sigma_{\mathrm{d}k}^2[\iota]$, for k from 1 to K.

Remark 2.3. The covariance factor of the source channel estimate, $\sigma_{\mathrm{s}k}^2[\iota]$, is independent of the coherence interval index ι. In addition, the factor of destination channel estimate, $\sigma_{\mathrm{d}k}^2[\iota]$, only depends on the source channel estimation errors, which are independent of ι. From this phenomenon, it is interesting to note that no error propagation exists for the proposed pilot-data transmission scheme in both duplex relaying systems, which differs from the semi-orthogonal pilot design proposed in [43] where CSI estimation errors accumulate as the increase of ι.

2.4 Achievable Rate Analysis

This section characterizes the performance of the proposed pilot-data transmission scheme by evaluating achievable rates of the considered massive MIMO relaying systems. For the multipair communication, the normalized system achievable rate is defined as an average of sum rates among all user pairs over the entire transmission time, that is

$$\mathcal{R} = \frac{1}{\mathcal{L} T_{\mathrm{c}}} \sum_{\iota=1}^{\mathcal{L}} \sum_{k=1}^{K} \mathcal{R}_k[\iota]. \tag{2.12}$$

The individual achievable rate in the DF relaying system is given by

$$\mathcal{R}_k[\iota] = \min\{\mathcal{R}_k^{\mathrm{UL}}[\iota], \mathcal{R}_k^{\mathrm{DL}}[\iota]\}, \tag{2.13}$$

where $\mathcal{R}_k^{\mathrm{UL}}[\iota]$ and $\mathcal{R}_k^{\mathrm{DL}}[\iota]$ denote the uplink and downlink rates between user pair k and the RS in the coherence interval ι, respectively. Here employ the technique developed by [22] to approximate the ergodic achievable rate for per-link communication, i.e., $\mathcal{R}_k^{\mathrm{UL}}[\iota]$ and $\mathcal{R}_k^{\mathrm{DL}}[\iota]$. In this technique, the received signal is separated into desired signal and

effective noise terms, where the former term is the product of transmitted signal and the expectation of channels while the latter one consists of uncorrelated interferences and AWGN. Hence, only the statistical, other than instantaneous, CSI is required to evaluate the achievable rate. The rate calculated by this technique is the lower bound of the exact one, and numerical results presented in both [14] and [30] show that it is tolerably close to the genie result produced by Monte-Carlo simulation. Consequently, the per-link ergodic achievable rate within a coherence interval is bounded by

$$\mathcal{R}_k^{\mathrm{PL}}[\iota] = \tau_{\mathrm{d}}[\iota] \log_2(1 + \gamma_k^{\mathrm{PL}}[\iota]), \tag{2.14}$$

where $\tau_{\mathrm{d}}[\iota]$ denotes the data transmission time within the coherence interval ι and the effective signal to noise ratio is defined as $\gamma_k^{\mathrm{PL}}[\iota] \triangleq \mathcal{S}_k^{\mathrm{PL}}[\iota]/(\mathcal{I}_k^{\mathrm{PL}}[\iota]+\mathcal{N}_k^{\mathrm{PL}}[\iota])$. Here, $\mathcal{S}_k^{\mathrm{PL}}[\iota]$, $\mathcal{I}_k^{\mathrm{PL}}[\iota]$ and $\mathcal{N}_k^{\mathrm{PL}}[\iota]$ represent the power of the desired signal, uncorrelated interference and AWGN, respectively.

2.4.1 Downlink Analysis

Here the downlink achievable rates for both HD and FD schemes are analyzed. By applying the MRT processing at the RS to the downlink data $\mathbf{X}[\iota] \in \mathbb{C}^{K\times T_{\mathrm{d}}}$ and transmitting to the destination channels with power ρ_{d}, the received signal at destination users is obtained, for user k ($k = 1, 2, \cdots, K$), as

$$\mathbf{y}_k[\iota] = \sqrt{\rho_{\mathrm{d}}}\alpha[\iota]\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{G}}_{\mathrm{d}}[\iota]\mathbf{X}[\iota] + \mathbf{z}_k[\iota], \tag{2.15}$$

where $\mathbf{z}_k[\iota] \in \mathbb{C}^{1\times T_{\mathrm{d}}}$ is the AWGN vector consisting of $\mathcal{CN}(0, 1)$ RVs and $\alpha[\iota]$ is the factor to normalize the average transmit power, i.e., letting $\mathbb{E}\{\|\alpha[\iota]\hat{\mathbf{G}}_{\mathrm{d}}[\iota]\|^2\} = 1$, thus $\alpha[\iota] = \sqrt{1/\left(M\sum_{i=1}^{K}\sigma_{\mathrm{d}i}^2[\iota]\right)}$. To separate the desired signal from the interference and noise, (2.15) can be rewritten as

$$\mathbf{y}_k[\iota] = \underbrace{\sqrt{\rho_{\mathrm{d}}}\alpha[\iota]\mathbb{E}\left\{\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}k}[\iota]\right\}\mathbf{x}_k[\iota]}_{\text{desired signal}} + \underbrace{\breve{\mathbf{z}}_k[\iota]}_{\text{effective noise}}, \tag{2.16}$$

where the effective noise is defined by

$$\breve{\mathbf{z}}_k[\iota] \triangleq \sqrt{\rho_{\mathrm{d}}}\alpha[\iota]\left(\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}k}[\iota] - \mathbb{E}\left\{\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}k}[\iota]\right\}\right)\mathbf{x}_k[\iota] + \sqrt{\rho_{\mathrm{d}}}\alpha[\iota]\sum_{i=1,i\neq k}^{K}\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}i}[\iota]\mathbf{x}_i[\iota] + \mathbf{z}_k[\iota].$$

Therefore, the effective SINR of the received signal at the kth destination user can be expressed as

$$\gamma_k^{\mathrm{DL}}[\iota] = \frac{\rho_{\mathrm{d}}\alpha^2[\iota]\left|\mathbb{E}\left\{\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}k}[\iota]\right\}\right|^2}{\rho_{\mathrm{d}}\alpha^2[\iota]\mathbb{V}\mathrm{ar}\left\{\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}k}[\iota]\right\} + \mathrm{MI}_k^{\mathrm{DL}}[\iota] + 1}, \tag{2.17}$$

where the power of the downlink multipair interference (MI) is defined by

$$\mathrm{MI}_k^{\mathrm{DL}}[\iota] \triangleq \rho_{\mathrm{d}}\alpha^2[\iota]\sum_{i=1,i\neq k}^{K}\mathbb{E}\left\{\left|\mathbf{g}_{\mathrm{d}k}^{\mathrm{H}}[\iota]\hat{\mathbf{g}}_{\mathrm{d}i}[\iota]\right|^2\right\}. \tag{2.18}$$

Theorem 2.1. *By employing the MRT processing, the achievable rate of the downlink data forwarded to the destination user k ($k = 1, 2, \cdots, K$) in both HD and FD, for a finite number of RS transmitter antennas M, can be characterized by*

$$\mathcal{R}_k^{\mathrm{DL}}[\iota] = T_d \log_2\left(1 + \gamma_k^{\mathrm{DL}}[\iota]\right), \tag{2.19}$$

where

$$\gamma_k^{\mathrm{DL}}[\iota] = \frac{M\sigma_{\mathrm{d}k}^4[\iota]}{(\beta_{\mathrm{d}k} + 1/\rho_{\mathrm{d}})\sum_{i=1}^{K}\sigma_{\mathrm{d}i}^2[\iota]}. \tag{2.20}$$

Proof: The results can be directly obtained by applying similar manipulations employed in [30]. ■

2.4.2 Uplink Analysis

In the proposed pilot-data overlay scheme, the uplink data is transmitted in two successive phases where the first part of data is transmitted during phase B and the remaining is sent within phase C. For the two duplex systems, the phase B communication is similar while the phase C differs. The following description first conducts the rate analysis of phase B for both duplex systems, and then perform the phase C analysis distinguished by each mode.

- **Analysis of Uplink Phase B**

  The kth row of $\tilde{\mathbf{S}}^{\mathrm{B}}[\iota]$ in (2.9) can be rewritten as

$$\tilde{\mathbf{s}}_k^{\mathrm{B}}[\iota] = \underbrace{\sqrt{\rho_{\mathrm{s}}}\mathbb{E}\left\{\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota]\right\}\mathbf{s}_k^{\mathrm{B}}[\iota]}_{\text{desired signal}} + \underbrace{\breve{\mathbf{n}}_k^{\mathrm{B}}[\iota]}_{\text{effective noise}}, \tag{2.21}$$

where the effective noise is

$$\breve{\mathbf{n}}_k^{\mathrm{B}}[\iota] \triangleq \sqrt{\rho_{\mathrm{s}}}\left(\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota] - \mathbb{E}\left\{\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota]\right\}\right)\mathbf{s}_k^{\mathrm{B}}[\iota] + \sqrt{\rho_{\mathrm{s}}}\sum_{i=1,i\neq k}^{K}\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}i}[\iota]\mathbf{s}_i^{\mathrm{B}}[\iota] + \sqrt{\rho_{\mathrm{p}}K}\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{G}_{\mathrm{d}}[\iota]\mathbf{\Psi} + \hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{N}^{\mathrm{B}}[\iota].$$

Then, the effective SINR of the received signal during phase B can be expressed by

$$\gamma_k^{\mathrm{B}}[\iota] = \frac{\rho_{\mathrm{s}}\left|\mathbb{E}\left\{\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota]\right\}\right|^2}{\rho_{\mathrm{s}}\mathbb{V}\mathrm{ar}\left\{\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota]\right\} + \mathrm{MI}_k^{\mathrm{UL}}[\iota] + \mathrm{PI}_k^{\mathrm{UL}}[\iota] + \mathrm{AN}_k^{\mathrm{UL}}[\iota]}, \tag{2.22}$$

where the power of uplink MI, destination pilot interference (PI) and AWGN are respectively defined as

$$\mathrm{MI}_k^{\mathrm{UL}}[\iota] \triangleq \rho_{\mathrm{s}}\sum_{i=1,i\neq k}^{K}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}i}[\iota]\right|^2\right\}, \tag{2.23}$$

$$\mathrm{PI}_k^{\mathrm{UL}}[\iota] \triangleq \rho_{\mathrm{p}}\mathbb{E}\left\{\left\|\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{G}_{\mathrm{d}}[\iota]\right\|^2\right\}, \tag{2.24}$$

$$\mathrm{AN}_k^{\mathrm{UL}}[\iota] \triangleq \mathbb{E}\left\{\left\|\hat{\mathbf{g}}_{\mathrm{s}k}[\iota]\right\|^2\right\}. \tag{2.25}$$

- **Analysis of Uplink Phase C**

  Regarding the uplink data transmission during phase C, only MI interferes for the HD communication while for the FD, LI also exists. The received signal at the RS during phase C can be expressed as

$$\mathbf{R}^{\mathrm{C}}[\iota] = \sqrt{\rho_{\mathrm{s}}}\mathbf{G}_{\mathrm{s}}[\iota]\mathbf{S}^{\mathrm{C}}[\iota] + \mathbf{L}^{\mathrm{C}}[\iota] + \mathbf{N}^{\mathrm{C}}[\iota], \tag{2.26}$$

  where $\mathbf{L}^{\mathrm{C}}[\iota] \triangleq \sqrt{\rho_{\mathrm{d}}}\alpha[\iota]\mathbf{G}_{\mathrm{LI}}[\iota]\hat{\mathbf{G}}_{\mathrm{d}}^{\mathrm{H}}[\iota]\mathbf{X}^{\mathrm{C}}[\iota]$ is the LI for the FD relaying while for HD, it is zero, and $\mathbf{N}^{\mathrm{C}}[\iota]$ is the AWGN matrix. By applying MRC to the kth source data with similar manipulations to the combined signal as in (2.21), the effective SINR of the received signal during phase C is obtained as

$$\gamma_k^{\mathrm{C}}[\iota] = \frac{\rho_{\mathrm{s}}\left|\mathbb{E}\left\{\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota]\right\}\right|^2}{\rho_{\mathrm{s}}\mathbb{V}\mathrm{ar}\left\{\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{g}_{\mathrm{s}k}[\iota]\right\} + \mathrm{MI}_k^{\mathrm{UL}}[\iota] + \mathrm{LI}_k^{\mathrm{UL}}[\iota] + \mathrm{AN}_k^{\mathrm{UL}}[\iota]}, \tag{2.27}$$

  where $\mathrm{MI}_k^{\mathrm{UL}}[\iota]$ and $\mathrm{AN}_k^{\mathrm{UL}}[\iota]$ are respectively defined by (2.23) and (2.25), and the

power of LI is defined as

$$\mathrm{LI}_k^{\mathrm{UL}}[\iota] \triangleq \begin{cases} 0, & \text{HD} \\ \rho_{\mathrm{d}}\alpha^2[\iota]\mathbb{E}\left\{\left\|\hat{\mathbf{g}}_{\mathrm{s}k}^{\mathrm{H}}[\iota]\mathbf{G}_{\mathrm{LI}}[\iota]\hat{\mathbf{G}}_{\mathrm{d}}[\iota]\right\|^2\right\}, & \text{FD}. \end{cases} \tag{2.28}$$

Finally, the uplink achievable rates of pilot-data overlay relaying systems are given by Theorem 2.2.

Theorem 2.2. *With MRC processing, the uplink achievable rate of source user k ($k = 1, 2, \cdots K$) during coherence interval ι in both HD and FD systems, for a finite antenna number M, can be characterized by*

$$\mathcal{R}_k^{\mathrm{UL}}[\iota] = K\log_2(1+\gamma_k^{\mathrm{B}}[\iota]) + (T_{\mathrm{d}} - K)\log_2(1+\gamma_k^{\mathrm{C}}[\iota]), \tag{2.29}$$

where

$$\gamma_k^{\mathrm{B}}[\iota] = \frac{M\sigma_{\mathrm{s}k}^2[\iota]}{\sum_{i=1}^{K}\beta_{\mathrm{s}i} + \left(\rho_{\mathrm{p}}\sum_{i=1}^{K}\beta_{\mathrm{d}i} + 1\right)/\rho_{\mathrm{s}}}, \tag{2.30}$$

and

$$\gamma_k^{\mathrm{C}}[\iota] = \begin{cases} \frac{M\sigma_{\mathrm{s}k}^2[\iota]}{\sum_{i=1}^{K}\beta_{\mathrm{s}i}+1/\rho_{\mathrm{s}}}, & \textit{HD} \\ \frac{M\sigma_{\mathrm{s}k}^2[\iota]}{\sum_{i=1}^{K}\beta_{\mathrm{s}i}+(\rho_{\mathrm{d}}\beta_{\mathrm{LI}}+1)/\rho_{\mathrm{s}}}, & \textit{FD}. \end{cases} \tag{2.31}$$

Proof: The results can be directly obtained by applying similar manipulations employed in [30]. ■

Remark 2.4. Having the numerator and denominator of the FD case in (2.31) both divided by M, it is observed that the LI term, $\rho_{\mathrm{d}}\beta_{\mathrm{LI}}/M$, vanishes as the number of RS antennas tending infinity, i.e., $M \to \infty$. This is consistent with the observation on LI discovered in [30] but for a different pilot-data transmission scheme. By applying the same manner to (2.20) and (2.31), it is evident that the multipair interferences in both uplink and downlink transmissions vanish if an infinite number of antennas is deployed on the RS.

2.5 Asymptotic Analysis

As the seminal investigation in [12], the effects of additive receiver noise, fast fading and intra-cellular interference vanish as the number of antennas grows without limit, which leads to an infinite SINR. Hence, the signal can be transmitted at a very low power by employing large-scale antenna arrays. In this section, the performance of the proposed scheme at asymptotic low SNR regions is analyzed and compared with the conventional

pilot scheme. In following asymptotic derivations, it is assumed that the transmit power of all data and pilot is identical, i.e., $\rho_{\mathrm{s}} = \rho_{\mathrm{d}} = \rho_{\mathrm{p}} = \rho$. To perform the asymptotic low SNR analyses, an issue is addressed here. In a practical system, the number of RS antennas is impossible to achieve infinity, i.e., M is a finite number. Such that, when $\rho \to 0$, the achievable rate will be zero with insignificance. Hence, very low SNR in the asymptotic analyses is considered where ρ tends to a positive infinitesimally small value denoted by $\rho \to \theta^{+}$. Finally, this chapter adopts $\tilde{(\cdot)}$ to indicate the corresponding symbols for the conventional scheme in the derivations throughout this section.

2.5.1 Half-duplex Relaying

Firstly, the asymptotic performance at the low SNR region is analyzed for the HD relaying with the proposed pilot-data overlay scheme. Note that the suffix $[\iota]$ is dropped for concision in the derivations of this subsection due to the consistence of the overlay scheme for HD relaying.

- **Uplink Analyses**

  The difference of the uplink rates between the proposed and conventional schemes can be expressed by

$$\begin{aligned}\mathcal{R}_k^{\mathrm{UL}} - \tilde{\mathcal{R}}_k^{\mathrm{UL}} =& K\log_2\left(1+\gamma_k^{\mathrm{B}}\right) + (T_d - K)\log_2\left(1+\gamma_k^{\mathrm{C}}\right) - \tilde{T}_d\log_2\left(1+\tilde{\gamma}_k^{\mathrm{UL}}\right)\\ =& K\log_2\left(1+\gamma_k^{\mathrm{B}}\right) - \frac{K}{2}\log_2\left(1+\gamma_k^{\mathrm{C}}\right) \qquad (2.32a)\\ =& \frac{K}{2}\log_2\left(1+\frac{2\gamma_k^{\mathrm{B}} - \gamma_k^{\mathrm{C}} + (\gamma_k^{\mathrm{B}})^2}{1+\gamma_k^{\mathrm{C}}}\right), \qquad (2.32b)\end{aligned}$$

  where $T_d = (T_c - K)/2$ for the proposed scheme while $\tilde{T}_d = (T_c - 2K)/2$ for the conventional scheme and (2.32a) is obtained due to $\gamma_k^{\mathrm{C}} = \tilde{\gamma}_k^{\mathrm{UL}}$. It is evident that $\lim\limits_{\rho\to\theta^{+}} \gamma_k^{\mathrm{B}} = \lim\limits_{\rho\to\theta^{+}} \gamma_k^{\mathrm{C}}$. Hence, (2.32b) is larger than zero when $\rho \to \theta^{+}$ which leads to

$$\lim_{\rho\to\theta^{+}} \mathcal{R}_k^{\mathrm{UL}} > \lim_{\rho\to\theta^{+}} \tilde{\mathcal{R}}_k^{\mathrm{UL}}. \qquad (2.33)$$

- **Downlink Analyses**

  To compare the downlink rates, this chapter first compares the covariance of the destination channel estimations as follows:

$$\lim_{\rho\to\theta^{+}} \frac{\sigma_{\mathrm{d}k}^2}{\tilde{\sigma}_{\mathrm{d}k}^2} = \lim_{\rho\to\theta^{+}} \frac{1+\rho K\beta_{\mathrm{d}k}}{\rho\sum_{i=1}^{K}\varepsilon_{\mathrm{s}i}^2[\iota] + 1 + \rho K\beta_{\mathrm{d}k}} = 1. \qquad (2.34)$$

Hence, $\lim_{\rho\to\theta^+} \sigma_{\mathrm{d}k}^2 = \lim_{\rho\to\theta^+} \tilde{\sigma}_{\mathrm{d}k}^2$. By substituting the covariances of destination channel estimates into (2.20) and noting that the effective downlink SNR for conventional scheme maintains the same structure as (2.20), it can derive the following conclusion as

$$\lim_{\rho\to\theta^+} \mathcal{R}_k^{\mathrm{DL}} > \lim_{\rho\to\theta^+} \tilde{\mathcal{R}}_k^{\mathrm{DL}}, \tag{2.35}$$

due to $T_d > \tilde{T}_d$.

2.5.2 Full-duplex Relaying

For the FD relaying, the asymptotic analyses for the first coherence interval, i.e., $\iota = 1$, are similar to the HD relaying and lead to the same results. Hence, it is only considered in this subsection that $\iota > 1$.

- **Uplink Analyses**

  Firstly, the difference of the uplink rate between the proposed and conventional schemes can be derived as follows:

$$\begin{aligned}
&\mathcal{R}_k^{\mathrm{UL}}[\iota] - \tilde{\mathcal{R}}_k^{\mathrm{UL}}[\iota] \\
=&K\log_2\left(1+\gamma_k^{\mathrm{B}}[\iota]\right) + (T_d - K)\log_2\left(1+\gamma_k^{\mathrm{C}}[\iota]\right) - \tilde{T}_d\log_2\left(1+\tilde{\gamma}_k^{\mathrm{UL}}[\iota]\right) \\
=&K\log_2\left(1+\gamma_k^{\mathrm{B}}[\iota]\right) + K\log_2\left(1+\gamma_k^{\mathrm{C}}[\iota]\right) + (T-2K)\log_2\frac{1+\gamma_k^{\mathrm{C}}[\iota]}{1+\tilde{\gamma}_k^{\mathrm{UL}}[\iota]},
\end{aligned} \tag{2.36}$$

  where $T_d = T_c - K$ for the proposed scheme while $\tilde{T}_d = T_c - 2K$ for the conventional scheme. It is straightforward to obtain that $\lim_{\rho\to\theta^+} \sigma_{\mathrm{s}k}^2[\iota] = \lim_{\rho\to\theta^+} \tilde{\sigma}_{\mathrm{s}k}^2[\iota]$ which induces $\lim_{\rho\to\theta^+} \gamma_k^{\mathrm{C}}[\iota] = \lim_{\rho\to\theta^+} \tilde{\gamma}_k^{\mathrm{UL}}[\iota]$. Hence, the third term of (2.36) tends to zero as $\rho \to \theta^+$. Therefore, (2.36) is greater than zero at very low SNR due to the first two positive terms, such that

$$\lim_{\rho\to\theta^+} \mathcal{R}_k^{\mathrm{UL}}[\iota] > \lim_{\rho\to\theta^+} \tilde{\mathcal{R}}_k^{\mathrm{UL}}[\iota]. \tag{2.37}$$

- **Downlink Analyses**

  Similar to the HD analyses, it also holds in the FD relaying that $\lim_{\rho\to\theta^+} \sigma_{\mathrm{d}k}^2[\iota] = \lim_{\rho\to\theta^+} \tilde{\sigma}_{\mathrm{d}k}^2[\iota]$, such that $\lim_{\rho\to\theta^+} \gamma_k^{\mathrm{DL}}[\iota] = \lim_{\rho\to\theta^+} \tilde{\gamma}_k^{\mathrm{DL}}[\iota]$. Therefore, the downlink rate of the proposed scheme is greater than the conventional one due to the same structure of downlink effective SNRs between the two schemes and $T_d > \tilde{T}_d$, which can be represented mathematically as follows:

$$\lim_{\rho\to\theta^+} \mathcal{R}_k^{\mathrm{DL}}[\iota] > \lim_{\rho\to\theta^+} \tilde{\mathcal{R}}_k^{\mathrm{DL}}[\iota]. \tag{2.38}$$

From the achievable rate comparisons between two pilot schemes denoted by (2.33), (2.35), (2.37) and (2.38), Corollary 2.1 can be drawn to confirm the advantages of the proposed pilot-data overlay scheme.

Corollary 2.1. *With a large number of user pairs served by the RS, the proposed pilot-data overlay transmission scheme outperforms the conventional one in both multipair HD and FD massive MIMO relaying systems for the low SNR scenario.*

Remark 2.5. According to Corollary 2.1, it is interesting to stress that the conclusion is independent of the coherence interval length T_{c} and the number of successive coherence intervals used for continuous communications $\mathcal{L}$. Therefore, the proposed scheme is always superior to, at least not worse than, the conventional scheme for low SNR, regardless of the coherence interval length.

2.6 Power Allocation for FD Relaying

By analyzing the effective SINRs of the FD relaying system, one can find that the self-interference between uplink and downlink data transmissions, and hence the impact upon pilots sending, is a major factor that limits the system achievable rate [62]. Therefore, making a tradeoff between uplink and downlink data transmission is necessary to improve the overall system performance. The power allocation is performed in the FD relaying system to optimize the transmit power of both source and forwarding data to control interference among uplink, downlink and pilot transmissions. Note that this chapter only considers balancing the source and the RS data transmission power level while the power allocation for each individual source user is not taken into account. In addition, this chapter assumes the total power consumption of all data transmission to be constrained under a fixed pilot transmit power. In other words, data transmission power ρ_{s} and ρ_{d} are balanced with respect to a given ρ_{p} to achieve the maximal overall system rate.

To begin with, the power allocation problem is formulated as

$$\begin{aligned} \underset{\rho_{\mathrm{s}},\rho_{\mathrm{d}}}{\text{maximize}} \quad & \sum_{\iota=1}^{\mathcal{L}} \sum_{k=1}^{K} \mathcal{R}_k[\iota] \\ \text{subject to:} \quad & \mathcal{R}_k[\iota] = \min\left\{\mathcal{R}_k^{\mathrm{UL}}[\iota], \mathcal{R}_k^{\mathrm{DL}}[\iota]\right\} \\ & \mathcal{L}T_d\left(\rho_{\mathrm{s}}K + \rho_{\mathrm{d}}\right) = E_{\mathrm{d}} \\ & k = 1, 2, \cdots, K \text{ and } \iota = 1, 2, \cdots, \mathcal{L}, \end{aligned} \tag{2.39}$$

where $\mathcal{R}_k^{\mathrm{UL}}[\iota]$, $\mathcal{R}_k^{\mathrm{DL}}[\iota]$ and $\mathcal{R}_k[\iota]$ denote the uplink, downlink and end-to-end (e2e) achievable rate between the kth user pair and the RS within the ιth coherence interval, respectively, and E_{d} is the total data transmission energy [30]. The operator $\min\{\cdot,\cdot\}$ in the

above formulation is due to the DF relaying expressed in (2.13), where the e2e achievable rate is determined by the minimum of the corresponding uplink and downlink rates. To remove the minimum operator, the first constraint of problem (2.39) can be reformulated equivalently to

$$\begin{aligned} \mathcal{R}_k[\iota] &\leq \mathcal{R}_k^{\mathrm{UL}}[\iota], \\ \mathcal{R}_k[\iota] &\leq \mathcal{R}_k^{\mathrm{DL}}[\iota]. \end{aligned} \tag{2.40}$$

By examining $\mathcal{R}_k^{\mathrm{UL}}[\iota]$ and $\mathcal{R}_k^{\mathrm{DL}}[\iota]$ with respect to ρ_{s} and ρ_{d}, it is clear that these two inequality constraints are non-convex and thus it is difficult to solve (2.39) directly.

To cope with this problem, a successive convex approximation (SCA) approach [63–65] is proposed in this chapter by replacing constraint (2.40) with linear inequalities. That is to approximate both uplink and downlink rates by their first-order Taylor series in the ith iteration with the known power allocation $\boldsymbol{\rho}^{(i)}$ via

$$\mathcal{R}_k^{\mathrm{UL,DL}}[\iota](\boldsymbol{\rho}) = \mathcal{R}_k^{\mathrm{UL,DL}}[\iota](\boldsymbol{\rho}^{(i)}) + \nabla^{\mathrm{T}}\mathcal{R}_k^{\mathrm{UL,DL}}[\iota](\boldsymbol{\rho}^{(i)})(\boldsymbol{\rho} - \boldsymbol{\rho}^{(i)}), \tag{2.41}$$

where $\boldsymbol{\rho} \triangleq [\rho_{\mathrm{s}}\ \rho_{\mathrm{d}}]^{\mathrm{T}}$ and the superscript $(\cdot)^{(i)}$ denotes the corresponding value at the ith iteration. Substituting (2.41) into (2.40), the optimization problem can be reformulated in the ith iteration as

$$\begin{aligned} \underset{\rho_{\mathrm{s}},\rho_{\mathrm{d}}}{\text{maximize}} \quad & \sum_{\iota=1}^{\mathcal{L}}\sum_{k=1}^{K}\mathcal{R}_k^{(i)}[\iota] \\ \text{subject to:} \quad & \mathcal{R}_k^{(i)}[\iota] \leq \mathcal{R}_k^{\mathrm{UL}}[\iota](\boldsymbol{\rho}^{(i)}) + \nabla^{\mathrm{T}}\mathcal{R}_k^{\mathrm{UL}}[\iota](\boldsymbol{\rho}^{(i)})(\boldsymbol{\rho} - \boldsymbol{\rho}^{(i)}) \\ & \mathcal{R}_k^{(i)}[\iota] \leq \mathcal{R}_k^{\mathrm{DL}}[\iota](\boldsymbol{\rho}^{(i)}) + \nabla^{\mathrm{T}}\mathcal{R}_k^{\mathrm{DL}}[\iota](\boldsymbol{\rho}^{(i)})(\boldsymbol{\rho} - \boldsymbol{\rho}^{(i)}) \\ & \mathcal{L}T_d\left(\rho_{\mathrm{s}}K + \rho_{\mathrm{d}}\right) = E_{\mathrm{d}} \\ & k = 1, 2, \cdots, K \text{ and } \iota = 1, 2, \cdots, \mathcal{L}. \end{aligned} \tag{2.42}$$

It is obvious that (2.42) is a linear programming (LP) which can be solved efficiently by utilizing, e.g., the conventional interior-point method [66].

The LP in (2.42) is solved repeatedly in iterations with i increased by 1 each time, until the sum rate $\bar{\mathcal{R}}^{(i)}$ is stable which can be characterized by $|\bar{\mathcal{R}}^{(i+1)} - \bar{\mathcal{R}}^{(i)}| < \epsilon$, where $\bar{\mathcal{R}}^{(i)} \triangleq \sum_{\iota=1}^{\mathcal{L}}\sum_{k=1}^{K}\mathcal{R}_k^{(i)}[\iota]$ and ϵ is a set error tolerance [67]. After all, the $\boldsymbol{\rho}^{(i)}$ in the last iteration is the optimizer of (40), $\boldsymbol{\rho}^{*}$, that leads to the local maximum of the sum rate. The SCA approach is summarized by Algorithm 1.

Algorithm 1 Successive Convex Approximation Approach

Require: Solve optimization problem (2.39)
Input: K, M, $\mathcal{L}$, ρ_{p}, ρ_{LI}, β_{sk}, $\beta_{\mathrm{d}k}$ and ϵ
Output: Optimized power allocation ρ_{s}^{*} and ρ_{d}^{*}

```
Initialize ϵ and set i = 0
Initialize ρ^(0)
loop
    Solve the LP in (2.42) to obtain ρ* and R̄*
    if |R̄* − R̄^(i)| / R̄^(i) < ϵ then
        Stop loop
    else
        R̄^(i+1) = R̄*
        i = i + 1
    end if
end loop
Output ρ_s* and ρ_d*.
```

2.7 Numerical and Simulation Results

In this section, the performance of the proposed pilot-data overlay transmission scheme and the power allocation approach are studied by simulation and numerical evaluation for both HD and FD relaying systems. Unless otherwise specified, by default, the RS is equipped with 128 antennas serving 40 pairs of users, the LI channel gain β_{LI} is set to be 3 dB, the processing delay of the FD relaying for the conventional pilot scheme is 1 symbol slot, the coherence interval is set to be 200[4], the transmission power of pilots, source data and forwarding data satisfy $\rho_{\mathrm{p}} = \rho_{\mathrm{s}} = \rho_{\mathrm{d}} = 20$ dB, the large-scale fading factor for each user is set to be 1. Note that the power and energy values mentioned in this section are all relative to the noise power and denoted in dB. The Monte Carlo results are calculated following (2.12) to (2.14) by taking the average of 1000 random simulations with instantaneous channels [30].

Firstly, the system achievable rates under different SNRs are evaluated and the performance is compared between the proposed and conventional schemes for both HD and FD relayings. The performance comparisons are shown in Fig. 2.3 where the SNR varies from -30 dB to 30 dB. In the figure, the lines represent the rates obtained by Monte Carlo simulation while the markers denote the ergodic achievable rate lower bound computed by the closed-form expressions with statistical CSI (refer to as Theorem 2.1 and 2.2). The comparison shows that the relative performance gap between Monte Carlo and the closed-form results is small, e.g., 4.19 bits/s/Hz for FD relative to 45.08 bits/s/Hz and 2.17 bits/s/Hz for HD relative to 23.55 bits/s/Hz at 0 dB of SNR, which implies that our

[4]Note that the coherence interval in this section is the normalized value relative to the symbol duration.

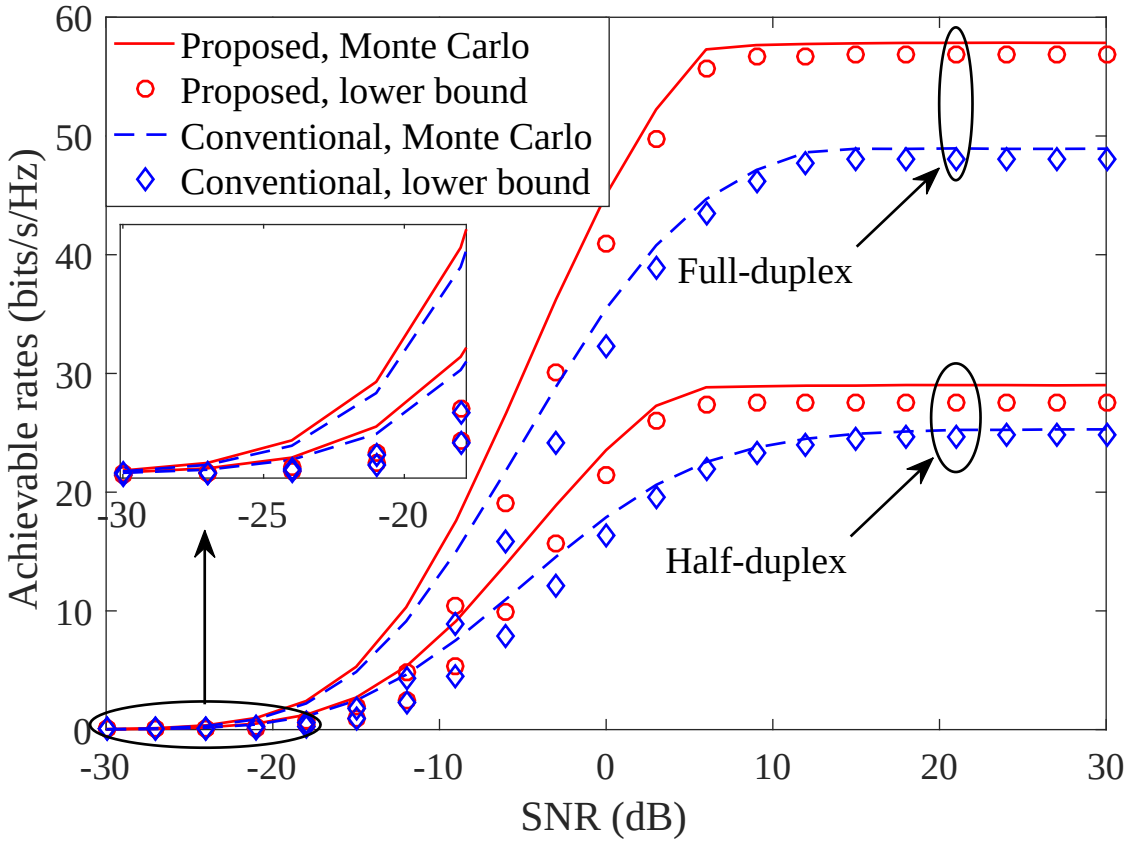


Figure 2.3: Comparisons of achievable rates between the proposed and conventional schemes under different SNRs.

closed-form ergodic achievable rate expression is a good predictor for the system performance. In addition, Fig. 2.3 shows that the proposed scheme outperforms the conventional one in both high and low SNR regions, which verifies Corollary 2.1, where about 8.91 bits/s/Hz and 3.73 bits/s/Hz improvements are observed in the high SNR region for the FD and HD systems, respectively. On top of that, the FD mode is shown exceeding the HD in system rates by about 28.82 bits/s/Hz in high SNRs with the proposed scheme, which verifies Remark 2.2.

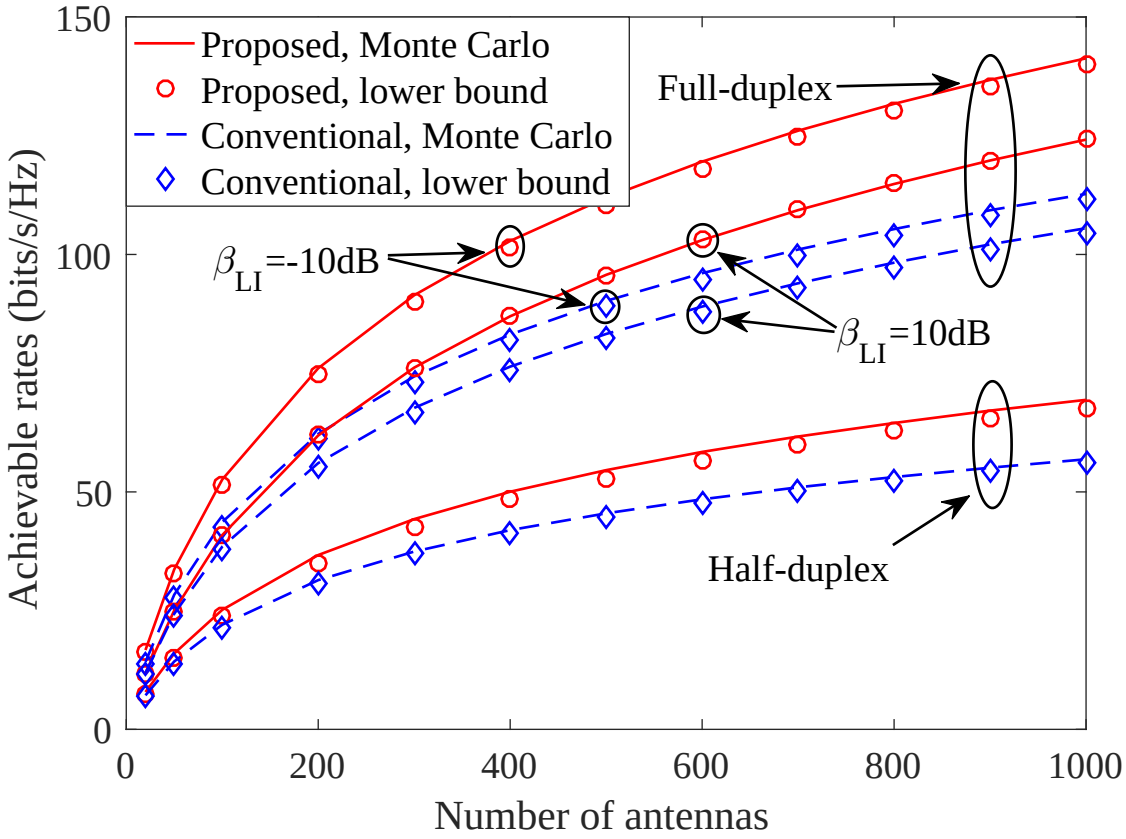


Figure 2.4: Comparisons between the proposed and conventional schemes versus the number of antennas equipped on the RS.

Next, the performance of the massive MIMO system versus the growing number of RS antennas is depicted in Fig. 2.4. It is obvious that the rate gap between the proposed and conventional schemes increases as the number of RS antennas grows. With more RS antennas, the source and destination channels are closer to be orthogonal to each other, and thus less interferences reside in the combined signal leading to improved performance of the proposed scheme. As expected, the increase of the LI power degrades the FD performance and Fig. 2.4 shows the superiority of the proposed scheme in the FD system as the LI power equals to both as small as 10 dB and as large as 30 dB, where the LI power is calculated by $\rho_{\mathrm{LI}} = \rho_{\mathrm{d}}\beta_{\mathrm{LI}}$. Further, even as few as 20 antennas are deployed on the RS, the proposed scheme still outperforms the conventional one, which verifies

that the proposed scheme is feasible and superior even in medium scale MIMO relaying systems.

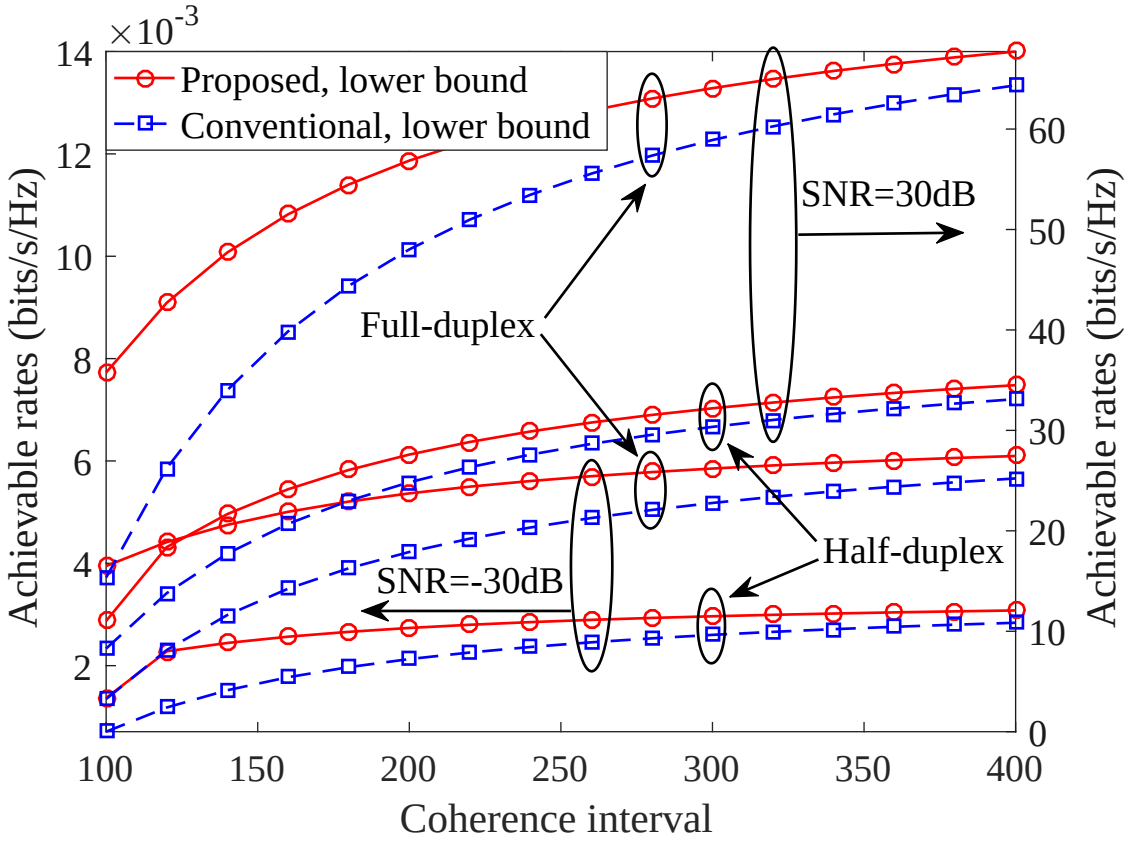


Figure 2.5: Impact of the coherence interval length on the performance of the proposed and conventional schemes.

As discussed in Section 2.6, further performance improvement of the proposed scheme can be obtained by balancing the tradeoff between pilot overhead and channel estimation accuracy where the estimation precision decreases due to the data interference as the percentage of data transmission time within a coherence interval increases. Therefore, the length of the coherence interval and the SNR of the received pilots are two crucial system parameters that affect the channel estimation overhead and accuracy, respectively. Here, the performance improvement of the proposed scheme is verified with respect to the length of coherence intervals. According to Fig. 2.5, the achievable rate of the proposed scheme always outperforms the conventional one for coherence interval from 100 to 400 at both high and low SNRs, which verifies the conclusion in Remark 2.5. Nonetheless, the gap between the two schemes decreases with the increase of the interval length due to the reduction of relative pilot transmission overhead within a longer coherence interval.

In practical systems, the network operator would prefer to serve more user pairs to improve the overall performance of the entire system at long coherence intervals, espe-

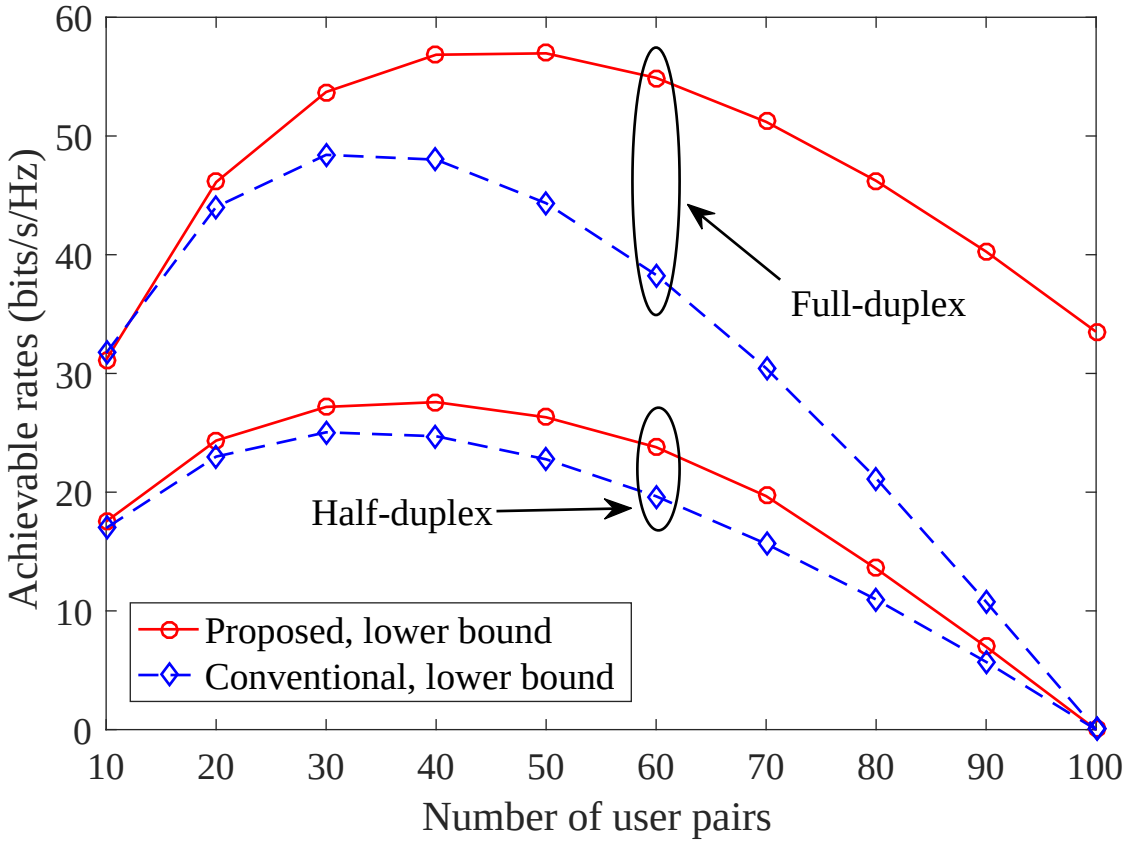


Figure 2.6: Comparisons between the proposed and conventional schemes versus the number of user pairs.

cially when a large number of antennas exist[42, 68]. Therefore, the performance of both pilot schemes by varying the number of user pairs is examined at a fixed coherence interval. As shown in Fig. 2.6, the maximal rate is achieved when 50 user pairs are communicating simultaneously in the FD system by utilizing the proposed transmission scheme while only 30 user pairs are served with the conventional scheme whose sum rate rapidly deceases with the increase of user numbers. Moreover, the proposed scheme performs almost the same as the conventional one even with 10 user pairs served by the FD relaying system, which affirms the robustness of the proposed scheme when small number of users access to the network. In the HD mode, similar performance comparisons are observed. As expected, the growing number of user pairs also enlarges the rate gap between two pilot schemes. Particularly, when the number of user pairs equals to the half length of coherence interval, i.e., $K = 100$, all HD and conventional FD curves touch zero rates due to no time resource spared for data transmission, except the FD overlay system still working in the high-throughput state. Therefore, the comparisons reveals that both HD and FD relaying systems employing the pilot-data overlay scheme achieve higher system rates and serve more user pairs than those with the conventional pilot scheme. Further, the FD overlay system emerges to be superior to all other systems in the extreme scenario.

Finally, the performance of the proposed power allocation approach is evaluated for the FD overlay relaying. In the simulation, the pilot power is fixed to be 20 dB per user and the total data transmission power per coherence interval varies from 0 dB to 70 dB for both optimal and equal power allocations. Additionally, $\rho_{\mathrm{d}} = \rho_{\mathrm{s}} K$ is set in the equal power allocation. Fig. 2.7 shows the comparison of the system rates between the two allocation schemes while the complexity and convergence of the proposed approach at various cases are depicted in Fig. 2.8. Fig. 2.7 shows that the optimal power allocation achieves better performance than the equal power allocation for both low and high transmission powers, especially at moderate power region, where the rate increment becomes more noticeable. The rate of equal power allocation starts decreasing when the data transmission power increases to an extreme large value while the proposed scheme maintains a fixed high performance. Such rate decrement of equal power allocation is due to more pilot contamination when data transmission power is large yet pilot power is fixed. In contrast, the optimal power allocation scheme can adaptively adjust the data transmission power to control the interference between pilot and data transmissions and therefore always achieve the best performance. From the observation of Fig. 2.7, it can be found that if only trivial power allocation, i.e., the equal one, is performed in the FD relaying system, the leakage power through the FD loop channel will introduce increasing LI when the downlink transmission power is becoming larger, which on the contrary leads to the decrement of system throughput. Hence, the RS should adjust the downlink

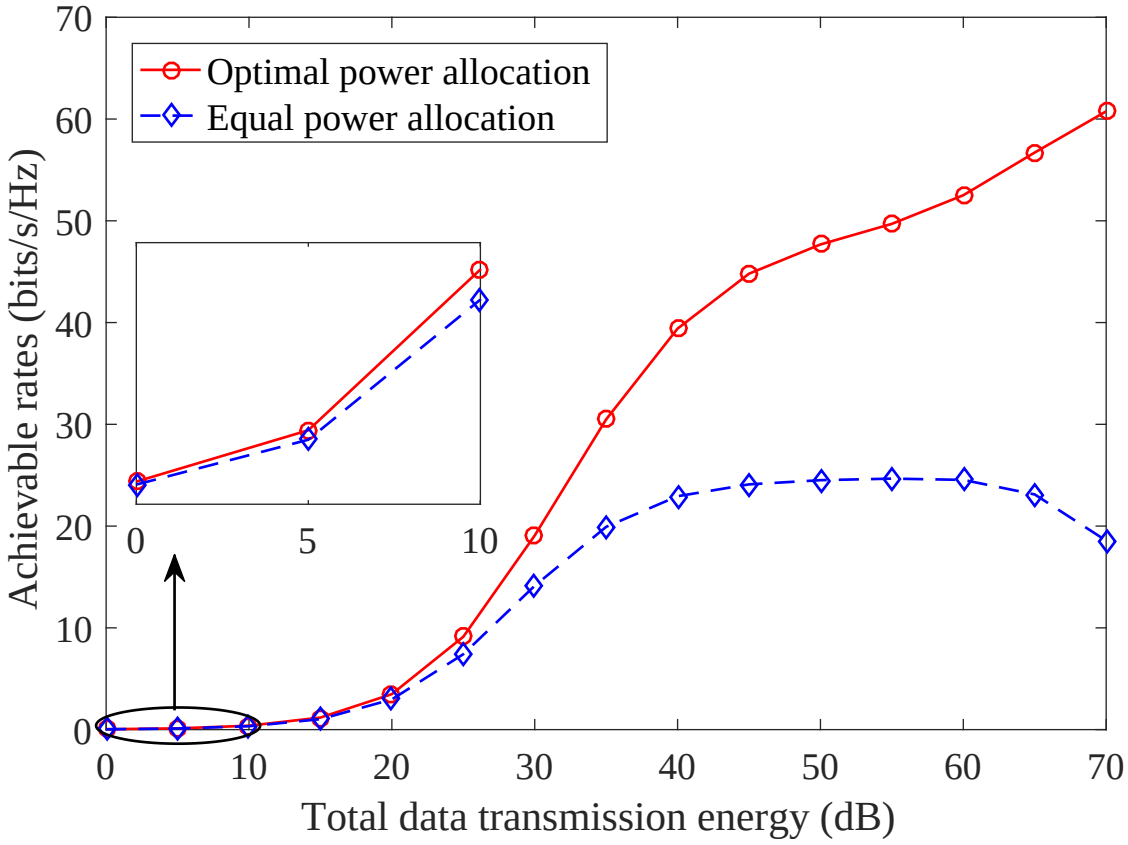


Figure 2.7: Performance comparison between two power allocation schemes in the FD overlay system.

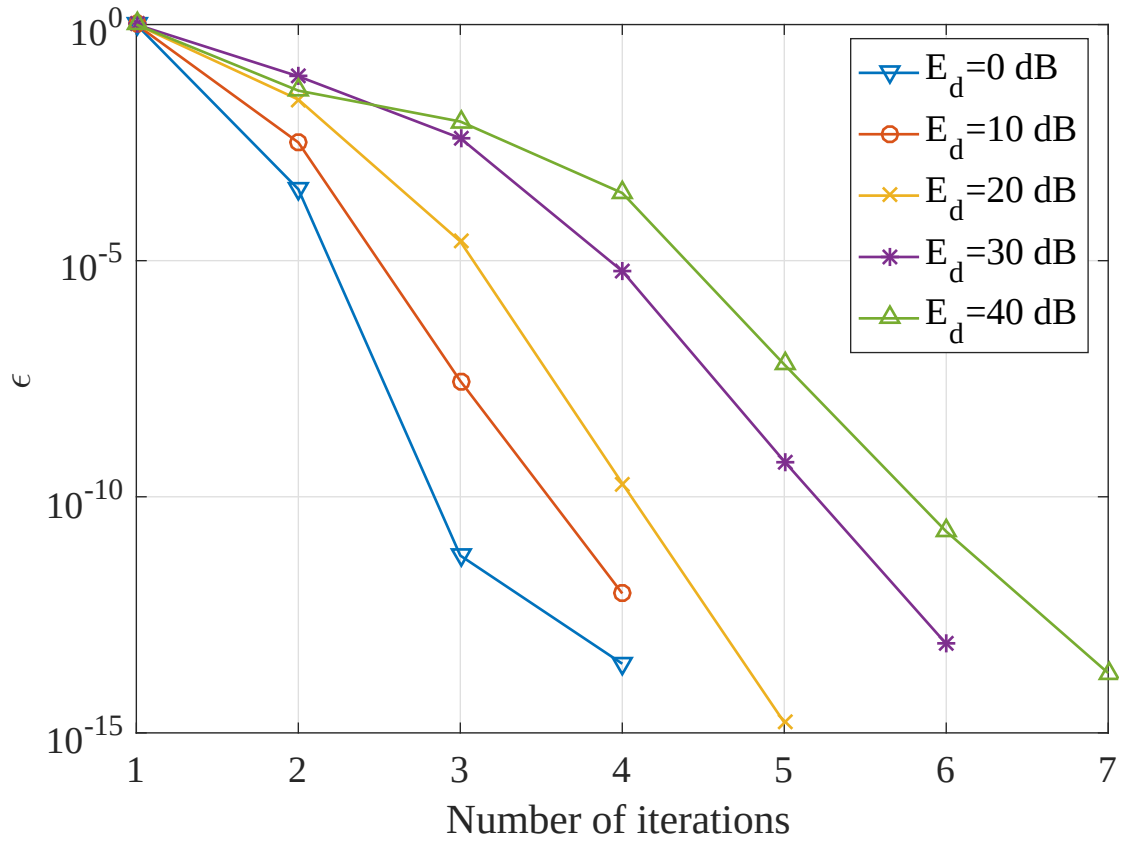


Figure 2.8: Convergence of the proposed SCA approach.

transmission power properly by applying the proposed SCA approach and then inform the source users to adjust their uplink data transmission power through the back-haul links (which are out of the scope of this chapter) to maximize the system throughput. If the RS does not consider power allocation, the system throughput may rapidly decrease at good channel scenarios, which is revealed by the blue dashed line in Fig. 7. This is an unacceptable performance loss in a practical system. Therefore, a proper power allocation is necessary for the FD relaying system. Fig. 2.8 shows the complexity and convergence of the proposed SCA approach. It is evident that the number of iterations is mostly below 5 if the relative error ϵ in Algorithm 1 is set to be 10^{-5}, which indicates the low complexity of the proposed approach. Moreover, the relative error ϵ decreases almost 5 orders if one more iterate is performed, which shows that the proposed SCA approach converges fast in solving the power allocation problem, which in turn leads low computational complexity in iterations. Finally, Fig. 2.8 also shows that the proposed SCA approach converges faster as the decreasing of total power allocated to data transmission, which implies that it performs with very low complexity and is a remarkable candidate for low power wireless communications.

2.8 Conclusions

Channel estimation overhead is a critical limitation on improving performance of the multiuser massive MIMO systems. To deal with this problem, this chapter has proposed a pilot-data overlay transmission scheme in both HD and FD one-way multipair massive MIMO relaying systems. The proposed scheme has exploited the orthogonality of source-relay and relay-destination channels with massive MIMO setups at the RS and redesigned the pilot and data transmission scheme to increase the effective data transmission period, which improves the achievable rate performance of the system in practice. The asymptotic analysis in both low and high SNRs with infinite number RS antennas has been conducted in this chapter and it has proven the superiority of the proposed scheme theoretically. Finally, a power allocation algorithm based on SCA has been proposed for the FD system, which further increases the achievable rate of the relaying system.

Chapter 3

Multipair Two-way Decode-and-Forward Relaying with Physical Layer Network Coding in Massive MIMO Systems

In the previous chapter, a pilot-data transmission overlay scheme in multipair one-way relaying systems has been proposed and analyzed. In contrast to the one-way relaying, the multipair two-way relaying (TWR) is a highly efficient scheme to realize the bi-directional information exchange. Different from the one-way relaying, the relay station (RS) employing TWR receives signals from all mobile stations (MSs) simultaneously and then broadcasts the forwarding signals to all MSs. To improve both the spectral and energy efficiency, this chapter investigates the efficient half-duplex decode-and-forward (DF) schemes in the multipair massive multiple-input multiple-output (MIMO) TWR system by exploiting the physical layer network coding (PNC) technique.

3.1 Introduction

MASSIVE multiple-input multiple-output (MIMO) has attracted significant interests from both academia and industry in the past decade, and is a promising technique for next-generation mobile communications, i.e., the 5G [12, 14, 15]. By employing the massive MIMO technique, multiple single-antenna mobile stations (MSs) can be served simultaneously by a base station (BS) equipped with massive antenna arrays using the same time-frequency resource and the system can achieve optimal spectral efficiency performance by employing less-complicated signal processing techniques, e.g., maximum-ratio combining/maximum-ratio transmission (MRC/MRT) [12, 42]. Massive

MIMO can also achieve high reliability and throughputs while the transmission power can be substantially reduced, which can be adopted to meet the higher demands of data rates and power efficiency in the future wireless communications [13, 16, 69]. Apart from the requirements of spectral and power efficiency, it also looks forward to satisfying the seamless coverage in 5G, which can be achieved by employing cell densification technologies, where the cooperative relaying is one of the most promising realization [9].

The relaying technology is an emerging cooperative technique which can significantly improve the throughputs, extend the coverage area and reduce the energy consumption of cellular communications [24]. Among all the realizations of relaying systems, the one-way and two-way relayings are the two most generic schemes in practice [33–37, 70, 71]. Both relaying schemes can significantly improve the system performance. However, the one-way relaying scheme has to take two more phases to transmit reverse-link signals in the bi-directional communication scenario which reduces the spectral efficiency of the entire system due to more channel uses. In contrast, the two-way relaying (TWR) scheme only adopts two communication phases to exchange information between MSs served by the relay station (RS), which significantly improves the spectral efficiency of the relaying system. In a typical decode-and-forwarding (DF) TWR system, all MSs transmit their signal to the RS in the multiple access (MA) phase. Then, the RS decodes the signal and broadcasts (BC) them back to MSs by applying the physical-layer network coding (PNC) technique [38]. The PNC is an apparatus similar to the conventional network coding (NC) while it adopts the proper modulation-and-demodulation technique at the RS to avoid the over-demodulation operations. Further, the adaptive channel-quantization PNC scheme with multiple antennas is proposed and analyzed by [39] to improve the performance of two-way relaying channels (TWRC). However, the reported analysis only considers the scenario of one-pair MSs.

By employing massive antenna arrays at RS, multipair MSs can exchange information within the pair through the shared RS [28]. However, the interpair interference (IPI) is introduced in such multipair relaying system. Nevertheless, the investigation in [26] shows that the IPI vanishes when the number of RS antennas tends infinity. The vanishing of IPI is due to the full channel state information (CSI) known by RS, which is obtained by channel estimations. Hence, the accurate and highly efficient channel estimation is one of the key enablers to realize the full potentials of the multipair massive MIMO TWR system. In previous literature, the perfect CSI is assumed to be known by the RS where the amplify-and-forward (AF) is adopted as the signal forwarding scheme in the multipair relaying [28, 58]. Few literature has reported the study of highly efficient channel estimation in the multipair massive MIMO TWR system with DF signal forwarding scheme.

In this chapter, we propose and investigate an efficient pilot transmission scheme,

namely, the sum decode-and-forward (SDF), in the multipair massive MIMO TWR system with DF signal forwarding technique at the RS which improves the spectral efficiency performance compared to conventional schemes [40, 72]. Recall the As comparison, a joint decode-and-forward (JDF) scheme applying to the multipair relaying is proposed and investigated. Different from the JDF, in the SDF, pilots from each pair of MSs are transmitted simultaneously to decrease the training durations which in turn extends the data transmission intervals. By employing the PNC technique at the RS, the broadcasted signal can be received and decoded by each MS correctly. In details, the main contributions of this chapter are summarized as follows:

- **Design the SDF and JDF schemes:** This chapter designs an efficient pilot transmission scheme, namely SDF, to decrease the resource consumption on pilot transmissions in the multipair massive MIMO TWR system. In the designed SDF, the inner-pair MSs transmit the same pilot to the RS and then the sum channel is estimated. With the estimated sum channel, the RS decodes the MA data in pairs, rather than in each user, and employs the PNC technique to perform signal detection and generates the BC signal for each MS pair. Then, the BC signal is precoded by the sum channel before transmission. Due to the asymptotic orthogonality of the massive MIMO channels, the IPI is well suppressed due to the precoding. For comparisons, the JDF scheme in the multipair massive MIMO relaying extended from the SISO system is also designed in this chapter.

- **Spectral efficiency analyses:** To analyze the performance of the proposed SDF and JDF schemes, the *closed-form expressions* of the approximate spectral efficiency of both schemes by employing the maximum ratio combining/maximum ratio transmission (MRC/MRT) processing are derived in this chapter. The derived expressions reveal that the PNC residual self-interference (RSI) is introduced in the detected signal with the SDF scheme which decreases the spectral efficiency performance. However, both theoretical analyses and numerical simulations show that the RSI can be rapidly reduced by deploying massive antenna arrays at the RS and the effective signal to interference and noise ratio (SINR) of the SDF can achieve that of the JDF scheme. Such that, the SDF outperforms the JDF in spectral efficiency due to longer data transmission durations. Further, the simulation results show that the derived *closed-form expressions* are quite accurate in characterizing spectral efficiency when large antenna arrays are deployed on the RS, thus providing useful guidelines in practical system designs.

- **Energy efficiency analyses:** Apart from the spectral efficiency, the energy efficiency of the proposed scheme is also analyzed in this chapter. From both theoretical anal-

yses and numerical evaluations, it is shown that the required transmission power can be rapidly decreased to achieve the desired spectral efficiency of each communication pair, e.g., about -25 dB data transmission power is required to achieve 1 bit/s/Hz with 1000 antennas deployed on the RS. Moreover, the simulation results also verify that the SDF scheme outperforms JDF in energy efficiency comparisons.

Organizations: The rest of this chapter is organized as follows. In Section 3.2, the system and signal transmission model of the multipair massive MIMO TWR system is illustrated. Section 3.3 proposes the JDF and SDF transmission schemes for the considered TWR system and the spectral efficiency of the proposed schemes are analyzed in Section 3.4. Section 3.5 conducts the asymptotic analyses of both spectral and energy efficiency with massive antenna arrays deployed on the RS. The numerical and simulation results are presented in Section 3.6. Finally, Section 3.7 concludes this chapter.

3.2 System Model

3.2.1 System and Channel Model

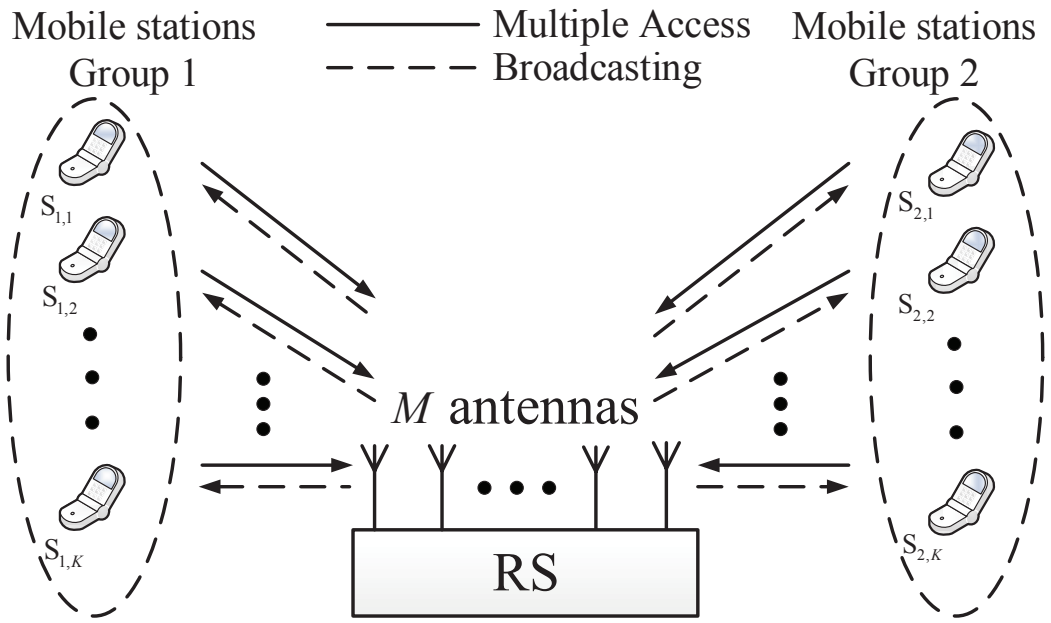


Figure 3.1: System diagram of multipair half-duplex DF two-way relaying.

As depicted in Fig. 3.1, this chapter considers the half-duplex multipair massive MIMO two-way relaying system, where K pairs of MSs ($\mathrm{S}_{1,k}$ and $\mathrm{S}_{2,k}$, $k = 1, 2, \cdots, K$) communicate to the corresponding pair mates with the help of a common RS by sharing

the same time-frequency resource, i.e., MS $\mathrm{S}_{1,k}$ wants to exchange information with the corresponding MS $\mathrm{S}_{2,k}$, via the RS. Each MS is equipped with a single antenna while the RS is equipped with M antennas operating in the time-division half-duplex mode. Without significant loss of generality, it is assumed that the total number of RS antennas is much greater than the total number of MSs, i.e., $M \gg 2K$. Throughout this chapter, the noise power is normalized to 1, and let ρ_{p}, ρ_{s} and ρd denote the pilot, source and RS transmission power, respectively. The channel matrix from MS group i ($i = 1, 2$) to the RS is denoted by $\mathbf{G}_i \in \mathbb{C}^{M\times K}$, where each column of the matrix, i.e., $\mathbf{g}_{i,k}$ for $k = 1, 2\cdots K$, stands for the channel vector from MS $\mathrm{S}_{i,k}$ to the RS. In this chapter, the Rayleigh fading channels are considered in the analyses. That is to say, the channel matrix $\mathbf{G}_i$ can be decomposed as $\mathbf{G}_i = \mathbf{H}_i\mathbf{D}_i^{1/2}$, where $\mathbf{H}_i$ denotes the small-scale fading constructed by independent identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ random variables (RVs), while $\mathbf{D}_i$ is formed by $\beta_{i,k}$ ($k = 1, 2, \cdots, K$) diagonally with the kth entry representing the large-scale fading of the channel between MS $\mathrm{S}_{i,k}$ and the RS, respectively. Further, it is assumed that the long-term parameters, such as the large-scale fading factors, MS numbers and pilot/data transmission power, are well known by the RS and all MSs[1].

3.2.2 Signal Transmission Model

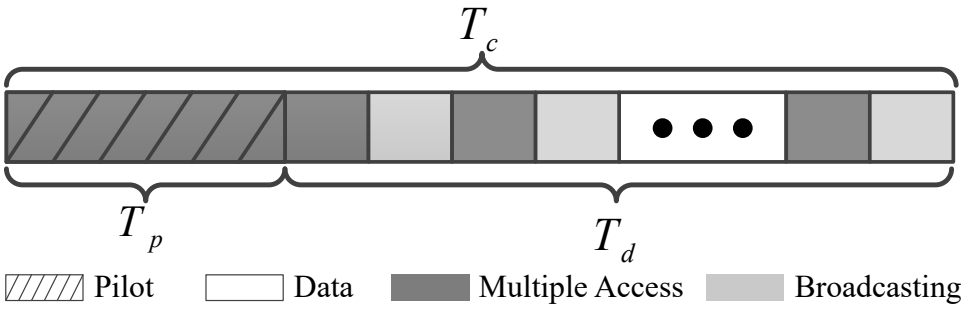


Figure 3.2: Signal transmission diagram of multipair half-duplex DF two-way relaying.

As shown by Fig. 3.2, there are two communication durations of the relaying system within a coherence interval T_{c}, namely, the pilot training duration T_{p} and data transmissions T_{d}. Obviously, there exists $T_{\mathrm{p}} + T_{\mathrm{d}} = T_{\mathrm{c}}$. During the pilot training period, the MSs transmit training pilots to the RS for channel estimations. Within the data transmission durations, all MSs first transmit data to the RS simultaneously in the MA phase. Then, the RS decodes the received MA data with the estimated CSI and performs preprocessing followed by broadcasting it to all MSs. The two data transmission phases, i.e.,

[1]This is a reasonable assumption in practice because the long-term parameters can be estimated by RS without sustainable resource consumption and exchanged with all MSs via control channels, which is out of the scope of this chapter.

MA and BC, are alternately repeated until the end of the coherence interval. Since the MA and BC are performed in different communication phases, the direct links between corresponding MSs are naturally ignored in the TWR systems.

In details, we present the signal transmission model mathematically in the following part of this subsection.

- *Pilot Training:* Within the training period, all MSs transmit the T_{p} orthogonal pilots to the RS simultaneously and the received signal at RS antennas can be expressed by

$$\mathbf{Y}_{\mathrm{p}} = \sqrt{\rho_{\mathrm{p}} T_{\mathrm{p}}}\mathbf{G}_1\mathbf{\Phi}_1 + \sqrt{\rho_{\mathrm{p}} T_{\mathrm{p}}}\mathbf{G}_2\mathbf{\Phi}_2 + \mathbf{N}_{\mathrm{p}}, \tag{3.1}$$

  where $\mathbf{\Phi}_i \in \mathbb{C}^{K\times T_{\mathrm{p}}}$ presents the pilot matrix whose kth row stands for the pilot sequence transmitted by $\mathrm{S}_{i,k}$, and the noise term $\mathbf{N}_{\mathrm{p}} \in \mathbb{C}^{M\times T_{\mathrm{p}}}$ is an additional white Gaussian noise (AWGN) matrix which is composed of i.i.d. $\mathcal{CN}(0,1)$ RVs.

- *Multiple Access:* During the MA phase, all $2K$ MSs $\mathrm{S}_{i,k}$ ($i = 1, 2$ and $k = 1, 2, \cdots, K$) transmit their own signals $s_{i,k}$, with the power ρ_{s}, to the RS. That is, the received signal at the RS can be expressed by

$$\mathbf{y}_{\mathrm{R}} = \sqrt{\rho_{\mathrm{s}}}\mathbf{G}_1\mathbf{s}_1 + \sqrt{\rho_{\mathrm{s}}}\mathbf{G}_2\mathbf{s}_2 + \mathbf{n}_{\mathrm{R}}, \tag{3.2}$$

  where the transmitted signal $\mathbf{s}_i \triangleq [s_{i,1}, s_{i,2}, \cdots, s_{i,K}]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ satisfy $\mathbb{E}\left[\mathbf{s}_i\mathbf{s}_j^{\mathrm{H}}\right] = \mathbf{I}_K$ if $i = j$, otherwise vanishes, and the AWGN term $\mathbf{n}_{\mathrm{R}} \in \mathbb{C}^{M\times 1}$ is constructed by i.i.d. $\mathcal{CN}(0,1)$ RVs.

- *Broadcasting:* Following the MA phase, the RS combines the received signal and decodes the data by employing the PNC technique. The decoded data is then broadcasted to all MSs after precoding with the estimated CSI. Hence, the received signal at each MS can be obtained as

$$r_{i,k} = \sqrt{\rho_{\mathrm{d}}}\mathbf{g}_{i,k}^{\mathrm{H}}\mathbf{F}\mathbf{x} + n_{i,k}, \tag{3.3}$$

  where $\mathbf{F} \in \mathbb{C}^{M\times K}$ denotes the precoding matrix, $\mathbf{x} \in \mathbb{C}^{K\times 1} \triangleq [x_1, x_2, \cdots, x_K]^{\mathrm{T}}$ represents the broadcasting signal transmitted by the RS, and $n_{i,k} \sim \mathcal{CN}(0,1)$ stands for the AWGN. Note that the precoding matrix follows $\mathbb{E}\{\|\mathbf{F}\mathbf{x}\|^2\} = 1$ to satisfy the transmission power constraint.

As shown in the signal transmission model described above, the pilot transmission is evidently the overhead of the relaying system. Hence, it can dramatically improve the average throughputs of the system with a shorter training period, i.e., decreasing T_{p},

even if the accuracy of channel estimations may be reduced. This is inspired by the structure of the channel capacity formula that the spectral efficiency increases linearly as the growing of data transmission duration T_{d} while deceases logarithmically as the degrading of the received SNR which can be affected by the inaccurate channel estimates. In the following section of this chapter, we investigate the TWR system shown above and propose an efficient pilot transmission scheme to decrease the training duration based on the asymptotic orthogonality of massive MIMO channels and employing the PNC manipulation for forwarding data processing.

3.3 Joint and Sum Decode and Forwarding

In this section, the joint decode-and-forward (JDF) scheme is proposed in the multi-pair massive MIMO two-way relaying system, which evolves from the conventional JDF scheme proposed in [40] for the single-pair single input single output (SISO) two-way relaying system. Further, a more efficient pilot training and data forwarding scheme, namely, the sum decode-and-forward (SDF), is first proposed in this chapter to diminish the pilot transmission duration which in turn increases the spectral efficiency performance of the multipair massive MIMO two-way relaying systems. In the following descriptions, the MRC/MRT processing is adopted in this chapter to combine the received MA signal and precoding the BC signal for transmission, respectively, which is viable and common in massive MIMO systems [12, 41].

3.3.1 Joint Decode-and-Forward (JDF) with Full CSI

In channel estimations of the JDF, all the pilot sequences are assumed to be orthogonal to avoid pilot contaminations, i.e., $\mathbf{\Phi}_i\mathbf{\Phi}_j^H = \mathbf{I}_K$ if $i = j$ while it vanishes if $i \neq j$. Evidently, $T_{\mathrm{p}} \geq 2K$ is required to hold the orthogonal conditions. Without loss of generalities, we select $T_{\mathrm{p}}^{\mathrm{JDF}} = 2K$ in the JDF scheme for ease of exposition. With the received pilot training signal expressed by (3.1) and following the MMSE criterion, the estimate of $\mathbf{G}_i$ $(i = 1, 2)$ can be respectively derived as

$$\hat{\mathbf{G}}_i^{\mathrm{JDF}} = \frac{1}{\sqrt{2\rho_{\mathrm{p}}K}}\mathbf{Y}_{\mathrm{p}}\mathbf{\Phi}_i^{\mathrm{H}}\tilde{\mathbf{D}}_i = \mathbf{G}_i\tilde{\mathbf{D}}_i + \frac{1}{\sqrt{2\rho_{\mathrm{p}}K}}\tilde{\mathbf{N}}_{\mathrm{p},i}\tilde{\mathbf{D}}_i, \tag{3.4}$$

where the MMSE channel estimation coefficient matrix $\tilde{\mathbf{D}}_i \triangleq \left(\mathbf{I}_K + \frac{1}{2\rho_{\mathrm{p}}K}\mathbf{D}_i^{-1}\right)^{-1}$ and the effective noise matrix $\tilde{\mathbf{N}}_{\mathrm{p},i} \triangleq \mathbf{N}_{\mathrm{p}}\mathbf{\Phi}_i^{\mathrm{H}}$. From the orthogonal properties of $\mathbf{\Phi}_i^{\mathrm{H}}$, the elements of $\tilde{\mathbf{N}}_{\mathrm{p},i}$ are also i.i.d. $\mathcal{CN}(0, 1)$ RVs. By recalling the property of the MMSE estimation, the channel matrix can be decomposed into two mutually independent com-

ponents expressed by

$$\mathbf{G}_i = \hat{\mathbf{G}}_i^{\mathrm{JDF}} + \mathbf{\Xi}_i^{\mathrm{JDF}}, \tag{3.5}$$

where $\mathbf{\Xi}_i^{\mathrm{JDF}}$ denotes the estimation error matrix and $i = 1, 2$. Considering each column of $\mathbf{G}_i$, namely, the channel between the RS and each MS, there is

$$\mathbf{g}_{i,k} = \hat{\mathbf{g}}_{i,k}^{\mathrm{JDF}} + \boldsymbol{\xi}_{i,k}^{\mathrm{JDF}}, \tag{3.6}$$

where $\hat{\mathbf{g}}_{i,k}^{\mathrm{JDF}}$ and $\boldsymbol{\xi}_{i,k}^{\mathrm{JDF}}$ denote the kth columns of $\hat{\mathbf{G}}_i^{\mathrm{JDF}}$ and $\mathbf{\Xi}_i^{\mathrm{JDF}}$, respectively, which follow $\hat{\mathbf{g}}_{i,k}^{\mathrm{JDF}} \sim \mathcal{CN}(\mathbf{0}, \sigma_{i,k}^2 \mathbf{I}_M)$ and $\boldsymbol{\xi}_{i,k}^{\mathrm{JDF}} \sim \mathcal{CN}(\mathbf{0}, \epsilon_{i,k}^2 \mathbf{I}_M)$, respectively. The covariance factors can be derived as $\sigma_{i,k}^2 = 2\rho_{\mathrm{p}} K \beta_{i,k}^2 / (1 + 2\rho_{\mathrm{p}} K \beta_{i,k})$ and $\epsilon_{i,k}^2 \triangleq \beta_{i,k} - \sigma_{i,k}^2$, respectively.

a) MRC Receiver

In the multipair JDF scheme, the received signal with stronger receiving strength is decoded in prior to the weaker one. Hence, the receiver starts from decoding the strongest signal, following which the decoded signal is canceled from the received signal, and then continues decoding the remaining signals sequentially as the same manner. To be simplified, here sort the $2K$ large-scale fading factors $\beta_{i,k}$ of the channels in the descending order as shown by $\beta_{m_1} \geq \beta_{m_2} \geq \cdots \geq \beta_{m_{2K}}$. The corresponding transmitted signal $s_{i,k}$, channel $\mathbf{g}_{i,k}$ and channel estimates $\hat{\mathbf{g}}_{i,k}$ are all arranged following the same order shown by $[s_{m_1}, s_{m_2}, \cdots, s_{m_{2K}}]$, $[\mathbf{g}_{m_1}, \mathbf{g}_{m_2}, \cdots, \mathbf{g}_{m_{2K}}]$ and $[\hat{\mathbf{g}}_{m_1}, \hat{\mathbf{g}}_{m_2}, \cdots, \hat{\mathbf{g}}_{m_{2K}}]$. To decode the m_kth ($m_k = 1, 2, \cdots, 2K$) data, the combined signal can be obtained by applying the MRC processing to the residual of the received signal $\mathbf{y}_{\mathrm{R}}$ after subtracting the decoded signal as

$$y_{m_k}^{\mathrm{JDF}} = \mathbf{w}_{m_k} \left(\mathbf{y}_{\mathrm{R}} - \sqrt{\rho_{\mathrm{s}}} \sum_{j=1}^{m_{k-1}} \hat{\mathbf{g}}_{m_j} \hat{s}_{m_j} \right), \tag{3.7}$$

where $\mathbf{y}_{\mathrm{R}}$ is expressed by (3.2) denoting the original received signal at RS, $\mathbf{w}_{m_k} \triangleq \hat{\mathbf{g}}_{m_k}^{\mathrm{H}}$ stands for the MRC combiner for MS S_{m_k}, and $\hat{s}_{m_j}$ represents the decoded signal of s_{m_j}. In the massive MIMO relaying system where the RS is deployed with large-scale antenna

arrays, the detection of s_{m_k} follows

$$\begin{aligned}\hat{s}_{m_k}^{\mathrm{JDF}} =& \frac{y_{m_k}^{\mathrm{JDF}}}{\sqrt{\rho_{\mathrm{s}}}\|\hat{\mathbf{g}}_{m_k}\|^2} \\ =& \frac{\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{m_k}+\boldsymbol{\xi}_{m_k}\right)s_{m_k}}{\|\hat{\mathbf{g}}_{m_k}\|^2} + \frac{\sum_{j=m_{k+1}}^{m_{2K}}\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_j}s_{m_j}}{\|\hat{\mathbf{g}}_{m_k}\|^2} \\ &+ \frac{\sum_{j=m_1}^{m_{k-1}}\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\left(\mathbf{g}_{m_j}s_{m_j}-\hat{\mathbf{g}}_{m_j}\hat{s}_{m_j}\right)}{\|\hat{\mathbf{g}}_{m_k}\|^2} + \frac{\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{n}_{\mathrm{R}}}{\sqrt{\rho_{\mathrm{s}}}\|\hat{\mathbf{g}}_{m_k}\|^2} \\ \overset{\text{a.s.}}{=}& s_{m_k},\end{aligned} \tag{3.8}$$

where the almost surely result can be obtained by applying the same manipulation presented in [43] when the number of RS antennas tends infinity and the fact that $\hat{\mathbf{g}}_{m_k}$ is mutually independent of the vectors $\mathcal{E}_{m_k}$, $\mathbf{g}_{m_j}$, $\hat{\mathbf{g}}_{m_j}$ $(j \neq k)$ and $\mathbf{n}_{\mathrm{R}}$. Hereafter, $\hat{s}_{m_k}^{\mathrm{JDF}}$ is assumed to be accurate and equal to s_{m_k} due to the massive MIMO configurations[2].

b) MRT Transmitter

After decoding the signals, the RS performs the PNC to the decoded two bit streams to generate the broadcasting signal, namely,$\mathbf{x} = \mathbf{s}_1 \oplus \mathbf{s}_2$, where $\oplus$ is the PNC operator applying to $\mathbf{s}_1$ and $\mathbf{s}_2$ element-wisely while the operation is not necessary the bit-wise XOR for each entry [38, 39]. Subsequently, the MRT precoding is applied to the broadcasting signal by multiplying it with the conjugate of the channel matrices and the precoding matrix is defined by

$$\mathbf{F}^{\mathrm{JDF}} \triangleq \alpha^{\mathrm{JDF}}\left(\hat{\mathbf{G}}_1 + \hat{\mathbf{G}}_2\right), \tag{3.9}$$

where α^{JDF} is the power normalization factor of the precoding matrix expressed as

$$\alpha^{\mathrm{JDF}} = \sqrt{\frac{1}{M\sum_{j=1}^{K}\left(\sigma_{1,j}^2+\sigma_{2,j}^2\right)}}. \tag{3.10}$$

By substituting (3.9) into (3.3), the received broadcasting signal at MS $\mathrm{S}_{i,k}$ can be obtained given by

$$r_{i,k}^{\mathrm{JDF}} = \alpha^{\mathrm{JDF}}\sqrt{\rho_{\mathrm{d}}}\left(\mathbf{g}_{i,k}^{\mathrm{H}}(\hat{\mathbf{g}}_{1,k}+\hat{\mathbf{g}}_{2,k})x_k + \sum_{j=1,j\neq k}^{2K}\mathbf{g}_{i,k}^{\mathrm{H}}(\hat{\mathbf{g}}_{1,j}+\hat{\mathbf{g}}_{2,j})x_j + n_{i,k}\right), \tag{3.11}$$

where x_k is the kth element of the transmission signal $\mathbf{x}$. With the same manner as (3.8), the broadcasting data x_k can be detected from $r_{i,k}^{\mathrm{JDF}}$ as $\hat{x}_k$, and the desired signal

[2]This assumption is reasonable in massive MIMO systems that the detection error vanishes when infinite RS antennas are deployed. When the number of RS antennas is large but finite, the small signal error can be suppressed by employing channel codings.

transmitted by $\mathrm{S}_{j,k}$ $(j \neq i)$ is received at $\mathrm{S}_{i,k}$ by applying the PNC again expressed by

$$\hat{s}_{j,k} = \hat{x}_k \oplus s_{i,k}, \tag{3.12}$$

where $s_{i,k}$ is the signal transmitted and thus perfectly known by $\mathrm{S}_{i,k}$.

3.3.2 Sum Decode-and-Forward (SDF) with Partial CSI

In this subsection, we propose a novel decode-and-forward scheme, namely, the sum decode-and-forward (SDF), with only partial CSI to be estimated, which can decrease the pilot training overhead for channel estimations and therefore improve the spectral efficiency performance of the multipair relaying system. As described in the previous parts of this section, the length of each pilot sequence must be equal to or greater than the total number of MSs, i.e., $T_{\mathrm{p}}^{\mathrm{JDF}} \geq 2K$, to satisfy the orthogonal condition and thus to avoid the pilot contaminations. Therefore, it consumes T_{p} time slots on pilot transmission, which suppresses the effectiveness of data transmissions, especially in short coherence interval occasions. Moreover, the MRT precoder performed at the RS is the summation of the two estimated channel matrices in the JDF scheme (refer to (3.9)), which implies that it is unnecessary to respectively estimate $\hat{\mathbf{G}}_1$ and $\hat{\mathbf{G}}_2$. Instead, it is sufficient to precode the forwarding data by estimating the sum of each paired channels. Inspired from this, the corresponding MS pairs can send the same pilots to RS, namely, setting $\mathbf{\Phi} \triangleq \mathbf{\Phi}_1 = \mathbf{\Phi}_2$ in (3.1). As such, it saves a half of pilot training periods for channel estimations compared to the JDF scheme, namely, $T_{\mathrm{p}}^{\mathrm{SDF}} \geq K$ is sufficient to hold the inter-pair pilot orthogonal conditions, i.e., $\mathbf{\Phi}\mathbf{\Phi}^{\mathrm{H}} = \mathbf{I}_K$. Without loss of generalities, it assumes $T_{\mathrm{p}}^{\mathrm{SDF}} = K$ throughout this chapter. Note that this chapter defines the sum channel matrix by $\mathbf{G} \triangleq \mathbf{G}_1 + \mathbf{G}_2$ for concise descriptions and its estimate is denoted by $\hat{\mathbf{G}}$.

By applying the MMSE criterion, the estimated channel of the SDF scheme can be derived from (3.1) expressed by

$$\hat{\mathbf{G}}^{\mathrm{SDF}} = \frac{1}{\sqrt{\rho_{\mathrm{p}} K}} \mathbf{Y}_{\mathrm{p}} \mathbf{\Phi}^{\mathrm{H}} \tilde{\mathbf{D}} = \mathbf{G}\tilde{\mathbf{D}} + \frac{1}{\sqrt{\rho_{\mathrm{p}} K}} \tilde{\mathbf{N}}_{\mathrm{p}} \tilde{\mathbf{D}}, \tag{3.13}$$

where the MMSE coefficient matrix is defined by $\tilde{\mathbf{D}} \triangleq \left(\mathbf{I}_K + \frac{1}{\rho_{\mathrm{p}} K}\mathbf{D}^{-1}\right)^{-1}$ and the effective noise term $\tilde{\mathbf{N}}_{\mathrm{p}} \triangleq \mathbf{N}_{\mathrm{p}}\mathbf{\Phi}^{\mathrm{H}}$ whose entries are i.i.d. $\mathcal{CN}(0,1)$ RVs. Similar to the channel decomposition in the JDF channel estimations, the channel estimates of the SDF scheme follow

$$\mathbf{G} = \hat{\mathbf{G}}^{\mathrm{SDF}} + \mathbf{\Xi}^{\mathrm{SDF}}, \tag{3.14}$$

where $\mathbf{\Xi}^{\mathrm{SDF}}$ denotes the error matrix of MMSE estimations. Regarding the channel esti-

mation of each MS pair, say the kth one, there exists

$$\mathbf{g}_k = \hat{\mathbf{g}}_k^{\mathrm{SDF}} + \boldsymbol{\varepsilon}_k^{\mathrm{SDF}}, \tag{3.15}$$

where the kth columns of $\hat{\mathbf{G}}^{\mathrm{SDF}}$ and $\boldsymbol{\Xi}^{\mathrm{SDF}}$ follow $\hat{\mathbf{g}}_k^{\mathrm{SDF}} \sim \mathcal{CN}(\mathbf{0}, \varsigma_k^2 \mathbf{I}_M)$ and $\boldsymbol{\xi}_k^{\mathrm{SDF}} \sim \mathcal{CN}(\mathbf{0}, \varepsilon_k^2 \mathbf{I}_M)$ with the covariance factors derived by $\varsigma_k^2 = \rho_{\mathrm{p}} K \beta_k^2 / (1 + \rho_{\mathrm{p}} K \beta_k)$ and $\varepsilon_k^2 \triangleq \beta_k - \varsigma_k^2$, respectively. Note that $\beta_k \triangleq \beta_{1,k} + \beta_{2,k}$ for simplicity.

Moreover, the sum channel estimate can also be used as the combiner to perform MRC on the received signal at the RS, where the PNC is employed to obtain the forwarding signal for broadcasting. In the following parts of this subsection, the signal processing details with the SDF scheme are illustrated where the MRC/MRT is employed as the combiner and precoder at the RS.

a) MRC Receiver

Here first consider the signal detection of the kth MS pair. At the RS, the MRC processing is performed to the received signal by multiplying it with the linear combiner matrix and the combined signal can be expressed as

$$y_k^{\mathrm{SDF}} = \mathbf{w}_k \mathbf{y}_{\mathrm{R}}, \tag{3.16}$$

where $\mathbf{w}_k \triangleq \hat{\mathbf{g}}_k^{\mathrm{H}}$ denotes the MRC combiner for MS pair k and $\hat{\mathbf{g}}_k$ is the kth column of $\hat{\mathbf{G}}$ standing for the estimate of the sum channel between RS and $\mathrm{S}_{i,k}$ ($i = 1, 2$). By substituting (3.2) and the MRC combiner into (3.16), the combined signal can be derived as follows:

$$y_k^{\mathrm{SDF}} = \underbrace{\sqrt{\rho_{\mathrm{s}}} \hat{\mathbf{g}}_k^{\mathrm{H}} \left(\mathbf{g}_{1,k} s_{1,k} + \mathbf{g}_{2,k} s_{2,k} \right)}_{\text{desired signal } \check{y}_k^{\mathrm{SDF}}} + \underbrace{\sqrt{\rho_{\mathrm{s}}} \sum_{j=1, j \neq k}^{K} \hat{\mathbf{g}}_k^{\mathrm{H}} \left(\mathbf{g}_{1,j} s_{1,j} + \mathbf{g}_{2,j} s_{2,j} \right)}_{\text{inter-pair interference}} + \underbrace{\hat{\mathbf{g}}_k^{\mathrm{H}} \mathbf{n}_{\mathrm{R}}}_{\text{noise}}, \tag{3.17}$$

where the desired signal term can be reformulated as

$$\check{y}_k^{\mathrm{SDF}} = \begin{cases} \sqrt{\rho_{\mathrm{s}}} \hat{\mathbf{g}}_k^{\mathrm{H}} \mathbf{g}_{1,k} \left(s_{1,k} + s_{2,k} \right) + \sqrt{\rho_{\mathrm{s}}} \hat{\mathbf{g}}_k^{\mathrm{H}} \left(\mathbf{g}_{2,k} - \mathbf{g}_{1,k} \right) s_{2,k} & \beta_{1,k} \geq \beta_{2,k} \\ \sqrt{\rho_{\mathrm{s}}} \hat{\mathbf{g}}_k^{\mathrm{H}} \mathbf{g}_{2,k} \left(s_{1,k} + s_{2,k} \right) + \sqrt{\rho_{\mathrm{s}}} \hat{\mathbf{g}}_k^{\mathrm{H}} \left(\mathbf{g}_{1,k} - \mathbf{g}_{2,k} \right) s_{1,k} & \beta_{1,k} < \beta_{2,k}. \end{cases} \tag{3.18}$$

In the signal detections, the RS detects the sum signal $s_{1,k} + s_{2,k}$ to obtain the BC signal following $x_k \triangleq s_{1,k} \oplus s_{2,k}$. Here the PNC technique is performed to $s_{1,k} + s_{2,k}$ following the illustration in [39] where the decoding can be treated as the case I with $L = 1$.

Remark 3.1. Note that the inter-pair interference in (3.17) vanishes as the increment of RS antennas, i.e., M, which can be proved following the same manipulation as (3.8). However, the desired signal term expressed by (3.18) contains both $s_{1,k} + s_{2,k}$ and the following effective noise term introduced by the large-scale fading differences between each MS pair, which results in the degradation of the received signal to interference and noise ratio (SINR) of the SDF scheme compared to the JDF one. Nevertheless, the data transmission phase increases due to the more efficient pilot trainings. As such, the spectral efficiency of the SDF scheme is still improved in the limited coherence interval scenario. The detailed spectral efficiency is derived and analyzed in Section 3.4.

b) MRT Transmitter

During the BC phase in the SDF scheme, MRT precoding can be applied to the BC signal by multiplying it with the conjugate of the estimated sum channel matrices and the precoding matrix can be derived as

$$\mathbf{F}^{\mathrm{SDF}} = \alpha^{\mathrm{SDF}} \hat{\mathbf{G}}, \tag{3.19}$$

where α^{SDF} is used to normalize the power of precoder, namely, $\mathbb{E}\left\{\|\mathbf{F}\mathbf{x}\|^2\right\} = 1$, defined by

$$\alpha^{\mathrm{SDF}} = \sqrt{\frac{1}{M\sum_{k=1}^{K}\sigma_k^2}}, \tag{3.20}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_K]^{\mathrm{T}}$ is constructed by the generated transmission signal of each pair of MSs. By substituting (3.19) into (3.3), the received signal of the MS $\mathrm{S}_{i,k}$ can be obtained given by

$$r_{i,k}^{\mathrm{SDF}} = \underbrace{\alpha^{\mathrm{SDF}}\sqrt{\rho_{\mathrm{d}}}\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k x_k}_{\text{desired signal}} + \underbrace{\alpha^{\mathrm{SDF}}\sqrt{\rho_{\mathrm{d}}}\sum_{j=1,j\neq k}^{K}\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_j x_j}_{\text{inter-pair interference}} + \underbrace{n_{i,k}}_{\text{noise}}, \tag{3.21}$$

where $n_{i,k}$ denotes the AWGN at MS $\mathrm{S}_{i,k}$.

At MS $\mathrm{S}_{i,k}$, the transmitted BC signal $\hat{x}_{i,k}$ can be detected from $r_{i,k}^{\mathrm{SDF}}$ and the desired signal transmitted by the inner-pair MS can be decoded by re-applying the PNC processing with the same manipulation as in the JDF scheme.

3.4 Spectral Efficiency Analysis

This section analyzes the spectral efficiency of the two-way relaying system by calculating the achievable rates of the end-to-end (e2e) communications with both the JDF

and SDF relaying schemes proposed above. As presented in [40], here first define the **two-way rate** as follows:

Definition 3.1. During the time of N symbols, let $\mathrm{S}_{1,k}$ reliably receives $\mathcal{R}_{21,k}$ bits from $\mathrm{S}_{2,k}$ and $\mathrm{S}_{2,k}$ reliably receives $\mathcal{R}_{12,k}$ bits from $\mathrm{S}_{1,k}$. Define the one-pair two-way rate of the relaying system by

$$\mathcal{R}_k = \frac{\mathcal{R}_{21,k} + \mathcal{R}_{12,k}}{N}. \tag{3.22}$$

Furthermore, the two-way sum rates of the whole relaying system are defined by

$$\mathcal{R} = \sum_{k=1}^{K} \mathcal{R}_k. \tag{3.23}$$

Since the two-way relaying data transmission between any MS pair involves two phases, the MA and BC via the relay, the e2e achievable rate depends on the rates of both links. Let $\mathcal{R}_{1\mathrm{R},k}$ and $\mathcal{R}_{2\mathrm{R},k}$ denote the achievable rates of the MA phase $\mathrm{S}_{1,k} \to \mathrm{RS}$ and $\mathrm{S}_{2,k} \to \mathrm{RS}$, respectively. And the achievable rates of the BC phase $\mathrm{RS} \to \mathrm{S}_{1,k}$ and $\mathrm{RS} \to \mathrm{S}_{2,k}$ are similarly denoted by $\mathcal{R}_{\mathrm{R}1,k}$ and $\mathcal{R}_{\mathrm{R}2,k}$, respectively. Based on the property of the DF schemes, the e2e achievable rate between the kth MS pair is limited by the weaker connection of the MA and BC links, namely,

$$\mathcal{R}_{12,k} = \min\{\mathcal{R}_{1\mathrm{R},k}, \mathcal{R}_{\mathrm{R}2,k}\}, \tag{3.24}$$

$$\mathcal{R}_{21,k} = \min\{\mathcal{R}_{2\mathrm{R},k}, \mathcal{R}_{\mathrm{R}1,k}\}. \tag{3.25}$$

Therefore, the average sum achievable rate of the two-way relaying is obtained by adding up the corresponding rates of all MS pairs and divided by the entire communication time, expressed by

$$\mathcal{R} = \frac{T_c - T_p}{2T_c} \sum_{k=1}^{K} \left(R_{12,k} + R_{21,k}\right), \tag{3.26}$$

where T_c and T_p denote the coherence interval in symbols and the number of pilot symbols, respectively. Note that there are in total $(T_c - T_p)/2$ two-way communications (MA and BC) within the coherence interval (refer to Fig. 3.2).

3.4.1 Achievable Rate of the JDF Scheme

a) MA Achievable Rate

In this subsection and the follows, the technique proposed in [73] is adopted to compute the achievable rate which is widely utilized in massive MIMO systems and yields a simple and insightful closed-form expression of achievable rates with the statistical CSI

[22, 30, 74, 75]. By considering (3.7), the received signal of the MA phase used for detections can be rewritten as follows:

$$y_{m_k}^{\mathrm{JDF}} = \underbrace{\sqrt{\rho_{\mathrm{s}}}\mathbb{E}\left\{\hat{\mathbf{g}}_{m_k}^{H}\mathbf{g}_{m_k}\right\}s_{m_k}}_{\text{desired signal}} + \underbrace{\tilde{n}_{m_k}^{\mathrm{JDF}}}_{\text{effective noise}}, \tag{3.27}$$

where $\tilde{n}_{m_k}^{\mathrm{JDF}}$ is the effective noise term given by

$$\begin{aligned}\tilde{n}_{m_k}^{\mathrm{JDF}} =& \sqrt{\rho_{\mathrm{s}}}\left(\hat{\mathbf{g}}_{m_k}^{H}\mathbf{g}_{m_k} - \mathbb{E}\left\{\hat{\mathbf{g}}_{m_k}^{H}\mathbf{g}_{m_k}\right\}\right)s_{m_k} + \sqrt{\rho_{\mathrm{s}}}\sum_{j=m_1}^{m_{k-1}}\hat{\mathbf{g}}_{m_k}^{H}\left(\mathbf{g}_{m_j} - \hat{\mathbf{g}}_{m_j}\right)s_{m_j} \\ &+ \sqrt{\rho_{\mathrm{s}}}\sum_{j=m_{k+1}}^{m_{2K}}\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_j}s_{m_j} + \hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{n}_{\mathrm{R}}.\end{aligned} \tag{3.28}$$

Note that the subscript index $(\cdot)_{\mathrm{J},m_k}$ follows the mapping definition in (3.7) which is determined by the order of $\beta_{i,k}$.

From (3.27) and (3.28), it is evident that the desired signal and the effective noise term are uncorrelated. Therefore, the effective signal to interference and noise ratio (SINR) can be obtained by

$$\gamma_{m_k}^{\mathrm{JDF}} = \frac{\left|\mathbb{E}\left\{\hat{\mathbf{g}}_{m_k}^{H}\mathbf{g}_{m_k}\right\}\right|^2}{\mathbb{V}\mathrm{ar}\left\{\hat{\mathbf{g}}_{m_k}^{H}\mathbf{g}_{m_k}\right\} + \mathrm{SRI}_{m_k}^{\mathrm{JDF}} + \mathrm{IPI}_{m_k}^{\mathrm{JDF}} + \mathrm{AN}_{m_k}^{\mathrm{JDF}}}, \tag{3.29}$$

where $\mathrm{SRI}_{i,k}^{\mathrm{JDF}}$, $\mathrm{IPI}_{i,k}^{\mathrm{JDF}}$ and $\mathrm{AN}_{i,k}^{\mathrm{JDF}}$ denote the subtraction residual interference, interpair interference and additive noise, respectively, defined by

$$\mathrm{SRI}_{m_k}^{\mathrm{JDF}} \triangleq \sum_{j=m_1}^{m_{k-1}}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\left(\mathbf{g}_{m_j} - \hat{\mathbf{g}}_{m_j}\right)\right|^2\right\}, \tag{3.30a}$$

$$\mathrm{IPI}_{m_k}^{\mathrm{JDF}} \triangleq \sum_{j=m_k+1}^{m_{2K}}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_j}\right|^2\right\}, \tag{3.30b}$$

$$\mathrm{AN}_{m_k}^{\mathrm{JDF}} \triangleq \frac{1}{\rho_{\mathrm{s}}}\mathbb{E}\left[\left|\hat{\mathbf{g}}_{m_k}^{H}\mathbf{n}_{\mathrm{R}}\right|^2\right]. \tag{3.30c}$$

By using the fact that the worst-case uncorrelated additive noise is the independent Gaussian noise with the same variance as the effective SINR γ [73], the lower bound of the achievable rate $C(\cdot)$ can be defined by

$$C(\gamma) = \log_2(1+\gamma)\ [\text{bits/s/Hz}]. \tag{3.31}$$

Hence, the closed-form expression for the uplink rates of the JDF scheme are provided as follows:

Theorem 3.1. *By employing the MRC combining with the JDF scheme, the achievable rate lower bound of the MA communication phase* $\mathrm{S}_{i,k} \rightarrow \mathrm{RS}$ *(*$i = 1, 2$ *and* $k = 1, 2, \cdots, K$*), with a finite number of antennas deployed on RS, is given by*

$$\mathcal{R}_{i\mathrm{R},k}^{\mathrm{JDF}} = C\left(\gamma_{i\mathrm{R},k}^{\mathrm{JDF}}\right), \tag{3.32}$$

where $\gamma_{i\mathrm{R},k}^{\mathrm{JDF}}$ *is inversely mapped from the sequence of* $\{\gamma_{m_k}^{\mathrm{JDF}}\}$ *according to the order of* $\{\beta_{m_k}\}$ *(refer to Section a) for the mapping details) and*

$$\gamma_{m_k} = \frac{M\sigma_{m_k}^2}{\sum_{j=1}^{2K}\beta_{m_j} - \sum_{j=1}^{k-1}\sigma_{m_j}^2 + 1/\rho_{\mathrm{p}}}.$$

Proof: See Appendix 3.8.1. ■

b) BC Achievable Rate

Following the similar methodology as the derivations of the MA phase, the achievable rate of the BC phase $\mathrm{RS} \rightarrow \mathrm{S}_{i,k}$ ($i = 1, 2$ and $k = 1, 2, \cdots, K$) can be obtained by rewriting (3.11) into the following two-term form:

$$r_{i,k}^{\mathrm{JDF}} = \underbrace{\alpha^{\mathrm{JDF}}\mathbb{E}\left[\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,k} + \hat{\mathbf{g}}_{2,k}\right)\right]x_k}_{\text{desired signal}} + \underbrace{\tilde{n}_{i,k}^{\mathrm{JDF}}}_{\text{effective noise}}, \tag{3.33}$$

where $\tilde{n}_{i,k}^{\mathrm{JDF}}$ is the effective noise term defined by

$$\begin{aligned}\tilde{n}_{i,k}^{\mathrm{JDF}} =& \alpha^{\mathrm{JDF}}\left(\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,k} + \hat{\mathbf{g}}_{2,k}\right) - \mathbb{E}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,k} + \hat{\mathbf{g}}_{2,k}\right)\right\}\right)x_k \\ &+ \alpha^{\mathrm{JDF}}\sum_{j=1, j\neq k}^{K}\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,j} + \hat{\mathbf{g}}_{2,j}\right)x_j + n_{i,k}.\end{aligned} \tag{3.34}$$

From (3.33) and (3.34), it is evident that the effective SINR can be formulated as follows:

$$\gamma_{i,k}^{\mathrm{JDF}} = \frac{\left|\mathbb{E}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,k} + \hat{\mathbf{g}}_{2,k}\right)\right\}\right|^2}{\mathbb{V}\mathrm{ar}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,k} + \hat{\mathbf{g}}_{2,k}\right)\right\} + \mathrm{IPI}_{i,k}^{\mathrm{JDF}} + \mathrm{AN}_{i,k}^{\mathrm{JDF}}}, \tag{3.35}$$

where $\mathrm{IPI}_{i,k}^{\mathrm{JDF}}$ and $\mathrm{AN}_{i,k}^{\mathrm{JDF}}$ denote the interpair interference and additive noise, respectively defined by

$$\mathrm{IPI}_{i,k}^{\mathrm{JDF}} \triangleq \sum_{j=1, j\neq k}^{K}\mathbb{E}\left\{\left|\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,j} + \hat{\mathbf{g}}_{2,j}\right)\right|^2\right\}, \tag{3.36}$$

$$\mathrm{AN}_{i,k}^{\mathrm{JDF}} \triangleq \frac{1}{(\alpha^{\mathrm{JDF}})^2}\mathbb{E}\left\{|n_{i,k}|^2\right\}. \tag{3.37}$$

By substituting the SINR expressed by (3.35) into the achievable rate expression defined by (3.31) and applying computations following the theory of the stochastic process, the following closed-form expression for the BC achievable rate of the JDF scheme can be derived.

Theorem 3.2. *By employing the MRT precoding with the JDF scheme, the achievable rate lower bound of the BC communication phase* $\mathrm{RS} \rightarrow \mathrm{S}_{i,k}$ *(*$i = 1, 2$ *and* $k = 1, 2, \cdots, K$*), with a finite number of antennas deployed on RS, is given by*

$$\mathcal{R}_{\mathrm{R}i,k}^{\mathrm{JDF}} = C\left(\gamma_{\mathrm{R}i,k}^{\mathrm{JDF}}\right), \tag{3.38}$$

where

$$\gamma_{\mathrm{R}i,k}^{\mathrm{JDF}} = \frac{M\sigma_{i,k}^4}{\left(\beta_{i,k} + 1/\rho_{\mathrm{d}}\right)\sum_{j=1}^{K}\left(\sigma_{1,j}^2 + \sigma_{2,j}^2\right)}.$$

Proof: The results can be directly obtained by applying the same manipulation employed in Theorem 3.1. ■

3.4.2 Achievable Rate of the SDF Scheme

a) MA Achievable Rate

Following the methodology taken by the JDF scheme, this subsection employs the same technique to derive the achievable rate of the SDF scheme by utilizing the statistical CSI. By substituting (3.18) into (3.17), the received signal used for detection can be rewritten as follows:

$$y_k^{\mathrm{SDF}} = \underbrace{\sqrt{\rho_{\mathrm{s}}}\mathbb{E}\left\{\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k}\right\}\left(s_{1,k} + s_{2,k}\right)}_{\text{desired signal}} + \underbrace{\tilde{n}_k^{\mathrm{SDF}}}_{\text{effective noise}}, \tag{3.39}$$

where $i = 1$ if $\beta_{1,k} \geq \beta_{2,k}$, otherwise $i = 2$, and $\tilde{n}_{i,k}^{\mathrm{SDF}}$ is the effective noise term given by

$$\begin{aligned}\tilde{n}_k^{\mathrm{SDF}} =& \sqrt{\rho_{\mathrm{s}}}\left(\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k} - \mathbb{E}\left\{\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k}\right\}\right)\left(s_{1,k} + s_{2,k}\right) + \sqrt{\rho_{\mathrm{s}}}\hat{\mathbf{g}}_k^{H}\left(\mathbf{g}_{i',k} - \mathbf{g}_{i,k}\right)s_{i',k} \\ &+ \sqrt{\rho_{\mathrm{s}}}\sum_{j=1,j\neq k}^{K}\hat{\mathbf{g}}_k^{\mathrm{H}}\left(\mathbf{g}_{1,j}s_{1,j} + \mathbf{g}_{2,j}s_{2,j}\right) + \hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{n}_{\mathrm{R}},\end{aligned} \tag{3.40}$$

where $i' = 1$ if $i = 2$, otherwise $i' = 2$. The effective noise denoted by (3.40) is obviously uncorrelated with the desired signal represented by (3.39) such that the effective SINR can be obtained as

$$\gamma_k^{\mathrm{SDF}} = \frac{2\left|\mathbb{E}\left\{\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k}\right\}\right|^2}{2\mathbb{V}\mathrm{ar}\left\{\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k}\right\} + \mathrm{RSI}_{i,k}^{\mathrm{SDF}} + \mathrm{IPI}_{i,k}^{\mathrm{SDF}} + \mathrm{AN}_{i,k}^{\mathrm{SDF}}}, \tag{3.41}$$

where $\mathrm{RSI}_{i,k}^{\mathrm{SDF}}$, $\mathrm{IPI}_{i,k}^{\mathrm{SDF}}$ and $\mathrm{AN}_{i,k}^{\mathrm{SDF}}$ denote the PNC residual self-interference, inter-pair interference and additive noise, respectively defined by

$$\mathrm{RSI}_{i,k}^{\mathrm{SDF}} \triangleq \mathbb{E}\left\{\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\left(\mathbf{g}_{i',k}-\mathbf{g}_{i,k}\right)\right|^2\right\}, \tag{3.42}$$

$$\mathrm{IPI}_{i,k}^{\mathrm{SDF}} \triangleq \sum_{j=1,j\neq k}^{K}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{1,j}\right|^2+\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{2,j}\right|^2\right\}, \tag{3.43}$$

$$\mathrm{AN}_{i,k}^{\mathrm{SDF}} \triangleq \frac{1}{\rho_{\mathrm{s}}}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{n}_{\mathrm{R}}\right|^2\right\}. \tag{3.44}$$

With the obtained SINR, the closed-form expression of the MA achievable rate for the SDF scheme can be obtained as the following Theorem 3.3.

Theorem 3.3. *By employing the MRC combing with the SDF scheme, the lower bound of the achievable rate of the MA communication phase* $\mathrm{S}_{i,k}\rightarrow\mathrm{RS}$*, with a finite number of antennas deployed on the RS, is given by*

$$\mathcal{R}_{1\mathrm{R},k}^{\mathrm{SDF}}=\mathcal{R}_{2\mathrm{R},k}^{\mathrm{SDF}}=C\left(\gamma_k^{\mathrm{SDF}}\right), \tag{3.45}$$

where

$$\gamma_k^{\mathrm{SDF}}=\frac{2M\varsigma_{i,k}^4}{2\beta_{i,k}\varsigma_k^2+M\left(\varsigma_{1,k}^2-\varsigma_{2,k}^2\right)^2+\varsigma_k^2\sum_{j=1}^{K}\beta_j+\varsigma_k^2/\rho_{\mathrm{s}}},$$

$\varsigma_{1,k}^2\triangleq\frac{\rho_{\mathrm{p}}K\beta_k\beta_{1,k}}{\rho_{\mathrm{p}}K\beta_k+1}$ *and* $\varsigma_{2,k}^2\triangleq\frac{\rho_{\mathrm{p}}K\beta_k\beta_{2,k}}{\rho_{\mathrm{p}}K\beta_k+1}$*. In the equation, if* $\beta_{1,k}\geq\beta_{2,k}$*,* $i=1$*, otherwise* $i=2$*.*

Proof: See Appendix 3.8.2. ■

b) BC Achievable Rate

Following the similar methodology as the derivation of the MA phase, the achievable rate of the downlink phase $\mathrm{RS}\rightarrow\mathrm{S}_k$ can be obtained by rewriting (3.21) into the following two-term form:

$$r_{i,k}^{\mathrm{SDF}}=\underbrace{\alpha^{\mathrm{SDF}}\mathbb{E}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k\right\}x_k}_{\text{desired signal}}+\underbrace{\tilde{n}_{i,k}^{\mathrm{SDF}}}_{\text{effective noise}}, \tag{3.46}$$

where $\tilde{n}_{i,k}^{\mathrm{SDF}}$ is the effective noise term defined by

$$\tilde{n}_{i,k}^{\mathrm{SDF}}=\alpha^{\mathrm{SDF}}\left(\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k-\mathbb{E}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k\right\}\right)x_k+\alpha^{\mathrm{SDF}}\sum_{j=1,j\neq k}^{K}\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_j x_j+n_{i,k}^{\mathrm{SDF}}. \tag{3.47}$$

From (3.46) and (3.47), the effective SINR can be derived as

$$\gamma_{i,k}^{\mathrm{SDF}} = \frac{\left|\mathbb{E}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k\right\}\right|^2}{\mathbb{V}\mathrm{ar}\left\{\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k\right] + \sum_{j=1,j\neq k}^{K}\mathbb{E}\left\{\left|\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_j\right|^2\right\} + 1/(\alpha^{\mathrm{SDF}})^2}. \tag{3.48}$$

By substituting the SINR given by (3.48) into the achievable rate expression (3.31) and applying the ensemble average computations, the closed-form expression for the BC achievable rate can be obtained as the following Theorem.

Theorem 3.4. *For the MRT precoding in the SDF scheme and with M antennas deployed on the RS, the lower bound of the achievable rate of the BC communication phase* $\mathrm{RS} \rightarrow \mathrm{S}_{i,k}$ *is given by*

$$\mathcal{R}_{\mathrm{R}i,k}^{\mathrm{SDF}} = C\left(\gamma_{\mathrm{R}i,k}^{\mathrm{SDF}}\right), \tag{3.49}$$

where

$$\gamma_{\mathrm{R}i,k}^{\mathrm{SDF}} = \frac{M\varsigma_{i,k}^4}{(\beta_{i,k}+1/\rho_{\mathrm{d}})\sum_{j=1}^{K}\varsigma_j^2}.$$

In above equations, both ς_j^2 and $\varsigma_{i,k}^2$ are defined in Theorem 3.3.

Proof: By applying the similar manipulation as Theorem 3.3, the results can be directly derived. ■

Finally, by substituting (3.45) and (3.49) into (3.26), the lower bounds of the sum achievable rates of the two-way relaying with the SDF scheme can be obtained.

Remark 3.2. The MA achievable rates of the SDF scheme in (3.45) are obtained by approximating the effective noise as the uncorrelated additive white Gaussian noise which was proved to be the worst case in [73]. Moreover, the PNC residual self-interference term also contains the signal information while it is treated as the noise in the achievable rate analysis. Therefore, the achievable uplink rate given in Theorem 3.3 is a lower bound of the considered relaying system.

Remark 3.3. It is worth to mention that the BC achievable rate in (3.49) is an exact value which is different from the MA achievable rate in (3.45). That is to say, all the useful information is combined into the desired signal term for decoding as expressed in (3.39) while (3.46) shows that the effective noise term also consists of useful information.

3.4.3 Accuracy Analysis

In this section, the accuracy of the achievable rates predictor obtained with only the statistical knowledge of the CSI is analyzed. To examine the accuracy of the achievable rates predictor derived above, here present the achievable rates formula expressed by

the instantaneous channel matrices $\mathbf{g}_{1,k}$ and $\mathbf{g}_{2,k}$. With instantaneous CSI, the ergodic achievable rates for MA and BC communications are given by

$$\tilde{\mathcal{R}}_{i\mathrm{R},k} = \mathbb{E}\left\{C(\tilde{\gamma}_{i\mathrm{R},k})\right\}, \tag{3.50}$$

$$\tilde{\mathcal{R}}_{\mathrm{R}i,k} = \mathbb{E}\left\{C(\tilde{\gamma}_{\mathrm{R}i,k})\right\}, \tag{3.51}$$

where $\tilde{\mathcal{R}}_{i\mathrm{R},k}$ ($i = 1, 2$ and $k = 1, 2, \cdots, K$) denotes the MA achievable rate from $\mathrm{S}_{i,k}$ to the RS while on the contrary $\tilde{\mathcal{R}}_{\mathrm{R}i,k}$ represents the BC achievable rate from the RS to $\mathrm{S}_{i,k}$. For the JDF scheme, the instaneous effective SINRs are expressed by

$$\tilde{\gamma}_{m_k}^{\mathrm{JDF}} = \frac{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_k}\right|^2}{\sum_{j=1}^{k-1}\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\left(\mathbf{g}_{m_j}-\hat{\mathbf{g}}_{m_j}\right)\right|^2 + \sum_{j=k+1}^{2K}\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_j}\right|^2 + \left\|\hat{\mathbf{g}}_{m_k}\right\|^2/\rho_{\mathrm{s}}}, \tag{3.52}$$

$$\tilde{\gamma}_{\mathrm{R}i,k}^{\mathrm{JDF}} = \frac{\left|\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,k}+\hat{\mathbf{g}}_{2,k}\right)\right|^2}{\sum_{j=1,j\neq k}^{K}\left|\mathbf{g}_{i,k}^{\mathrm{H}}\left(\hat{\mathbf{g}}_{1,j}+\hat{\mathbf{g}}_{2,j}\right)\right|^2 + 1/(\alpha^{\mathrm{JDF}})^2}, \tag{3.53}$$

where the subscript index m_k follows the ordering definition described in Section a). Similarly, the effective SINRs of the SDF scheme can be expressed as

$$\tilde{\gamma}_{1\mathrm{R},k}^{\mathrm{SDF}} = \tilde{\gamma}_{2\mathrm{R},k}^{\mathrm{SDF}} = \frac{2\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i',k}\right|^2}{2\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\left(\mathbf{g}_{2,k}-\mathbf{g}_{1,k}\right)\right|^2 + \sum_{j=1,j\neq k}^{K}\left(\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{1,j}\right|^2 + \left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{2,j}\right|^2\right) + \left\|\hat{\mathbf{g}}_k\right\|^2/\rho_{\mathrm{s}}}, \tag{3.54}$$

$$\tilde{\gamma}_{\mathrm{R}i,k}^{\mathrm{SDF}} = \frac{\left|\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_k\right|^2}{\sum_{j=1,j\neq k}^{K}\left|\mathbf{g}_{i,k}^{\mathrm{H}}\hat{\mathbf{g}}_j\right|^2 + 1/(\alpha_{\mathrm{SDF}})^2}, \tag{3.55}$$

where if $\beta_{1,k} \geq \beta_{2,k}$, $i' = 1$, otherwise $i' = 2$. Therefore, the achievable sum rates of both JDF and SDF schemes can be obtained by

$$\tilde{R}^{\mathrm{JDF}} = \frac{T-2K}{2T}\sum_{k=1}^{K}\left(\min\{\tilde{R}_{2\mathrm{R},k}^{\mathrm{JDF}}, \tilde{R}_{\mathrm{R}1,k}^{\mathrm{JDF}}\} + \min\{\tilde{R}_{1\mathrm{R},k}^{\mathrm{JDF}}, \tilde{R}_{\mathrm{R}2,k}^{\mathrm{JDF}}\}\right), \tag{3.56}$$

$$\tilde{R}^{\mathrm{SDF}} = \frac{T-K}{2T}\sum_{k=1}^{K}\left(\min\{\tilde{R}_{2\mathrm{R},k}^{\mathrm{SDF}}, \tilde{R}_{\mathrm{R}1,k}^{\mathrm{SDF}}\} + \min\{\tilde{R}_{1\mathrm{R},k}^{\mathrm{SDF}}, \tilde{R}_{\mathrm{R}2,k}^{\mathrm{SDF}}\}\right). \tag{3.57}$$

Section 3.6 will show that the gap of the achievable rates between calculated by the statistical CSI and given by (3.56) and (3.57) is almost negligible, especially in the case that large antenna arrays are deployed on the RS. This phenomenon indicates that the achievable rates obtained by the statistical CSI are accurate enough to predict the achievable rates attained with the instantaneous CSI, which can provide useful guidelines for system analyses.

3.5 Properties with Massive Antenna Arrays

In this section, the asymptotic performance of both JDF and SDF schemes are investigated by extending the number of RS antennas to infinity where the asymptotic sum achievable rate and the power efficiency of the massive MIMO two-way relaying system are analyzed.

3.5.1 Asymptotic Achievable Rate ($M \to \infty$)

In the following parts of this subsection, the asymptotic achievable rate of both MA and BC communication phases is derived with the number of RS antennas tends infinity. To simplify the descriptions, here define the total pilot transmission power by $P_{\mathrm{p}} \triangleq \tau\rho_{\mathrm{p}}$ in the following where $\tau = 2K$ for JDF and $\tau = K$ for SDF. Further, two auxiliary variables, namely, $P_{\mathrm{s}} = M\rho_{\mathrm{s}}$ and $P_{\mathrm{d}} = M\rho_{\mathrm{d}}$, are defined to reveal the insight of massive MIMO relaying system with the proposed forwarding schemes.

a) Asymptotic MA achievable rates

As the number of RS antennas M tending infinity, the required signal transmission power of each MS can be accordingly decreased to persist in the asymptotic achievable rate performance. To reveal such property of the massive MIMO system in the two-way relaying system with the proposed forwarding schemes, P_{s} and the pilot transmission power ρ_{p} are fixed in the following derivations. As $M \to \infty$, the effective SINR of the MA phase with the JDF scheme is asymptotically approaching

$$\gamma_{i\mathrm{R},k}^{\mathrm{JDF}} \to \sigma_{i,k}^2 P_{\mathrm{s}}, \tag{3.58}$$

while with the SDF scheme, the asymptotic achievable rate is given by

$$\gamma_{1\mathrm{R},k}^{\mathrm{SDF}} = \gamma_{2\mathrm{R},k}^{\mathrm{SDF}} \to \frac{2\varsigma_{i',k}^4}{\left(\varsigma_{1,k}^2 - \varsigma_{2,k}^2\right)^2 + \varsigma_k^2/P_{\mathrm{s}}}, \tag{3.59}$$

where if $\beta_{1,k} \geq \beta_{2,k}$, $i' = 1$, otherwise $i' = 2$. Comparing (3.58) with (3.59) results in the following discussions.

- *Case I:* $P_{\mathrm{p}} \gg 1$

  Without loss of generalities, it is assumed that $\beta_{1,k} \geq \beta_{2,k}$ in this case. With $P_{\mathrm{p}} \gg 1$, there exists $\sigma_{i,k}^2 \approx \varsigma_{i,k}^2 \approx \beta_{i,k}$ and thus $\gamma_{i\mathrm{R},k}^{\mathrm{JDF}} \to \beta_{i,k}P_{\mathrm{s}}$ and $\gamma_{i\mathrm{R},k}^{\mathrm{SDF}} \to \frac{2\beta_{1,k}^2}{(\beta_{1,k}-\beta_{2,k})^2+\left(\beta_{1,k}+\beta_{2,k}\right)/P_{\mathrm{s}}}$. If $1/P_{\mathrm{s}} \gg \beta_{1,k} - \beta_{2,k}$, it is obtained that the limitation of $\gamma_{i\mathrm{R},k}^{\mathrm{SDF}}$ can be approximated as $\gamma_{i\mathrm{R},k}^{\mathrm{SDF}} \to \frac{2\beta_{1,k}}{\left(\beta_{1,k}+\beta_{2,k}\right)/P_{\mathrm{s}}} \geq \beta_{1,k}P_{\mathrm{s}} \geq \beta_{2,k}P_{\mathrm{s}}$. Such

comparison implies that with large pilot transmission power (i.e., $P_{\mathrm{p}} \gg 1$) and small signal transmission power (i.e., $P_{\mathrm{s}} \ll 1/(\beta_{1,k} - \beta_{2,k})$), the effective SINR in the MA phase of the SDF scheme is greater than or equal to that of the JDF one. The scenario discussed above is realistic in the practical massive MIMO systems where the signal transmission power can be rapidly decreased with a large number of antennas equipped on the RS.

- *Case II:* $\beta_{1,k} \approx \beta_{2,k}$ $(k = 1, 2, \cdots, K)$

  In this case, it is assumed that the communication pair $\mathrm{S}_{1,k}$ and $\mathrm{S}_{2,k}$ locate at almost the same distances to the RS, namely, $\beta_{1,k} \approx \beta_{2,k}$. Hence, $\sigma_{1,k}^2 \approx \sigma_{2,k}^2 \approx \frac{2\rho_{\mathrm{p}} K \beta_{1,k}^2}{2\rho_{\mathrm{p}} K \beta_{1,k}+1}$ and $\varsigma_{1,k}^2 \approx \varsigma_{2,k}^2 \approx \frac{2\rho_{\mathrm{p}} K \beta_{1,k}^2}{2\rho_{\mathrm{p}} K \beta_{1,k}+1}$. Such that, it can be derived that $\gamma_{i\mathrm{R},k}^{\mathrm{JDF}} \approx \gamma_{i\mathrm{R},k}^{\mathrm{SDF}} \approx \beta_{i,k} P_{\mathrm{s}}$. Due to larger data transmission duration within the coherence interval of the SDF scheme than the JDF one, i.e., $T_{\mathrm{d}}^{\mathrm{SDF}} > T_{\mathrm{d}}^{\mathrm{JDF}}$, it concludes that $\mathcal{R}_{i\mathrm{R},k}^{\mathrm{SDF}} > \mathcal{R}_{i\mathrm{R},k}^{\mathrm{JDF}}$ which can be stated that the SDF scheme outperforms the JDF one in the MA communications when MS pairs locate at similar distances to the RS.

b) Asymptotic BC achievable rates

In the BC phase, as the increasing of the number of antennas deployed on RS, the achievable rates of the considered TWR grows without bound while the transmission power of the RS also tends asymptotically to infinity. Actually, the transmission power of the RS can not keep growing without bound in the practical systems, which limits the sum rates performance of the whole system. In this discussion, it is assumed that the total transmission power of the RS is constrained to be P_{d}, which is fixed regardless of the RS antenna number M. Therefore, as $M \to \infty$ and fixing the pilot transmission power ρ_{p}, the effective SINR of the BC communication with the JDF scheme is bounded by

$$\gamma_{\mathrm{R}i,k}^{\mathrm{JDF}} \to \frac{\sigma_{i,k}^4}{\sum_{j=1}^{K} \left(\sigma_{1,j}^2 + \sigma_{2,j}^2\right) / P_{\mathrm{d}}}, \tag{3.60}$$

while with the SDF scheme, it is bounded by

$$\gamma_{\mathrm{R}i,k}^{\mathrm{SDF}} \to \frac{\varsigma_{i,k}^4}{\sum_{j=1}^{K} \left(\varsigma_{1,j}^2 + \varsigma_{2,j}^2\right) / P_{\mathrm{d}}}. \tag{3.61}$$

By comparing (3.60) with (3.61), the following discussions can be drawn out:

- *Case I:* $P_{\mathrm{p}} \gg 1$

  Evidently, there exist $\sigma_{i,k}^2 \approx \varsigma_{i,k}^2 \approx \beta_{i,k}$ if $P_{\mathrm{p}} \gg 1$, and thus $\gamma_{\mathrm{R}i,k}^{\mathrm{JDF}} \approx \gamma_{\mathrm{R}i,k}^{\mathrm{SDF}} \approx \frac{\beta_{i,k}}{\sum_{j=1}^{K}(\beta_{1,j}+\beta_{2,j})/P_{\mathrm{d}}}$. With lower pilot training overhead, the SDF scheme outperforms

the JDF one in BC achievable rate performance when the pilot transmission power of each MS ρ_{p} or the number of MS pairs K is large.

- *Case II:* $\beta_{1,k} \approx \beta_{2,k}$ $(k = 1, 2, \cdots, K)$.

  In this case, it is assumed that the communication pairs locate at almost the same distances to the RS, respectively, namely, the large scale fading factors of each MS pair are almost identical to each other. Thus, there exists $\sigma_{i,k} \approx \varsigma_{i,k} \approx \frac{2\rho_{\mathrm{p}} K \beta_{1,k}^2}{2\rho_{\mathrm{p}} K \beta_{1,k}+1}$. Hence, $\gamma_{\mathrm{R}i,k}^{\mathrm{JDF}} \approx \gamma_{\mathrm{R}i,k}^{\mathrm{SDF}}$. Therefore, the SDF scheme outperforms the JDF one in asymptotic achievable rate of BC communications due to its lower overhead of pilot transmissions.

Consequently, by summarizing the asymptotic achievable rate comparisons of both the MA and BC communication phases, it is evident that the SDF scheme outperforms the JDF one in both considered cases.

3.5.2 Power Efficiency ($M \to \infty$)

In the previous part of this section, the very large pilot transmission power is considered where $P_{\mathrm{s}} \gg 1$. In this part, the pilot transmission power of each MS is assumed to be fixed as a not very large value, namely, ρ_{p} is a finite constant. Therefore, $\sigma_{i,k}$ and $\varsigma_{i,k}$ are all constants if the number of MS pairs K is also fixed which is reasonable within a coherence interval. In addition, it is also assumed that P_{s} and P_{d} are all independent of the RS antenna number. In this occasion, the sum achievable rate of each MS pair is achieving

$$\begin{aligned}\mathcal{R}_k^{\mathrm{JDF}} \to & \min\left\{C\left(\sigma_{2,k}^2 P_{\mathrm{s}}\right), C\left(\frac{\sigma_{1,k}^4}{\sum_{j=1}^{K}\left(\sigma_{1,j}^2+\sigma_{2,j}^2\right)/P_{\mathrm{d}}}\right)\right\} \\ & + \min\left\{C\left(\sigma_{1,k}^2 P_{\mathrm{s}}\right), C\left(\frac{\sigma_{2,k}^4}{\sum_{j=1}^{K}\left(\sigma_{1,j}^2+\sigma_{2,j}^2\right)/P_{\mathrm{d}}}\right)\right\},\end{aligned} \tag{3.62}$$

$$\begin{aligned}\mathcal{R}_k^{\mathrm{SDF}} \to & \min\left\{C\left(\frac{2\varsigma_{i',k}^4}{\left(\varsigma_{1,k}^2-\varsigma_{2,k}^2\right)^2+\varsigma_k^2/P_{\mathrm{s}}}\right), C\left(\frac{\varsigma_{1,k}^4}{\sum_{j=1}^{K}\varsigma_j^2/P_{\mathrm{d}}}\right)\right\} \\ & + \min\left\{C\left(\frac{2\varsigma_{i',k}^4}{\left(\varsigma_{1,k}^2-\varsigma_{2,k}^2\right)^2+\varsigma_k^2/P_{\mathrm{s}}}\right), C\left(\frac{\varsigma_{2,k}^4}{\sum_{j=1}^{K}\varsigma_j^2/P_{\mathrm{d}}}\right)\right\},\end{aligned} \tag{3.63}$$

where $i' = 1$ when $\beta_{1,k} \geq \beta_{2,k}$, otherwise $i' = 2$. With above assumptions, both $\mathcal{R}_k^{\mathrm{JDF}}$ and $\mathcal{R}_k^{\mathrm{SDF}}$ in (3.62) and (3.63) are asymptotically approaching constant achievable rates, respectively, which implies that the transmission power of each MS, i.e., ρ_{s}, can be reduced proportionally to $1/M$ with large antenna arrays deployed on RS, while the achievable rates performance of the relaying system persists which in turn rapidly improves the

power efficiency of the relaying system. In Section 3.6, the simulation results of the system power efficiency are evaluated in details.

3.6 Numerical and Simulation Results

This section evaluate the performance of the proposed schemes by examining both the spectral and power efficiency in numerical and simulation results. The spectral efficiency is defined by the sum achievable rates per channel use denoted by (3.26) in bits/s/Hz. As defined in [30], the coherence interval is set as $T_{\mathrm{c}} = 200$ throughout this section, while if not specifically mentioned, the large scale fading parameters are all chosen as $\beta_{i,k} = 1$, ($i = 1, 2$ and $k = 1, 2, \cdots, K$). Note that the coherence interval in this section is the normalized value relative to the symbol duration. Moreover, all power parameters are denoted in dB relative to the noise power which is normalized to 1 and this section defines $\mathrm{SNR} \triangleq \rho_{\mathrm{s}}$. Finally, this section assumes $\rho_{\mathrm{p}} = \rho_{\mathrm{s}}$ and $\rho_{\mathrm{d}} = 2\rho_{\mathrm{s}}K$ for numerical results calculation and simulations without specifications.

3.6.1 Accuracy of Achievable Rates Derived from Statistical CSI

This part examines the accuracy of the achievable rates obtained by detecting the received signal with the statistical CSI against that with the instantaneous CSI, which are illustrated in section 3.4, by (3.32),(3.38),(3.45) and (3.49) against (3.56) and (3.57), respectively. To evaluate the spectral efficiency with the instantaneous CSI, 1000 random channel implements are generated and the achievable rates are obtained by taking the average results. Fig. 3.3 and Fig. 3.4 plot the comparisons of the spectral efficiency which are calculated and simulated with the statistical and instantaneous CSI for the JDF and SDF schemes, respectively. From the comparison, it is obvious to find that the corresponding spectral efficiency gap between the statistical and instantaneous CSI is tolerable small with both JDF and SDF schemes. Particularly, there are almost no differences in the low SNR region. Furthermore, with larger antenna arrays, the performance gap also decreases. The phenomenon observed above indicates that the derived closed-form expression of the achievable rates are quite accurate in characterizing the simulated spectral efficiencies achieved by both the JDF and SDF schemes throughout the whole SNR range, thus providing useful guidelines in practical system designs and performance analyses. With such benefits, the followed performance comparisons are all evaluated by the closed-form expressions calculated with the statistical CSI.

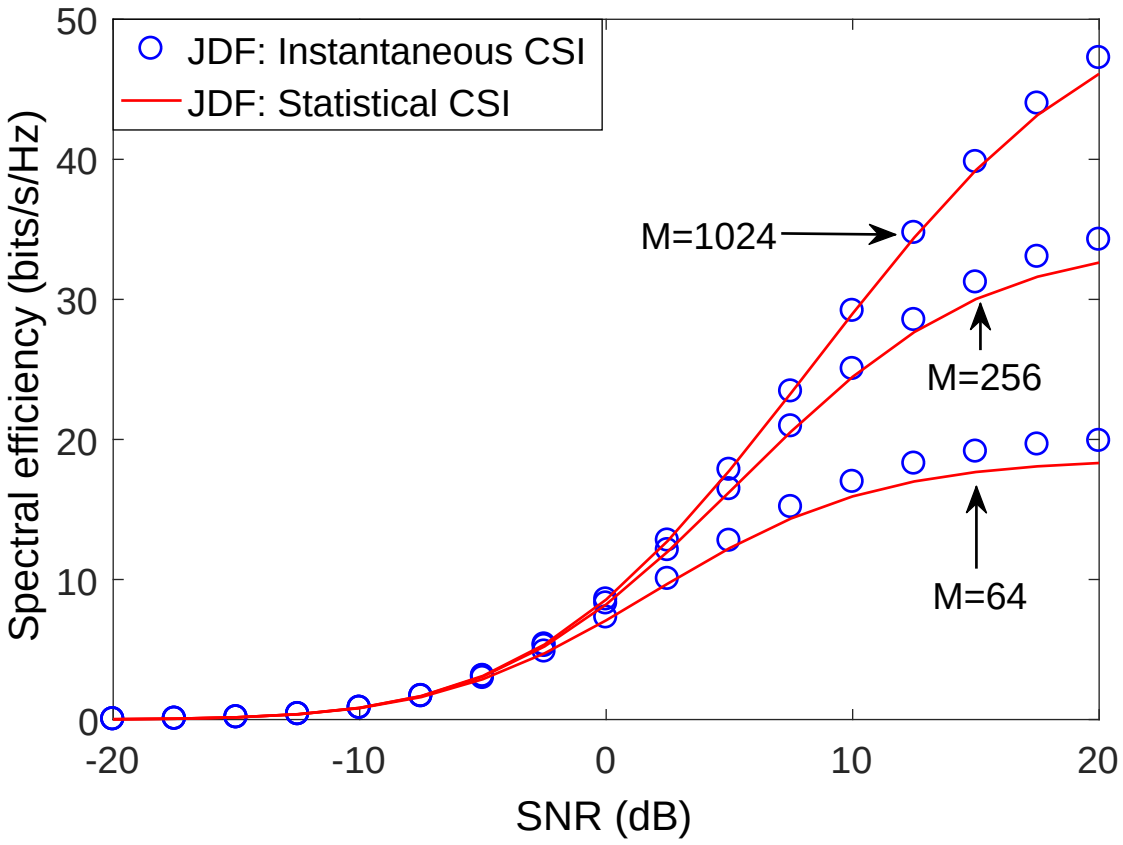


Figure 3.3: Spectral efficiency comparison of the JDF scheme between the closed-form and simulation results. ($K = 10$)

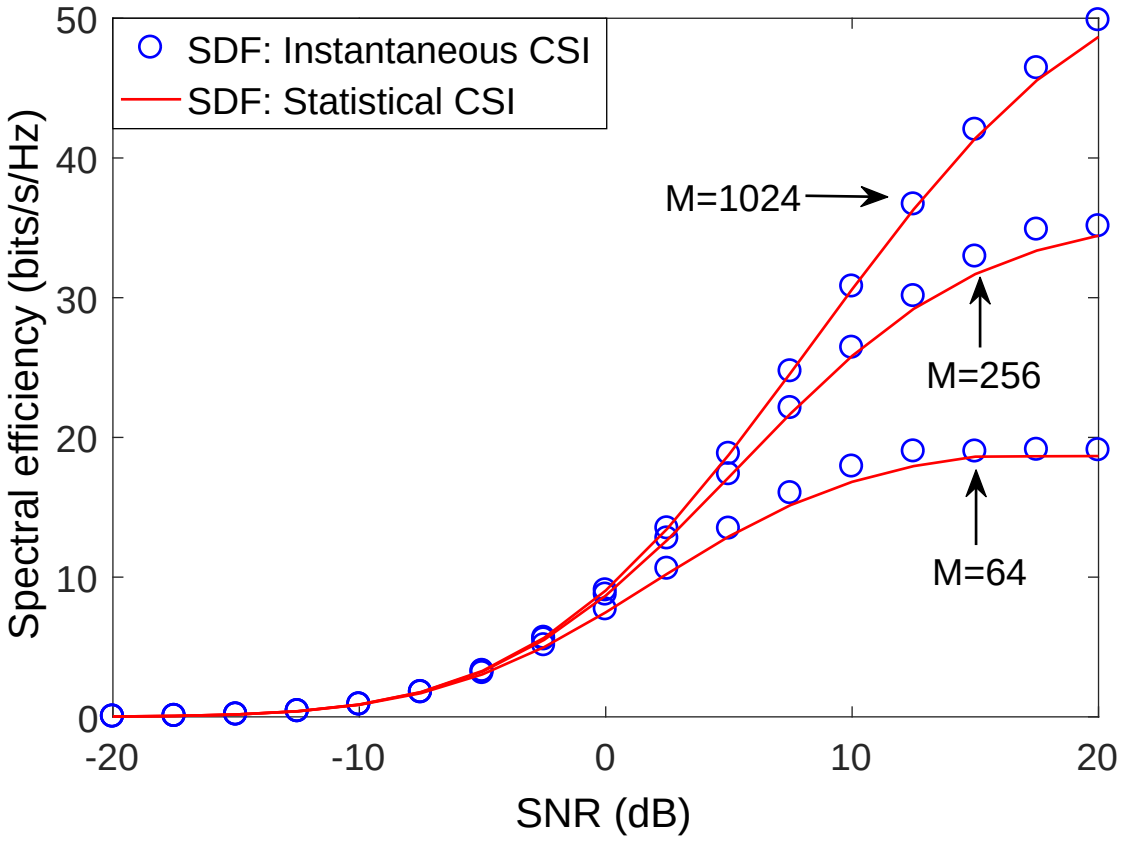


Figure 3.4: Spectral efficiency comparison of the SDF scheme between the closed-form and simulation results. ($K = 10$)

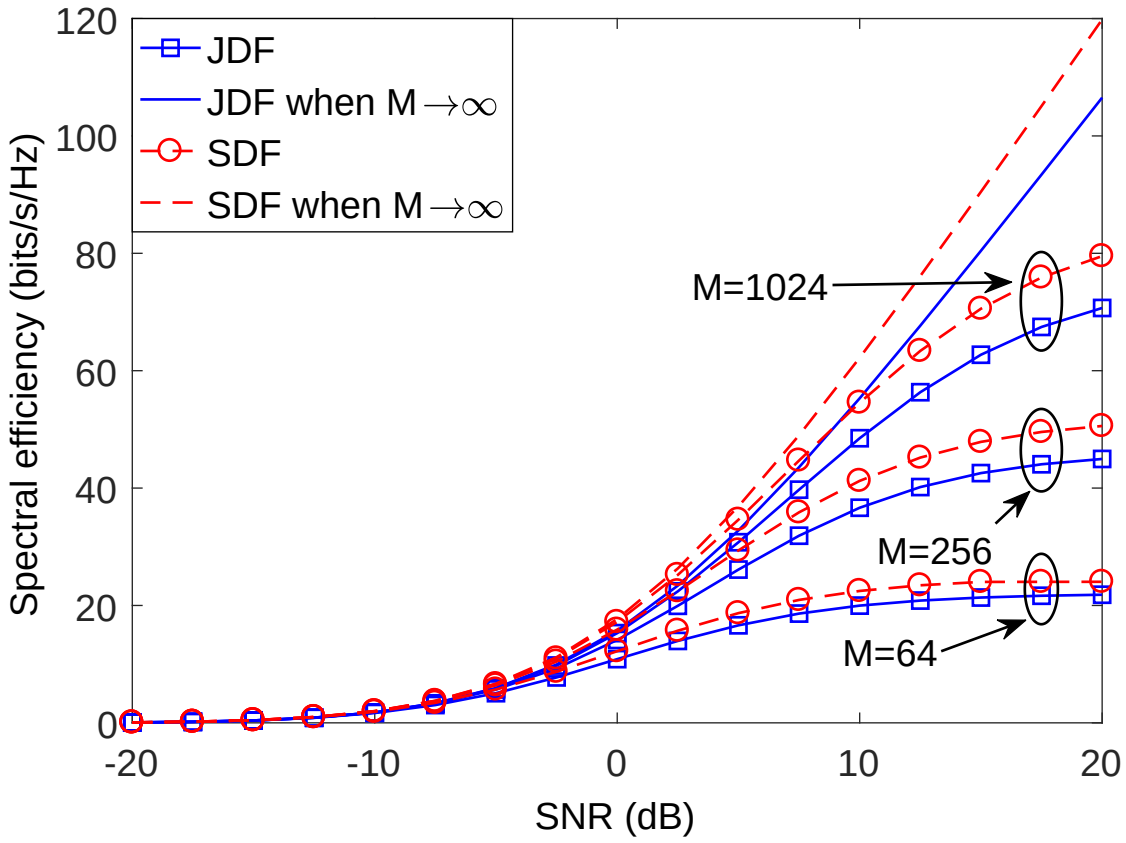


Figure 3.5: Spectral efficiency comparison between JDF and SDF schemes. ($K = 20$)

3.6.2 Spectral Efficiency Comparisons between JDF and SDF

Firstly, this subsection compares the performance of the SDF scheme with the JDF scheme by evaluating their spectral efficiencies against different number of RS antennas. Fig. 3.5 shows the improvements of the spectral efficiency performance of the SDF scheme compared to the JDF one. In the numerical results, the number of users pairs are set as $K = 20$ and SNR is ranging from -20 dB to 20 dB. From the comparison, it can be noticed that the spectral efficiency of the SDF scheme is almost the same as the JDF scheme at the low SNR region while the SDF outperforms the JDF at the high SNR region. The theoretical analyses present that the dominated BC effective SINR between the two schemes are almost identical while the pilot transmission overhead used for channel estimations of the SDF scheme $\tau^{\mathrm{SDF}} = K$ is smaller than that of the JDF scheme $\tau^{\mathrm{JDF}} = 2K$. Hence, the SDF scheme achieves a higher spectral efficiency than the JDF one in high SNR regions. Further, the performance gap increases when more antennas are deployed on the RS which implies that the SDF is a more superior forwarding scheme in massive MIMO relaying systems compared to the JDF.

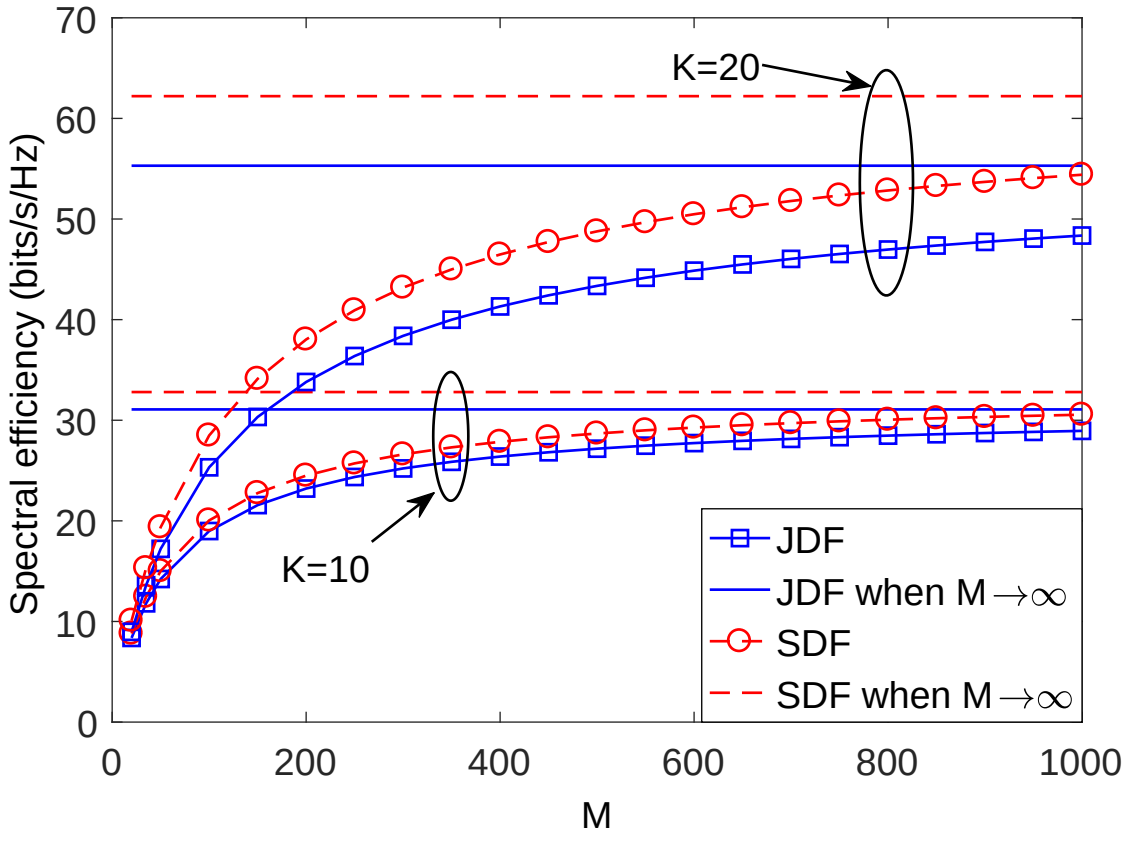


Figure 3.6: Spectral efficiency comparison between JDF and SDF schemes with asymptotic performance. (SNR=10 dB)

Moreover, the spectral efficiencies of both SDF and JDF schemes with respect to the

increasing of the number of RS antennas are plotted in Fig. 3.6, where the SNR is set to be 10 dB. As expected, the spectral efficiencies of both the SDF and JDF schemes are growing as the increasing of M. Comparing the performance of the proposed two schemes, it is evident that the SDF always outperforms the JDF no matter M is small or large. Further, the spectral efficiency gap between these two schemes keeps increasing as M becoming larger. This is explained that with large number of antennas deployed on the RS, the additional RSI of the SDF scheme reduces, thus the effective SINRs of the two schemes become closer while the differences of pilot transmission overheads dominate the spectral efficiency gap. Hence, the SDF achieves larger spectral efficiencies than the JDF, which verifies the analyses in Section 3.5. This phenomenon can also be confirmed by observing the gap of the asymptotic spectral efficiencies between these two schemes, i.e., the lines without markers in Fig. 3.6, where the RSI of the SDF scheme diminishes as $M \to \infty$. Therefore, the SDF scheme achieves a much better asymptotic performance than the JDF scheme, especially in the regime with more MSs served by the RS, which again confirms the superiority of the SDF in massive MIMO systems.

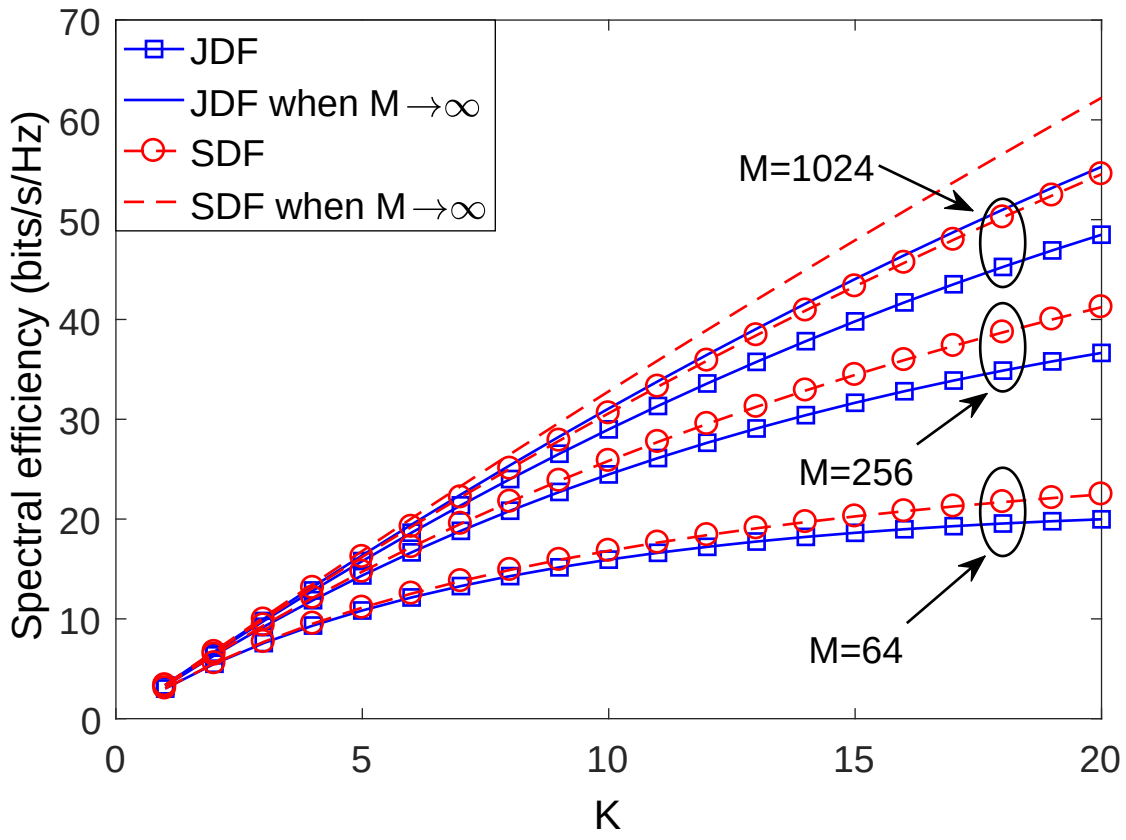


Figure 3.7: Spectral efficiency comparison between JDF and SDF schemes versus the number of MSs. (SNR=10 dB)

Finally, the performances of both JDF and SDF schemes are compared by evaluating their spectral efficiencies as the function of the number of MSs. Fig. 3.7 shows that the

achievable rates of both JDF and SDF schemes increases with more MSs served by the RS. In the meanwhile, the performance gap between the two schemes also enlarges with larger K. It explicitly indicates that the SDF scheme can achieves better performance than the JDF when more MSs are served by the RS.

3.6.3 Power Efficiency Comparisons between JDF and SDF

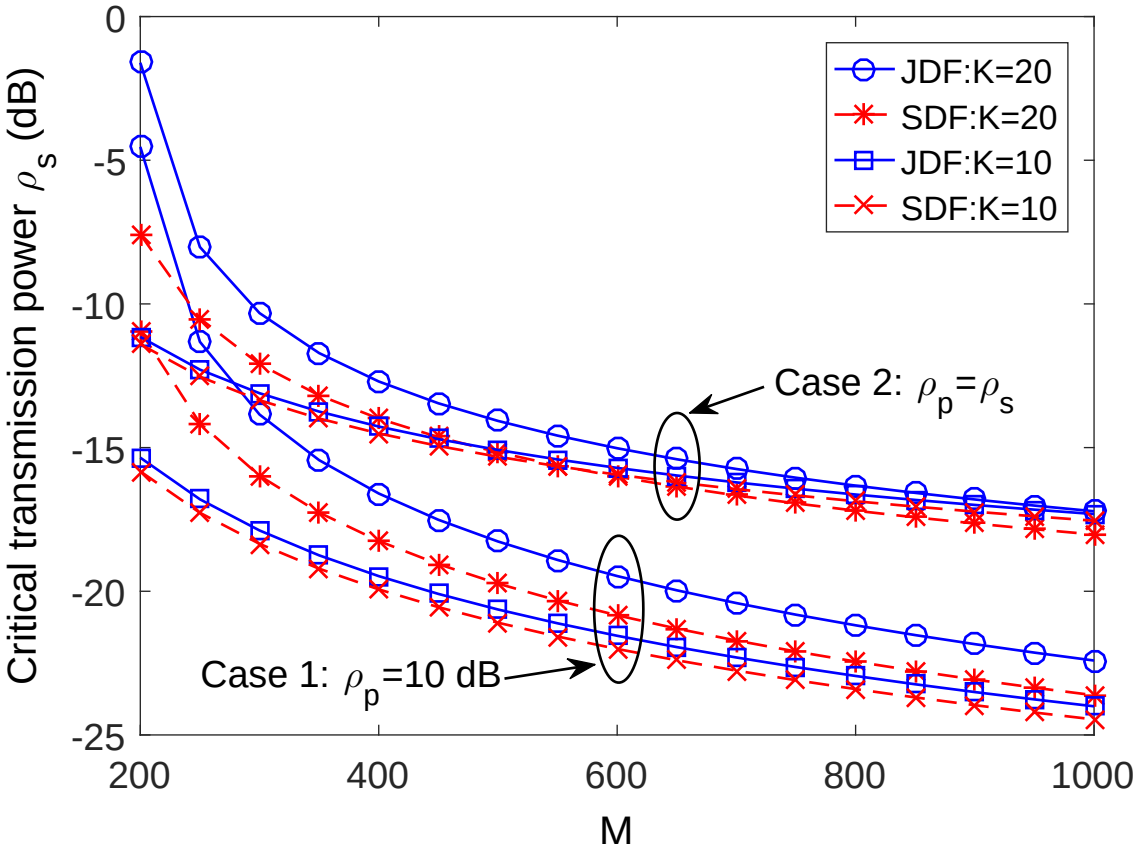


Figure 3.8: Required power ρ_s to achieve 1 bit/s/Hz per user. (SNR=10 dB and $\rho_d = 2\rho_s K$)

Apart from the spectral efficiency comparisons, the power efficiency[3] of the massive MIMO two-way relaying system with both the SDF and JDF schemes is examined in this subsection. Here, two cases are considered to evaluate the power efficiency performance. One of which is to fix ρ_p and the other one is to set $\rho_p = \rho_s$. Fig. 3.8 shows the required lowest transmission power of each MS, namely, the critical transmission power, to maintain the QoS of each pair of MSs, e.g., each communication pair can achieve 1 bit/s/Hz spectral efficiency. From the numerical results, it can be found that as the increasing of the antenna number M, the critical transmission power is significantly reduced. On

[3]Note that the power considered here includes only signal transmission power while excludes the signal processing power consumptions.

one hand, if the transmission power of the pilot transmission ρ_{p} is fixed to be 10 dB, the critical signal transmission power can be reduced to about -15 dB when $K = 10$ and $M = 200$ while to about -25 dB if $M = 1000$, which shows that the power efficiency of the system can be dramatically improved with massive antenna arrays deployed on the RS for both schemes. On the other hand, if the pilot transmission power is set to be identical to the signal transmission power, i.e., $\rho_{\mathrm{p}} = \rho_{\mathrm{s}}$, the more MSs are served by the RS, the lower critical transmission power is required to achieve the desired spectral efficiency. Finally, it can be observed that the critical transmission power of the SDF scheme is evidently lower than that of the JDF scheme in both cases throughout all the antenna number range, which essentially reveals the superiority of the former scheme to the latter one in power efficiency performance comparisons.

3.7 Conclusions

In this chapter, two efficient decode-and-forwarding schemes, namely, the JDF and SDF, have been proposed in the massive MIMO two-way relaying systems. Both the theoretical analyses and numerical evaluations are presented in this chapter to examine the performance of the two schemes. To investigate the spectral efficiencies, the PNC and MRC/MRT processings are adopted at the RS side where the CSI is obtained by the MMSE channel estimations. Further, the asymptotic performance of the proposed schemes when the number of RS antennas tending infinity, i.e., $M \to \infty$ is investigated and the results show that the SDF scheme outperforms the JDF in massive MIMO scenarios. The theoretical closed-form expressions of the spectral efficiency are derived in the chapter by using the statistical CSI where the simulation results show that the expressions are quite accurate to characterize spectral efficiency performance of the two schemes. Finally, the power efficiency of the MSs and RS deployed with massive antenna arrays is also discussed by theoretical analyses and evaluated by numerical results.

3.8 Appendix

3.8.1 Proof of Theorem 3.1

Proof: As described in (3.29), we have

$$\mathbb{E}\left\{\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_k}\right\} = \mathbb{E}\left\{\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\hat{\mathbf{g}}_{m_k} + \hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\boldsymbol{\xi}_{m_k}\right\} = M\sigma_{m_k}^2 \tag{3.64}$$

$$\mathbb{V}\mathrm{ar}\left\{\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_k}\right\} = \mathbb{E}\left\{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_k}\right|^2\right\} - \left|\mathbb{E}\left\{\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_k}\right\}\right|^2 = M\sigma_{m_k}^2\beta_{m_k} \tag{3.65}$$

From (3.30), the power of interference can be deduced as follows:

$$\mathrm{SRI}_{m_k}^{\mathrm{JDF}} \triangleq \sum_{j=1}^{k-1} \mathbb{E}\left\{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\left(\mathbf{g}_{m_j}-\hat{\mathbf{g}}_{m_j}\right)\right|^2\right\} = M\sigma_{m_k}^2 \sum_{j=1}^{k-1}\left(\beta_{m_j}-\sigma_{m_j}^2\right) \tag{3.66}$$

$$\mathrm{IPI}_{m_k}^{\mathrm{JDF}} \triangleq \sum_{j=k+1}^{2K} \mathbb{E}\left\{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{g}_{m_j}\right|^2\right\} = M\sigma_{m_k}^2 \sum_{j=k+1}^{2K}\beta_{m_j} \tag{3.67}$$

$$\mathrm{AN}_{m_k}^{\mathrm{JDF}} \triangleq \frac{1}{\rho_{\mathrm{s}}}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_{m_k}^{\mathrm{H}}\mathbf{n}_{\mathrm{R}}\right|^2\right\} = \frac{M\sigma_{m_k}^2}{\rho_{\mathrm{s}}}. \tag{3.68}$$

By substituting (3.64) – (3.68) into (3.29), the effective SINR of the JDF MA phase can be derived as

$$\gamma_{\mathrm{R}_{i,k}}^{\mathrm{JDF}} = \frac{M\delta_{i,k}}{\sum_{j=1}^{2K}\beta_{\mathrm{J},m_j}-\sum_{j=1}^{m_k-1}\delta_{\mathrm{J},m_j}+\sigma_n^2/P_{\mathrm{S}}}. \tag{3.69}$$

Consequently, Theorem 3.1 is obtained by substituting (3.69) into (3.31). ■

3.8.2 Proof of Theorem 3.3

Proof: From (3.41) to (3.44), there exist the following derivations:

$$\mathbb{E}\left\{\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k}\right\} = \varsigma_{i,k}^2 \tag{3.70}$$

$$\mathbb{V}\mathrm{ar}\left\{\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{i,k}\right\} = M\beta_{i,k}\varsigma_k^2 \tag{3.71}$$

$$\mathrm{RSI}_{i,k}^{\mathrm{SDF}} \triangleq \mathbb{E}\left\{\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\left(\mathbf{g}_{2,k}-\mathbf{g}_{1,k}\right)\right|^2\right\} = M^2\left(\varsigma_{1,k}^2-\varsigma_{2,k}^2\right)^2 + M\beta_k\varsigma_k^2 \tag{3.72}$$

$$\mathrm{IPI}_{i,k}^{\mathrm{SDF}} \triangleq \sum_{j=1,j\neq k}^{K} \mathbb{E}\left\{\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{1,j}\right|^2+\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{g}_{2,j}\right|^2\right\} = \sum_{j=1,j\neq k}^{K} M\beta_j\varsigma_k^2 \tag{3.73}$$

$$\mathrm{AN}_{i,k}^{\mathrm{SDF}} \triangleq \frac{1}{P_{\mathrm{S}}}\mathbb{E}\left\{\left|\hat{\mathbf{g}}_k^{\mathrm{H}}\mathbf{n}_{\mathrm{R}}\right|^2\right\} = \frac{M\varsigma_k^2}{\rho_{\mathrm{s}}}. \tag{3.74}$$

Hence, by substituting (3.70) – (3.74) into (3.41), there derives

$$\gamma_{\mathrm{R}_{1,k}}^{\mathrm{SDF}} = \gamma_{\mathrm{R}_{2,k}}^{\mathrm{SDF}} = \frac{2M\varsigma_{i,k}^4}{2\beta_{i,k}\varsigma_k^2+M\left(\varsigma_{1,k}^2-\varsigma_{2,k}^2\right)^2+\varsigma_k^2\sum_{j=1}^{K}\beta_j+\varsigma_k^2/\rho_{\mathrm{s}}} \tag{3.75}$$

where if $\beta_{1,k} \geq \beta_{2,k}$, $i = 1$, otherwise $i = 2$. Therefore, Theorem 3.3 is obtained by substituting (3.75) into (3.31). ■

Chapter 4

Framework of Channel Estimation for Hybrid Analog-and-Digital Processing Enabled Massive MIMO Communications

The above two chapters have focused on the investigation of multipair massive multiple-input multiple-output (MIMO) relaying systems. In this chapter, a peer-to-peer signal transmission system is considered which employs the hybrid analog/digital precoding/combining structure in a massive MIMO system to reduce hardware cost and baseband signal processing complexity. Due to limited radio-frequency (RF) chains between antennas and the baseband processing module, it is more challenging to perform channel estimation in the considered hybrid-structured system compared to the conventional full-chain massive MIMO system. Dealing with such issue, this chapter focuses on investigating the general channel estimation problem in a hybrid-structured massive MIMO system and designing the optimal RF combiners by exploiting the channel covariance matrix.

4.1 Introduction

MASSIVE multiple-input multiple-output (MIMO) is potentially one of the most promising and key technologies to meet the stringent performance requirements for the next-generation, i.e., 5G, wireless communications [9, 16]. It has attracted considerable research interests from both academia and industry since the seminal work [12] was published in 2010 [14, 15, 29, 47]. Most notably, with an excessive amount of antennas mounted at the transmitter and/or receiver, signal processing, including both

transmission precoding and receiving combining, can be greatly simplified while achieving highly optimal performance [13, 68, 76]. Simple linear precoding schemes, such as zero-forcing (ZF), are virtually optimal and comparable to the capacity-achieving nonlinear dirty paper coding (DPC). Thanks to the employment of millimeter wave (mmWave) bands, a large number of antennas can be packed into a small area due to the short wavelength [19].

However, the conventional signal processing is typically performed at the baseband, which means that the signal received from each antenna port needs to be properly filtered, down-converted, and then sampled, where the hardware module performing such tasks is normally referred to as a radio-frequency (RF) chain. Analogous procedure exists for signal transmission. Given a large number of antennas in a massive MIMO system, it would be formidable to feed each antenna with a dedicated RF chain due to high cost and power consumption. To circumvent the challenging requirement of massive RF chains, an analog/digital hybrid structure has been proposed for massive MIMO systems operated on the mmWave bands [20]. On the transmitter side of a hybrid system, low-dimensional baseband signals (after digital processing) are converted to the RF domain, feeding a phase-shifting network to properly adjust the phases of transmission signals which are then transmitted to wireless channels by antennas [19]. The design of the RF and baseband precoding/combining matrices in a hybrid structure is a challenging problem and has been extensively studied in recent years, e.g., [17, 20, 77–79]. However, these papers all assume the availability of channel state information (CSI), and the hybrid processing matrices are dependent on the MIMO channel realization. Channel estimation is one of the most important elements in reaping all the advantages and gains of massive MIMO, and unsurprisingly has been under extensive investigation. It is a challenging task, yet is even more difficult to fulfill under the limited RF-chain hybrid structure constraint.

The conventional MIMO channel estimation methods in the literature cannot be applied in a hybrid structure to obtain full CSI [80–82]. Some previous papers have studied channel estimation with limited RF chains in hybrid structured mmWave communications [18, 83–85]. An adaptive compressed sensing solution for the hybrid channel estimation was proposed in [18] from the perspective of angular sparsity of mmWave channels. Scanning in the angular domain is performed at both the transmitter and receiver sides. As such, the complexity and resource consumption of the proposed channel estimation scheme is dominated by the sparsity of channel scatterings. For instance, it will consume much more resource with higher complexity to achieve a desired angular resolution in channel estimation when the scattering paths are rich, compared to the case of very sparse channels. A similar method was developed in a frequency-selective channel scenario [23]. In [83], channel estimation with phase-shifters or switches in a hybrid structured system were discussed. Still, it employs the high-complexity compressed

sensing method with angular domain scanning to estimate the angle of arrival/angle of departure (AoA/AoD) of each scattering path. Few prior literature has considered the general problem of channel estimation in a hybrid structured massive MIMO system without significant dependence on the channel sparsity. Since the degree of freedom (DoF) of the received signals at baseband is limited by the number of RF chains, it becomes prohibitively difficult to obtain satisfactory higher-dimensional vector channel estimates, especially for rich scattering environments, in a hybrid precoding system using conventional channel estimation approaches.

In this chapter, we consider the uplink channel estimation of a massive MIMO system with the hybrid RF-baseband processing structure. Note that the considered system structure is also applicable to communications in mmWave bands. An efficient channel estimation scheme is proposed by exploiting the spatial correlation of massive MIMO channels. The main task of the proposed channel estimation in the hybrid precoding system is to recover the M-dimensional channel vector from L observations at the baseband where L ($< M$) is the limited number of RF chains. Following the well-known result in [73], one pilot symbol is optimal, and in fact enough, for a single-antenna user to assist uplink estimation of uncorrelated MIMO channels under some mild assumptions. In a massive MIMO system with limited RF chains, only a much smaller L-dimensional signal is finally captured at baseband from the massive M antennas. It is, therefore, evident that a single training symbol becomes insufficient for the BS to conduct full-dimensional channel estimation of M (independent) channel coefficients. Further, considering the additional constant-magnitude constraint on the receiving combiners due to the use of phase shifters, the conventional MIMO channel estimator does not ever apply even when the same one-pilot-per-user strategy is used. To approach the performance of full-chain channel estimation, multiple trainings are required to achieve the DoF of full-chain baseband signal measurements. That is, TL observations can be utilized to estimate the M-dimensional MIMO channels, where T is the number of trainings. Hence, optimal RF combiners for different training phases need to be properly designed to capture the channel energy and then recover the CSI as accurately as possible. Theoretically, $T = M/L$ training symbols are required to achieve the DoF of a full-chain training in the estimation of uncorrelated channels. However, such conclusions may not hold for correlated channels. As revealed by both [60] and [86], the optimal number of training symbols can be reduced due to the fact that the dimension of statistically dominant subspaces is less than the number of antennas. In practice, massive MIMO including mmWave channels are inevitably spatially correlated due to the limitation in either the number of propagation paths or angular spreads [18, 80, 87]. Thus, less than M/L trainings could be sufficient to achieve the full-chain channel estimation performance in such spatially correlated channels. We summarize the main contributions of this chapter as follows:

- **RF Combiner Design for Single Training:** We investigate the channel estimation in a hybrid precoding system and formulate an optimization problem on channel estimation following the minimum mean square error (MMSE) criteria. To solve the problem, the constant-magnitude constraint of the RF combiner is temporarily relaxed. By employing the properties of *Block Generalized Rayleigh Quotient*, the theoretical optimizer of the formulated optimization problem is solved and the corresponding optimal RF combiner is designed in the single training scenario. Analyses of the designed combiner show that the mean square error (MSE) of the channel estimation decreases when *(1) the channel is more spatially correlated, (2) more RF chains for channel estimation are deployed, (3) larger transmission power for training pilots is utilized*. Moreover, the *closed-form* expression of the MSE is derived and verified to be quite accurate compared to the simulation results, thus providing useful guidelines in practical system designs.

- **RF Combiner Design for Multiple Trainings:** We formulate the optimization problem to design RF combiners for multiple trainings and solve it by *Sequential* and *Alternating* methods with the assistance of the single-training result. The *Sequential* method can achieve step-wisely minimum MSE with low complexity while the *Alternating* one solves the joint optimization problem iteratively with high complexity, however, achieving the local optimum. The performance and complexity of the proposed RF combiners are examined in simulations under both nonparametric and parametric channel models [88].

- **Propose a Covariance Matching Method to Generate Channel Correlation:** To perform the proposed RF combiner design and channel estimation, the channel covariance matrix should be known by the BS. In this chapter, a covariance matching method used to generate the channel covariance matrix is proposed based on the angular spread of the arrived RF rays which can be readily estimated in a hybrid precoding/combining system with few resource consumption [18, 89]. From the simulation results, it achieves comparable spectral efficiency performance with the CSI estimated from the matched covariance matrix compared to that with the perfect one, which proves the effectiveness and robustness of the proposed methods.

Apart from the contributions, we present the main characteristics differentiating this chapter from the existing works summarized below.

- Compared to the typical full-chain massive MIMO systems, the signal dimension in the considered system is reduced after the processing of the phase-shifting network. That is, it converts the received M-dimensional RF signal to L-dimensional baseband signal where M denotes the number of antenna elements and L represents the

number of RF chains typically satisfying $L < M$ in hybrid precoding/combining systems. As such, the channel estimation performed in baseband cannot utilize the complete training information as the full-chain massive MIMO does.

- Compared to existing work on channel estimation in a hybrid-structure mmWave system, no explicit usage of channel nature, e.g., sparsity, is required by our proposed methods to perform the estimation. In addition, the complexity of our proposed scheme is determined by the number of RF chains and training phases which are fixed in an online system, rather than the scattering circumstance which is time-variant and difficult to be guaranteed.

Organization: The rest of this chapter is organized as follows. Section 4.2 illustrates the system model of the massive MIMO system with the hybrid precoding structure and the channel models adopted in this chapter. In Section 4.3, the channel estimation schemes in the hybrid structure are proposed and analyzed. Section 4.4 proposes the method to generate the channel covariance matrices which can be employed in practical system implementations. The numerical and simulation results are presented in Section 4.5. Finally, Section 4.6 concludes our work in this chapter.

4.2 System Model

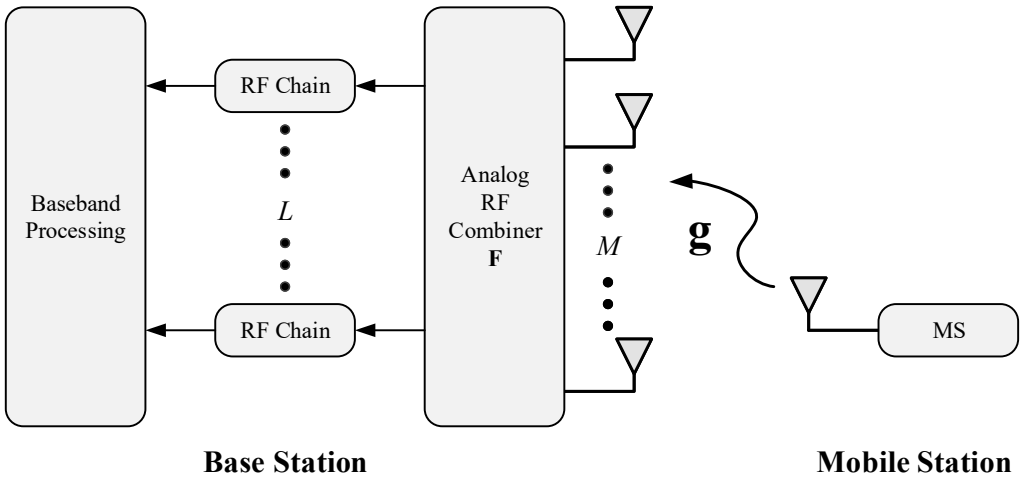


Figure 4.1: Block diagram of massive MIMO with a hybrid structure.

We investigate a single-cell hybrid massive MIMO communication structure as shown in Fig. 4.1 where only one mobile station (MS) is considered for the hybrid channel estimation in this chapter. Here, we restrict ourselves in the single MS scenario to simplify the elaboration on channel estimations with limited RF chains. Note that the proposed methods are readily extended to the multi-user scenario by simply allowing each MS to

sequentially/alternately conduct the channel estimation as proposed. In the considered system, the base station (BS) is equipped with M antennas and L RF chains interfaced by an analog RF phase-shifting network converting the high-dimensional received signal to a low-dimensional signal through phase-only linear combining, where $M > L$ for limited RF chains and $M = L$ for full chains. The combined RF signal after the phase-shifting network is then down-converted to baseband and is sampled by analog-to-digital converter (ADC) for further full-digital processing [19].

At the user side, a single-antenna MS is considered in the channel estimation period where the MS transmits T training pilots to the BS. Considering the ith ($i = 1, 2, \cdots, T$) pilot training operated by the MS, the received signal at BS baseband after pilot compensation can be expressed as

$$\mathbf{y}_i = \mathbf{F}_i \left(\sqrt{\rho}\mathbf{g}\varphi_i + \tilde{\mathbf{n}}_i\right)\varphi_i^* = \sqrt{\rho}\mathbf{F}_i\mathbf{g} + \mathbf{F}_i\mathbf{n}_i, \tag{4.1}$$

where $\mathbf{g} \in \mathbb{C}^{M\times 1}$ represents the uplink channel from the MS to BS which is assumed to remain static throughout a coherence interval, φ_i stands for the ith training symbol with $\varphi_i\varphi_i^* = 1$, ρ is the pilot power per transmission, $\tilde{\mathbf{n}}_i \in \mathbb{C}^{M\times 1}$ denotes the additive white Gaussian noise (AWGN) with zero mean and unit variance, $\mathbf{n}_i \triangleq \tilde{\mathbf{n}}_i\varphi_i^*$ which persists as $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ AWGN, and $\mathbf{F}_i \in \mathbb{C}^{L\times M}$ is the RF phase-shifting matrix constructed by unit-magnitude elements. For T pilot transmissions, the received signal can be stacked as a concatenated vector $\mathbf{y}_\mathrm{c} = [\mathbf{y}_1^\mathrm{T}, \mathbf{y}_2^\mathrm{T}, \cdots, \mathbf{y}_T^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{TM\times 1}$ given by

$$\mathbf{y}_\mathrm{c} = \sqrt{\rho}\mathbf{F}_\mathrm{c}\mathbf{g} + \mathbf{F}_\mathrm{d}\mathbf{n}_\mathrm{c}, \tag{4.2}$$

where the pilot power is identical for each training period, $\mathbf{F}_\mathrm{c} = [\mathbf{F}_1^\mathrm{T}, \cdots, \mathbf{F}_T^\mathrm{T}]^\mathrm{T}$, $\mathbf{F}_\mathrm{d} = \mathrm{blkdiag}\{\mathbf{F}_1, \cdots, \mathbf{F}_T\}$, and $\mathbf{n}_\mathrm{c} = [\mathbf{n}_1^\mathrm{T}, \cdots, \mathbf{n}_T^\mathrm{T}]^\mathrm{T}$.

We consider a popular MIMO channel model with spatial correlation, which is a typical case in most MIMO scenarios due to the limited numbers of incident paths and angular spreads on BS [80], expressed by

$$\mathbf{g} = \mathbf{R}^{1/2}\mathbf{h}, \tag{4.3}$$

where $\mathbf{h} \in \mathbb{C}^{M\times 1}$ is the small-scale fading of the channel with elements being independent and identically distributed (i.i.d.) Gaussian random variables (RVs), i.e., $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, and $\mathbf{R} = \mathbb{E}\{\mathbf{g}\mathbf{g}^\mathrm{H}\}$ denotes the slowly varying covariance matrix of the channel which is assumed to be known by the BS. We normalize $\mathrm{tr}\,(\mathbf{R}) = M$ without loss of generality and the estimation of the covariance matrix in practice will be discussed in Section 4.4.

4.3 Channel Estimation with Hybrid Structure

For channel estimation at the BS, the task is to obtain the channel estimate in the MMSE sense with the designed phase-shifting matrix. By employing the MMSE aimed linear estimator, the channel estimation problem can be formulated as follows:

$$\begin{aligned} \underset{\mathbf{F}_\mathrm{c},\mathbf{W}}{\text{minimize}} \quad & \mathbb{E}\left\{\|\mathbf{g}-\hat{\mathbf{g}}\|^2\right\} \\ \text{subject to:} \quad & \hat{\mathbf{g}} = \mathbf{W}\mathbf{y}_\mathrm{c} \\ & \mathbf{y}_\mathrm{c} = \sqrt{\rho}\mathbf{F}_\mathrm{c}\mathbf{g} + \mathbf{F}_\mathrm{d}\mathbf{n}_\mathrm{c} \\ & \mathbf{F}_i \in \mathcal{F}, i = 1, \cdots, T \end{aligned} \tag{4.4}$$

where $\mathbf{W}$ represents the baseband processing matrix, and $\mathcal{F}$ denotes the set of all feasible phase-only RF combiners. In order to minimize MSE, the optimal solution to $\mathbf{W}$ is the well-known Weinner filter [59] given by

$$\mathbf{W} = \sqrt{\rho}\mathbf{R}\mathbf{F}_\mathrm{c}^\mathrm{H}\left(\rho\mathbf{F}_\mathrm{c}\mathbf{R}\mathbf{F}_\mathrm{c}^\mathrm{H} + \mathbf{R}_{\mathbf{F}_\mathrm{d}}\right)^{-1}, \tag{4.5}$$

where $\mathbf{F}_\mathrm{c}$ is assumed to be known as a prior and $\mathbf{R}_{\mathbf{F}_\mathrm{d}} \triangleq \mathbf{F}_\mathrm{d}\mathbf{F}_\mathrm{d}^\mathrm{H}$ for notation simplification. Using (4.5), the objective function in (4.4) equals

$$\begin{aligned} \mathrm{MSE} =& \mathbb{E}\left\{\|\mathbf{g}-\hat{\mathbf{g}}\|^2\right\} \\ =& \mathrm{tr}\left(\mathbb{E}\left\{(\mathbf{g}-\hat{\mathbf{g}})(\mathbf{g}-\hat{\mathbf{g}})^\mathrm{H}\right\}\right) \\ =& \mathrm{tr}\left(\mathbf{R} - \rho\mathbf{R}\mathbf{F}_\mathrm{c}^\mathrm{H}\left(\rho\mathbf{F}_\mathrm{c}\mathbf{R}\mathbf{F}_\mathrm{c}^\mathrm{H} + \mathbf{R}_{\mathbf{F}_\mathrm{d}}\right)^{-1}\mathbf{F}_\mathrm{c}\mathbf{R}\right) \\ =& \mathrm{tr}\left(\left(\mathbf{R}^{-1} + \rho\mathbf{F}_\mathrm{c}^\mathrm{H}\mathbf{R}_{\mathbf{F}_\mathrm{d}}^{-1}\mathbf{F}_\mathrm{c}\right)^{-1}\right), \end{aligned} \tag{4.6}$$

where the last equality holds due to the Woodbury matrix identity [90] and it is assumed that $\mathbf{R}$ is in full rank. If $\mathbf{R}$ is rank deficient, remedies similar to the one discussed in [60] can be applied which does not affect the MSE. Combining (4.4)–(4.6), the primal channel estimation problem is readily equivalent to the following problem as

$$\begin{aligned} \underset{\mathbf{F}_\mathrm{c}}{\text{minimize}} \quad & \mathrm{tr}\left(\left(\mathbf{R}^{-1} + \rho\mathbf{F}_\mathrm{c}^\mathrm{H}\mathbf{R}_{\mathbf{F}_\mathrm{d}}^{-1}\mathbf{F}_\mathrm{c}\right)^{-1}\right) \\ \text{subject to:} \quad & \mathbf{F}_i \in \mathcal{F}, i = 1, \cdots, T. \end{aligned} \tag{4.7}$$

The optimization problem (4.7) aims to find the optimal and feasible RF combiners $\mathbf{F}_i$ ($i = 1, \cdots, T$) for the T-length training sequence. However, it is difficult to directly obtain the optimal solution in either closed-form expressions or through numerical approaches without sustainable resource consumption due to the non-convex constraint.

To facilitate further analysis, we temporarily drop the constant-magnitude constraint in the following design. This is a good way to begin with because it is capable to derive some closed-form expressions and provide guidelines for the design of phase-only combiners. Note that the design of phase-only combiners will be considered in Section 4.3.3.

Without the phase-only constraint, it can be shown that the optimal RF unconstrained combiner has a structure characterized in the following lemma.

Lemma 4.1. *For any optimal solution to the unconstrained RF combiners given their singular value decompositions (SVD) as*

$$\mathbf{F}_i^* = \mathbf{U}_i[\mathbf{\Sigma}_i\ \mathbf{0}][\mathbf{V}_{i,\mathrm{L}}\ \mathbf{V}_{i,\mathrm{R}}]^{\mathrm{H}}, \tag{4.8}$$

it is safely to reconstruct the optimal combiners according to

$$\mathbf{F}_i^{\mathrm{opt}} = \mathbf{V}_{i,\mathrm{L}}^{\mathrm{H}}, \tag{4.9}$$

which yields the same MSE performance as $\mathbf{F}_i^*$ *does, where* $i = 1, \cdots, T$

Proof: See Appendix 4.7.1 ■

Lemma 4.1 indicates that the optimal unconstrained RF combiner $\mathbf{F}_i^{\mathrm{opt}}$ can always be restricted as row-unitary and is thus *independent* of both $\mathbf{U}_i$ and $\mathbf{\Sigma}_i$. Hence, we can set both of them to be identity matrices and equivalently simplify the MSE in (4.6) as

$$\mathrm{MSE} = \mathrm{tr}\left(\left(\mathbf{\Lambda}^{-1} + \rho\sum_{i=1}^{T}\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^{\mathrm{H}}\right)^{-1}\right), \tag{4.10}$$

where $\tilde{\mathbf{V}}_i \triangleq \mathbf{U}^{\mathrm{H}}\mathbf{V}_{i,\mathrm{L}}$, $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \cdots, \lambda_M\}$ and $\mathbf{U}$ are the eigenvalues and corresponding eigenvectors of $\mathbf{R}$, i.e., $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}$, respectively. Without loss of generality, it is assumed that the eigenvalues are arranged in the decreasing order, i.e., $\lambda_1 \geq \cdots \geq \lambda_M$, and hence $\sum_{i=1}^{M}\lambda_i = \mathrm{tr}(\mathbf{R}) = M$. In addition, there exists $\tilde{\mathbf{V}}_i^{\mathrm{H}}\tilde{\mathbf{V}}_i = \mathbf{I}_L$ due to the unitary property of both $\mathbf{U}$ and $\mathbf{V}_{i,\mathrm{L}}$. To design the optimal RF combiners, we first investigate the single-pilot strategy, i.e., $T = 1$, and then extend to the multiple pilot trainings where $T > 1$.

4.3.1 Optimal Combiner Design of Single Training

When $T = 1$, the MSE expressed in (4.10) reduces to

$$\text{MSE} = \text{tr}\left(\left(\mathbf{\Lambda}^{-1} + \rho\tilde{\mathbf{V}}_1\tilde{\mathbf{V}}_1^{\text{H}}\right)^{-1}\right) \tag{4.11a}$$

$$= \text{tr}\left(\mathbf{\Lambda} - \mathbf{\Lambda}\tilde{\mathbf{V}}_1\left(\tilde{\mathbf{V}}_1^{\text{H}}\mathbf{\Lambda}\tilde{\mathbf{V}}_1 + \rho^{-1}\mathbf{I}_M\right)^{-1}\tilde{\mathbf{V}}_1^{\text{H}}\mathbf{\Lambda}\right) \tag{4.11b}$$

$$= \text{tr}(\mathbf{\Lambda}) - \text{tr}\left(\left(\tilde{\mathbf{V}}_1^{\text{H}}\left(\mathbf{\Lambda} + \rho^{-1}\mathbf{I}_M\right)\tilde{\mathbf{V}}_1\right)^{-1}\tilde{\mathbf{V}}_1^{\text{H}}\mathbf{\Lambda}^2\tilde{\mathbf{V}}_1\right), \tag{4.11c}$$

where (4.11b) follows from the Woodbury matrix identity [90] and (4.11c) is derived on the fact that $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$. From (4.11), the design of the unconstrained optimal RF combiner in (4.7) can be reformulated as the following optimization problem:

$$\begin{aligned} \underset{\tilde{\mathbf{V}}_1}{\text{maximize}} \quad & \text{tr}\left(\left(\tilde{\mathbf{V}}_1^{\text{H}}\left(\mathbf{\Lambda} + \rho^{-1}\mathbf{I}_M\right)\tilde{\mathbf{V}}_1\right)^{-1}\tilde{\mathbf{V}}_1^{\text{H}}\mathbf{\Lambda}^2\tilde{\mathbf{V}}_1\right) \\ \text{subject to:} \quad & \tilde{\mathbf{V}}_1^{\text{H}}\tilde{\mathbf{V}}_1 = \mathbf{I}_L. \end{aligned} \tag{4.12}$$

To solve (4.12), we employ the *block generalized Rayleigh quotient* which is presented as Lemma 4.2 in Appendix 4.7.2. Applying Lemma 4.2, the objective function of (4.12) is the *block generalized Rayleigh quotient* of $\tilde{\mathbf{V}}_1$ with respect to the pencil $(\mathbf{\Lambda}^2, \mathbf{\Lambda} + \rho^{-1}\mathbf{I}_M)$. The property of *block generalized Rayleigh quotient* described by Lemma 4.2 indicates that the maximizer of (4.12) is the matrix spanned by the L generalized eigenvectors corresponding to the L largest eigenvalues of the pencil $(\mathbf{\Lambda}^2, \mathbf{\Lambda} + \rho^{-1}\mathbf{I}_M)$ if the matrix happens to satisfy the constraint in (4.12). As $\mathbf{\Lambda}^2$ and $\mathbf{\Lambda} + \rho^{-1}\mathbf{I}_M$ are both diagonal, the generalized eigenvalues and corresponding eigenvectors can be directly obtained as $\tilde{\mathbf{\Lambda}} = \left[\tilde{\lambda}_1, \cdots, \tilde{\lambda}_M\right]$ and $\tilde{\mathbf{U}} = \mathbf{I}_M$, respectively, where $\tilde{\lambda}_i = \lambda_i^2/(\lambda_i + \rho^{-1})$ $(i = 1, \cdots, M)$ and $\tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_M$ according to the decreasing order of λ_i. Therefore, the maximizer of (4.12) is constructed by the eigenvectors corresponding to the largest L eigenvalues of the pencil $(\mathbf{\Lambda}^2, \mathbf{\Lambda} + \rho^{-1}\mathbf{I}_M)$, denoted as

$$\tilde{\mathbf{V}}_1^{\text{opt}} = \tilde{\mathbf{U}}_{[1:L]} = \mathbf{I}_{M[1:L]}, \tag{4.13}$$

where $(\tilde{\mathbf{V}}_1^{\text{opt}})^{\text{H}}\tilde{\mathbf{V}}_1^{\text{opt}} = \mathbf{I}_L$ satisfies the constraint in (4.12). Now, we can arrive at the following Theorem 4.1 which presents the optimal design of the unconstrained single-training RF combiner.

Theorem 4.1. *The optimal unconstrained RF combiner for the single pilot strategy can be designed as follows:*

$$\mathbf{F}_1^{\text{opt}} = (\mathbf{U}_{[1:L]})^{\text{H}}, \tag{4.14}$$

where $\mathbf{U}$ *is the matrix spanned by the eigenvectors of* $\mathbf{R}$.

Proof: By substituting (4.13) into (4.9) with the transition defined by $\tilde{\mathbf{V}}_i \triangleq \mathbf{U}^{\mathrm{H}}\mathbf{V}_{i,\mathrm{L}}$, the optimal unconstrained single-training RF combiner can be obtained. ■

Theorem 4.1 implies that the RF combiner should receive the training pilots along the largest L dominant eigen-directions of $\mathbf{R}$ to minimize the MSE of channel estimations. More explicitly, we substitute (4.13) into (4.11a) and obtain the optimal MSE as follows:

$$\begin{aligned}
\mathrm{MSE} =& \mathrm{tr}\left(\left(\mathbf{\Lambda}^{-1} + \rho\begin{bmatrix}\mathbf{I}_L & \\ & \mathbf{0}_{M-L}\end{bmatrix}\right)^{-1}\right) \\
=& \mathrm{tr}\left(\begin{bmatrix}\mathbf{\Lambda}^{-1}_{[1:L]} + \rho\mathbf{I}_L & \\ & \mathbf{\Lambda}^{-1}_{[L+1:M]}\end{bmatrix}^{-1}\right) \\
=& \sum_{l=1}^{L}\frac{\lambda_l}{1+\rho\lambda_l} + \sum_{l=L+1}^{M}\lambda_l \\
=& M - \sum_{l=1}^{L}\lambda_l + \sum_{l=1}^{L}\frac{\lambda_l}{1+\rho\lambda_l} \qquad (4.15a) \\
=& M - \sum_{l=1}^{L}\frac{\lambda_l^2}{\lambda_l + 1/\rho}, \qquad (4.15b)
\end{aligned}$$

where (4.15a) is obtained by applying the power constraint defined by $\sum_{i=1}^{M}\lambda_i = M$. Prior to stating more insights obtained from (4.15b), we define a useful concept of one channel being *more spatially correlated* than another in the following way.

Definition 4.1. Let channel vectors $\mathbf{g}_1 \in \mathbb{C}^M$ and $\mathbf{g}_2 \in \mathbb{C}^M$ have covariance matrices $\mathbf{R}_1 \in \mathbb{R}^{M\times M}$ and $\mathbf{R}_2 \in \mathbb{R}^{M\times M}$, respectively. We say $\mathbf{g}_1$ is *more spatially correlated* than $\mathbf{g}_2$ if and only if $\boldsymbol{\lambda}_1 \succ \boldsymbol{\lambda}_2$, where $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are composed of the eigenvalues sorted in a descending order of $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively.

By Definition 4.1 and observing (4.15b), we can conclude the following corollaries on the design insights of the optimal unconstrained RF combiner revealed in Theorem 4.1.

Corollary 4.1. *Given the number of RF chains and pilot power fixed, the MSE of channel estimation with the optimal unconstrained single-training RF combiner decreases if the channel is more spatially correlated.*

Proof: See Appendix 4.7.3. ■

Corollary 4.2. *Given the channel correlation and pilot power fixed, the MSE of channel estimation with the optimal unconstrained single-training RF combiner decreases with more RF chains deployed.*

Corollary 4.3. *Given the channel correlation and number of RF chains fixed, the MSE of channel estimation with the optimal unconstrained single-training RF combiner decreases with increasing pilot power.*

Corollary 4.2 and 4.3 can be directly proved by checking the fact that the MSE in (4.15) monotonously decreases with the increase of ρ and/or L.

4.3.2 Combiner Design of Multiple Trainings

In a typical full-chain MIMO system, the M-dimensional signal can be observed at baseband to estimate the M-dimensional MIMO channels. However, the hybrid structure MIMO system can only capture L-dimensional signal from the phase-shifting network fed by M antennas for each training. Note that $L < M$ in the limited RF chain scenario. Hence, the observed low-dimensional signal is not sufficient to recover the high-dimensional channel information for a single training. To approach the performance of full-chain channel estimation, multiple trainings can be employed to achieve the DoF of full-chain baseband signal measurements, namely, TL observations can be utilized to estimate the M-dimensional MIMO channels. Typically, it is assumed that $TL \leq M$. Therefore, the RF combiner for each training phase needs to be properly designed to capture the channel energy for channel recovery as accurate as possible.

In this subsection, we investigate the design of unconstrained combiners for multiple trainings, i.e., $\mathbf{F}_1, \cdots, \mathbf{F}_T$. Intuitively, the multiple trainings can be performed by simply repeating the single-pilot training for multiple times. Inspired by Theorem 4.1 where the most dominated L eigen-directions of $\mathbf{R}$ are selected to construct the combiner, it is heuristic to select the most dominated TL eigen-directions to establish the T combiners for multiple trainings, namely, the *Block Selection* method, where the RF combiner for the ith ($i = 1, 2, \cdots, T$) training can be composed of the ith most dominated L-eigenvector block of $\mathbf{R}$, as expressed by

$$\mathbf{F}_i^{\text{opt}} = (\mathbf{U}_{[(i-1)L+1:iL]})^{\text{H}}. \tag{4.16}$$

Evidently, the *Block Selection* method has low complexity, yet leading to non-optimal performance.

In order to improve the performance of channel estimation with multiple trainings, we formulate the optimization problem to design combiners for multiple trainings by

recalling (4.10) shown as follows:

$$\begin{aligned} \underset{\tilde{\mathbf{V}}_i}{\text{maximize}} \quad & \text{MSE} \\ \text{subject to:} \quad & \tilde{\mathbf{V}}_i^{\text{H}}\tilde{\mathbf{V}}_i = \mathbf{I}_L, \quad i = 1, \cdots, T \end{aligned} \tag{4.17}$$

where the MSE of the objective function is expressed by (4.10). The designed $\tilde{\mathbf{V}}_i$ can be employed to construct the RF combiners according to Theorem 4.1. Hence, it is necessary to find the optimizer of (4.17) to design RF combiners for multiple trainings. Dealing with the optimization problem, we propose two methods to solve it in the following part of this section.

a) Sequential Optimization

It has been investigated in Subsection 4.3.1 that the closed-form expression of the optimal combiner can be obtained in the single-training scenario. For multiple trainings, however, it is difficult to obtain the global optimal solution directly. Nevertheless, we propose a sequential approach, namely, the *Sequential Optimization* (short as *Sequential*) method, to minimize MSE step-wisely when $T > 1$. To illustrate the *Sequential Optimization* method, recall (4.10) and reformulate it in the following form:

$$\text{MSE} = \text{tr}\left(\left(\mathbf{\Gamma}_T^{-1} + \rho\tilde{\mathbf{V}}_T\tilde{\mathbf{V}}_T^{\text{H}}\right)^{-1}\right), \tag{4.18}$$

where

$$\mathbf{\Gamma}_i^{-1} = \begin{cases} \mathbf{\Gamma}_{i-1}^{-1} + \rho\tilde{\mathbf{V}}_{i-1}\tilde{\mathbf{V}}_{i-1}^{\text{H}}, & i > 1 \\ \mathbf{\Lambda}^{-1}, & i = 1. \end{cases} \tag{4.19}$$

Following the iterative definition of Γ_i^{-1} denoted by (4.19), the MSE can be minimized by the *Sequential Optimization* method as follows.

Firstly, by setting $T = 1$ the problem reduces to the single-pilot training case which can be solved as presented in Section 4.3.1.

Subsequently, with $T = 2$, the MSE denoted by (4.18) can be reformulated as

$$\text{MSE} = \text{tr}\left(\left(\mathbf{\Gamma}_2^{-1} + \rho\tilde{\mathbf{V}}_2\tilde{\mathbf{V}}_2^{\text{H}}\right)^{-1}\right). \tag{4.20}$$

By applying the same manipulations as in Section 4.3.1, the optimal $\tilde{\mathbf{V}}_2$ to minimize the MSE expressed by (4.20) can be obtained. Similar to (4.13), the optimal $\tilde{\mathbf{V}}_2$, say $\tilde{\mathbf{V}}_2^{\text{opt}}$, is constructed by the most dominated L eigen-directions of the pencil $(\mathbf{\Gamma}_2^2, \mathbf{\Gamma}_2 + \rho^{-1}\mathbf{I}_M)$. In other words, $\tilde{\mathbf{V}}_2^{\text{opt}}$ is combined by L generalized eigenvectors corresponding to the L

largest generalized eigenvalues of $(\mathbf{\Gamma}_2^2, \mathbf{\Gamma}_2 + \rho^{-1}\mathbf{I}_M)$. For ease of exposition, we denote $\mathbf{\Gamma}_2 = \mathrm{diag}\{\gamma_{1,2}, \cdots, \gamma_{M,2}\}$. The generalized eigenvalues and eigenvectors of pencil $(\mathbf{\Gamma}_2^2, \mathbf{\Gamma}_2 + \rho^{-1}\mathbf{I}_M)$ can be obtained as $\tilde{\boldsymbol{\gamma}}_2 = [\tilde{\gamma}_{1,2}, \cdots, \tilde{\gamma}_{M,2}]^{\mathrm{T}}$ and $\mathbf{I}_M$, respectively, where $\tilde{\gamma}_{j,2} = \gamma_{j,2}^2/(\gamma_{j,2} + \rho^{-1})$ for $j = 1, \cdots, M$. Therefore,

$$\tilde{\mathbf{V}}_2^{\mathrm{opt}} = \mathbf{I}_{M[j_1, \cdots, j_L]}, \tag{4.21}$$

where the indices $[j_1, \cdots, j_L]$ are the first L numbers of $[j_1, \cdots, j_M]$ which follow from the descending order of $\tilde{\gamma}_{i,2}$ shown as $\tilde{\gamma}_{j_1,2} \geq \cdots \geq \tilde{\gamma}_{j_M,2}$.

In the similar manner as $T = 2$, we can obtain each $\tilde{\mathbf{V}}_T^{\mathrm{opt}}$ for any $T > 2$ successively according to the generalized eigenvalues and eigenvectors of pencil $(\mathbf{\Gamma}_T^2, \mathbf{\Gamma}_T + \rho^{-1}\mathbf{I}_M)$.

Finally, the combiners are derived by applying the definition of $\tilde{\mathbf{V}}_i$ and Lemma 4.1. For instance, the optimal $\mathbf{F}_{2,\mathrm{opt}}$ can be expressed by

$$\mathbf{F}_2^{\mathrm{opt}} = \mathbf{U}_{[j_1, \cdots, j_L]}^{\mathrm{H}}. \tag{4.22}$$

By reviewing (4.19) and the *Sequential Optimization* method, a natural question arises here: Whether $\mathbf{\Gamma}_i$ is invertible or not? To check this question, we start with considering a two-stage training with $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$. Note that the known covariance matrix $\mathbf{R}$ is assumed to be symmetric and full-rank[1], thus it is positive definite with all eigenvalues positive. Hence, $\mathbf{\Gamma}_1$ is composed of positive diagonal elements which indicates that it is invertible. By setting $T = 2$ in (4.19), we have $\mathbf{\Gamma}_2^{-1} = \mathbf{\Gamma}_1^{-1} + \rho\tilde{\mathbf{V}}_1\tilde{\mathbf{V}}_1^{\mathrm{H}}$. In addition to positive diagonal $\mathbf{\Gamma}_1$, the optimal $\tilde{\mathbf{V}}_1$ solved in Section 4.3.1 is shown to be $\mathbf{I}_{M[1:L]}$. Hence, it is guaranteed that $\mathbf{\Gamma}_1^{-1} + \rho\tilde{\mathbf{V}}_1\tilde{\mathbf{V}}_1^{\mathrm{H}}$ is always diagonal with positive entries for $\rho > 0$. So is $\mathbf{\Gamma}_2$ invertible. Using the similar manipulation, one can successively verify that $\mathbf{\Gamma}_{T-1}$ is composed of positive diagonal entries and hence $\mathbf{\Gamma}_T$ is diagonal and invertible for any $T > 2$.

Remark 4.1. In the aforementioned illustrations, each $\tilde{\mathbf{V}}_i^{\mathrm{opt}}$ $(i = 1, \cdots, T)$ during the *Sequential Optimization* has the same structure in which they are constructed by column-wise permutations of $\mathbf{I}_M$ which denotes the set of generalized eigenvectors of pencil $(\mathbf{\Gamma}_i^2, \mathbf{\Gamma}_i + \rho^{-1}\mathbf{I}_M)$, and the selected columns correspond to the L largest generalized eigenvalues of this pencil. Accordingly, each combiner for the multiple trainings, namely $\mathbf{F}_i^{\mathrm{opt}}$, is the Hermitian of the matrix constructed by the consistent column-wise permutations of $\mathbf{U}$ as the construction of $\tilde{\mathbf{V}}_i^{\mathrm{opt}}$. Note that the eigenvectors of the above pencil always compose an identity matrix due to diagonal $\mathbf{\Gamma}_i$.

Remark 4.2. Inevitably, the optimum $\tilde{\mathbf{V}}_i^{\mathrm{opt}}$ $(i = 1, \cdots, T)$ obtained by the *Sequential Optimization* method is not a global minimizer of the MSE in (4.10). Nevertheless, it

[1] If $\mathbf{R}$ is rank deficient, the method mentioned below (4.6) can be employed for adjustment.

minimizes the MSE reformulated by (4.18) for each step of iterations, namely the step-wise minimizer. The performance of such sub-optimal combiners will be evaluated by numerical results in Section 4.5.

b) Alternating Optimization

For further performance enhancement, we solve (4.17) via joint optimization over $\tilde{\mathbf{V}}_i$'s. It in many cases can achieve near optimal performance via the alternating optimization to solve non-convex problems. More specifically, we consider fixing all the other $\tilde{\mathbf{V}}_i$ $(i \neq k)$ while optimizing only a single $\tilde{\mathbf{V}}_k$. And then, iterations are taken to update each $\tilde{\mathbf{V}}_i$ $(i = 1, \cdots, T)$ alternatively until convergence. This solution is named as *Alternating Optimization* (short as *Alternating*) throughout this chapter.

For an explicit exposure of semi-joint selection method, we reformulate the MSE expressed by (4.10) to separate $\tilde{\mathbf{V}}_k$ from the sum as given by

$$\mathrm{MSE}(\tilde{\mathbf{V}}_k) = \mathrm{tr}\left(\left(\mathbf{Q}_k^{-1} + \rho\tilde{\mathbf{V}}_k\tilde{\mathbf{V}}_k^{\mathrm{H}}\right)^{-1}\right), \quad k = 1, \cdots, T \tag{4.23}$$

where $\mathbf{Q}_k^{-1} = \mathbf{\Lambda}^{-1} + \rho\sum_{i=1, i\neq k}^{T}\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^{\mathrm{H}}$. With given values of $\tilde{\mathbf{V}}_i$ $(i \neq k)$, $\tilde{\mathbf{V}}_k$ can be solved in the same manner as Subsection 4.3.1 where Lemma 4.2 is applied. In this chapter, the relative MSE increment of each iteration round, namely,

$$\epsilon_n = \left|\mathrm{MSE}\left(\tilde{\mathbf{V}}_k^{(n)}\right) - \mathrm{MSE}\left(\tilde{\mathbf{V}}_k^{(n-1)}\right)\right| / \mathrm{MSE}\left(\tilde{\mathbf{V}}_k^{(n-1)}\right),$$

will be taken as the error measurement. The iterations continue until ϵ_n falls below a prescribed tolerance ϵ and the last iterate $\tilde{\mathbf{V}}_k^{(n)}$ is taken as the solution.

Finally, the unconstrained near-optimal RF combiners are computed from the solved $\tilde{\mathbf{V}}_i^{\mathrm{opt}}$ $(i = 1, \cdots, T)$ by applying Lemma 4.1. The alternating optimization method is summarized step-by-step in Algorithm 2.

4.3.3 Design of Phase-only RF Combiners

In the previous subsections, we first omit the phase-only constraints when designing the RF combiners. However, the RF combiners are implemented using phase shifters in practical hybrid systems, which can only perform phase rotations on the received RF signals. Hence, the magnitude of each element of a RF combiner matrix must keep constant. In this subsection, we propose a method to design the phase-only RF combiners based on the unconstrained optimal ones. Moreover, the performance loss incurred by the constant-magnitude constraint is characterized by numerical evaluation shown in part c) of Subsection 4.5.1.

Algorithm 2 Alternating Optimization to Design Optimal Unconstrained Combiners for Multiple Trainings

Input: M, T and $\mathbf{R}$

Output: Optimal unconstrained RF combiners for multiple trainings, $\mathbf{F}_i^{\text{opt}}$ $(i = 1, \cdots, T)$

Initialize $n = 0$ and $\tilde{\mathbf{V}}_i^{(0)}$ for $i = 1, \cdots, T$
loop
  Use $\tilde{\mathbf{V}}_i^{(n)}$ $(i \neq k)$ to obtain $\tilde{\mathbf{V}}_k^{(n)}$ which minimizes MSE expressed by (4.23) for $k = 1, \cdots, T$ sequentially and update $\tilde{\mathbf{V}}_k^{(n)}$ with the new solved values
  if $\frac{\left|\text{MSE}\left(\tilde{\mathbf{V}}_k^{(n)}\right) - \text{MSE}\left(\tilde{\mathbf{V}}_k^{(n-1)}\right)\right|}{\text{MSE}\left(\tilde{\mathbf{V}}_k^{(n-1)}\right)} < \epsilon$ **then**
    Stop **loop**
  else
    Update n by 1 (i.e., $n = n + 1$)
  end if
end loop
Compute $\mathbf{F}_i^{\text{opt}}$ with $\tilde{\mathbf{V}}_i^{\text{opt}}$ by applying Lemma 4.1
Output $\mathbf{F}_i^{\text{opt}}$

Inspired by the fact that the phase shifters only perform phase adjustments on the received signal, we take out the phase of each unconstrained combiner and construct the constant-magnitude RF combiners according to

$$\mathbf{F}_k^{(i,j)} = e^{\mathrm{j}\phi_k^{(i,j)}}, \tag{4.24}$$

where $\phi_k^{(i,j)}$ denotes the phase of the (i,j)th element of $\mathbf{F}_{k,\text{opt}}$, i.e., $\mathbf{F}_{k,\text{opt}} = a_k^{(i,j)} e^{\mathrm{j}\phi_k^{(i,j)}}$.

It is interesting to find that such heuristic phase-only RF combiners achieve desirable performance via simulation verifications in Section 4.5. Besides its effectiveness, the very low complexity is another advantage of this design. Furthermore, the performance of the phase-only combiners designed above will be examined and analyzed in part c) and d) of Subsection 4.5.1 in terms of both MSE of channel estimation and spectral efficiency in hybrid precoded multiuser massive MIMO communications.

4.4 Spatial Correlations Estimation by Covariance Matching

In the proposed channel estimation, the spatial correlation is assumed to be known by the BS. Practically, the BS has to estimate the covariance matrix which characterizes the channel correlation. However, it is infeasible to estimate the covariance matrix $\mathbf{R}$ in the hybrid precoding system due to limited RF chains by employing conven-

tional approaches [91–94]. Dealing with the covariance estimation in the hybrid precoding system, an online method by employing a parallel switch network was proposed in [95], which increases the hardware cost in practice and consumes sustainable training resources. Specifically, we here propose a covariance matching method for channel covariance estimation which can be fortunately performed offline to rapidly decrease the estimation complexity compared to online methods.

In our previous descriptions, the proposed channel estimation scheme is based on the nonparametric channel mode denoted by (4.3) which consists of a covariance matrix and an uncorrelated Rayleigh fading channel vector. A theoretical channel correlation model, namely the exponential model, can be adopted to match the real covariance matrix, which is described in detail as follows.

Nonparametric Channel Model: The covariance matrix of the exponential model used to construct the nonparametric channel model denoted by (4.3) can be expressed by

$$[\mathbf{R}^{(a)}]_{m,n} = a^{|m-n|}, \tag{4.25}$$

where $[\mathbf{R}^{(a)}]_{m,n}$ denotes the (m,n)th element of $\mathbf{R}^{(a)}$ and $0 \leq a < 1$ is a real number that controls the correlation introduced to the channel model [80, 81]. Here, a larger a corresponds to more highly correlated channels. When $a = 0$, $\mathbf{g}$ degenerates to an i.i.d. Rayleigh fading channel.

Recalling Definition 4.1 and the proof of Corollary 4.1, the spatial correlation of channel vectors can be characterized by the eigenvalue distribution of the channel covariance matrix and analyzed with the majorization theory [86]. Hence, the eigenvalue distribution can be utilized to match the real channel covariance matrix to (4.25). To describe the matching method, we first introduce the parametric channel model as follows:

Parametric Channel Model: The parametric model for the channel considered in this chapter can be expressed according to

$$\mathbf{g} = \sqrt{\frac{M}{N_R}} \sum_{l=1}^{N_R} \alpha_l \mathbf{a}_{\mathrm{BS}}(\theta_l), \tag{4.26}$$

where N_R represents the number of scatter rays, α_l and θ_l denote the complex gain and azimuth AoA of the lth ray, respectively. Here, the complex gain $\alpha_l \sim \mathcal{CN}(0,1)$ and $\theta_l \in [0, 2\pi]$ follows the truncated Laplacian distribution with angular spread σ_{s} [17, 18]. Note that the MS is equipped with a single antenna which implies that the steering vector at the MS degenerates to a unit scalar while the BS is assumed to be equipped with a uniform linear array (ULA) of half wavelength spacing. The

steering vector at the BS can be expressed as

$$\mathbf{a}_{\mathrm{BS}}(\theta) \triangleq [1, e^{\mathrm{j}\pi \sin\theta}, \cdots, e^{\mathrm{j}\pi(M-1)\sin\theta}]^{\mathrm{T}}. \tag{4.27}$$

The channel generated according to (4.26) can be used to easily capture the scattering nature where the spatial correlation is determined by the angular spreading σ_{s}, i.e., the richer scattering corresponds to weaker antenna correlations and vice versa [96].

For a specific angular spreading σ_{s}, a large amount of channel vectors $\mathbf{g}_i$ $(i = 1, \cdots, N)$ can be generated and the maximum likelihood estimator of the covariance matrix is given by

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{g}_i \mathbf{g}_i^{\mathrm{H}}. \tag{4.28}$$

Thereafter, the eigenvalues of $\hat{\mathbf{R}}$ is calculated which can be matched to the exponential model by applying the least square error criterion as follows:

$$\hat{a} = \arg\min \|\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^{(a)}\|^2, \tag{4.29}$$

where $\hat{\boldsymbol{\lambda}} = [\hat{\lambda}_1, \cdots, \hat{\lambda}_M]^{\mathrm{T}}$ and $\boldsymbol{\lambda}^{(a)} = [\lambda_1^{(a)}, \cdots, \lambda_M^{(a)}]^{\mathrm{T}}$ denote the eigenvalue vectors of $\hat{\mathbf{R}}$ and $\mathbf{R}^{(a)}$, respectively. Hence, a mapping between the estimated $\hat{a}$ and corresponding σ_{s} can be established[2].

In practice, the angular spread of the incident rays can be estimated from their AoA [89], which can be obtained by angle quantization method proposed by [18]. And then, the corresponding $\hat{a}$ is utilized to generate the covariance matrix according to (4.25). The generated covariance matrix will be used for hybrid channel estimation. Note that the mapping between $\hat{a}$ and σ_{s} can be generated offline and used for online channel estimation which decreases the complexity and saves the hardware cost (no parallel switch-based RF chains) for covariance estimation compared to the online method [95]. In cellular communications, the angular spread characterizes the second-order statistic of the channel which usually varies at a much slower scale than fast fadings. Hence, the estimation of such information does not consume a substantial amount of resources.

4.5 Numerical and Simulation Results

In this section, we evaluate the performance of the proposed channel estimation scheme by employing both unconstrained and phase-only RF combiners to perform up-

[2]Note that the ULA channel model is adopted here for illustrative purpose. The method proposed in this section also applies to other structures of channel models such as the uniform planar array (UPA) model which is a better option for practical realizations.

link channel estimation in the massive MIMO system with the hybrid precoding structure. The evaluations are performed with nonparametric channel models followed by parametric ones.

In simulations, we consider both single-user and multi-user scenarios. In the single-user scenario, a single-antenna MS is served by a BS deployed in the hybrid analog/digital system structure which is shown by Fig. 4.1. While the multi-user scenario adopts K single-antenna MSs served by the BS with the same structure as that employed in the single-user case. By default, the BS is equipped with $M = 64$ antennas followed by $L = 8$ RF chains. The RF phase shifters are assumed to have continuous phase shifting within $[0, 2\pi)$. As the power of noise is normalized to 1, the transmission power ρ can be used to denote the received signal-to-noise ratio (SNR). Each simulation result is obtained by averaging over $10,000$ channel realizations.

4.5.1 Performance with Nonparametric Channel Model

In this subsection, the nonparametric channel model denoted by (4.3) with the exponential channel correlation model given by (4.25) is considered.

a) Single Training Performance Evaluation

Fig. 4.2 shows the normalized MSE performance of the hybrid channel estimations with the designed single-training unconstrained combiner at different numbers of RF chains. As a benchmark, the performance of the full-chain estimation is also presented. Note that the full-chain estimation is performed with single training throughout this section. From the comparison, it is evident to find that the MSE decreases with the increase of the pilot training power, which verifies Corollary 4.3. Moreover, with more RF chains equipped at the BS, the channel estimation performance is also improved, which confirms Corollary 4.2. Note that in Fig. 4.2, we also plot the estimation performance with RF combiners composed of the columns randomly selected from a discrete Fourier transform (DFT) matrix. Its poor performance as compared to the designed combiners proposed in this chapter justifies the necessity to properly design RF combiners and the effectiveness of our design. Furthermore, it is observed that the derived closed-form expressions are quite accurate in characterizing the MSE performance of the channel estimations throughout the whole SNR range, thus providing valuable guidelines in practical system designs. Finally, the figure also shows that the performance gap between the full-chain estimation and the limited-chain hybrid channel estimations is tolerably small at low SNR regions. There is only about 0.5 dB MSE gap between the hybrid 16-chain and the conventional full-chain estimations at SNR=0 dB, where three-quarter RF chains are saved with the hybrid structure. However, one may argue: how about the performance gap at

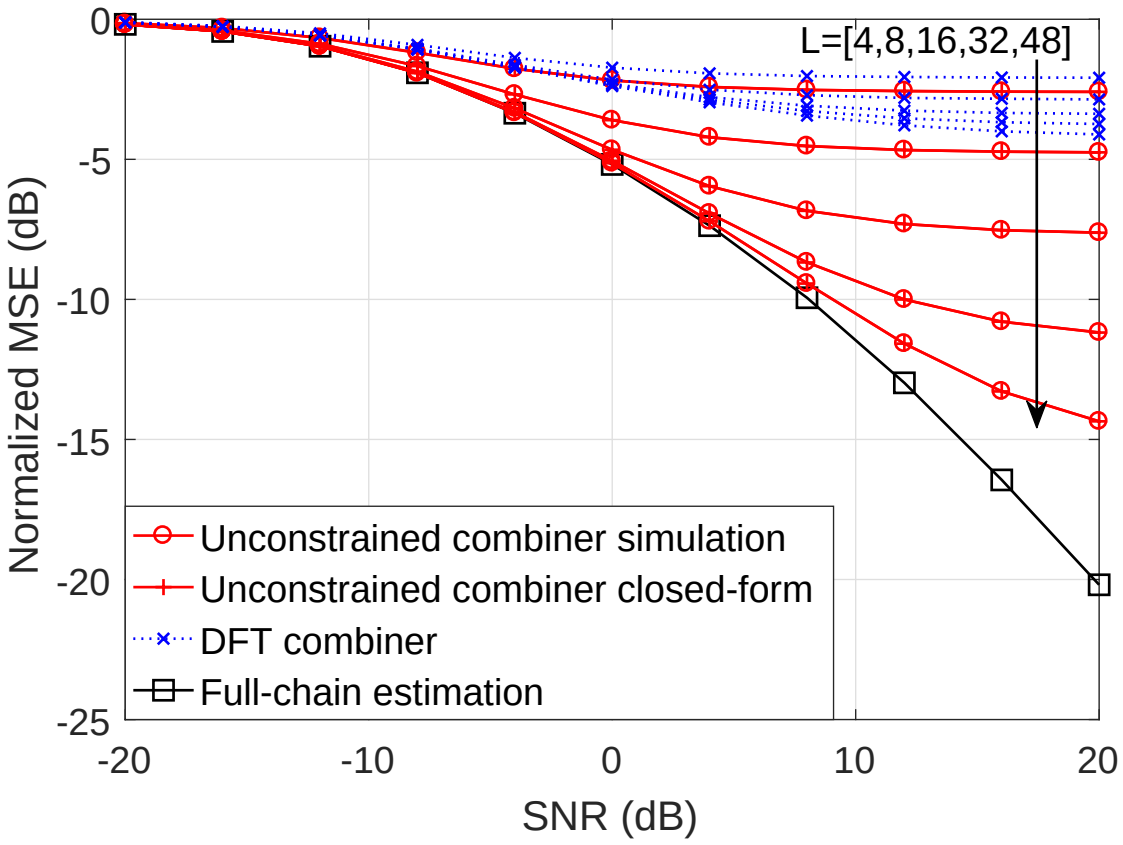


Figure 4.2: Performance comparison of the single-training hybrid channel estimation with different RF chains. ($M = 64$, $a = 0.8$)

high SNR regions? Due to limited chains, the single training is not sufficient to achieve the performance of full-chain channel estimation. To improve the performance of the hybrid channel estimations at full SNR regions, multiple trainings are introduced in this chapter and the performance is examined in the next subsection.

b) Multiple Trainings Performance Evaluation

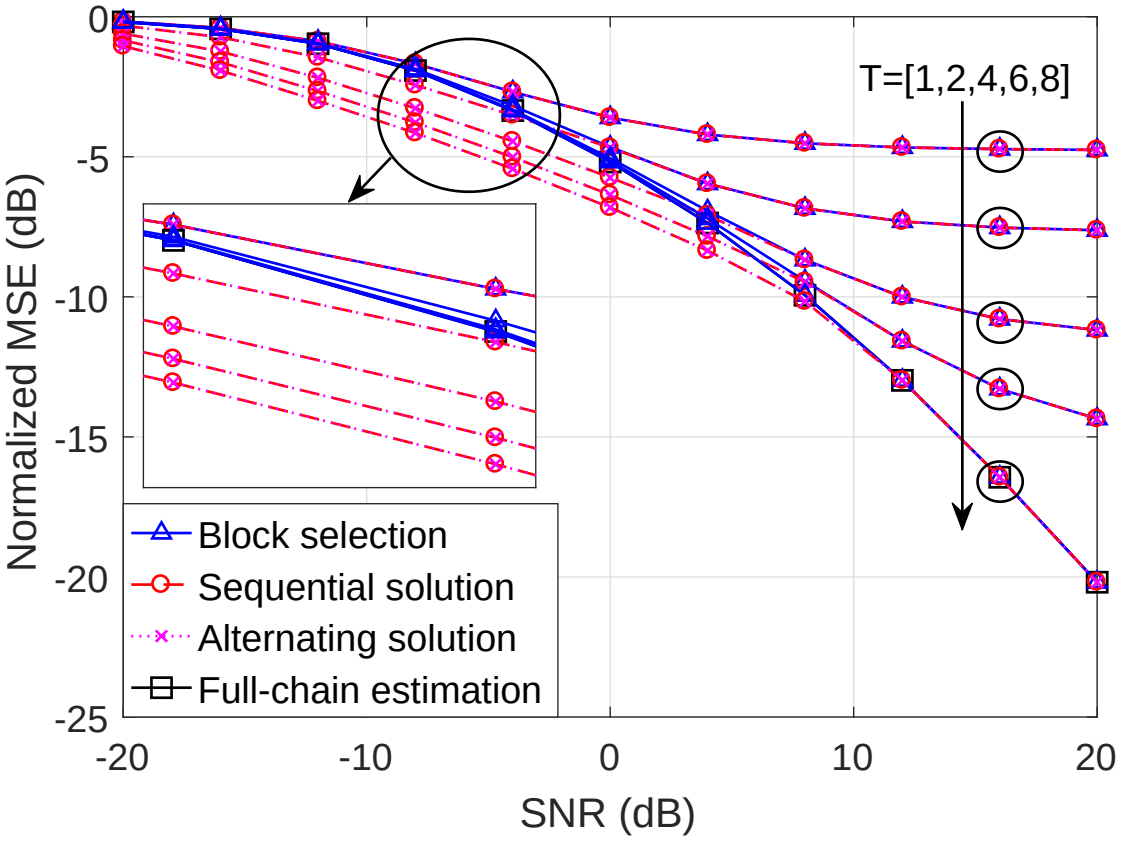


Figure 4.3: Performance comparison of the multiple-training hybrid channel estimation with different training times. ($M = 64$, $L = 8$, $a = 0.8$)

In this chapter, the design of unconstrained combiners for multiple trainings is formulated as (4.17), which are solved by both sequential and alternating approaches along side with the intuitive Block Selection method. The performance is examined in Fig. 4.3 with $T = 1, 2, 4, 6$ and 8. The figure shows that the normalized MSE of the hybrid channel estimation decreases as the increasing of training times due to the increased DoF of baseband observations. Additionally, it is interesting to find that the MSE performance with $L = 8$ and $T = 8$ almost achieves that of the full-chain estimation at high SNRs. This phenomenon verifies that the performance of full-chain channel estimation can be achieved by the limited-chain hybrid channel estimation with sufficient DoF of trainings, i.e., $T \times L = M$. In the low SNR regions, it is evident to find that even $T = 2$ trainings

can outperform the full-chain estimation. This phenomenon reveals the fact that, when the dimension of the dominated subspaces of the correlated channels is far less than M, the limited-chain estimation can suppress noise better than the full-chain one due to multiple trainings, which results in the distinct performance gap in the low SNR regions. On top of that, more performance differences can also be compared from the observation of Fig. 4.3. In the figure, the normalized MSE of the *Block Selection* method denoted by the triangle curve is always larger than or equal to that of the other two methods, i.e., the *Sequential* and *Alternating* solutions. The performance gap is particularly obvious at the low SNR regions. Such performance loss of *Block Selection* is caused by the fact that the noise is not considered when designing the combiners for the second and later pilot trainings. However, little performance gap can be observed between *Block Selection* and the others at high SNR regions due to the less significance of noise to channel estimation. By comparing the performances between the *Sequential* and *Alternating* solutions, it is evidenced that they achieve almost the same MSE no matter at high or low SNR regions which implies that the *Sequential* method can achieve the local optimum in solving (4.17). On the other hand, the lower designing complexity of the *Sequential* designates its superiority than the *Alternating* method.

c) Performance Evaluation of Phase-only RF Combiners

In the previous two subsections, the performance of the hybrid channel estimation with unconstrained RF combiners are evaluated. However, only phase shifters can be employed in the current practical applications with hybrid precoding structure, which means that only phase-only combiners are applicable in practice. In this subsection, we examine the performance in hybrid channel estimation with the phase-only combiners and compare it to that of the corresponding unconstrained ones, where the combiners for multiple trainings are designed by employing the *Sequential* method. In Fig. 4.4, the dashed lines marked by triangles denote the MSE performance of the hybrid channel estimation with phase-only combiners which are designed according to (4.24), where the performance of the unconstrained combiners are denoted by the circled solid lines. From comparisons, it is obvious that the performance loss caused by the phase-only combiners is tolerable at high SNR regions and is negligible throughout the low SNR regions. For instance, there is only 1.5 dB MSE loss at SNR=20 dB while 0.08 dB loss at SNR=-20 dB when $T = 8$. The figure also shows that the performance loss is smaller with fewer trainings, i.e., smaller T. Therefore, the hybrid channel estimation with phase-only combiners achieve the desirable performance measured by the normalized MSE. Furthermore, we also present the performance of hybrid channel estimation at different channel correlation shown in Fig. 4.5. It is evident that the normalized MSE of the hybrid

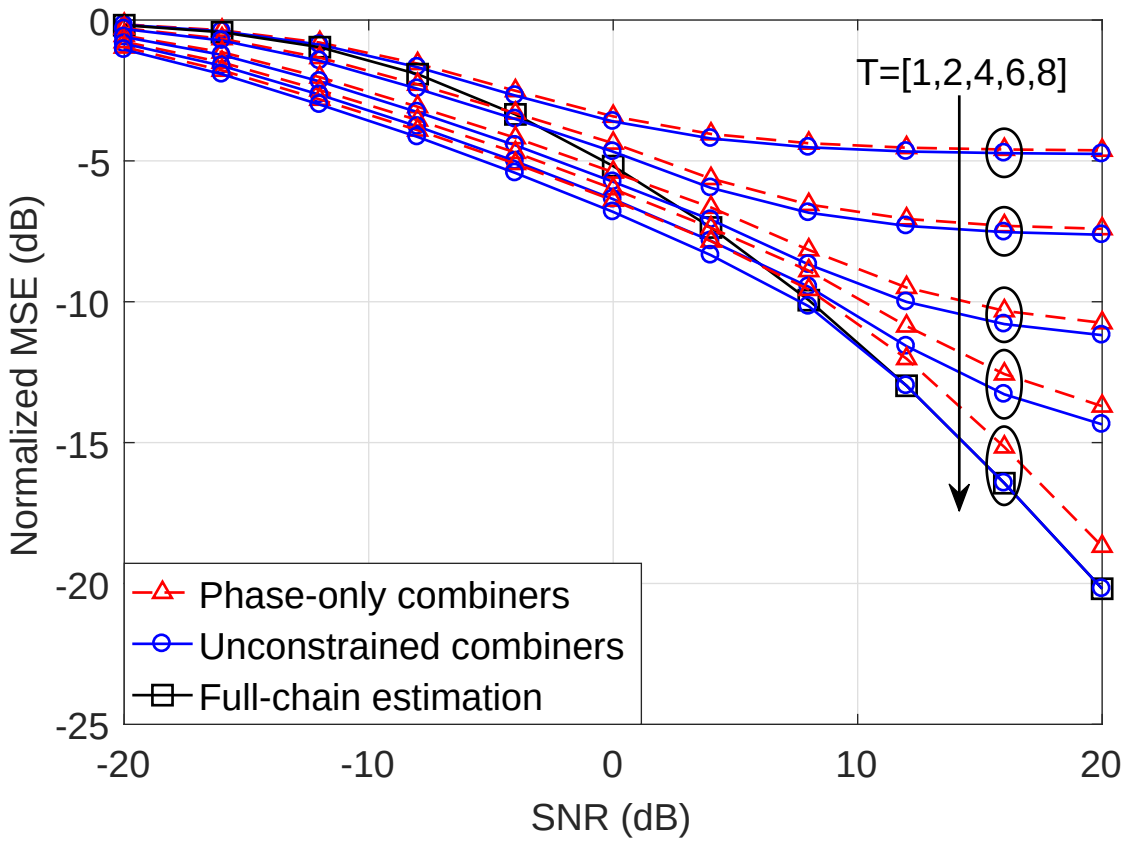


Figure 4.4: Performance comparison of channel estimation between unconstrained and phase-only combiners with the *Sequential* method under different pilot trainings. ($M = 64$, $L = 8$, $a = 0.8$)

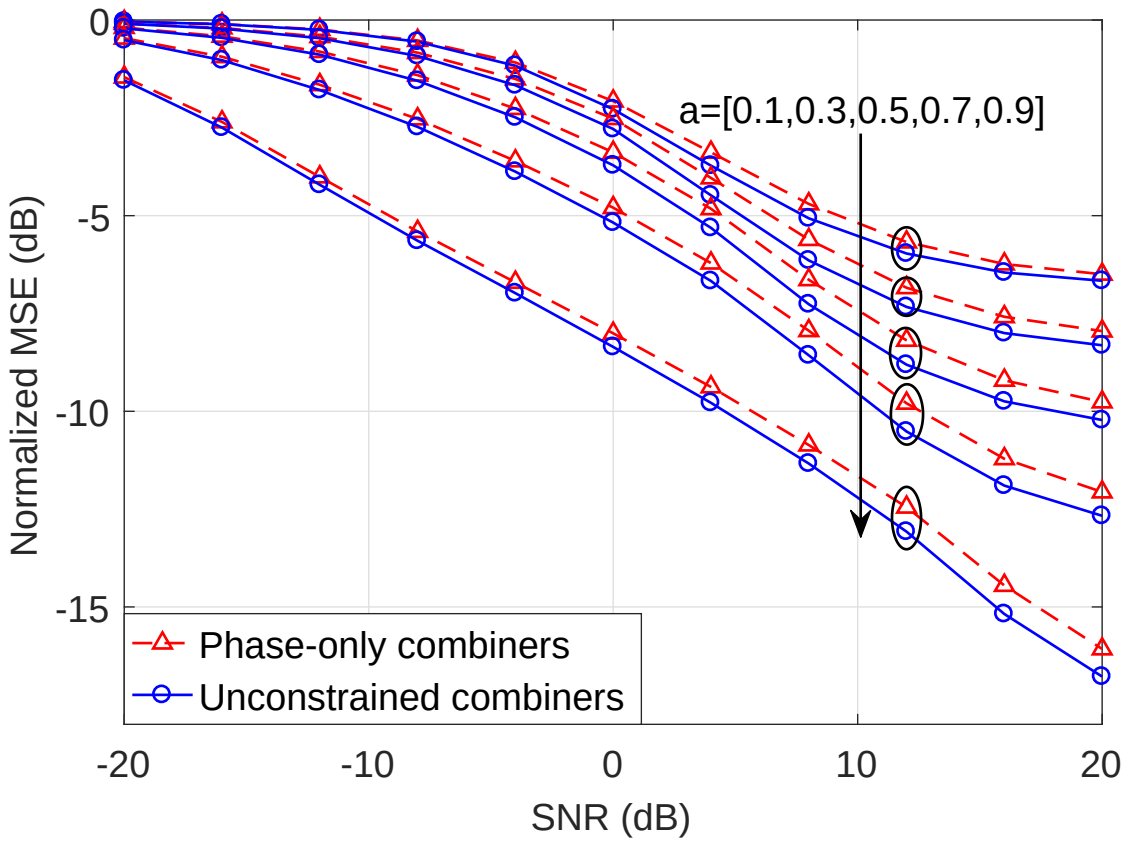


Figure 4.5: Channel estimation performance of both unconstrained and phase-only combiners with the *Sequential* method under different channel correlations. ($M = 64$, $L = 8$, $T = 6$)

channel estimation decreases, thus the performance increases, as the increasing of a, namely, with more spatially correlated channels, which verifies Corollary 4.1. Moreover, it is obvious that the performance of the corresponding phase-only and unconstrained combiners is tight throughout the SNR regions, e.g., there exist 0.7 dB of MSE gap at SNR=20 dB when $a = 0.9$.

d) Spectral Efficiency Evaluation with Hybrid Channel Estimation

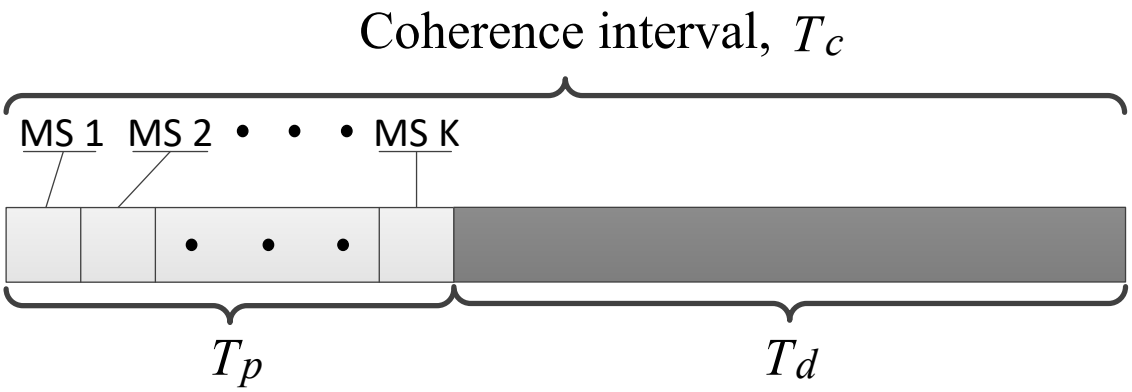


Figure 4.6: Diagram of multi-user massive MIMO communications with hybrid channel estimation. It consists of T pilot transmission slots within the training period of each MS and T_{d} downlink data transmission slots. Hence, $T_{\mathrm{c}} = T_{\mathrm{p}} + T_{\mathrm{d}}$ and $T_{\mathrm{p}} = KT$.

Apart from the normalized MSE performance presented in the previous subsections, the spectral efficiency of the hybrid channel estimation is examined. Here, K single-antenna MSs are served by a hybrid-structured BS equipped with M antennas followed by L RF chains. In uplink communications, K MSs transmit training pilots to the BS independently for channel estimations, i.e., using mutually orthogonal time slots, which avoids pilot contaminations. Note that the uplink pilot transmission is performed T times each user to improve the accuracy of channel estimation. Following the training phases, the BS broadcasts data to all MSs during the downlink communication phase by employing the hybrid precoding with the estimated channels. See Fig. 4.6 for the signal transmission diagram, where T_{p}, T_{d} and T_{c} denote the length of pilot training, downlink data transmission and coherence intervals, respectively. Note that the uplink and downlink channels are assumed reciprocal in this simulation. We calculate the spectral efficiency over the data transmission interval in Fig. 4.7 and the entire coherence interval in Fig. 4.8, i.e., $\mathcal{S} = \sum_{\tau=1}^{T_{\mathrm{d}}} \sum_{k=1}^{K} \mathcal{R}_k[\tau]/\mathcal{T}$, where $\mathcal{S}$ denotes the spectral efficiency, $\mathcal{R}_k[\tau]$ represents the downlink data rate of MS k at time slot τ, and $\mathcal{T} = T_{\mathrm{d}}$ in Fig. 4.7 while $\mathcal{T} = T_{\mathrm{c}}$ in Fig. 4.8. In the simulation, we adopt the low-complexity hybrid precoding scheme proposed by Liang *et al.* in [77] for the multi-user massive MIMO system to pre-process the transmission signal in downlink communications. In both figures, we ob-

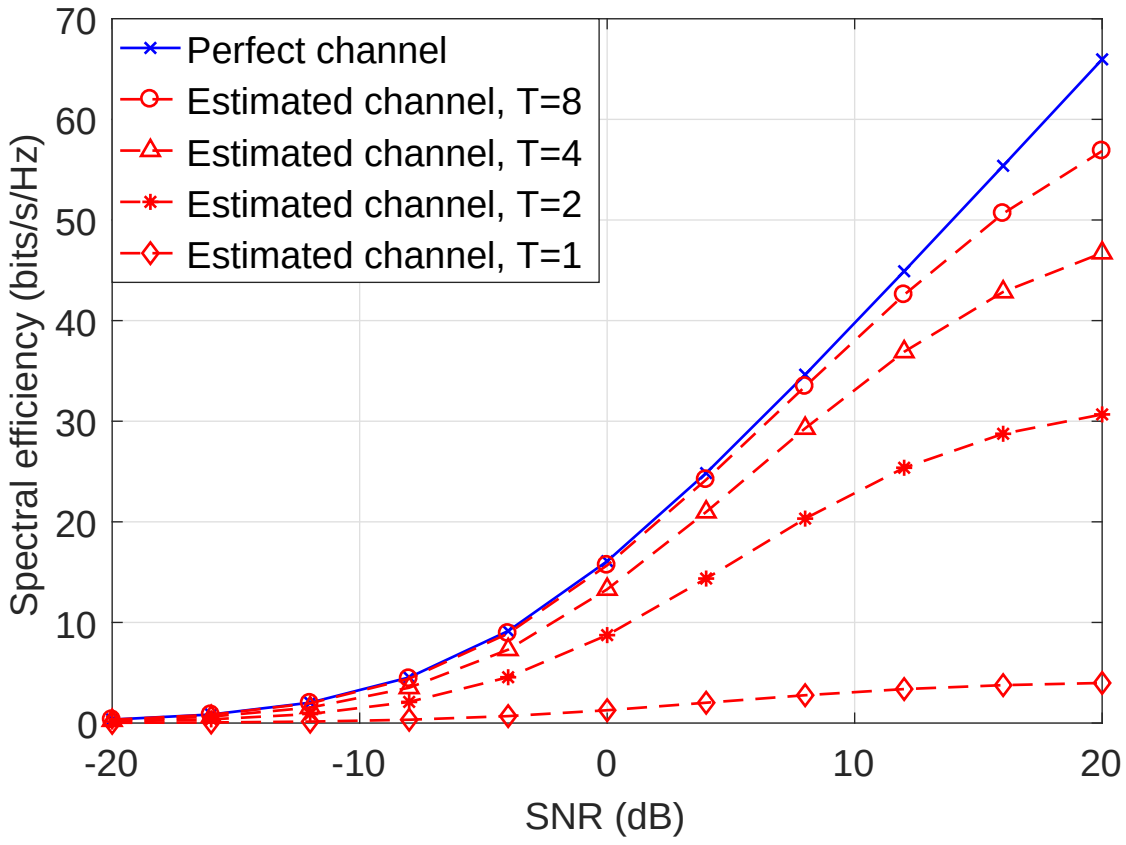


Figure 4.7: Spectral efficiency achieved by the hybrid precoding scheme using estimated and perfect CSI. The *Sequential* method is adopted for multiple-training design. The spectral efficiency is calculated over the downlink data transmission interval T_{d}. ($M = 64$, $L = 8$ and $K = 8$)

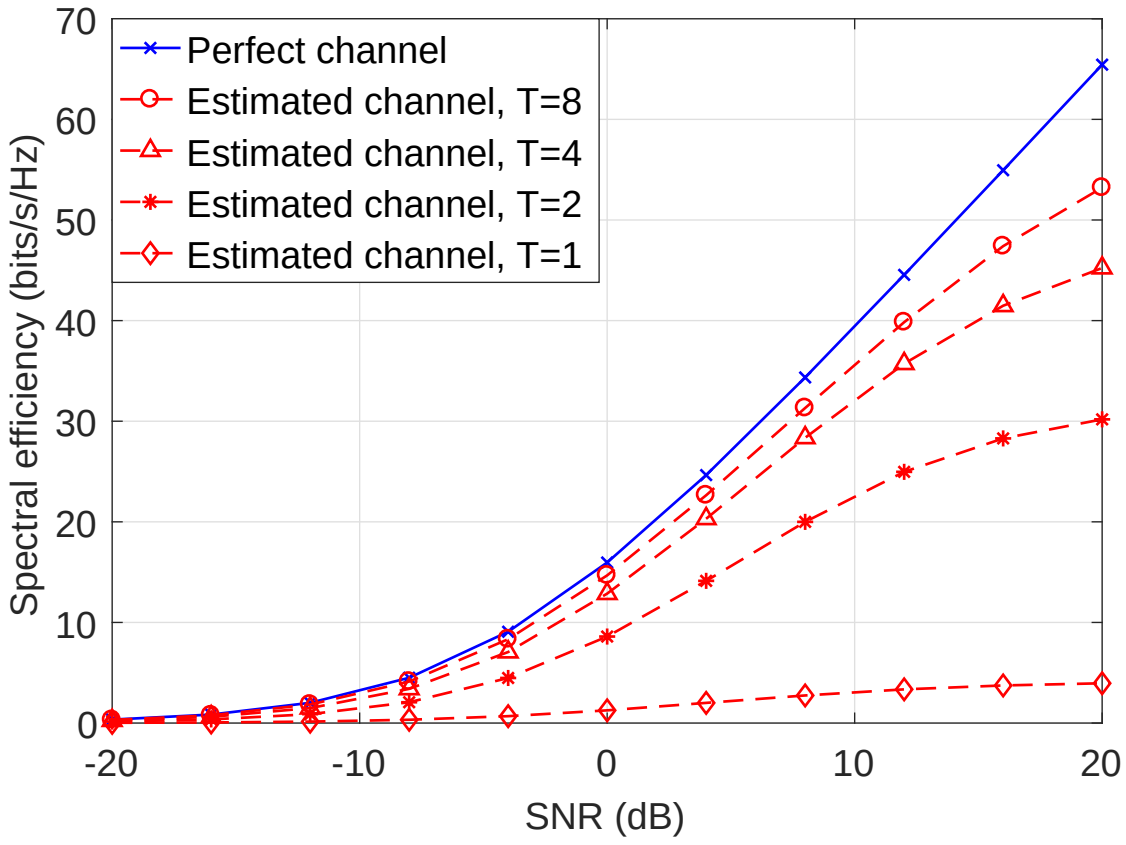


Figure 4.8: Spectral efficiency of the hybrid precoding scheme using estimated and perfect CSI. The *Sequential* method is adopted for multiple-training design. The spectral efficiency is calculated over the entire coherence interval T_{c}. ($M = 64$, $L = 8$, $K = 8$ and $T_{\mathrm{c}} = 1000$)

serve that the spectral efficiency of the precoding with estimated channels increases when multiple trainings are performed which shows that the spectral efficiency performance with the hybrid channel estimation improves by employing more trainings. It is further shown that with a small portion of channel coherence interval dedicated to the channel estimation, the hybrid precoding scheme with 8 pilot trainings achieves highly desirable performance as compared to the gene-assisted system where perfect CSI is available to the transmitter, particularly in the low SNR regions, where a negligible performance gap is introduced by the constant-magnitude constraint. Consequently, the proposed channel estimation scheme achieves comparable spectral efficiency performance in the hybrid-structured multiuser massive MIMO system compared to the perfect CSI case, especially with an increasing T.

4.5.2 Performance with Parametric Channel Model

In the previous subsection, the nonparametric channel models are employed in simulations to evaluate the performance of the proposed channel estimation design in the hybrid precoding system, where the spatial channel correlation is assumed to be perfectly known by the BS. In practice, the BS can only adopt the estimated channel correlation to estimate real-time channel vectors. In this subsection, the parametric channel model denoted by (4.26) is considered. The covariance matrix is estimated by the proposed covariance matching method which is generated offline and used for online channel estimation in the hybrid precoding system.

The simulation results of the spectral efficiency performance of the hybrid system with both perfect and estimated CSI are presented in Figs. 4.9 and 4.10. In both figures, it is obvious that the spectral efficiency of the hybrid precoding scheme with the estimated CSI is measurably close to that with the perfect CSI. At the low SNR regions, the performance of the precoding scheme between the estimated and perfect CSI is especially tight, say 0.24 bits/s/Hz spectral efficiency gap in Fig. 4.9 and 1.37 bits/s/Hz in Fig. 4.10 when $\sigma_{\mathrm{s}} = 90^{\circ}$. Furthermore, the performance gap between the proposed design and the perfect CSI keeps consistent for spread angles ranging from 5° to 90° which verifies the robustness of the proposed channel estimation. Finally, by comparing the corresponding curves between Figs. 4.9 and 4.10, it is evident that the spectral efficiency with the estimated channel moderately reduces when it is calculated over the entire coherence interval due to $T_{\mathrm{c}} > T_{\mathrm{d}}$. However, the performance reduction is negligible, especially in the relatively low SNR regions.

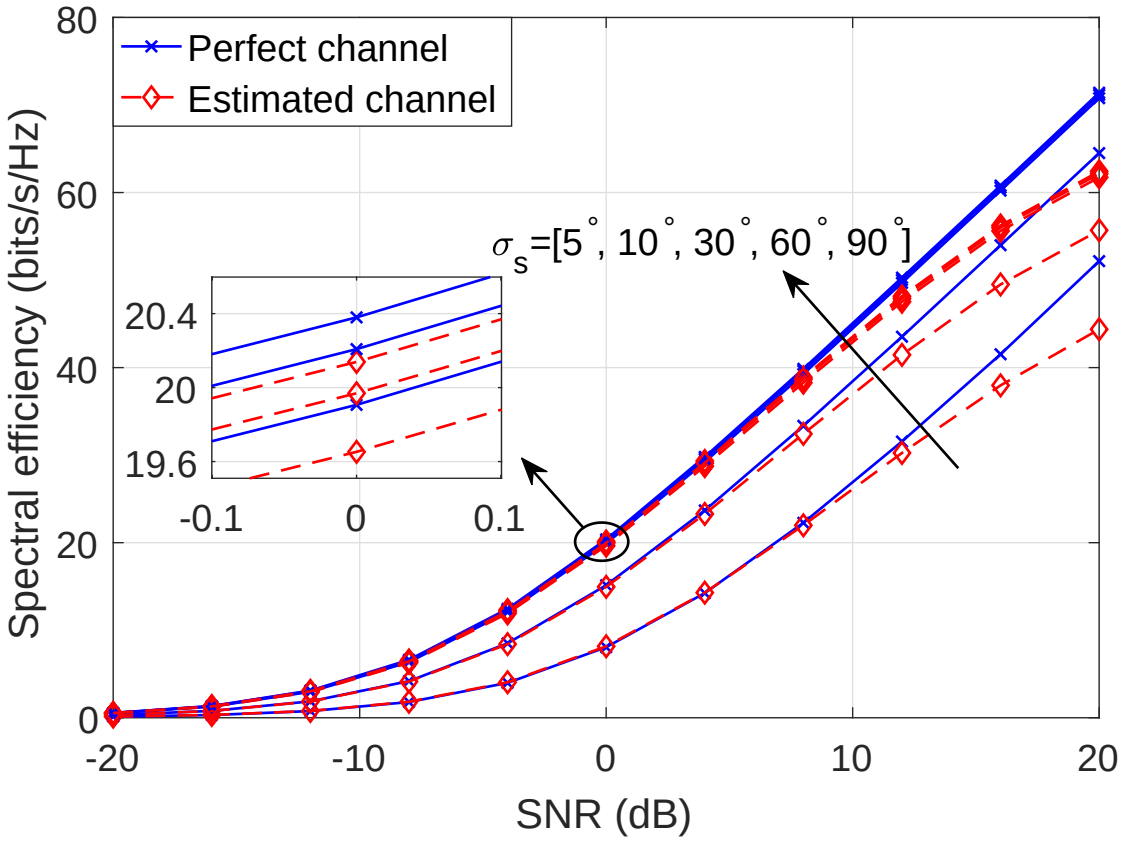


Figure 4.9: Spectral efficiency achieved by the hybrid precoding scheme using estimated and perfect CSI. The *Sequential* method is adopted multiple-training design. The spectral efficiency is calculated over the downlink data transmission interval T_{d}. The parametric channel models are employed in the simulations. ($M = 64$, $L = 8$, $T = 8$, $N_R = 10$ and $T_{\mathrm{c}} = 1000$)

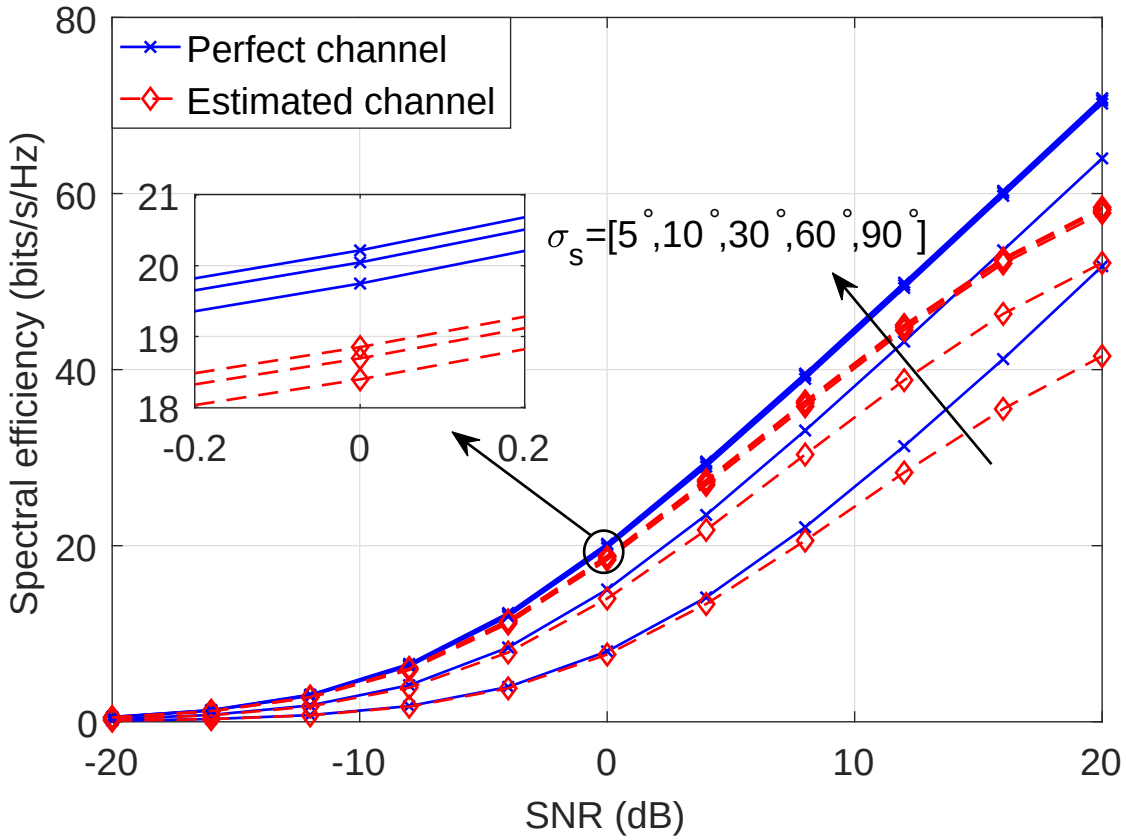


Figure 4.10: Spectral efficiency achieved by the adopted hybrid precoding scheme using estimated and perfect CSI. The *Sequential* method is adopted in hybrid channel estimation for multiple training. The spectral efficiency is calculated over the entire coherence interval T_{c}. The parametric channel models are employed in the simulations. ($M = 64$, $L = 8$, $T = 8$, $N_R = 10$ and $T_{\mathrm{c}} = 1000$)

4.6 Conclusions

This chapter has investigated and proposed a channel estimation framework for massive MIMO systems employing a hybrid transceiver structure with a limited number of RF chains and a phase shifting network. The optimal RF combiners have been designed in the case of single training by exploiting the channel covariance matrix and the complete CSI is estimated in the hybrid structured massive MIMO system. The theoretical analyses have shown that the channel estimation performance can improve when the channel is more spatially correlated, the BS deploys more RF chains or the MSs increase training power, which has been verified by simulation results. In addition, the RF combiners have been designed in the case of multiple trainings to increase the DoF of the received signal measurements at BS. The simulation results reveal that multiple trainings can approach the performance of full-chain estimation with single training and even outperform it at low SNR regions when the channel is highly correlated. Finally, this chapter has proposed a covariance matching method to generate the channel covariance matrix offline for the online channel estimation which rapidly decreases the resource consumption of the covariance matrix estimation. The simulation results have shown the effectiveness of the proposed method in hybrid channel estimation under various scenarios in terms of both the MSE of channel estimates and the spectral efficiency of the hybrid precoding.

4.7 Appendix

4.7.1 Proof of Lemma 4.1

By applying the eigenvalue decomposition of a positive-definite $\mathbf{R}$, i.e., $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}$, the objective function in (4.7) can be derived as

$$\begin{aligned}
\mathrm{MSE} =& \mathrm{tr}\left(\left(\mathbf{R}^{-1} + \rho\mathbf{F}_{\mathrm{c}}^{\mathrm{H}}\mathbf{R}_{\mathbf{F}_{\mathrm{d}}}^{-1}\mathbf{F}_{\mathrm{c}}\right)^{-1}\right) \\
=& \mathrm{tr}\left(\left(\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^{\mathrm{H}} + \rho\mathbf{F}_{\mathrm{c}}^{\mathrm{H}}\mathbf{R}_{\mathbf{F}_{\mathrm{d}}}^{-1}\mathbf{F}_{\mathrm{c}}\right)^{-1}\mathbf{U}\mathbf{U}^{\mathrm{H}}\right) \\
=& \mathrm{tr}\left(\left(\mathbf{\Lambda}^{-1} + \rho\mathbf{U}^{\mathrm{H}}\mathbf{F}_{\mathrm{c}}^{\mathrm{H}}\mathbf{R}_{\mathbf{F}_{\mathrm{d}}}^{-1}\mathbf{F}_{\mathrm{c}}\mathbf{U}\right)^{-1}\right) \\
=& \mathrm{tr}\left(\left(\mathbf{\Lambda}^{-1} + \rho\mathbf{U}^{\mathrm{H}}\left(\sum_{i=1}^{p}\mathbf{F}_i^{\mathrm{H}}(\mathbf{F}_i\mathbf{F}_i^{\mathrm{H}})^{-1}\mathbf{F}_i\right)\mathbf{U}\right)^{-1}\right),
\end{aligned} \tag{4.30}$$

where the last equality uses

$$\mathbf{R}_{\mathbf{F}_{\mathrm{d}}}^{-1} = \mathrm{blkdiag}\left\{(\mathbf{F}_1\mathbf{F}_1^{\mathrm{H}})^{-1}, \cdots, (\mathbf{F}_T\mathbf{F}_T^{\mathrm{H}})^{-1}\right\}. \tag{4.31}$$

Substituting (4.8) into the above equation yields

$$\mathrm{MSE} = \mathrm{tr}\left(\left(\mathbf{\Lambda}^{-1} + \rho\sum_{i=1}^{T}\mathbf{U}^{\mathrm{H}}\mathbf{V}_{i,L}\mathbf{V}_{i,L}^{\mathrm{H}}\mathbf{U}\right)^{-1}\right), \tag{4.32}$$

which indicates that the MSE is independent of both $\mathbf{U}_i$ and $\mathbf{\Sigma}_i$. Hence, we can safely set both of them to identity matrices and obtain $\mathbf{F}_i^{\mathrm{opt}} = \mathbf{V}_{i,\mathrm{L}}^{\mathrm{H}}$ from (4.8) without loss of the optimality, which completes the proof.

4.7.2 Block Generalized Rayleigh Quotient

Lemma 4.2. *(Block Generalized Rayleigh Quotient)[97] Given a $N \times T$ matrix of full column rank, $\mathbf{V}$, the block generalized Rayleigh quotient with respect to the pencil ($\mathbf{A}$, $\mathbf{B}$) is defined as follows:*

$$\mathrm{GRQ}(\mathbf{V}) = \mathrm{tr}\left(\left(\mathbf{V}^{\mathrm{T}}\mathbf{B}\mathbf{V}\right)^{-1}\mathbf{V}^{\mathrm{T}}\mathbf{A}\mathbf{V}\right). \tag{4.33}$$

Suppose the generalized eigenvalues and the corresponding eigenvectors of the pencil ($\mathbf{A}$, $\mathbf{B}$) are denoted by $\lambda_1 \geq \cdots \geq \lambda_N$ and $(\mathbf{v}_1, \cdots, \mathbf{v}_N)$, respectively, where the eigenvalues are arranged in decreasing order, without loss of generality. For a matrix containing a subset of distinct generalized eigenvectors, the generalized Rayleigh quotient evaluates to the sum of the associated eigenvalues:

$$\mathrm{GRQ}([\mathbf{v}_{i_1}, \cdots, \mathbf{v}_{i_T}]) = \sum_{j=1}^{T}\lambda_{i_j} \tag{4.34}$$

and bounds can be placed on the Rayleigh quotient of an $N \times T$ matrix as

$$\lambda_1 + \lambda_2 + \cdots + \lambda_T \geq \mathrm{GRQ}(\mathbf{V}) \geq \lambda_{N-T+1} + \cdots + \lambda_N, \tag{4.35}$$

where an equality in this bound indicates that $\mathbf{V}$ is composed of the extreme eigenvectors as $\mathrm{colsp}(\mathbf{V}) = \mathrm{colsp}([\mathbf{v}_1, \cdots, \mathbf{v}_T])$ or $\mathrm{colsp}(\mathbf{V}) = \mathrm{colsp}([\mathbf{v}_{N-T+1}, \cdots, \mathbf{v}_N])$. Here $\mathrm{colsp}(\cdot)$ represents the column space spanned by the specified matrix.

4.7.3 Proof of Corollary 4.1

We start with the definition of function $\phi : \mathbb{R}^L \to \mathbb{R}$ as

$$\phi(\mathbf{x}) = \sum_{l=1}^{L}\psi(x_l), \tag{4.36}$$

where $\psi(x_l) \triangleq \frac{x_l^2}{x_l+1/\rho}$. Note that the MSE shown in (4.15b) can be denoted by $\mathrm{MSE} = M - \phi(\boldsymbol{\lambda})$. It is straightforward to show that $\psi(x_l)$ is convex and hence $\phi(\mathbf{x})$ is *Schur-convex* according to [98].

Suppose channel $\mathbf{g}_1 \in \mathbb{C}^M$ is *more spatially correlated* than $\mathbf{g}_2 \in \mathbb{C}^M$ which can be mathematically expressed as $\boldsymbol{\lambda}_1 \succ \boldsymbol{\lambda}_2$, where $\boldsymbol{\lambda}_1 = [\lambda_{1,1}, \cdots, \lambda_{M,1}]^{\mathrm{T}}$ and $\boldsymbol{\lambda}_2 = [\lambda_{1,2}, \cdots, \lambda_{M,2}]^{\mathrm{T}}$ denote the eigenvalues sorted in descending order of the covariance matrices of channel $\mathbf{g}_1$ and $\mathbf{g}_2$, respectively. Applying the majorization theory on the Schur-convex function $\phi(\mathbf{x})$, there exists

$$\phi(\boldsymbol{\lambda}_1) \geq \phi(\boldsymbol{\lambda}_2). \tag{4.37}$$

Hence, $\mathrm{MSE}(\mathbf{g}_1) \leq \mathrm{MSE}(\mathbf{g}_2)$, which proves Corollary 4.1.

Chapter 5

Conclusions and Future Work

The previous three chapters investigate the efficient pilot-data transmission schemes in both multipair massive multiple-input multiple-output (MIMO) one-way and two-way relaying communications and the channel estimations in the hybrid-structured massive MIMO systems, respectively. This chapter concludes the above research contributions and draws out the remaining issues and expectations as the guidelines for the future research.

5.1 Multipair Massive MIMO Relaying with Pilot-data Transmission Overlay

THIS dissertation has investigated the pilot and data transmission scheme in multipair massive MIMO relaying systems for both half-duplex (HD) and full-duplex (FD) communications. Due to a large amount of antennas equipped on relay station (RS), it is found that the source-relay and relay-destination channels are asymptotically orthogonal to each other, and thereby the transmission phases of pilots and data can be shifted to overlap each other to reduce the overhead of pilot transmissions and thus to improve the system performance. Based on this consideration, a transmission scheme with pilot-data overlay in both HD and FD multipair massive MIMO relaying systems is proposed and properly designed. In the HD overlay scheme, destination pilots are transmitted simultaneously with source data transmission, such that the effective data transmission duration is increased. Moreover, both source and destination pilots are transmitted along with data transmission in the FD system and thus the effective data transmission duration can be further increased. The closed-form expressions of the ergodic achievable rates of the considered relaying systems with the proposed scheme are derived in this dissertation. From the derived expressions, it is found that the loop interference (LI) in the FD overlay scheme can be effectively suppressed when the number of RS antennas is large and

no error propagation exists with the proposed scheme in contrast to the semi-orthogonal pilot design proposed by [43]. Both theoretical analyses and numerical evaluations are performed to support the superiority of the proposed scheme to conventional ones. Finally, an optimal power allocation problem is formulated for the FD overlay scheme to further increase the system performance and a successive convex approximation (SCA) approach to solve the non-convex optimization problem is proposed in this dissertation.

5.2 Multipair Two-way Decode-and-Forward Relaying with Physical Layer Network Coding in Massive MIMO Systems

This dissertation has proposed and analyzed an efficient multipair HD decode-and-forward (DF) scheme, namely sum decode-and-forward (SDF), with physical layer network coding (PNC) in the two-way relaying system. As a comparison, the performance of the conventional joint decode-and-forward (JDF) scheme has also been investigated in the considered relaying system. Compared with the JDF scheme, this proposal has shown that the SDF employed the PNC technique at the RS can improve the performance of the DF scheme. In the analyses, the minimum mean square error (MMSE) channel estimation has been adopted to obtain the channel state information (CSI) which is more practical than the case that the ideal CSI was employed in the previous literature. The theoretical analyses have revealed the superiority of the proposed SDF scheme to the JDF one, where only half number of the training pilots are needed to obtain the CSI and hence increase the system throughput. Finally, the asymptotic energy efficiency of the proposed scheme has been investigated and the numerical results support the higher efficiency of the proposed scheme compared with that of the conventional JDF scheme.

5.3 Channel Estimations with Hybrid Precoding in Massive MIMO Communications

This dissertation investigates and proposes a channel estimation scheme to meet the requirements of CSI to perform hybrid precoding in massive MIMO systems, and the proposed scheme can also be applied to the mmWave communications. The hybrid precoding structure employs limited radio frequency (RF) chains to suppress the hardware cost and baseband signal processing complexity. Due to the phase-shifting networks interfacing the high-dimensional antennas with the low-dimensional baseband processing modules, the channel estimation of the hybrid structure is different from the typical full-

chain MIMO systems. With the properly designed RF combiners, the BS can estimate the complete channel information and achieve the performance of the full-chain channel estimations by employing multiple training phases. Both the theoretical analyses and numerical simulations are performed to evaluate the performance of the proposed scheme and the results show its superiority. Finally, a covariance matching method is proposed to obtain the channel correlations which are used to design RF combiners for channel estimations.

5.4 Future Research Issues

There are many open issues for further research related to the topics in this dissertation which are summarized as follows:

- In the peer-to-peer millimeter wave communications, both the transmitter and receiver can be equipped with multiple antennas followed by the limited RF chains as shown by the system model of [17]. The channel estimations for such kind of system model should also be investigated with the proposed method in this dissertation to decrease the complexity compared to the method proposed in [18]. The research on how to estimated the channels of such kind of system model is left as the future work.

- Considering the efficient pilot-data transmission overlay, it can also be adopted in the hybrid massive MIMO systems by properly designing. Due to the limited RF chains, the multiple pilot trainings can be utilized to improve the channel estimation performance. However, the multiple pilot trainings decrease the data transmission intervals in the multiuser system. Hence, it is a promising design to transmit the pilots and user data simultaneously to regain the data transmission efficiency where the channels between different users are asymptotically orthogonal to each other in massive MIMO systems. However, a proper pilot-data transmission scheme in such system should be designed. This dissertation leaves this work for the future research.

Chapter 6

Publications

Published papers:

L. Pan and T. Lu, "Highly efficient boundary element analysis of whispering gallery microcavities," *IEEE Photonics Technology Letters*, vol. 26, no. 24, pp. 2465–2468, December 2014.

L. Pan, Y. Dai, W. Xu, and X. Dong, "A novel block-shifted pilot design for multipair massive MIMO relaying," in *Proc. European Signal Processing Conference (EUSIPCO)*, 848–852, August 2016.

Y. Dai, **L. Pan**, and X. Dong, "Physical-layer network coding aided bi-directional cooperative relays for transmitted reference pulse cluster UWB systems," in *Proc. IEEE International Conference on Communications (ICC)*, June 2014, pp. 5825–5830.

Submitted papers:

L. Pan, Y. Dai, W. Xu, and X. Dong, "Multipair massive MIMO relaying with pilot-data transmission overlay," accepted by *IEEE Transactions on Wireless Communications*, March 2017.

L. Pan, L. Liang, W. Xu, and X. Dong, "Framework of Channel Estimation for Hybrid Analog-and-Digital Processing Enabled Massive MIMO Communications," submitted to *IEEE Transactions on Communications*.

Prepared manuscripts:

L. Pan, Y. Dai, and X. Dong, "Multipair two-way decode-and-forward relaying with physical layer network coding in massive MIMO systems," prepared for submission.

Bibliography

[1] R. Muharar, "Multiuser precoding in wireless communication systems," Ph.D. dissertation, University of Melbourne, October 2012.

[2] Wikipedia, "IMT Advanced," November 2016. [Online]. Available: https://en.wikipedia.org/wiki/IMT_Advanced

[3] M. Mouly, M.-B. Pautet, and T. Foreword By-Haug, *The GSM System for Mobile Communications*. Telecom Publishing, 1992.

[4] H. Holma, A. Toskala *et al.*, *WCDMA for UMTS*. Wiley Online Library, 2000, vol. 2.

[5] V. K. Garg, *IS-95 CDMA and CDMA2000: Cellular/PCS systems implementation*. Pearson Education, 1999.

[6] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS Long Term Evolution*. Wiley Online Library, 2009.

[7] Wikipedia, "LTE (telecommunication)," February 2017. [Online]. Available: https://en.wikipedia.org/wiki/LTE_(telecommunication)

[8] ——, "LTE Advanced," February 2017. [Online]. Available: https://en.wikipedia.org/wiki/LTE_Advanced

[9] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[10] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November 1999.

[11] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1998.

[12] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[13] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, January 2013.

[14] J. Hoydis, S. Ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, February 2013.

[15] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 30, no. 1, pp. 186–195, February 2014.

[16] W. 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, April 2014.

[17] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[18] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, October 2014.

[19] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, April 2016.

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

[21] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, October 2006, pp. 359–363.

[22] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, August 2011.

[23] 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 Communications Letters*, vol. 20, no. 6, pp. 1259–1262, June 2016.

[24] L. K. S. Jayasinghe, N. Rajatheva, and M. Latva-Aho, "Robust precoder-decoder design for physical layer network coding-based MIMO two-way relaying system," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, pp. 1–16, May 2013.

[25] M. Dohler and Y. Li, *Cooperative Communications: Hardware, Channel and PHY*. John Wiley & Sons, January 2010.

[26] H. Q. Ngo and E. G. Larsson, "Spectral efficiency of the multipair two-way relay channel with massive arrays," in *Proc. Asilomar Conference on Signals, Systems, and Computers (ACSSC)*, November 2013, pp. 275–279.

[27] H. Q. Ngo, E. G. Larsson, S. Member, and S. Member, "Large-scale multipair two-way relay networks with distributed AF beamforming large-scale multipair two-way relay networks with distributed AF beamforming," *IEEE Communications Letters*, vol. 17, no. 12, pp. 2288–2291, December 2013.

[28] H. A. Suraweera, H. Q. Ngo, T. Q. Duong, C. Yuen, and E. G. Larsson, "Multi-pair amplify-and-forward relaying with very large antenna arrays," in *Proc. IEEE International Conference on Communications (ICC)*, June 2013, pp. 4635–4640.

[29] Y. Dai and X. Dong, "Power allocation for multi-pair massive MIMO two-way AF relaying with linear processing," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 5932–5946, September 2016.

[30] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1721–1737, September 2014.

[31] H. Q. Ngo, H. a. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair massive MIMO full-duplex relaying with MRC/MRT processing," in *Proc. IEEE International Conference on Communications (ICC)*, June 2014, pp. 4818–4824.

[32] L. J. Rodriguez, N. H. Tran, and T. Le-Ngoc, "Optimal power allocation and capacity of full-duplex AF relaying under residual self-interference," *IEEE Wireless Communications Letters*, vol. 3, no. 2, pp. 233–236, April 2014.

[33] P. Larsson, N. Johansson, and K.-E. Sunell, "Coded bi-directional relaying," in *Proc. IEEE Vehicular Technology Conference (VTC) Spring*, vol. 2, May 2006, pp. 851–855.

[34] P. Popovski and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," in *Proc. IEEE Vehicular Technology Conference (VTC) Spring*, vol. 2, May 2006, pp. 588–593.

[35] ——, "Wireless network coding by amplify-and-forward for bi-directional traffic flows," *IEEE Communications Letters*, vol. 11, no. 1, pp. 16–18, January 2007.

[36] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1238–1250, March 2010.

[37] R. H. Louie, Y. Li, and B. Vucetic, "Practical physical layer network coding for two-way relay channels: performance analysis and comparison," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 764–777, February 2010.

[38] S. Zhang, S. Liew, and P. Lam, "Physical-layer network coding," in *Proc. The Annual International Conference on Mobile Computing and Networking (MobiCom)*, September 2006.

[39] S. Zhang, Q. Zhou, C. Kai, and W. Zhang, "Full diversity physical-layer network coding in two-way relay channels with multiple antennas," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4273–4282, August 2014.

[40] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *Proc. IEEE International Conference on Communications (ICC)*, June 2007, pp. 707–712.

[41] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[42] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293 – 1308, February 2016.

[43] H. Zhang, X. Zheng, W. Xu, and X. You, "On massive MIMO performance with semi-orthogonal pilot-assisted channel estimation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, pp. 1–14, December 2014.

[44] X. Zheng, H. Zhang, W. Xu, and X. You, "Semi-orthogonal pilot design for massive MIMO systems using successive interference cancellation," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, December 2014, pp. 3719–3724.

[45] L. You, X. Gao, X.-G. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 1276, no. 6, pp. 1–15, June 2015.

[46] F. Fernandes, A. Ashikhmin, and T. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, February 2013.

[47] S. Jin, X. Wang, Z. Li, and K.-K. Wong, "Zero-forcing beamforming in massive MIMO systems with time-shifted pilots," in *Proc. IEEE International Conference on Communications (ICC)*, June 2014, pp. 4801–4806.

[48] S. Jin, X. Wang, Z. Li, K.-K. Wong, Y. Huang, and X. Tang, "On massive MIMO zero-forcing transceiver using time-shifted pilots," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 59–74, January 2016.

[49] S. Jin, X. Liang, K.-K. Wong, X. Gao, and Q. Zhu, "Ergodic rate analysis for multipair massive MIMO two-way relay networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1480–1491, March 2015.

[50] J. S. Lemos, F. A. Monteiro, I. Sousa, and A. Rodrigues, "Full-duplex relaying in MIMO-OFDM frequency-selective channels with optimal adaptive filtering," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, December 2015, pp. 1081–1085.

[51] J. S. Lemos and F. A. Monteiro, "Full-duplex massive MIMO with physical layer network coding for the two-way relay channel," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, July 2016.

[52] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12–25, November 2004.

[53] A. Zappone, E. A. Jorswieck, and S. Buzzi, "Energy efficiency and interference neutralization in two-hop MIMO interference channels," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6481–6495, December 2014.

[54] J. Zhang and M. Haardt, "Energy efficient two-way non-regenerative relaying for relays with multiple antennas," *IEEE Signal Processing Letters*, vol. 22, no. 8, pp. 1079–1083, August 2015.

[55] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1668–1683, April 2016.

[56] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5983–5993, December 2011.

[57] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, April 2014, pp. 359–372.

[58] H. Cui, L. Song, and B. Jiao, "Multi-pair two-way amplify-and-forward relaying with very large number of relay antennas," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2636–2645, May 2014.

[59] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1st ed. Prentice Hall PTR, April 1993, vol. 1.

[60] J. H. Kotecha and A. M. Sayeed, "Transmit signal design for optimal estimation of correlated MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 546–557, February 2004.

[61] H. Cramér, *Random Variables and Probability Distributions*. Cambridge University Press, June 2004, vol. 36.

[62] J. S. Lemos, F. Rosário, F. A. Monteiro, J. Xavier, and A. Rodrigues, "Massive MIMO full-duplex relaying with optimal power allocation for independent multipairs," in *Proc. IEEE Signal Processing Advances in Wireless Communications (SPAWC)*, June 2015, pp. 306–310.

[63] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall New York, 2000, vol. 1.

[64] M. O. Hasna and M.-S. Alouini, "Optimal power allocation for relayed transmissions over Rayleigh-fading channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1999–2004, November 2004.

[65] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, July 2007.

[66] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, March 2004.

[67] B. R. Marks and G. P. Wright, "Technical note – A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, August 1978.

[68] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059 – 3075, June 2015.

[69] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293 – 1308, Feb. 2016.

[70] B. Xia, Y. Fan, J. Thompson, and H. V. Poor, "Buffering in a three-node relay network," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4492–4496, November 2008.

[71] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4381–4393, December 2012.

[72] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, July 2006, pp. 1668–1672.

[73] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, April 2003.

[74] H. Yang and T. L. Marzetta, "Performance of Conjugate and Zero-Forcing Beamforming in Large-Scale Antenna Systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, February 2013.

[75] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "On the optimality of single-carrier transmission in large-scale antenna systems," *IEEE Wireless Communications Letters*, vol. 1, no. 4, pp. 276–279, August 2012.

[76] L. Pan, Y. Dai, W. Xu, and X. Dong, "A novel block-shifted pilot design for multipair massive MIMO relaying," in *Proc. European Signal Processing Conference (EUSIPCO)*, August 2016, pp. 848–852.

[77] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 653–656, December 2014.

[78] A. Alkhateeb, G. Leus, and R. W. Heath Jr., "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[79] A. Alkhateeb and R. W. Heath Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[80] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: open-loop and closed-loop training with memory," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 802–814, October 2014.

[81] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884 – 893, March 2006.

[82] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1356–1368, March 2015.

[83] R. Méndez-Rial, C. Rusut, A. Alkhateeb, N. González-Prelcic, and R. W. Heath Jr., "Channel estimation and hybrid combining for mmWave: Phase shifters or switches?" in *Proc. Information Theory and Applications Workshop (ITA)*, Feb. 2015, pp. 90–97.

[84] L. Yang, Y. Zeng, and R. Zhang, "Efficient channel estimation for millimeter wave MIMO with limited RF chains," in *Proc. IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[85] G. C. Alexandropoulos and S. Chouvardas, "Low complexity channel estimation for millimeter wave systems with hybrid A/D antenna processing," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, December 2016, pp. 1–6.

[86] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, March 2010.

[87] A. Alkhateebt, O. E. Ayach, G. Leus, and R. W. Heath, Jr., "Single-sided adaptive estimation of multi-path millimeter wave channels," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2014, pp. 125–129.

[88] A. Forenza, D. J. Love, and R. W. Heath Jr., "Simplified spatial correlation models for clustered MIMO channels with different array configuration," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1924–1934, Jul. 2007.

[89] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 1872–1882, Jul. 2000.

[90] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed. SIAM, 2002.

[91] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, pp. 397–401, Jul. 1988.

[92] Y.-C. Liang and F. P. S. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 222–232, Feb. 2001.

[93] W. L. Melvin and G. A. Showman, "An approach to knowledge-aided covariance estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 3, pp. 1021–1042, Jul. 2006.

[94] R. Méndez-Rial, N. González-Prelcic, and R. W. Heath Jr., "Augmented covariance estimation with a cyclic approach in DOA," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2015, pp. 2784–2788.

[95] ——, "Adaptive hybrid precoding and combining in mmWave multiuser MIMO systems based on compressed covariance estimation," in *Proc. IEEE International*

Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Dec. 2015, pp. 213–216.

[96] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications*, vol. 5, no. 1, pp. 10–22, February 1998.

[97] C. G. Baker, "Riemannian manifold trust-region methods with applications to eigenproblems," Ph.D. dissertation, Florida State University, May 2008.

[98] A. W. Marshal, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, 2nd ed. Springer, 2010.