

The
University
Of
Sheffield.

Massive MIMO and Full-duplex Relaying Systems

Mengxue TANG

Department of Electronic & Electrical Engineering

University of Sheffield

This thesis is submitted for the degree of

Doctor of Philosophy

November 2020

To Yanchun Li, Shi-Chang Zhang and Hao Song, who always love, support and
inspire me.

Acknowledgements

I would like to express my very great appreciation to my supervisor Prof. Xiaoli Chu for keeping my progress on schedule. She has provided me with very valuable suggestions not only to my research works but also to my future career.

I would like to offer my special thanks to Dr. Mikko Vehkaperä for his patient supervision, support and encouragement. I really appreciate the large amount of time and effort he has spent on the discussions of fundamentals, innovative ideas and technical details with me. My skills of scientific paper and simulation code writing has greatly improved under his guidance. Without his continuous support, accomplishing this work would have been impossible.

I would like to acknowledge Prof. Risto Wichman for host me as a research visitor to the Aalto University. The advices given by Risto have been great help in the problem formulation of this work.

I would like to send my genuine gratitude to Dr. Hao Song, who has proofread this work. He has continuously provided support and encouragement throughout my PhD study.

I would like to express my deep gratitude to my parents, Yanchun Li and Prof. Shi-Chang Zhang, for their unsparing support and love.

I would like to thank my colleagues, Muntadher Alsabah, Nurulhuda Ismail, Zaid Mujaiyid Putra Ahmad Baidowi, Wenfei Yang, Fei He, Haojiang Hua, Fuyou Li, Bo Ma, Tingzhao Yang, Hao Fu, Mengxin Zhou, and Yixin Zhang for having technical discussions with me over research problems.

Finally, I would like to acknowledge my department for providing the great education opportunity with a tuition fee waiver.

Abstract

In this thesis, we study how massive multiple-input and multiple-output (MIMO) can be employed to mitigate loop-interference (LI), multi-user interference and noise in a full-duplex (FD) relaying system. For a FD relaying system with massive MIMO deployed at both source and destination, we investigate three FD relaying schemes: co-located, distributed cooperative, and distributed non-cooperative relaying. Asymptotic analysis shows that the three schemes can completely cancel multi-user interference and LI when the number of antennas at the source and destination grows without bound, in the case where the relay has a finite number of antennas. For the system with massive MIMO deployed at the FD relay, we propose a pilot protocol for LI channel minimum-mean-square-error estimation by exploiting the channel coherence time difference between static and moving transceivers. To maximize the end-to-end achievable rate, we design a novel power allocation scheme to adjust the transmit power of each link at the relay in order to equalize the achievable rate of the source-to-relay and relay-to-destination links. The analytical and numerical results show that the proposed pilot protocol and power allocation scheme jointly improve both spectral and energy efficiency significantly. To enable the use of low resolution analog-to-digital converters (ADCs) at relays for energy saving, we propose a novel iterative power allocation scheme to mitigate the resulting quantization noise via reducing the received LI power and numerically identify the optimum resolutions of ADCs for maximizing throughput and energy efficiency. For massive MIMO receivers employing one-bit ADCs, we propose three carrier frequency (CFO) offset estimation schemes for dual-pilot and multiple-pilot cases. The three schemes are developed under different scenarios: large but finite number of antennas at the receiver, infinite number of antennas at the receiver, and very small CFO, respectively.

Table of contents

List of figures	viii
List of tables	x
Nomenclature	xi
1 Introduction	1
1.1 Background	1
1.2 Research Aim and Objectives	3
1.3 Key Contributions	4
1.4 Publications	5
1.5 Thesis Outline	6
1.6 Notations	8
2 Literature Review	10
2.1 Massive MIMO	10
2.1.1 Precoding	11
2.1.2 Pilot Transmission and Channel Estimation	12
2.2 Full Duplex	14
2.2.1 Self-interference Cancellation	15
2.2.2 Full-duplex Relaying	17
2.3 Massive MIMO Full-duplex Relaying Systems	18
2.3.1 Relaying Schemes	18
2.3.2 Pilot Protocol	19
2.3.3 Power Allocation	20
2.4 Analog-to-Digital Converters	21
2.4.1 Low Resolution ADCs	21
2.4.2 One-bit ADCs	22

2.5	Frequency Synchronization	23
3	Full-Duplex Relaying Schemes	24
3.1	Introduction	24
3.2	System Model	25
3.2.1	Channel Estimation	27
3.2.2	Data Transmission	29
3.2.3	Loop Interference Cancellation	31
3.3	Asymptotic Interference Analysis	33
3.4	Achievable Rate Analysis	34
3.5	Numerical Results	37
3.6	Conclusion	41
4	Loop Interference Cancellation and Power Allocation	42
4.1	Introduction	42
4.2	System Model	43
4.2.1	Novel Pilot Protocol and Channel Estimation	45
4.2.2	Data Transmission	47
4.3	Novel Power Allocation Scheme and Performance Analysis	48
4.3.1	Achievable Rate Analysis	48
4.3.2	Novel Power Allocation Scheme	50
4.3.3	Spectral Efficiency and Energy Efficiency	52
4.4	Numerical Results	52
4.5	Conclusion	56
5	Full-duplex Massive MIMO Relay Systems with Low Resolution ADCs	57
5.1	Introduction	57
5.2	System Model	58
5.2.1	Quantization with low-resolution ADCs	59
5.2.2	Channel Estimation	60
5.2.3	Data Transmission	61
5.3	Achievable Rate Analysis	62
5.3.1	Delay Constrained Achievable Rate	64
5.4	Novel Power Allocation Scheme	65

5.5	Energy Efficiency	66
5.6	Numerical Results	67
5.7	Conclusion	71
6	Carrier Frequency Offset Estimation	73
6.1	Introduction	73
6.2	System Model	74
6.3	CFO Estimation	75
6.3.1	Method 1: Phase Noise Analysis	79
6.3.2	Method 2: Infinite Number of Antennas Assumption	83
6.3.3	Method 3: Small CFO Assumption	85
6.3.4	CFO Update	85
6.3.5	Multiple Pilots Extension	86
6.4	CFO Compensation and Channel Estimation	86
6.5	Numerical Results	88
6.6	Conclusion	94
7	Conclusions and Future Work	95
7.1	Conclusions	95
7.2	Future Work	98
	References	99
	Appendix A Proof of Proposition 3.1	108
	Appendix B Proof of Proposition 3.2	111
	Appendix C Proof of Proposition 4.1	115
	Appendix D Proof of Proposition 6.2	118
	Appendix E Proof of Lemma D.1	123
	Appendix F Proof of Proposition 6.3	127

List of figures

3.1	Full-duplex relaying with distributed relay nodes.	25
3.2	LI cancellation schemes for FD relaying	31
3.3	Spectral efficiency vs. SNR for different relaying schemes. The lines depict analytical results and markers correspond to Monte Carlo simulations. ($\beta_{kk} = 20$ dB, $\beta_{kj} = 10$ dB, $M = 100$ and $K = 10$)	38
3.4	Spectral efficiency vs. β_{kj} for different relaying schemes. ($\rho = 10$ dB, $\beta_{kk} = 20$ dB, $M = 100$ and $K = 10$)	39
3.5	Optimal K/M vs. M for different relaying schemes. ($\rho = 10$ dB, $\beta_{kk} = 20$ dB and $\beta_{kj} = 10$ dB)	40
4.1	Multipair full-duplex relaying system.	43
4.2	Pilot protocol (As $ST - M_{Tx} > 2K$).	45
4.3	Performance of the proposed pilot protocol and power allocation scheme, analytical results of FD w/ LI cancellation (solid), FD w/o LI cancellation (dashed), FD w/ LI cancellation and fixed relay power $m_{tot} = 23$ dBm (dash-dotted) and HD (dotted) presented with curves and Monte Carlo simulations with markers (stars, circles, diamonds and crosses, respectively).	53
4.4	Empirical PDF of the total power consumption of the proposed power allocation scheme at the relay station with ZF processing.	54
4.5	Convergence of the proposed power allocation scheme at the relay station as in (4.17).	55
5.1	Throughput of the system vs. the resolution of ADC, analytical results of the proposed power allocation scheme in FD mode (solid), the proposed power allocation scheme in HD mode (dotted) and fixed relay power $m_{tot} = 23$ dBm in FD mode (dash-dotted) presented with curves and Monte Carlo simulations with markers (stars, crosses and diamonds, respectively).	69
5.2	Empirical CDF of total transmit power at the FD relay station with proposed power allocation scheme.	69

5.3	Energy efficiency of the proposed power allocation scheme vs. ADC resolution N at the relay station with a circle at the peak of each curve.	70
5.4	Energy efficiency vs. the number of receive antennas at the FD relay station of the proposed power allocation scheme with a circle at the peak of each curve.	70
5.5	Throughput vs. code length τ with $T = 500$	71
6.1	Transmission Protocol.	75
6.2	Derivative of $f'(\omega_k)$ for different pilot power.	83
6.3	Mean square error (MSE) of method 1 vs. ω_k for different pilot power ($\tau_p = 2$).	90
6.4	Empirical PDF of $\tilde{\omega}_k$ in method 1 ($\omega_k = \frac{\pi}{8}$, $\rho = 10$ dB, $\tau_p = 2$).	90
6.5	MSE of CFO estimation vs. ω_k for different CFO estimation methods ($\tau_p = 2$, $\rho = 5$ dB).	91
6.6	MSE of CFO estimation vs. M for different CFO estimation methods ($\tau_p = 2$, $\rho = 5$ dB, $\omega_k = \frac{\pi}{20}$).	91
6.7	MSE of CFO estimation vs. pilot power for different CFO estimation methods ($\tau_p = 2$, $\omega_k = \frac{\pi}{20}$).	92
6.8	MSE of CFO estimation vs. pilot power for different multiple-pilot CFO estimation methods ($\tau_p = 24$, $\omega_k = \frac{\pi}{20}$).	93
6.9	MSE of CFO estimation vs. pilot length τ_p for different CFO estimation methods ($\rho = 0$ dB, $\omega_k = \frac{\pi}{20}$).	93

List of tables

5.1 θ for different resolutions of ADCs 59

Nomenclature

Acronyms / Abbreviations

5G	Fifth Generation
ADC	Analog-to-Digital Converter
AF	Amplify-and-Forward
BER	Bit-Error-Rate
BS	Base Station
CFO	Carrier Frequency Offset
CSCG	Circularly Symmetric Complex Gaussian
CSI	Channel State Information
D	Destination
DAC	Digital-to-Analog Converter
DF	Decode-and-Forward
EVM	Error Vector Magnitude
FD	Full-Duplex
FDD	Frequency-Division Duplex
HD	Half-Duplex
LA	Limiting Amplifier
LI	Loop Interference
LMMSE	Linear Minimum Mean Square Error
LOS	Line-of-Sight
LTE	Long-Term Evolution
MF	Matched Filter

MIMO	Multiple-input-multiple-output
MISO	multiple-input-singular-output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
MSE	Mean-Square Error
OFDM	Orthogonal frequency-division multiplexing
PAPR	Peak-to-Average Power Ratio
R	Relay
RF	Radio Frequency
S	Source
SI	Self Interference
SISO	Singular-input-singular-output
SNR	Signal-to-Noise Ratio
TDD	Time-Division Duplex
UE	User Equipment
URLLC	Ultra-reliable and low-latency communications
VGA	Variable Gain Amplifier
ZF	Zero Forcing

Symbols

α	Path loss exponent
\bar{g}	Small scale fading
\bar{x}	Information symbol
β	Large scale fading factor
\mathbf{C}	Correlation matrix
\mathbf{D}	Large scale fading matrix
\mathbf{g}	Channel vector

H	Small scale fading
h	CFO-combined channel vector
I	Identity matrix
m	Per-link power
P	Power allocation matrix
R	Quantized received matrix
r	Processed / Quantized received vector
u_{Tx}	transmit-side hardware impairments
u_{Rx}	receive-side hardware impairments
V	Precoder matrix
W	Detector matrix
x	Transmit signal
z	Locations of user terminal
ε	Code word error probability
ε_{SD}	Decoding error
γ	A design parameter for transmit power control
\hat{g}	Estimated Channel matrix
κ	Shadow fading factor
λ	Normalization factor of precoder
μ	Indicator for level of hardware impairments
ω	Carrier frequency offset
ϕ	Pilot symbol
ρ	Power of the transmit signal / noise / distortion / interference
σ^2	Variance
τ	Signal processing delay / Delay constraint (code word length)
τ_p	The length of pilot symbols for an individual user
θ	Coefficient related to the resolution of ADCs

$\tilde{\mathbf{g}}$	Channel estimation error
ξ	Remaining loop interference
ζ	Noise terms in “unbiased” average correlation model
$\zeta'_{*,Q}$	Phase noise terms in “unbiased” average correlation model
a	Number of blocks used for pilot transmission
c	LI cancellation factor
d	distance between a user cluster and the BS
E_{tot}	Total energy consumption
f_B	System bandwidth
f_{samp}	Sampling rate
K	Number of UE / distributed relay stations / co-located relay stations
M	Number of antennas of massive MIMO systems
m	Per-link transmit power of relay station
N	Resolution of ADCs / Quantization bits of ADCs
n	Noise
P	Power of a hardware component
q	Quantization noise
R	Achievable data rate
r	Quantized signal
r_{RD}	Radius of the destination-user circle
r_{SR}	Radius of the source-user circle
S	Number of coherence blocks
T	Coherence time
t	Time instance / Number of symbols
y	Received signal

Chapter 1

Introduction

1.1 Background

Global mobile data traffic is projected to significant increase in 5G era. It is predicted that the global mobile subscriber will reach 5.7 billion and the average 5G connection speed will increase to 575 Mbps by 2023 [1]. The capacity of mobile networks is required to be significantly enhanced to meet the demand of the global mobile data traffic in the future. As the most straightforward and simple approach, spectrum extension is able to enhance system capacity without aggravating system complexity [2]. Radio spectrum is, however, a limited, scarce and costly natural resources, the license of which needs to be authorized by governments [3]. It is difficult for wireless communication operators to obtain more licensed frequency bands even with huge cost [4]. With limited frequency band resources, the network capacity could also be enhanced by reusing frequency bands, where a large amount of micro base stations (BSs), like Pico and Femto, need to be deployed in mobile networks to provide the coverage for dense small cells [5]. Unfortunately, this approach also has many technical challenges. With dense small cells, resource managements are critical for the network performance, but also challenging and complicated, which need to

consider interference coordination and performance optimization [6]. To realize that, a powerful infrastructure is required, such as macro BSs, to provide centralized control and carry out a series of system procedures, like measurement, signal processing, and resource allocation. This will tremendously aggravate the system complexity with considerable cost and energy consumption. Therefore, novel and robust physical layer techniques are expected to be implemented in mobile networks, aimed at boosting spectral efficiency without increasing the system complexity [7].

A multi-user MIMO is a promising technology in capacity enhancement, which allows BS to broadcast in the same time-frequency resources for downlink transmission to multiple independent users and users to transmit different data streams by spacial multiple access [8]. Operating in the limited spectral resources, massive multiple-input-multiple-output (MIMO) brings the advantages of increased data rate, enhanced reliability, improved energy efficiency (EE), and reduced inter-pair interference, having thus potential to be used in multiple practical scenarios [9], [10]. 5G networks will rely on massive MIMO techniques to take advantage of the enormous multiplexing and array gain offered by rich scattering wireless environments.

Full-duplex (FD) relaying enlarges the coverage of BS, while nearly doubles the achievable rate by receiving and transmitting signals simultaneously [11]. The most crucial problem in FD relaying systems is how to reduce the loop-interference (LI). Recent research has already proved that applying massive MIMO techniques to FD relaying systems can significantly reduce the power of LI [12]. However, with practical imperfect hardware deployed at the relaying, which would bring transmit- and receive-side hardware impairments to the system, methods should be proposed to reduce LI and multi-user interference before FD relaying with massive number of antennas can be implemented.

Another problem in 5G network is frequency synchronization. Since the static BS and relay stations typically serve mobile user devices, the received signals at the BS

and relaying suffers from Doppler shift, which would cause a mismatch in oscillators between user device and the BS/relaying [13]. The mismatch has severe impact on throughput and would cause inter-carrier interference especially in frequency sensitive systems such as orthogonal frequency-division multiplexing (OFDM) [14]. Frequency synchronization methods should be simple and not cause system overhead.

1.2 Research Aim and Objectives

Deploying large number of antennas at a node in a wireless system enables multi-link transmission via the same time and frequency resources. The large number of antennas can be deployed at the base stations, and different nodes in FD relaying systems. Applying large number of antennas helps the FD relaying reduce the multi-user interference as well as LI. Meanwhile, to save circuit power consumption of the systems with large number antennas, low-resolution or even one-bit analog-to-digital converters (ADCs) can be used, although they bring extra noise and distortion into the system.

The research aim is to solve interference cancellation, power allocation and frequency synchronization problems in massive MIMO FD relaying systems. Specifically, the research objectives of the thesis are listed as follows.

- Investigating co-located, distributed FD relaying schemes, with or without cooperative relaying nodes, when a massive number of antennas are deployed at the source and destination.
- Enabling pilot-based digital LI cancellation for FD relaying systems with massive antenna arrays.
- Designing power allocation schemes to mitigate LI while maximizing spectral efficiency (SE).

- Enabling the use of low-resolution ADCs in massive MIMO FD relaying systems for energy consumption reduction.
- Proposing carrier frequency offset (CFO) estimation and compensation schemes for frequency synchronization in massive MIMO systems with one-bit ADCs.

1.3 Key Contributions

- FD relaying schemes

We study a FD relaying system with a massive number of antennas at the source and destination. By using the law of large numbers, we prove that the interference vanishes asymptotically for co-located, distributed cooperative and distributed non-cooperative relay nodes. The achievable rate of the system is derived for matched-filtering (MF) and zero-forcing (ZF) processing.

- Novel pilot protocol for digital LI cancellation

In a multi-pair FD relaying system with massive number of antennas at the relay node, we propose a new pilot protocol to enable digital LI cancellation by exploiting the coherence time difference of the LI channel and the channels between the relay station and the moving terminals. The numerical results show that our pilot protocol does not cause any extra overhead within practical parameter ranges.

- Power allocation scheme

We propose a novel power allocation scheme that adjusts the power of each link at the FD relay to maximize the E2E achievable rate via equalizing the achievable rate of source-to-relay and relay-to-destination links. The power allocation scheme only uses channel statistics and can be implemented by a low-complexity, fast-converging iterative algorithm. The numerical results show

that the proposed FD power allocation scheme outperforms the fixed relaying power and half-duplex (HD) cases.

- FD systems with low-resolution ADCs

Low-resolution ADCs are used to reduce the circuit power consumption for a decode-and-forward FD relaying with massive number of antennas. The SE of the system is derived considering the impact of low-resolution ADCs. We model the energy consumption of the FD relay as a function of the energy consumption of the ADCs and analyze the EE of the relay. From the numerical results we find that 4-7 bits ADCs are preferable for maximizing the EE.

- CFO estimation in massive MIMO systems with one-bit ADCs

We propose three CFO estimation methods for individual users in massive MIMO systems with one-bit ADCs by using only two pilot symbols. The CFO estimation methods utilize the statistics of the “unbiased” average correlation of two pilot symbols. These methods are expanded for multi-pilot symbol scenarios, which give a lower mean-square error (MSE). The numerical results show that the MSE of the proposed CFO estimation can reach 10^{-5} with only 24 pilot symbols.

1.4 Publications

The research has led to the following publications:

- **Mengxue Tang**, and Mikko Vehkaperä, “On the Performance of Full-Duplex Relaying Schemes for Point-to-Point MIMO with Large Antenna Arrays,” *International Symposium on Wireless Communication System (ISWCS)*, pp. 246-251, 2017.

- **Mengxue Tang**, Mikko Vehkaperä, Xiaoli Chu, and Risto Wichman, “LI cancellation and Power Allocation for Multipair FD Relay Systems with Massive Antenna Arrays,” *IEEE Wireless Communication Letters*, vol. 8, no. 4, pp. 1077-1081, Mar 2019.
- **Mengxue Tang**, Mikko Vehkaperä, Xiaoli Chu, and Risto Wichman, “Power Allocation for Multipair Massive MIMO FD Relay Systems with Low Resolution ADCs,” *International Symposium on Wireless Communication System (ISWCS)*, pp. 505-510, 2019.
- **Mengxue Tang**, Mikko Vehkaperä, Risto Wichman and Xiaoli Chu, “Carrier Frequency Offset and Channel Estimation in Uplink Massive MIMO Systems with One-bit ADCs,” *IEEE Transactions on Wireless Communications*. (In preparation)

1.5 Thesis Outline

This thesis is composed of seven chapters.

Chapter 2 first provides a broad literature review on the existing techniques in massive MIMO and full duplex systems. We explore the possibility of deploying massive number of antennas at different nodes in a FD relaying network. We then point out the current passive and analog domain LI cancellation schemes does not provide sufficient level of LI power mitigation for FD relaying systems with large but finite number of antennas. Digital LI cancellation schemes should be proposed for such systems. The necessity of adapting power allocation schemes for maximizing spectral efficiency (SE) is described. In the spirit of “green” communication, we explain how the resolution of ADCs affect the circuit power consumption of a massive MIMO system and the advantages of adapting low-resolution even one-bit ADCs. The state-of-

the-art frequency synchronization techniques for massive MIMO system are reviewed. Simple methods for massive MIMO with one-bit ADCs are yet to be proposed.

Based on the discussion of the possibility of deploying massive number of antennas at different nodes in a FD relaying network in the literature review, in Chapter 3, we investigate three types of FD relay (co-located, distributed cooperative and distributed non-cooperative relaying), where massive number of antennas are deployed at the source and destination devices. It is theoretically proved that as the number of antennas at the source and destination devices grows to infinity, the systems are multi-user interference, LI and noise free. A pilot-based digital LI cancellation scheme is proposed for finite number of antennas case. Based on MF and ZF precoding/detection, the E2E achievable rates of the system is analyzed. Numerical results show that cooperative FD relaying almost doubles the SE of HD relaying.

In Chapter 4, we investigate a system that is the “mirror” image of the system described in Chapter 3, where the large number of antennas are only deployed at the FD relaying. We propose a novel pilot protocol and an iterative power allocation algorithm. The novel pilot protocol enables the digital LI channel estimation by exploiting the coherence time difference between wireless mobile channels and static LI channel. The low complexity iterative power allocation algorithm that uses only the statistics of the channel matrices is proposed to mitigate residual LI and maximize E2E achievable rate. Numerical results show that the proposed iterative power allocation algorithm converges fast and outperforms fixed relay power case in terms of SE and EE.

In order to reduce circuit power consumption, Chapter 5 investigates a system similar to the one studied in Chapter 4, but with low-resolution ADCs applied at the decode-and-forward FD relay. The SE and EE of the decode-and-forward FD relay are studied. A power allocation scheme controls the link-wise transmit powers at the relay by taking into account the resolution of the ADCs and the transmit-side hardware impairments is

proposed. We then carry out detailed EE investigation and examine the effects of finite code word length. Numerical results shows that 4 to 7 bits ADCs benefits EE the most. To further reduce circuit power consumption, one-bit ADCs could be adopted in the systems with a large number of antennas. Chapter 6 studies frequency synchronization problems in massive MIMO systems with one-bit ADCs. Three CFO estimation methods has been proposed just using two pilots for each user. The first method utilize the statistics of the phase noise from the “unbiased” average correlation of two received pilots. The second and third method make the infinite number of antennas and extremely small CFO assumptions, respectively. These CFO estimation methods are then expanded to multiple pilot cases. The channel estimation and CFO estimation share the same pilot sequences thus does not generate extra redundancy and allows CFO updates between different coherence blocks. Numerical results show the performance of the proposed estimator. It is shown that the mean-square-error of the proposed CFO estimation as a function of CFO is cyclic. By using a small number of pilots, proposed CFO estimator can achieve good performance.

The last chapter first concludes the thesis: our proposed power allocation scheme for FD relaying has better SE and EE than the maximum relaying power and HD relaying cases; the MSE of the proposed CFO estimation method can reach 10^{-5} with just 24 pilot symbols for massive MIMO systems with one-bit ADCs. We suggest the main focus of the future work be proposing joint source-relay power allocation schemes for FD relaying systems, extending the massive MIMO FD systems millimeter wave scenarios and proposing CFO estimation methods for Massive MIMO-OFDM systems.

1.6 Notations

In this thesis, scalars are denoted by italicized letters/symbols, while vectors and matrices are represented by bold lower and upper case letters/symbols. Matrix/vector

operations of inversion, transpose, conjugate and Hermitian are $(\cdot)^{-1}$, $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$, respectively. Trace operation of a matrix, denoted by $\text{tr}(\cdot)$, sums the diagonal elements of the matrix. Euclidean norm of a vector is denoted by $\|\cdot\|$, while absolute value of a scalar is $|\cdot|$. The expectation and variance of a variable are denoted by $E\{\cdot\}$ and $\text{Var}\{\cdot\}$, respectively. Normal distribution is denoted by \mathcal{N} , while \mathcal{CN} denotes circularly symmetric complex Gaussian distribution. We use $\mathcal{Q}(\cdot)$ and $\text{sign}(\cdot)$ to denote one-bit quantization and the sign function. A tilde and a hat over a vector denote the estimate and estimation error of the vector, respectively. Q-function is denoted by $Q(\cdot)$. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ take the real and imaginary part of the signal, while subscripts $(\cdot)_I$ and $(\cdot)_Q$ denote the in-phase and quadrature of a complex random variable, respectively. The vectorization of a matrix, denoted by $\text{vec}(\cdot)$, means stacking the columns of the matrix into one vector. $\angle(\cdot)$ takes the angle. $\text{cov}(\cdot, \cdot)$ calculates the co-variance between two variables.

Chapter 2

Literature Review

2.1 Massive MIMO

In order to achieve higher gains and simplify signal processing, massive MIMO has received extensive attention [15]. Massive MIMO simultaneously serves multiple terminals with same frequency and time resource. Large number of antennas helps focusing energy to specific areas to active terminals, resulting in a boost in throughput and energy efficiency. It can simultaneously serve significantly more terminals comparing to multi-user MIMO. Massive MIMO inherits from multi-user MIMO the advantages of simple resource allocation and serving users in same time-frequency resource [10], and simplifies the signal processing both at the terminals and the massive MIMO base station via low complexity precoding schemes. It relies on spatial resources to separate transmit streams, which is often named as spatial-division multiplexing. How many streams/terminals a BS can serve simultaneously depends on the degrees of freedom of the channel [10]. Massive MIMO provides large degrees of freedom which can also be used in signal shaping to transmit small peak-to-average ratio signals, in order to reduce the non-linear distortion generated by analog-to-digital converters [16].

From the energy efficiency point of view, the power RF output of each stream reduced to one third of current techniques, which significantly increases energy efficiency [9]. Consequently, it makes hardware easy to implement as low-cost RF amplifier can be used. If we denote the number of antennas at BS as M , massive MIMO have same performance as single-input-single-output (SISO) system by using M times less transmit power when perfect channel state information (CSI) is known, or by using \sqrt{M} times less power as imperfect CSI is acquired [17]. This boots the energy efficiency of the future wireless networks that is exhausted in finding solutions for reducing energy consumption. If orthogonal frequency-division multiplexing (OFDM) is used, the law of large numbers also guarantees channel gains of each subcarrier to be at similar level, simplifying frequency domain scheduling [10]. Multipath propagation brings an effect, termed Rayleigh fading, would cause deep fading, during which data can not be transmitted in the standard systems [10]. A typical way to avoid data loss is to wait until channel changes to a better condition, which would cause a pause in transmission. This phenomenon is natrually diminished in massive MIMO system because of spacial diversity from precoding.

2.1.1 Precoding

A key baseband signal processing technique for massive MIMO to reduce energy consumption as well as mitigate inter-user/stream interference is precoding. Optimal non-linear precoding schemes like dirty paper coding [18] and vector perturbation [19] are considered inappropriate for massive MIMO due to complexity. Near optimal non-linear precoding techniques have been studied to adapt massive array scenario with a lower complexity. For instance, by adding an adjustable stopping searching threshold, complexity of vector perturbation scheme becomes controllable at the cost of acceptable sum rate loss [20], and by exploiting the channel correlation matrix.

Based on channel estimates, both linear and non-linear digital baseband precoding techniques can be applied in massive MIMO systems. Unlike multi-user MIMO systems, simple linear precoding schemes provides near-optimal performance in massive array scenario, such as standard schemes matched filtering (MF) and zero-forcing (ZF) [9]. The performance of ZF surpasses MF at high signal-to-noise ratio (SNR) region, while in low SNR region, it is slightly worse than its counterpart. This is due to the sensitiveness of ZF to additive white noise. Meanwhile, ZF mitigates inter-user interference to a extremely low level [11] and outperforms MF for high spectral-efficiency and low energy-efficiency regions [21]. Although ZF shows higher complexity in matrix pseudo-inverse operation, the complexity can be reduced by regularized dual decomposition [22], gradient approach [23] and sequential interference cancellation [24].

Other precoding techniques are proposed for more practical scenarios. The peak to average power ratio (PAPR) is defined as the ratio of the peak to average power or envelope of the transmitted signal. High PAPR of the transmitted signal causes non-linear distortion both at the power amplifier and digital-to-analog converters (DACs), which severely degrades the spectral efficiency and system performance [25]. Constant envelope precoding schemes [26, 27], for example, are proposed to reduce PAPR. Another aspect of implementation problem is grow-with-number-of-antenna complexity. A precoding method has been proposed to reduce computational complexity by norm descent search [28], but this scheme only works for multiple-input-single-output (MISO) system. Discovering low-complexity precoding schemes for practical scenarios are still of great importance.

2.1.2 Pilot Transmission and Channel Estimation

In contrast to standard multi-user MIMO and point-to-point MIMO which can operate in both frequency division duplex (FDD) and time division duplex (TDD) mode,

massive MIMO usually uses TDD operation. The reason for choosing TDD operation is directly related to spatial multiplexing, of which the essence is acquiring instantaneous CSI for both uplink and downlink before data transmission [9]. A typical way of getting CSI, in current systems like LTE, is transmitting pilots in both directions. In massive MIMO systems, for uplink, users transmit pilot sequences to BS, and CSI is estimated in the BS. For downlink, however, it is not feasible for BS with massive antennas to transmit pilots to terminals, the reason being [10]:

- For an arbitrary user device, it has to estimate M CSI from M antennas of massive MIMO BS. Since all the antennas at the BS uses the same frequency-time resources to transmit signals, the received signal at the user device at a certain time t is a combination of all the signals from M antennas. In order to use the low-complexity estimators like minimum-mean-square-error (MMSE) estimator, it is crucial that the target random variable is not interfered with other random variables [29]. To estimate the CSI of an arbitrary antenna at the BS to the user device, we need to eliminate the interference from other antennas. A straight-forward way would be transmitting a diagonal pilot matrix, which means only one antennas is transmitting pilot symbol at one time while the rest of the antennas keep silent. In this case, the received signal at the terminal is not interfered and it takes M pilot symbols to obtain all the CSI. However, it will cause high PAPR of the pilot symbol as at any time, the transmit power of the symbol is focused on one antenna. For optimum performance while maintaining relatively low PAPR of the pilot symbols, pilot sequences can be designed to be pair-wise orthogonal, which requires the length of the pilot sequence to be at least M . Thus the pilot length for downlink channel estimation, scales with the number of antennas at BS, which would require ten or hundred times more resources than standard systems for pilot transmission.

- Increasing the number of antennas at BS leads to more channel responses that should be estimated at user devices. This would cause processing delay or be even impossible to process.

From the discussion above, we can conclude that in massive MIMO systems, downlink pilot transmission causes severe overhead. It is not feasible for implementation. Utilizing the reciprocity of the channel, which means the channel responses are identical for both directions, guarantees downlink CSI acquisition with only uplink pilot transmission [30].

After receiving pilots, channel estimation can be carried out at the BS. Minimum mean-square error (MMSE) estimators are widely used in theoretical analysis of massive MIMO, aiming at minimizing mean-square error (MSE) between transmitted pilots and their estimates [31].

2.2 Full Duplex

Most of nodes in wireless communication networks, such as BS, terminals and relay stations, have the function of transmitting and receiving signals. Generally, these nodes are transmitting and receiving signals at different time or frequency resources, which is termed half-duplex (HD) mode. Full duplex (FD) is a concept in contrast to half duplex, which is to simultaneously transmit and receive signals in the same time and frequency resource [32]. It provides potentials to double the spectral efficiency. The concept was first brought up a long time ago, around 1940s [32]. Since what is transmitted can have huge impact on received signals, namely self-interference (SI) or loop-interference (LI), it has not been widely applied in cellular and WiFi systems. Analysis in [33] provides data in a contemporary femto-cell system to state the reason: even after SI cancellation, the gap between received residual interference and noise floor is still 52 dB, and not to mention larger transmit power causes more interference,

if implemented in a larger-size cell. Transmit power of the BS of a small-size cell is usually 20 to 35 dBm according to LTE standard [34], which is significantly smaller than that of a large-size cell, thus reducing the SI at the same time [32]. It is the reason that FD has been extensively studied to meet the requirements of increasing throughput in small-cell systems.

2.2.1 Self-interference Cancellation

The SI is composed as linear and non-linear components [35]. The linear components majorly depend on the multipath propagation between the transceivers and could be easily mitigated via channel estimation [36]. The non-linear components are generally caused by the saturation of power amplifier in high-power region, mixer non-linearity, and other transmitter imperfection from the RF chain as *I/Q* imbalance, oscillator imperfection, ADC quantization noise, etc [35].

Consider a small-cell FD BS with the average transmit power of 23 dBm operating in 20 MHz bandwidth. The thermal noise floor is -101 dBm. If the distortion of the RF chain is taking into account, the noise level can be raised to -90 dBm [34]. The SI cancellation schemes should be at least 113 dBm to reduce the SI to the noise floor.

According to [32], there are currently three categories of SI cancellation schemes, namely passive propagation-domain, active analog-circuit-domain and active digital-domain schemes.

- Propagation-domain cancellation, also called “passive suppression”, focus on electromagnetically separating transmit streams from received streams [32]. Results shows that physically isolating transmit antennas from receive antennas provides larger path loss, thus reducing the received interference level up to 15 – 30 dB [37, 34]. Placing copper or aluminum plates between transmit and receive antennas absorbs the electromagnetic waves that achieves around

10 dB mitigation for 2.4 GHz according to the experiment in [38]. If the same antenna is used for FD operation, using dual-polarization for receiving and transmitting could bring 15 – 30 dB SI cancellation [38]. Another typical way is to use directional antenna to perform beamforming [35]. Propagation-domain cancellation schemes cannot provide good performance of cancelling interference from rich-scattering environment.

- Active analog-circuit-domain cancellation only works well for narrow-band signals [32]. As SI can be represented by a complex signal, it can be subtracted at receiver directly in analog domain, via creating an auxiliary transmit chain to regenerate the transmit signals as analog cancellation signals [35]. This can be achieved via tapping the transmit signal and passing the tapped signal through a inverter to feed the receive-side circuit with the delay and attenuation matching the wireless environment [37, 39]. Since it is very sensitive to the delay and building a programmable delay with high precision is very challenging, there is a limited cancellation ability of around 25 dB [35]. An alternative method is using the linear combination of multiple taps with different delay and attenuation factors for SI cancellation [36]. The experiments in [36] show that with eight delay taps, it can achieve 48 dB SI cancellation.
- Active digital-domain cancellation schemes do not have good performance if used alone, as the nonlinear ADC causes severe distortion. It has to be applied with aforementioned schemes to further mitigate interference. Similar to analog-domain cancellation, the linear components of the SI could be subtracted from the baseband signals using the channel estimates. Some of the non-linear components, such as power amplifier distortion and oscillator noise, could be modeled as a polynomial or parallel Hammerstein model [40, 41] and phase noise [42], respectively. Other non-linear components such as distortions caused ADC are generally treated random and difficult to model accurately. It should

be carefully dealt with in the system design phase [35]. If multiple antennas are used at the transmitter and receiver, spacial suppression from beamforming could be used for SI suppression via digital signal processing [43, 44]. The architecture proposed in [36] has achieved 55 dB digital domain interference cancellation.

With the SI cancellation schemes mentioned above, combining FD with other techniques becomes feasible. FD MIMO two-way communication systems are studied in [45], where SI cancellation together with joint beamforming and power allocation schemes provides optimized throughput. An antenna selection scheme of the same scenario is proposed in [46]. There are several papers combine massive MIMO and FD for cellular [47] and WLAN systems [48]. Among all these systems, probably the most hot topics lie in relaying systems.

2.2.2 Full-duplex Relaying

First been brought up in 1700s, relaying is not a new topic to wireless communications. Facilitating relying in wireless systems can be beneficial in following aspects [49]. It enlarges the coverage of the system by improving SNR. And by introducing a relay node between transmitter and receiver, reliability of the link is hugely increased. Moreover, it brings higher throughput, if both source-to-relay ($S \rightarrow R$) and relay-to-destination ($R \rightarrow D$) links have higher SNR than direct link.

As FD brings the potential to double the spectral efficiency and relaying topology is relatively easier than other systems, FD relaying has received significant interest. In [50], with perfect LI cancellation, inserting FD relay nodes in a network can increase degree of freedom to the order of $\log(\text{SNR})$. With the developement of wireless communication systems, relay stations are widely equipped with multiple antennas. In [51], covariance optimization problem in FD MIMO relaying with perfect

CSI is studied. Information-theoretical analysis of FD MIMO relaying with imperfect LI cancellation is studied in [52]. Moreover, massive arrays has been applied to full-duplex relaying [12], where it was shown that when number of antennas grows large, LI disappears.

2.3 Massive MIMO Full-duplex Relaying Systems

Although massive MIMO and full-duplex relaying solves some traditional problems of wireless systems, they produce completely new research problems to be investigated [10, 36].

2.3.1 Relaying Schemes

The combination of massive MIMO and FD techniques for relay channels has been an active field of research recently (see, e.g. [53, 12, 54] and references therein).

Configuring more antennas in relaying will provide higher multiplexing gain. By introducing unlimited antennas in relay station, loop interference vanishes, which makes time-consuming channel estimation unnecessary for loop interference cancellation [12]. However, with large but limited number of antennas, LI still exists, and more time will be spent for accurate channel estimation during a block fading period, which may in turn degrade spectral efficiency. Thus, effective LI cancellation schemes should be carried out.

In the communication network, massive arrays are not necessarily equipped in relay nodes. In some scenarios, source and destination with large antenna arrays need to communicate through relay, where the relaying is carried out either by a single relay node with multiple co-located antennas, or a set of individual distributed nodes that may operate cooperatively or fully independently. In the case of geographically

distributed nodes, inter-relay cooperation facilitates the relay stations to exchange information that can be used, for example for LI cancellation, at the cost of processing delay and additional bandwidth [55]. On the other hand, single relay node equipped with multiple antennas will suffer from more serious initial LI since the antennas are closely placed, but does not suffer from difficulties in acquiring information needed for LI cancellation.

2.3.2 Pilot Protocol

As mentioned in Section 2.3.1, for FD relaying with massive MIMO, LI vanishes when number of antenna goes to infinity. However, if number of antenna is finite, LI still limits achievable rate. Thus, effective instantaneous CSI of LI channel acquisition for LI channel estimation should be carried out. Since the relaying is equipped with a massive number of antennas, sending pilots for LI channel estimation would cause severe pilot overhead. A potential way to design a pilot protocol trying to bypass the pilot overhead problem is by using the fact of channel changes slowly for static transceivers, especially with line-of-sight (LOS) propagation. A measurement for coherence time of static transceiver with multiple arrays at 2.9 GHz is conducted by Nokia Bell Labs campus in Germany [56]. As the transceivers are fixed, Doppler frequency from the movement of transceivers is not the dominant factor that affects coherence time. The temporal changes in channel vectors, which is called subspace drift, however, directly determines the coherence time. Threshold 0.985 is chosen for channel matrix correlation coefficients, giving the average coherence time of 0.2-0.4 s for massive MIMO transmission. An example given in [15] shows that terminals moving at 40 km/h in a cellular system, coherence time is around 500 ms, which implies that the coherence time of fixed terminals is nearly five to ten times that of standard moving terminal setups.

Pilot transmission protocol can be designed to adjust different coherence time. An example of pilot transmission protocol design and channel estimation for FD relaying can be found in [57], where the pilots for LI channel are repeated several times during the first block (a block fading time is chosen by user-to-relay channel), interfered with users' pilots. The acquired LI CSI can be used for LI cancellation for 6 following blocks, as it is assumed that coherence time for LI channel is 6 times longer than user-to-relay channels. The scheme does not fit for relaying with massive arrays, as the pilot sequence might be longer than a block. An efficient pilot transmission protocol should thus be designed for very large antenna array systems.

2.3.3 Power Allocation

Power allocation problems have been widely studied for various purpose, for example, to improve energy efficiency, spectral efficiency and total system throughput [58, 12, 59, 60]. In multi-cell massive MIMO scenario, power allocation among each stream or user has been studied for both uplink and downlink to maximize the global energy efficiency [58]. The power allocation problem becomes more complicated when considering FD relaying case as due to multi-hop transmission and SI. For a multi-pair FD relay station with massive number of antennas, power allocation and precoder design would be required at the same time for optimum E2E sum rate [61, 62]. Relay stations with ideal hardware and massive antenna arrays are known to have the ability to asymptotically cancel LI without LI CSI when the number of antenna elements grows without bound [12]. With non-ideal hardware and large but finite number of antennas, due to the limited coherence time of the mobile wireless channels, obtaining LI CSI in massive FD relay systems is often considered infeasible and has lead to unrealistic assumptions on the level of passive and analog cancellation in the literature (see e.g. [12, 63]). Although from 2.3.2 a pilot enabled digital channel estimation could be accomplished by exploiting the coherence time difference between wireless

mobile channels and LI channel, the digital channel estimation does not have very high accuracy due to the non-linear distortion generated by transmit-side hardware impairments. The residual LI [63, 64] is strong enough to have severe impact on the end-to-end (E2E) data rate. Building on this, it is crucial to propose a power allocation scheme that mitigates the residual SI via reducing the transmit power of the relay station while balancing the E2E data rate.

2.4 Analog-to-Digital Converters

Among all the components of the radio frequency (RF) chain, analog-to-digital converters (ADCs) consume significant power [65]. The ADCs convert the sampled, usually at a rate much higher than the Nyquist sampling rate, continuous analog signals to discrete digital bits [66]. An N -bit ADC requires $2^N - 1$ voltage comparators to generate quantization bins simultaneously [67]. It divides its output value range into 2^N segments and decides which of the segments the sampled analog signal amplitude falls. It results in an exponential power consumption growth with the increasing resolution N of the ADC [68]. Meanwhile, to ensure that the sampled analog signal amplitudes are within the spurious-free dynamic range of the ADCs, a variable gain amplifier (VGA), which consumes around 2 mW power, should be deployed for each antenna [65].

2.4.1 Low Resolution ADCs

The use of low resolution ADCs in massive MIMO systems has been investigated in several works, i.e., 4- to 6-bit ADCs can achieve the trade-off of spectral and energy efficiency for systems with a massive number of antennas through precoder design [69], power allocation [70, 71] and interference cancellation [72]. To the best of our knowledge, however, the research in massive MIMO FD relaying with

low resolution ADCs is limited to the case of amplify-and-forward mode, where the residual LI is assumed to be weak and independent of the quantization noise [73]. This assumption is reasonable if ideal hardware with full resolution ADCs is employed so that both analog and digital domain LI mitigation techniques can be used. With low resolution ADCs, however, the power of the quantization noise scales with the power of the received signal after passive isolation and analog circuit domain cancellation [74, 75]. Since the power of residual LI at this stage is high, the quantization noise has severe impact on the performance.

2.4.2 One-bit ADCs

With an extremely large number of antennas at the BS, even low-resolution ADCs, e.g., 4- to 6- bit ADCs, can still consume a significant amount of power, because the total power consumption of ADCs and VGAs grows linearly with the number of antennas, assuming a separate RF chain for each antenna is deployed. Power and chip area occupation savings could be achieved by adopting a one-bit ADC for the real and imaginary parts of the sampled analog signal separately, where the reference quantization bins of the comparators are at zero volt [68, 76]. In stead of a VGA, a limiting amplifier (LA) that consumes negligible power comparing to other hardware components can be deployed for each antenna [65]. The quantization noise generated by one-bit ADCs is non-linear in practical systems, it can be modeled as an additive noise that is uncorrelated with quantized signal, of which the second-order statistics are affected by the power of the sampled received analog signals [77], via the Bussgang decomposition [78]. Based on the statistics of the quantization noise, pilot sequences can be designed for channel estimation and power control can be applied for data transmission, in order to achieve better performance [79, 80].

2.5 Frequency Synchronization

Carrier frequency offset (CFO) describes the frequency mismatch between the oscillators of transmitter and receiver caused by hardware imperfection and Doppler shift [13]. CFO would cause phase rotation that accumulates over time. It can cause bit-error rate degradation and severely degrades the system performance, especially in the orthogonal frequency-division multiplexing (OFDM) systems that require strict orthogonality among sub-carriers [81]. The CFO estimation and cancellation methods for small-scale MIMO systems, e.g. the maximum likelihood (ML) estimators with training sequence design [82–84] and the blind CFO estimators without training sequences [85, 86], are not applicable for massive MIMO systems due to the rapidly growing computational complexity with the increasing number of antennas. Assuming with perfect ADCs, meaning that infinite quantization bits are used, the authors of [87] proposed an ML estimator for massive MIMO based on coarse grid search. A joint spatial-frequency alignment CFO estimator for individual users with the help of fast Fourier transform was proposed in [13]. For multi-user CFO estimation, the blind CFO estimator in [88] first mitigates the multi-user interference, and then performs single user CFO estimation iteratively. However, most of the existing CFO estimators have been proposed for massive MIMO systems, not all of them are applicable to the systems with one-bit ADCs due to the cross-correlation of quantization noise among pilot symbols. Only a few works have been reported on CFO estimation for massive MIMO systems with one-bit quantization [42], [89], [90]. In [42], CFO estimation was studied for one-bit digital-to-analog converters (DACs) deployed at the BS. For massive MIMO system with one-bit ADCs, the message-passing algorithm for channel and CFO estimation [89], and the one-dimensional optimization problem formulation for CFO estimation [90], both requires very long training sequences (up to thousands of training symbols) to achieve good performance.

Chapter 3

Full-Duplex Relaying Schemes

3.1 Introduction

Given the three FD relaying schemes (co-located, distributed cooperative and distributed non-cooperative relaying) described in 2.3.1, some interesting questions arise: 1) Does the attenuation provided by the propagation environment mitigate LI sufficiently, so that a simple distributed solution with no cooperation works efficiently in FD mode; 2) Are there any scenarios where simple HD relaying is more effective than using FD; 3) How many antennas should the relay node(s) be equipped with, given that the source and destination have (relatively) large antenna arrays. Especially the last question is not trivial since while it is clear that configuring more antennas at the relay will provide higher multiplexing gains, transmitting large numbers of pilots for channel estimation in turn degrades the spectral efficiency and potentially increases the LI in FD mode. This chapter analyzes the above questions under the assumption that the source and destination employ linear processing adhering to the “Massive MIMO” principles [91, 17, 9].

3.2 System Model

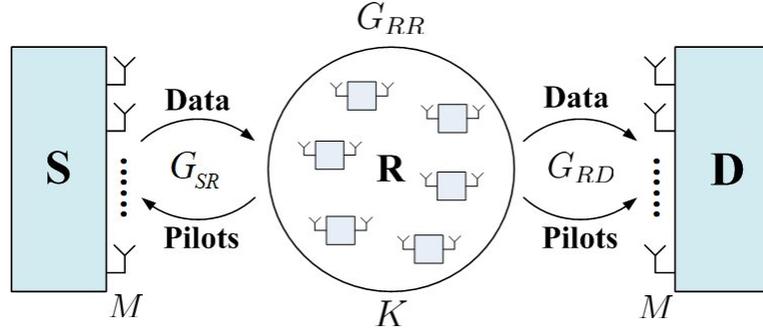


Fig. 3.1 Full-duplex relaying with distributed relay nodes.

Consider a decode-and-forward relaying system depicted in Fig. 3.1. where the source (S) and destination (D) are equipped with M antennas each. We assume that the direct link between the source and destination is blocked, so that a relay node (R) is needed in-between to facilitate the transmission. Herein two types of FD relaying strategies are considered: 1) a single relay station with $2K$ antennas¹ (co-located) and 2) a set of K individual relay stations equipped with 2 antennas each (distributed). Under both scenarios, the relay(s) operate in FD mode by using each antenna either for transmitting or receiving but not both at the same time. The co-located relay station in this study can always process signals jointly, while the distributed setup can be either cooperative or non-cooperative. The source operates always in HD mode.

At time instant i , the received signals at the relay and destination are given by²

$$\mathbf{y}_R = \sqrt{\rho_S} \mathbf{G}_{SR}^T \mathbf{x}_S[i] + \sqrt{\rho_R} \mathbf{G}_{RR} (\mathbf{x}_R[i] + \mathbf{u}_{Tx}) + \mathbf{n}'_R, \quad (3.1)$$

$$\mathbf{y}_D = \sqrt{\rho_R} \mathbf{G}_{RD} (\mathbf{x}_R[i] + \mathbf{u}_{Tx}) + \mathbf{n}'_D, \quad (3.2)$$

respectively. The transmitted signals $\mathbf{x}_S[i] \in \mathbb{C}^{M \times 1}$, $\mathbf{x}_R[i] \in \mathbb{C}^{K \times 1}$ satisfy power constraints $\mathbb{E}\{|\mathbf{x}_S[i]|^2\} = \mathbb{E}\{|\mathbf{x}_R[i]|^2\} = 1$ so that ρ_S and ρ_R are the total transmit powers

¹The co-located relaying can also be the case where K individual relay station with 2 antennas each locate close enough to each other that could be treated as a single device.

²For notational simplicity, the time index is explicit only for transmitted signals.

of the source and relay. $\mathbf{G}_{\text{SR}}^T \in \mathbb{C}^{K \times M}$ and $\mathbf{G}_{\text{RD}} \in \mathbb{C}^{M \times K}$ represent the S→R and R→D channel matrices, respectively. We assume TDD operation and reciprocity so that the R→S channel is \mathbf{G}_{SR} . For the S→R and R→D links, $\mathbf{G}_* = \bar{\mathbf{G}}\mathbf{D}_*^{1/2}$, where the entries of $\bar{\mathbf{G}}$ are i.i.d. standard complex Gaussian random variables (RVs) representing small scale fading and the large scale attenuation are modeled by diagonal matrices \mathbf{D}_{SR} and \mathbf{D}_{RD} whose k th diagonal entries are denoted $\beta_{\text{SR},k}$ and $\beta_{\text{RD},k}$, respectively. Since the transmit antenna(s) and receive antenna(s) at the same relay station are very close, LI channel follows the near-field propagation model, while the LI from different distributed relay stations are still considered far-field propagation³. For simplicity, we still model the LI matrix as $\mathbf{G}_{\text{RR}} \in \mathbb{C}^{K \times K}$ that has independent circularly symmetric complex Gaussian (CSCG) elements, where the kj th element has variance $\beta_{kj} = \text{E}\{\|g_{kj}\|^2\}$. Throughout the chapter we assume that the variance of diagonal elements g_{kk} of the LI channel are the same for all nodes, i.e. $\beta_{kk} = \beta \forall k$ since they are always located in same device and, thus, close to each other. By the same argument, in the co-located case (single relay node) we have $\beta_{kj} = \beta \forall k, j$.

Finally, the elements of the receive-side hardware impairments vectors \mathbf{n}'_{R} and \mathbf{n}'_{D} are i.i.d. standard complex Gaussian RVs. The system model also encompasses transmit-side hardware impairments that is modeled by vector \mathbf{u}_{Tx} whose entries are i.i.d. CSCG RVs $\mathcal{CN}(0, \sigma_u^2)$ [94]. For simplicity, we denote the combined transmit- and receive-side hardware impairment terms at the destination and relay as

$$\mathbf{n}_{\text{D}} = \mathbf{n}'_{\text{D}} + \sqrt{\rho_{\text{R}}}\mathbf{G}_{\text{RD}}\mathbf{u}_{\text{Tx}}, \quad (3.3)$$

$$\mathbf{n}_{\text{R}} = \mathbf{n}'_{\text{R}} + \sqrt{\rho_{\text{R}}}\mathbf{G}_{\text{RR}}\mathbf{u}_{\text{Tx}}, \quad (3.4)$$

³In the far-field propagation model, the pathloss follows the Friis' Law [92] where the transmit power rolls off as the inverse of the squared distance between transmit and receive antennas ($1/d^2$). The near-field transmit power rolls off as inverse forth ($1/d^4$) or even higher [93]. This means the SNR is generally higher than the far-field propagation, which means LI is stronger. The delay spread in the near-field is different and usually depend on the echo response of the environment [93].

respectively. Since the receiver knows only the statistics of the channel a priori, we take the worst case scenario for the channel estimation, where the noise is spatially uncorrelated CSCG and independent of the channel. The variance of the k th element in \mathbf{n}_D and \mathbf{n}_R are given by $\sigma_{R,k}^2 = \rho_R \sigma_u^2 (\beta + \sum_{j \neq k} \beta_{kj}) + 1$ and $\sigma_{D,k}^2 = \rho_R \sigma_u^2 \sum_K \beta_{RD,k} + 1$, respectively.

3.2.1 Channel Estimation

Block fading channel with coherence time of T symbols is considered. Pilot-aided channel estimation is used to obtain the instantaneous channel state information (CSI) needed for precoding at the source, loop interference cancellation at the relay and detection at the source. If t_p symbols are allocated for pilots, then $t_d = T - t_p$ symbols are available for data transmission. There are two phases in pilot transmission, which take t_{pS} , t_{pD} symbols each, so that $t_p = t_{pS} + t_{pD}$.

Channel estimation for precoding at the source

The antennas in the relay station that are used to receive the data transmissions from the source send orthogonal pilot sequences $\Phi_S \in \mathbb{C}^{K \times t_{pS}}$ to the source. To satisfy the power constraint and guarantee orthogonality, we require $\Phi_S \Phi_S^H = \frac{1}{K} \mathbf{I}$ with $t_{pS} \geq K$. The received pilot matrix at the source reads

$$\mathbf{Y}_S = \sqrt{t_{pS} \rho_{pS}} \mathbf{G}_{SR} \Phi_S + \mathbf{N}_S, \quad (3.5)$$

where $t_{pS} \rho_{pS}$ is the total energy consumed in this training phase at the relay(s). The source estimates \mathbf{G}_{SR} and uses it for precoding, owing to the reciprocity provided by the TDD operation. Analogous to the R→D link and (3.3), the noise matrix \mathbf{N}_S in (3.5) has independent CSCG entries with the elements in the k th row having variance $\sigma_{S,k}^2 = \frac{t_{pS}}{K} \rho_{pS} \sigma_u^2 \sum_K \beta_{SR,k} + 1$.

Channel estimation for decoding at the destination and LI cancellation at the relay(s)

Just like the source node above, the destination estimates the channel \mathbf{G}_{RD} based on orthogonal pilot sequences $\Phi_{\text{D}} \in \mathbb{C}^{K \times t_{\text{pD}}}$ transmitted by the relay. We let $t_{\text{pD}}\rho_{\text{pD}}$ be the total energy consumed in this training phase. At the same time, the relay estimates the LI channel from the same pilots. By (3.3) and (3.4), the noise matrices during this training phase are as in (3.5), i.e., independent CSCG entries with k th row's elements having variances $\sigma_{\text{D},k}^2 = \frac{t_{\text{pD}}}{K}\rho_{\text{pD}}\sigma_u^2 \sum_K \beta_{\text{RD},k} + 1$ for destination and $\sigma_{\text{R},k}^2 = \frac{t_{\text{pD}}}{K}\rho_{\text{pD}}\sigma_u^2(\beta + \sum_{l \neq j} \beta_{kj}) + 1$ for the relay.

Having described the signal model in the training phase, we now assume that the minimum mean square error (MMSE) channel estimator is used at all nodes to obtain the instantaneous channel estimates $\hat{\mathbf{G}}_{\text{SR}}$, $\hat{\mathbf{G}}_{\text{RD}}$ and $\hat{\mathbf{G}}_{\text{RR}}$. By the properties of the MMSE estimator [29], the error $\tilde{\mathbf{G}}_* = \mathbf{G}_* - \hat{\mathbf{G}}_*$ is uncorrelated with the estimate and they both have independent CSCG entries. Following similar derivations as in [12], the LI error matrix $\tilde{\mathbf{G}}_{\text{RR}}$ has then independent CSCG elements with the variance of the kj th element being

$$\tilde{\beta}_{kj} = \frac{\beta_{kj}}{t_{\text{pD}}\rho_{\text{pD}}\beta_{kj}/K\sigma_{\text{R},k}^2 + 1}. \quad (3.6)$$

The error matrices $\tilde{\mathbf{G}}_{\text{SR}}$ and $\tilde{\mathbf{G}}_{\text{RD}}$ have independent CSCG elements, the variance of the entries in the k th column being

$$\tilde{\beta}_{\text{SR},k} = \frac{\beta_{\text{SR},k}}{t_{\text{pS}}\rho_{\text{pS}}\beta_{\text{SR},k}/K\sigma_{\text{S},k}^2 + 1}, \quad (3.7)$$

$$\tilde{\beta}_{\text{RD},k} = \frac{\beta_{\text{RD},k}}{t_{\text{pD}}\rho_{\text{pD}}\beta_{\text{RD},k}/K\sigma_{\text{D},k}^2 + 1}, \quad (3.8)$$

respectively. The properties of the MMSE estimator also guarantee that $\hat{\beta}_* = \beta_* - \tilde{\beta}_*$ holds for all channels.

3.2.2 Data Transmission

After training, data transmission phase ensues and the source uses $\hat{\mathbf{G}}_{\text{SR}}$ to carry out linear precoding [91, 9]

$$\mathbf{x}_S[i] = \mathbf{V}\bar{\mathbf{x}}[i], \quad (3.9)$$

of the information vector $\bar{\mathbf{x}}[i] \in \mathbb{C}^K$ intended for the destination. Here we focus on MF precoding

$$\mathbf{V}_{\text{MF}} = \lambda_{\text{MF}} \hat{\mathbf{G}}_{\text{SR}}^*, \quad (3.10)$$

$$\lambda_{\text{MF}} = \frac{1}{\sqrt{\frac{M}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}}, \quad (3.11)$$

and ZF precoding

$$\mathbf{V}_{\text{ZF}} = \lambda_{\text{ZF}} \hat{\mathbf{G}}_{\text{SR}}^* (\mathbf{G}_{\text{SR}}^T \mathbf{G}_{\text{SR}}^*)^{-1} \quad (3.12)$$

$$\lambda_{\text{ZF}} = \frac{1}{\sqrt{\frac{1}{(M-K)K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}^{-1}}} \quad (3.13)$$

where λ_{MF} and λ_{ZF} guarantee that the long-term transmit power constraint $\mathbb{E}\{|\mathbf{x}_S[i]|^2\} = 1$ is satisfied.

By (3.1) and (3.9), the the k th relay node (or k th antenna element in the co-located setup) receives the signal

$$y_{\text{R},k}[i] = \underbrace{\sqrt{\rho_S} \mathbf{g}_{\text{SR},k}^T \mathbf{v}_k \bar{x}_k[i]}_{\text{desired signal}} + \underbrace{\sum_{j=1, j \neq k}^K \sqrt{\rho_S} \mathbf{g}_{\text{SR},k}^T \mathbf{v}_j \bar{x}_j[i]}_{\text{inter-pair interference}} + \underbrace{\sqrt{\rho_R} \mathbf{g}_{\text{RR},k}^T \mathbf{x}_{\text{R}}[i]}_{\text{loop interference}} + \underbrace{n_{\text{R},k}}_{\text{noise}} \quad (3.14)$$

where $\mathbf{g}_{\text{SR},k}$, $\mathbf{g}_{\text{RR},k}$ and \mathbf{v}_k are the k th column of \mathbf{G}_{SR} , \mathbf{G}_{RR} and \mathbf{v} , respectively, and $\bar{x}_k[i]$ is the k th elements of $\bar{\mathbf{x}}[i]$. For all FD cases, some form of LI cancellation is then applied before detection and decoding. Note that throughout the thesis, we discuss

only digital-domain LI cancellation. We assume that analog-domain LI cancellation has already been carried out before ADCs. More precisely, we express the digital LI cancellation at the k th node / antenna as

$$y'_{R,k}[i] = y_{R,k}[i] - c_k[i], \quad (3.15)$$

where $c_k[i]$ is a function of $\hat{\mathbf{g}}_{RR,k}$. After LI cancellation, the details of which will be presented in the next subsection, information from $y'_{R,k}[i]$ is decoded⁴ to $x_{R,k}[i]$ and sent forward (after re-encoding) to the destination. We note that following the common assumption in decode-and-forward relaying there is a processing delay $\tau \geq 1$ symbols at the relay

$$\mathbf{x}_R[i] = \bar{\mathbf{x}}[i - \tau], \quad (3.16)$$

so that for any time i , transmit signal signals at relay stations are uncorrelated with receive signals [95].

After receiving the signals from the relay node(s), the destination uses linear estimation to separate the streams, so that the k th information stream after estimation reads

$$y_{D,k}[i] = \sqrt{\rho_R} \mathbf{w}_k^H \mathbf{g}_{RD,k} x_{R,k}[i] + \sum_{j=1, j \neq k}^K \sqrt{\rho_S} \mathbf{w}_k^H \mathbf{g}_{RD,j} x_{R,j}[i] + \mathbf{w}_k^H \mathbf{n}_D, \quad (3.17)$$

where $\mathbf{g}_{RD,k}$ represents the k th column of matrix \mathbf{G}_{RD} . For simplicity, we concentrate in this study on the MF based estimation so that $\mathbf{w}_k = \hat{\mathbf{g}}_{RD,j}$, and for ZF based estimation, \mathbf{w}_k represents the k th column of $\mathbf{W}_{ZF} = \hat{\mathbf{G}}_{RD} \left(\hat{\mathbf{G}}_{RD}^H \hat{\mathbf{G}}_{RD} \right)^{-1}$.

⁴The relay does not know the instantaneous (pre-coded) channel since it is not estimated at any stage. Indeed, estimating the S→R channel at relay would require another training period that could be prohibitively long if $M \gg 1$. We follow the method proposed in [91], where the relay knows only the average precoded channel coefficient(s) $E\{\mathbf{g}_{SR,k}^T \mathbf{v}_k\}$. See also Section 3.4.

3.2.3 Loop Interference Cancellation

Let us now consider the details of the two LI cancellation schemes used in this chapter for relaying.

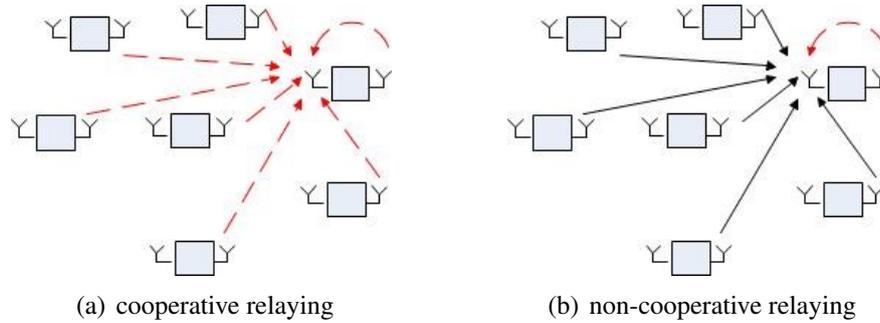


Fig. 3.2 LI cancellation schemes for FD relaying

Cancellation scheme for cooperative relay stations

If one relay station with multiple antennas is used, it is clear that the device has full knowledge of the received data \mathbf{y}_R , the transmitted signal $\mathbf{x}_R[i]$ given in (3.16) as well as the estimated CSI $\hat{\mathbf{G}}_{RR}$ discussed in Section 3.2.1. All of the above can be obtained virtually without delay using internal circuitry. In principle, cooperative distributed relays can also obtain the same information if dedicated control channel with sufficient capacity is available for them. In practice, the information needs to be quantized and there might be delays due to transmissions, albeit the latter should not cause problems if the relays operate with buffers and the initial transmissions are “ramped up” appropriately. Here we assume for simplicity that the control channel can be used instantaneously and perfectly to share the transmitted symbols and estimated CSI. However, it should be noted that this is a highly optimistic scenario that provides an upper bound for the performance of a practical system using quantization and finite capacity control channel.

For the cooperative relaying, as shown in 3.2(a), the k th relay node can use the knowledge from itself as well as other stations to conduct LI cancellation

$$c_k[i] = \sqrt{\rho_S} \hat{\mathbf{g}}_{RR,k}^T \mathbf{x}_R[i], \quad (3.18)$$

so that by (3.15), the received signal at the k th node, or antenna stream in colocated case, after cancellation reads

$$\begin{aligned} y'_{R,k}[i] = & \sqrt{\rho_S} \mathbf{g}_{SR,k}^T \mathbf{v}_k \bar{x}_k[i] + \sum_{j=1, j \neq k}^K \sqrt{\rho_S} \mathbf{g}_{SR,k}^T \mathbf{v}_j \bar{x}_j[i] \\ & + \underbrace{\sqrt{\rho_R} \tilde{\mathbf{g}}_{RR,k}^T \mathbf{x}_R[i]}_{\text{LI}_k^{\text{coop}}} + n_{R,k}, \end{aligned} \quad (3.19)$$

where $\tilde{\mathbf{g}}_{RR,k}$, $\hat{\mathbf{g}}_{RR,k}$ are the k th columns of the estimation error matrix $\tilde{\mathbf{G}}_{RR}$ and the channel estimation matrix $\hat{\mathbf{G}}_{RR}$, respectively, both of which have i.i.d. CSCG elements whose variances are as discussed in Section 3.2.1. We denote the remaining LI as $\text{LI}_k^{\text{coop}} \triangleq \sqrt{\rho_R} \tilde{\mathbf{g}}_{RR,k}^T \mathbf{x}_R[i]$ for later use.

Cancellation scheme for non-cooperative relay stations

While non-cooperating relay stations can estimate the channels that cause LI, they cannot obtain the transmit signals of other stations. Thus, only loop interference from the k th station itself can be cancelled by using the estimated CSI and the interference from the other stations remains, as shown in 3.2(b). This means that

$$c_k[i] = \sqrt{\rho_S} \hat{\mathbf{g}}_{RR,kk} x_{R,k}[i] \quad (3.20)$$

is used to cancel the loop interference caused by the node's own transmission while the inter-node interference remains unaffected. The received signal at the k th relay

node reads

$$\begin{aligned}
y'_{R,k}[i] &= \sqrt{\rho_S} \mathbf{g}_{SR,k}^T \mathbf{v}_k \bar{x}_k[i] + \sum_{j=1, j \neq k}^K \sqrt{\rho_S} \mathbf{g}_{SR,k}^T \mathbf{v}_j \bar{x}_j[i] \\
&+ \underbrace{\sqrt{\rho_R} \tilde{g}_{RR,kk} x_{R,k}[i] + \sum_{l=1, l \neq k}^K \sqrt{\rho_R} g_{RR,kl} x_{R,l}[i]}_{\text{LI}_k^{\text{non}}} + n_{R,k}, \tag{3.21}
\end{aligned}$$

where $g_{RR,ij}$ and $\tilde{g}_{RR,ij}$ are the ij th elements of \mathbf{G}_{RR} and $\tilde{\mathbf{G}}_{RR}$, respectively. The transmit symbol of the k th relay station is $x_{R,k}[i]$. As before, we write

$$\text{LI}_k^{\text{non}} \triangleq \sqrt{\rho_R} \tilde{g}_{RR,kk} x_{R,k}[i] + \sum_{l=1, l \neq k}^K \sqrt{\rho_R} g_{RR,kl} x_{R,l}[i]$$

for the residual LI in the case of non-cooperative relaying. It is clear that $\text{LI}_k^{\text{non}} > \text{LI}_k^{\text{coop}}$ and the question whether the LI is dominating in non-cooperative setup depends mostly on how strong the channels $g_{RR,kl}$ between the relay nodes are.

3.3 Asymptotic Interference Analysis

We now investigate the asymptotic region where the number of antennas at the source and destination grow without bound while the number of relay antennas is fixed and finite.

Proposition 3.1. *Assume that MF/ZF processing with estimated channel is used for precoding at the source, and estimation at the destination. Assume further that the LI cancellation schemes described in Section 3.2.3 are not used. Then, as $K/M \rightarrow 0$ for*

fixed K ,

$$\frac{y_{R,k}^{\text{MF}}[i]}{\sqrt{M}} \xrightarrow{a.s.} \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}} \hat{\beta}_{\text{SR},k} \bar{x}_k[i], \quad (3.22)$$

$$\frac{y_{D,k}^{\text{MF}}[i]}{M} \xrightarrow{a.s.} \sqrt{\rho_R} \hat{\beta}_{\text{RD},k} x_{R,k}[i], \quad (3.23)$$

$$\frac{y_{R,k}^{\text{ZF}}[i]}{\sqrt{M}} \xrightarrow{a.s.} \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}^{-1}}} \bar{x}_k[i], \quad (3.24)$$

$$y_{D,k}^{\text{ZF}}[i] \xrightarrow{a.s.} \sqrt{\rho_R} x_{R,k}[i], \quad (3.25)$$

where $\xrightarrow{a.s.}$ denotes almost sure convergence and $y_{R,k}[i]$ is the received signal at the k th station before loop interference cancellation for any of the considered relaying schemes.

Proof: See Appendix A.

The proposition implies that regardless of the relaying scheme, if massive antenna arrays are used at the source and destination then the system is free of both LI and inter-stream or inter-pair interference. This is an analogous result to the one obtained in [12] for a system that was a “mirror image” of the one considered in the present chapter. ■

3.4 Achievable Rate Analysis

The asymptotic result in Proposition 3.1 shows that the considered systems can operate free of noise and interference in the limit $M \rightarrow \infty$. However, the question how the performance of finite sized systems behaves still remains. In particular, we are interested in the achievable rate of the relaying schemes when MF processing and the LI cancellation schemes proposed in Section 3.2.3 are used.

Since the relay station does not know the instantaneous CSI and instead uses statistical channel gains, the “effective” received signal at relay station k can be expressed as [91]

$$\begin{aligned}
y_{R,k}[i] &= \underbrace{\sqrt{\rho_S} \mathbf{E}\{\mathbf{g}_{SR,k}^T \mathbf{v}_k\} \bar{x}_k[i]}_{\text{desired signal}} + \underbrace{\sqrt{\rho_S} (\mathbf{g}_{SR,k}^T \mathbf{v}_k - \mathbf{E}\{\mathbf{g}_{SR,k}^T \mathbf{v}_k\}) \bar{x}_k[i]}_{\text{effective interference + noise}} \\
&+ \underbrace{\sum_{j=1, j \neq k}^K \sqrt{\rho_S} \mathbf{g}_{SR,k}^T \mathbf{v}_j \bar{x}_j[i]}_{\text{effective interference + noise}} + \xi_k + n_{R,k},
\end{aligned} \tag{3.26}$$

where $\mathbf{E}\{\mathbf{g}_{SR,k} \mathbf{v}_k\}$ is the statistical channel gain and $\xi_k = \xi_k^{\text{non}}$ for non-cooperative relaying and $\xi_k = \xi_k^{\text{coop}}$ for cooperative relaying. The destination, on the other hand, has an estimate of the instantaneous CSI, i.e. $\hat{\mathbf{g}}_{RD,k}$ that it uses for detection, so that we have the received signal at the k th destination as

$$\begin{aligned}
y_{D,k}[i] &= \underbrace{\sqrt{\rho_R} \mathbf{w}_k^H \hat{\mathbf{g}}_{RD,k} x_{R,k}[i]}_{\text{desired signal}} + \underbrace{\sqrt{\rho_R} \mathbf{w}_k^H \tilde{\mathbf{g}}_{RD,k} x_{R,k}[i]}_{\text{effective interference + noise}} \\
&+ \underbrace{\sum_{j=1, j \neq k}^K \sqrt{\rho_S} \mathbf{w}_k^H \mathbf{g}_{RD,j} x_{R,j}[i]}_{\text{effective interference + noise}} + \mathbf{w}_k^H \mathbf{n}_D.
\end{aligned} \tag{3.27}$$

By using the fact that the worst case noise is when the additive noise and interference terms are independent of data with CSCG distribution of the same variance [96], the achievable rates for the k th stream of S→R and R→D links, denoted by $R_{SR,k}$ [bit/s/Hz] and $R_{RD,k}$ [bit/s/Hz] respectively, are lower bounded respectively by

$$R_{SR,k} = \log_2 \left(1 + \frac{\frac{\rho_S}{K} |\mathbf{E}\{\mathbf{g}_{SR,k}^T \mathbf{v}_k\}|^2}{\frac{\rho_S}{K} \text{Var}(\mathbf{g}_{SR,k}^T \mathbf{v}_k) + \sum_{j=1, j \neq k}^K \frac{\rho_S}{K} \mathbf{E}\{|\mathbf{g}_{SR,k}^T \mathbf{v}_j|^2\} + \rho_{\xi_k} + \sigma_{R,k}^2} \right), \tag{3.28}$$

$$R_{RD,k} = \mathbf{E} \left\{ \log_2 \left(1 + \frac{\frac{\rho_R}{K} |\mathbf{w}_k^H \hat{\mathbf{g}}_{RD,k}|^2}{\sum_{j=1, j \neq k}^K \frac{\rho_R}{K} \mathbf{E}\{|\mathbf{w}_k^H \mathbf{g}_{RD,j}|^2\} + \frac{\rho_R}{K} \mathbf{E}\{|\mathbf{w}_k^H \tilde{\mathbf{g}}_{RD,k}|^2\} + \mathbf{E}\{\mathbf{n}_D^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{n}_D\}} \right) \right\}, \tag{3.29}$$

where we denoted $\rho_{\xi_k} = \mathbb{E}\{|\xi_k|^2\}$ for notational convenience. It should be pointed out that these equations are valid also for other precoders / estimators.

Since the ergodic achievable rate of the k th stream is limited by the rate of the weaker link rate, the end-to-end rate of the k th stream is given by

$$R_k = \min\{R_{\text{SR},k}, R_{\text{RD},k}\}. \quad (3.30)$$

While the rates can be obtained through Monte Carlo simulation, it can still be time consuming. For this reason, we provide next simple approximations for the achievable rates when MF and ZF are used both for precoding and detection.

Proposition 3.2. *Assume that MF/ZF processing with estimated channel is used at the source and destination. Assume further that the relay uses the LI cancellation schemes described in Section 3.2.3. The end-to-end achievable rates for the source-to-destination link with DF relaying for MF/ZF, denoted by R_k^{MF} [bit/s/Hz] and R_k^{ZF} [bit/s/Hz] respectively, are then approximated as*

$$R_k^{\text{MF}} = \log_2 \left(1 + \min \left\{ \frac{\rho_S \hat{\beta}_{\text{SR},k}^2 M}{(\rho_S \beta_{\text{SR},k} + \rho_{\xi_k} + \sigma_{\text{R}}^2) \sum_K \hat{\beta}_{\text{SR},k}}, \frac{\rho_{\text{R}} \hat{\beta}_{\text{RD},k} M}{\rho_{\text{R}} \sum_K \beta_{\text{RD},k} - \rho_{\text{R}} \hat{\beta}_{\text{RD},k} + K} \right\} \right) \quad (3.31)$$

$$R_k^{\text{ZF}} = \log_2 \left(1 + \min \left\{ \frac{\rho_S (M - K)}{(\rho_S \tilde{\beta}_{\text{SR},k} + \rho_{\xi_k} + \sigma_{\text{R}}^2) \sum_K \hat{\beta}_{\text{SR},k}^{-1}}, \frac{\rho_{\text{R}} \hat{\beta}_{\text{RD},k} (M - K)}{\rho_{\text{R}} \sum_K \tilde{\beta}_{\text{RD},k} + K} \right\} \right) \quad (3.32)$$

where the remaining power of loop interference ρ_{ξ_k} is given for the different relaying strategies as

$$\rho_{\xi_k} = \begin{cases} \rho_{\text{R}} \tilde{\beta}, & \text{co-located;} \\ \frac{\rho_{\text{R}}}{K} (\tilde{\beta}_k + \sum_{j=1, j \neq k}^K \tilde{\beta}_{kj}), & \text{cooperative;} \\ \frac{\rho_{\text{R}}}{K} (\tilde{\beta}_k + \sum_{j=1, j \neq k}^K \beta_{kj}), & \text{non-cooperative.} \end{cases} \quad (3.33)$$

Spectral efficiency of the system reads then

$$\text{SE} = \frac{T - t_{\text{pS}} - t_{\text{pD}}}{T} \sum_{k=1}^K R_k. \quad (3.34)$$

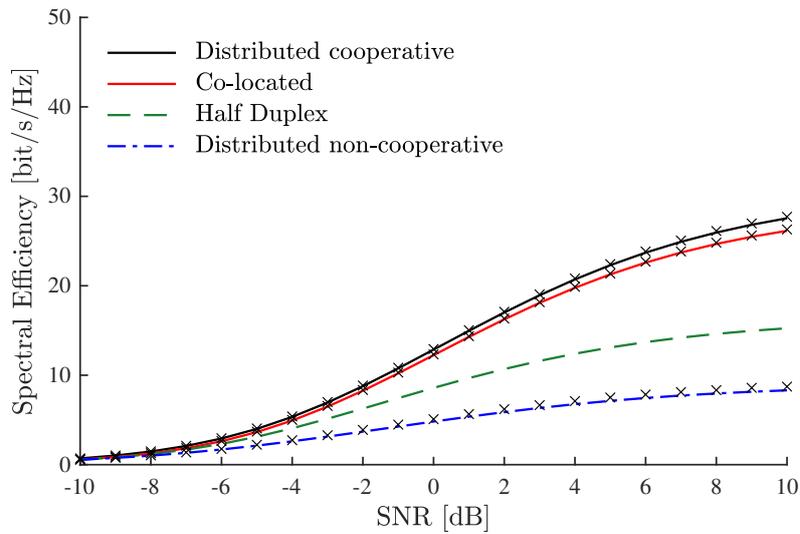
Proof: See Appendix B. ■

Note that the remaining power of loop interference ρ_{ξ_k} does not depend on the percoding and detection techniques, since percoding and detection are not performed at the relay station(s).

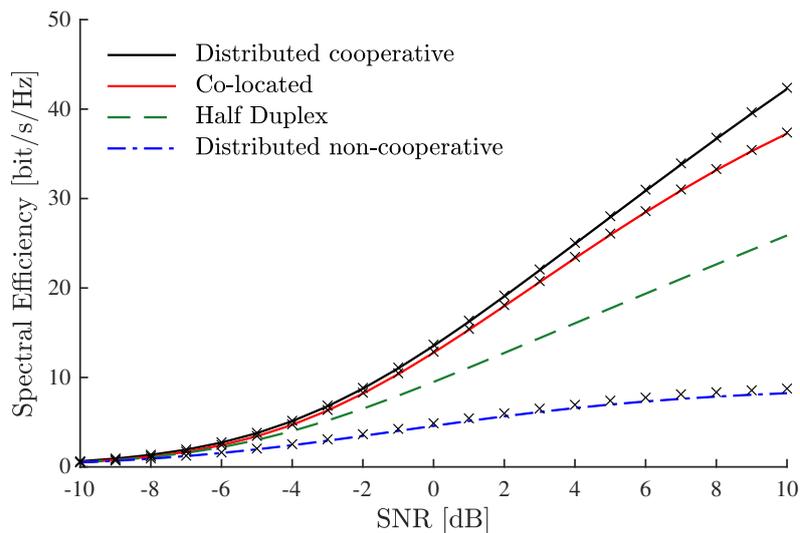
3.5 Numerical Results

Here the performance of finite-sized systems is investigated via numerical examples. Unless stated otherwise, the normalized transmit-side hardware impairments variance is $\sigma_u^2 = 10^{-3}/K$ that corresponds to Error Vector Magnitude (EVM) -30 dB, and the coherence time is set to $T = 200$ symbols. For simplicity, we assume that the large scale fading of all S \rightarrow R and R \rightarrow D links are equal, i.e. $\beta_{\text{SR},k} = \beta_{\text{RD},k} = 1 \forall k$. For co-located relay the self-interference channel strength is constant between all antenna pairs $\beta_{kj} = 20$ dB. In distributed case $\beta_{kk} = 20$ dB and the inter-node interference power $\beta_{kj} \forall k \neq j$ is assumed to be identical between all node pairs for simplicity. We set equal average transmit power constraint for source and relay (shared between the distributed nodes) and let the pilots and data have the same average powers so that $\rho = \rho_{\text{S}} = \rho_{\text{R}} = \rho_{\text{pS}} = \rho_{\text{pD}}$ represents also the signal-to-noise ratio (SNR) of the system.

Fig. 3.3 shows the spectral efficiency of the system as a function of transmit power ρ (or SNR), where in Fig. 3.3(a) and Fig. 3.3(b), MF and ZF processing are used respectively. The analytical results based on Proposition 3.2 match well with the Monte Carlo simulations depicted by the markers. The result demonstrates that if the



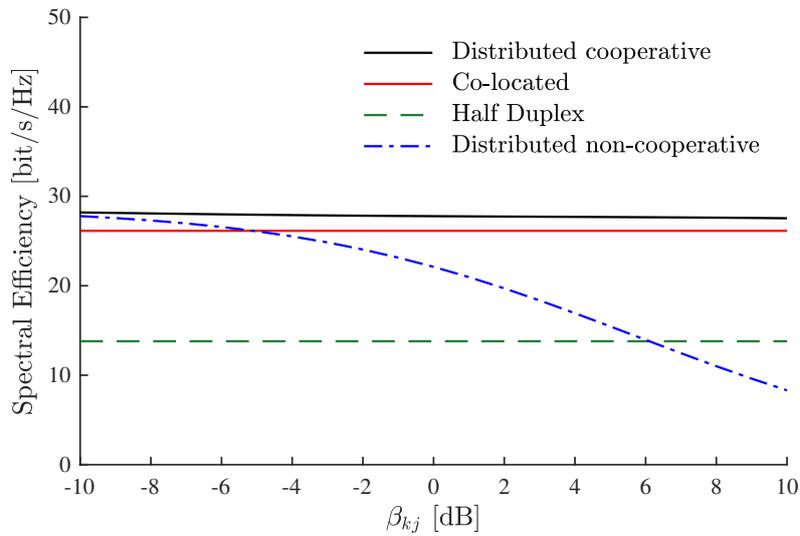
(a) MF processing



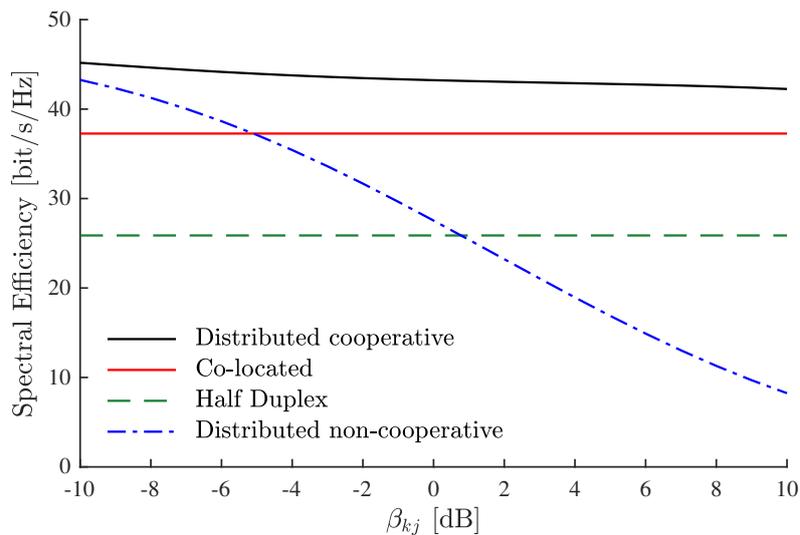
(b) ZF processing

Fig. 3.3 Spectral efficiency vs. SNR for different relaying schemes. The lines depict analytical results and markers correspond to Monte Carlo simulations. ($\beta_{kk} = 20$ dB, $\beta_{kj} = 10$ dB, $M = 100$ and $K = 10$)

channel estimation based LI cancellation can be done jointly over all antennas, FD has significant advantage over HD at moderate-to-high SNR. ZF processing outperforms MF processing in high SNR area, except for distributed setup where only device's own LI can be cancelled, the remaining inter-node interference channel strength of $\beta_{kj} = 10$ dB is too strong in this scenario for effective operation in FD mode.



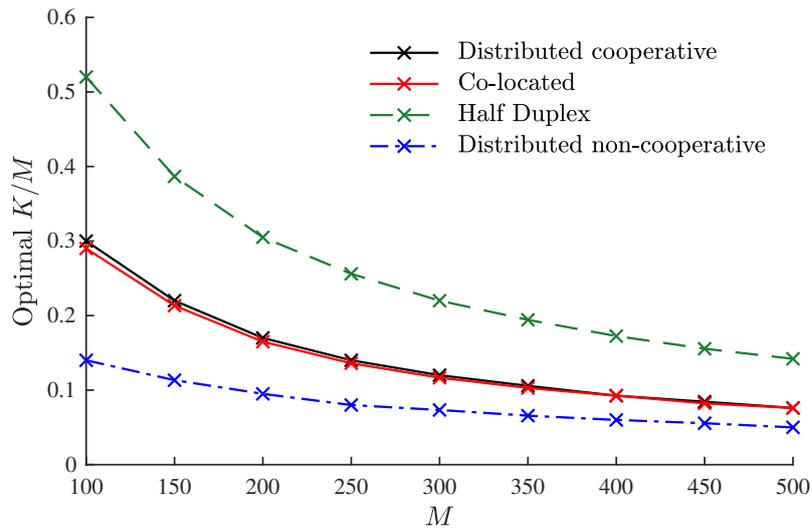
(a) MF processing



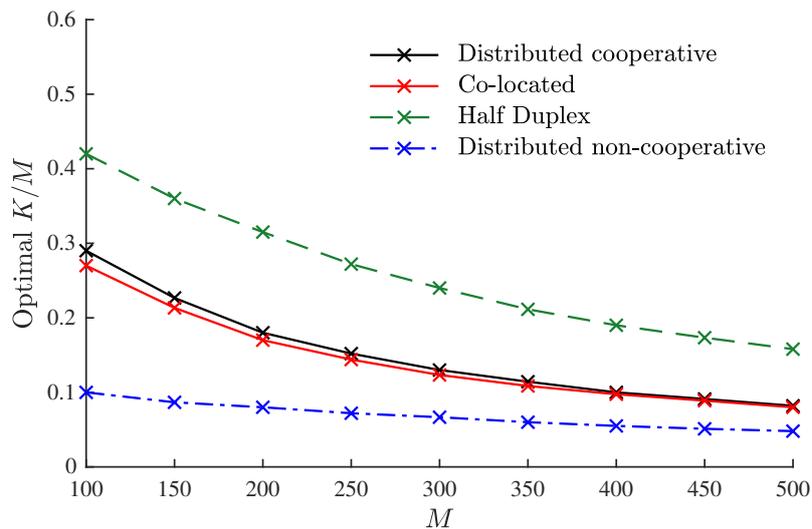
(b) ZF processing

Fig. 3.4 Spectral efficiency vs. β_{kj} for different relaying schemes. ($\rho = 10$ dB, $\beta_{kk} = 20$ dB, $M = 100$ and $K = 10$)

The effect of inter-node interference channel power β_{kj} on the spectral efficiency is illustrated in Fig. 3.4. The co-located and HD setups are independent of β_{kj} and distributed coordinated system depends on it only weakly. As expected, the fully distributed setup with independent FD nodes is superior to HD only when β_{kj} is small, here around 5 – 6 dB for MF processing and 0 – 1 dB for ZF processing. This implies



(a) MF processing



(b) ZF processing

Fig. 3.5 Optimal K/M vs. M for different relaying schemes. ($\rho = 10$ dB, $\beta_{kk} = 20$ dB and $\beta_{kj} = 10$ dB)

that independent FD nodes should be spaced very far apart of each other, or blocked by large obstacles. If this is not possible and using distributed non-cooperative devices is not an option, then distributed HD relaying is a more design effective option.

The optimal number of antennas at the relay(s) for a given size of the antenna arrays at the source and destination is illustrated in Fig. 3.5. The results are obtained through

exhaustive search using Proposition 3.2, which has low computational complexity since the rate is given in closed form. The results show that HD relaying benefits from large antenna numbers, up to 40 – 50% of the antennas used by the source and destination. As expected, the ratio diminishes as M increases since the coherence time $T = 200$ is fixed. It is worth noting that the FD relaying uses more than 10%–30% of antennas even in cooperative case for both MF and ZF processing, the reason being training overhead and additional LI caused by larger K . Thus, for relaying between two nodes with large antenna arrays, relatively small sized relay nodes achieve optimal performance in FD case. On the other hand, for HD relaying up to twice the number of antennas are required at the relay to obtain the best performance. Further improvements in the FD case may be achieved by optimizing the power allocation between data and pilots [59] or using spatial LI suppression [95, 97] in case of cooperative relaying.

3.6 Conclusion

In this chapter, the performance of co-located, distributed cooperative and distributed non-cooperative full-duplex relaying schemes for point-to-point MIMO with large antennas were investigated. Asymptotic analysis showed that all considered schemes are interference and noise free when number of antennas at the source and destination grows without bound, even when no LI cancellation is performed at the relay. Analytical achievable rate analysis and numerical examples for MF and ZF processing showed that given effective LI cancellation, FD operation can provide significant gains over HD relaying with a smaller antenna array at the relay node.

Chapter 4

Loop Interference Cancellation and Power Allocation

4.1 Introduction

Massive antenna arrays are capable of cancelling out the loop interference (LI) at the relay node in multipair full-duplex (FD) relay networks even without LI channel knowledge if the number of antennas is allowed to grow without a bound. For large but finite antenna numbers, however, channel estimation based LI cancellation is required. In this chapter, we study a FD system that is a “mirror” image of the system in last chapter, where the FD relaying is deployed with massive MIMO. LI cancellation and power allocation schemes will be investigated. A pilot protocol that utilizes the coherence time difference of the LI channel and the channels between the relay station and the moving terminals is proposed for digital LI cancellation. Furthermore, a novel power allocation scheme that adjusts the power for each link at the relay by using a simple iterative algorithm is presented. The power allocation algorithm is observed to converge fast and requires only knowledge of the channel statistics, which makes

the complexity of the proposed solution very low. The numerical results also provide insights into appropriate setting of system, such as relay transmit power.

4.2 System Model

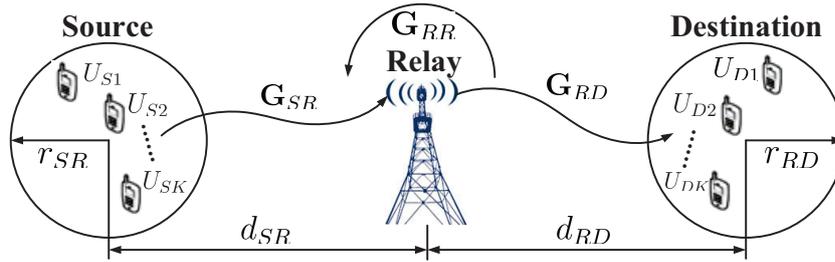


Fig. 4.1 Multipair full-duplex relaying system.

The system model is depicted in Fig. 4.1, where K source terminals, $U_{S1}, U_{S2}, \dots, U_{SK}$, each with one antenna, transmit signals to K single-antenna destination terminals $U_{D1}, U_{D2}, \dots, U_{DK}$, using the same frequency and time resources. The direct links between the source and destination terminals are assumed to be blocked. A decode-and-forward relay station with M_{Tx} transmit and M_{Rx} receive antennas is used to establish the connections between the terminals. We assume that all the source (destination) terminals are located inside a circle of radius r_{SR} (r_{RD}), which is d_{SR} (d_{RD}) meters away from the relay station. Note that the relay station works in FD mode, while all the terminals operate in HD mode.

At time instant i , the source terminals and the relay station transmit signals $\mathbf{x}_S[i] = (x_{S1}[i], x_{S2}[i], \dots, x_{SK}[i])^T$ and $\mathbf{x}_R[i] = (x_{R1}[i], x_{R2}[i], \dots, x_{RM_{Tx}}[i])^T$ over the channels \mathbf{G}_{SR} and \mathbf{G}_{RD}^T , respectively. Due to FD operation, the received and transmitted signals at the relay are coupled through the LI channel \mathbf{G}_{RR} . The received signals at the relay

station and the destinations are respectively [12, 63]

$$\mathbf{y}_R = \mathbf{G}_{SR}\mathbf{x}_S[i] + \mathbf{G}_{RR}\mathbf{x}_R[i] + \overbrace{\mathbf{G}_{RR}\mathbf{u}_{Tx} + \mathbf{u}_{Rx} + \mathbf{n}'_R}^{\triangleq \mathbf{n}_R}, \quad (4.1)$$

$$\mathbf{y}_D = \mathbf{G}_{RD}^T\mathbf{x}_R[i] + \mathbf{n}_D. \quad (4.2)$$

In addition to the LI term $\mathbf{G}_{RR}\mathbf{x}_R[i]$, the system model takes into account the combined effects of hardware impairments at the relay via the CSCG distributed transmit- and receive-side hardware impairment vectors [64, 63],

$$\mathbf{u}_{Tx} \sim \mathcal{CN}(\mathbf{0}, \mu_{Tx} \text{diag}(\mathbb{E}_{\{\mathbf{g}\}}\{\mathbf{x}_R\mathbf{x}_R^H\}))$$

$$\mathbf{u}_{Rx} \sim \mathcal{CN}(\mathbf{0}, \mu_{Rx} \text{diag}(\mathbb{E}_{\{\mathbf{g}\}}\{\mathbf{y}_R\mathbf{y}_R^H\})),$$

respectively. Coefficients $\mu_{Tx} > 0$ and $\mu_{Rx} > 0$ indicate the level of hardware impairments, the range of which are related to the error vector magnitude (EVM) requirements of the system.

The elements of the thermal noise vectors \mathbf{n}'_R and \mathbf{n}_D are i.i.d. CSCG RVs $\mathcal{CN}(0, \sigma_w^2)$. For simplicity, we denote the combined thermal noise and distortion terms induced by the hardware impairments at the relay station as \mathbf{n}_R . Note that \mathbf{n}_R is neither Gaussian distributed nor has i.i.d. elements in general. The elements of the LI channel $\mathbf{G}_{RR} \in \mathbb{C}^{M_{Rx} \times M_{Tx}}$ are assumed to be i.i.d. CSCG with equal variances β_R since the antennas are closely spaced and the passive and analog cancellation schemes can effectively mitigate the direct path of LI¹, with the residual LI coming from the rich scattering environment [36]. The source-to-relay (S→R) channel \mathbf{G}_{SR} and relay-to-destination (R→D) channel \mathbf{G}_{RD} are decomposed as² $\mathbf{G}_* = \bar{\mathbf{G}}_*\mathbf{D}_*^{1/2}$, where the entries of $\bar{\mathbf{G}}_*$ are i.i.d. standard CSCG RVs and model small scale fading. The

¹Rayleigh fading is considered herein when the direct path of LI is blocked. If there is a direct path, Rician fading model could be used and it leads only to a minor modification in the residual LI term.

²Whenever * is used, the actual subscript can be inferred from the context.

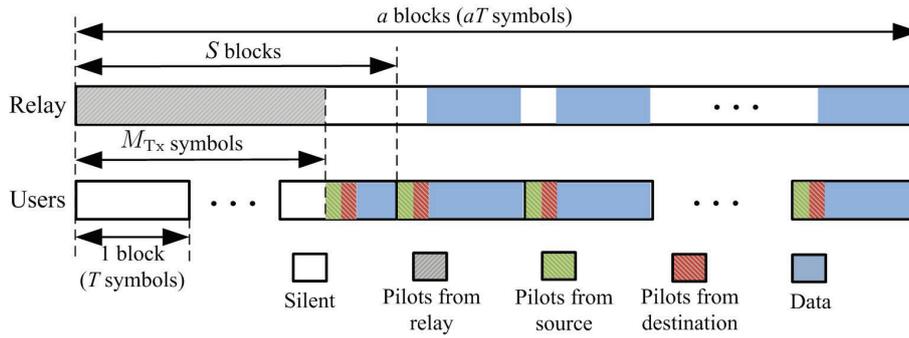


Fig. 4.2 Pilot protocol (As $ST - M_{Tx} > 2K$).

diagonal matrix \mathbf{D}_{SR} (\mathbf{D}_{RD}) with k th diagonal entries $\beta_{SR,k}$ ($\beta_{RD,k}$) represents the large scale attenuation. Assuming the relay station is located at the origin, $\beta_{SR,k}$ ($\beta_{RD,k}$) is modeled as $\beta_{SR,k} = \kappa_{SR,k} \|\mathbf{z}_{S,k}\|^{-\alpha}$ ($\beta_{RD,k} = \kappa_{RD,k} \|\mathbf{z}_{D,k}\|^{-\alpha}$), where α is the path loss exponent, vector $\mathbf{z}_{S,k}$ ($\mathbf{z}_{D,k}$) defines the location of the k th source (destination) terminal and $\kappa_{SR,k}$ ($\kappa_{RD,k}$) represents shadow fading between the relay and k th source (destination) terminal. Since $\beta_{SR,k}$ ($\beta_{RD,k}$) is changing slowly, it is assumed to be known at k th source (destination) terminal and the relay station.

4.2.1 Novel Pilot Protocol and Channel Estimation

As the relay station is static while the terminals are moving, the coherence time of the LI channel is typically several times longer than that of the $S \rightarrow R$ and $R \rightarrow D$ channels [98]. We model this by considering a block fading channel with coherence time of T symbols (one block) for $S \rightarrow R$ and $R \rightarrow D$ channels, and aT symbols (a blocks) for the LI channel. Pilot-aided channel estimation is used to obtain instantaneous channel state information (CSI) for detection, LI cancellation and precoding at the relay station. We design a pilot protocol as shown in Fig. 4.2 and described below.

In $S = \lceil M_{Tx}/T \rceil$ consecutive blocks, where $\lceil x \rceil$ denotes the smallest integer which is not less than x , first M_{Tx} symbols are used by the relay station to transmit pilots $\Phi_R \in \mathbb{C}^{M_{Tx} \times M_{Tx}}$ for estimating the LI channel. If the length of the remaining symbols

in the S th block is greater than $2K$, i.e. $ST - M_{\text{Tx}} > 2K$, the terminals transmit pilots as described below, followed by data transmission from the source and the relay station. Otherwise, all nodes keep radio silence for the rest of the block. From the $(S + 1)$ th block onwards, pilots $\Phi_S \in \mathbb{C}^{K \times K}$ and $\Phi_D \in \mathbb{C}^{K \times K}$ are transmitted by the terminals at the beginning of each block.

Channel estimation for LI cancellation

We require $\Phi_R \Phi_R^H = \mathbf{I}_{M_{\text{Tx}}}$ to satisfy pilot orthogonality and power constraint. The received pilot matrix at the receive-side antennas of the relay station is given by

$$\mathbf{Y}_{\text{RR}} = \sqrt{\rho_\xi} \mathbf{G}_{\text{RR}} \Phi_R + \mathbf{N}_{\text{RR}},$$

where ρ_ξ is the transmit power of one pilot symbol during the training phase. \mathbf{N}_{RR} is a combination of transmit- and receive-side hardware impairment terms and is in general not Gaussian. It is, however, uncorrelated with the LI channel \mathbf{G}_{RR} and a pessimistic prediction of the channel estimator performance can be obtained by treating the elements of \mathbf{N}_{RR} as being independent with equal variance $\sigma_{\text{RR}}^2 = (\mu_{\text{Tx}} + \mu_{\text{Rx}}) \rho_\xi \beta_{\text{R}} + \sigma_{\text{w}}^2 (1 + \mu_{\text{Rx}})$.

Channel estimation for detection and precoding

We assume that all terminals know their own channel statistics. To satisfy pilot orthogonality, we require $\Phi_S \Phi_S^H$ and $\Phi_D \Phi_D^H$ to be diagonal matrices with $K\rho_{\text{pS},k}$ and $K\rho_{\text{pD},k}$ on their k th diagonal. The received pilots at the relay station are given by

$$\mathbf{Y}_{\text{R}} = \mathbf{G}_{\text{SR}} \Phi_S + \mathbf{N}_{\text{R}}, \quad (4.3)$$

$$\mathbf{Y}_{\text{D}} = \mathbf{G}_{\text{RD}} \Phi_D + \mathbf{N}_{\text{D}}, \quad (4.4)$$

where the entries of the noise matrices \mathbf{N}_R and \mathbf{N}_D are i.i.d. complex Gaussian RVs $\mathcal{CN}(0, \sigma_w^2)$.

After receiving all pilots, the relay uses linear minimum mean squared error (LMMSE) estimator to get instantaneous channel estimates $\hat{\mathbf{G}}_{SR}$, $\hat{\mathbf{G}}_{RD}$ and $\hat{\mathbf{G}}_{RR}$. We denote $\tilde{\mathbf{G}}_* = \mathbf{G}_* - \hat{\mathbf{G}}_*$ the error matrix, which is uncorrelated with the estimate [29]. Note that for the LI channel, the LMMSE estimator does not yield optimal MMSE, which leads to a lower bound for E2E spectral efficiency [64]. The per-element variance of estimation error $\tilde{\beta}_*$ can be obtained from the knowledge of pilot energy E_p , noise power σ_*^2 and channel gain β_* [29], as $\tilde{\beta}_* = \frac{\beta_*}{E_p \beta_* / \sigma_*^2 + 1}$. Thus, the entries of $\tilde{\mathbf{G}}_{RR}$ have the same variance $\tilde{\beta}_R = \frac{\beta_R}{\rho_\xi \beta_R / \sigma_{RR}^2 + 1}$, while the error matrices $\tilde{\mathbf{G}}_{SR}$ and $\tilde{\mathbf{G}}_{RD}$ have independent CSCG elements, the variance of the entries in the k th column being $\tilde{\beta}_{SR,k} = \frac{\beta_{SR,k}}{K \rho_{PS,k} \beta_{SR,k} / \sigma_w^2 + 1}$ and $\tilde{\beta}_{RD,k} = \frac{\beta_{RD,k}}{K \rho_{PD,k} \beta_{RD,k} / \sigma_w^2 + 1}$ respectively. The properties of the estimator also guarantee that $\hat{\beta}_* = \beta_* - \tilde{\beta}_*$ holds for all channels.

4.2.2 Data Transmission

At time instant i , the source terminals transmit information vector $\mathbf{x}_S[i] = \text{diag}(\sqrt{\rho_{S1}}, \sqrt{\rho_{S2}}, \dots, \sqrt{\rho_{SK}}) \bar{\mathbf{x}}[i]$ directly to the relay station. For all time instants i , the entries of $\bar{\mathbf{x}}[i]$ are assumed to be i.i.d. standard CSCG. After subtracting LI by using the knowledge of $\hat{\mathbf{G}}_{RR}$ and $\mathbf{x}_R[i]$, the received signal at the relay station reads $\mathbf{y}_R[i] = \mathbf{G}_{SR} \mathbf{x}_S[i] + \boldsymbol{\xi} + \mathbf{n}_R$, where we denoted $\boldsymbol{\xi} = \tilde{\mathbf{G}}_{RR} \mathbf{x}_R[i]$ for the residual LI due to imperfect CSI of the LI channel. Assuming the system employs linear detection and precoding by using matrices \mathbf{W} and \mathbf{V} , which are functions of $\hat{\mathbf{G}}_{SR}$ and $\hat{\mathbf{G}}_{RD}$, respectively, the k th decoded signal at the relay reads

$$\mathbf{y}_{R,k}[i] = \underbrace{\mathbf{w}_k^H \mathbf{g}_{SR,k} x_{S,k}[i]}_{\text{desired signal}} + \underbrace{\sum_{j \neq k} \mathbf{w}_k^H \mathbf{g}_{SR,j} x_{S,j}[i]}_{\text{inter-pair interference}} + \underbrace{\mathbf{w}_k^H (\boldsymbol{\xi} + \mathbf{n}_R)}_{\text{residual LI + noise}} \quad (4.5)$$

where $x_{S,k}$, $\mathbf{g}_{SR,k}$, $\mathbf{g}_{RR,k}$ and \mathbf{w}_k are the k th column (element) of \mathbf{x}_S , \mathbf{G}_{SR} , \mathbf{G}_{RR} and \mathbf{W} , respectively. Following the common assumption in decode-and-forward relaying, there is a processing delay of $\tau \geq 1$ symbols at the relay, $\mathbf{x}_R[i] = \mathbf{V}\bar{\mathbf{x}}[i - \tau]$, and thus the transmit signal at the relay station is uncorrelated with the received signal [95]. Finally, the received signal at k th destination terminal is given by

$$y_{D,k}[i] = \underbrace{\mathbf{g}_{RD,k}^T \mathbf{v}_k \bar{x}_k[i - \tau]}_{\text{desired signal}} + \underbrace{\sum_{j \neq k} \mathbf{g}_{RD,k}^T \mathbf{v}_j \bar{x}_j[i - \tau]}_{\text{inter-pair interference}} + \underbrace{n_{D,k}}_{\text{noise}} \quad (4.6)$$

where $\mathbf{g}_{RD,k}$, \mathbf{v}_k , $\bar{x}_k[i - \tau]$ and $n_{D,k}$ denote the k th column (elements) of \mathbf{G}_{RD} , \mathbf{V} , $\bar{\mathbf{x}}[i - \tau]$ and \mathbf{n}_D , respectively.

4.3 Novel Power Allocation Scheme and Performance Analysis

4.3.1 Achievable Rate Analysis

While the residual LI and noise + distortion term $\boldsymbol{\xi} + \mathbf{n}_R$ is uncorrelated with the signal, it is not Gaussian, making straightforward achievable rate analysis of the system difficult. We thus consider an auxiliary system to find a lower bound on the achievable rate. More precisely, we treat the sum of inter-pair interference, residual LI and noise in (4.5) as additive Gaussian noise of the same variance [96], which is independent of the desired signal. For the k th S→R and R→D links, the lower bound

of achievable rates³ are then given by

$$R_{\text{SR},k} = \log_2 \left(1 + \frac{\rho_{\text{Sk}} |\mathbb{E}_{\{\mathbf{g}\}} \{\mathbf{w}_k^H \mathbf{g}_{\text{SR},k}\}|^2}{\rho_{\text{Sk}} \text{Var}_{\{\mathbf{g}\}} (\mathbf{w}_k^H \mathbf{g}_{\text{SR},k}) + \sum_{j=1, j \neq k}^K \rho_{\text{Sj}} \mathbb{E}_{\{\mathbf{g}\}} \{|\mathbf{w}_k^H \mathbf{g}_{\text{SR},j}\|^2\} + \mathbb{E}_{\{\mathbf{g}, \boldsymbol{\xi}, \mathbf{n}_R\}} \{\|\mathbf{w}_k^H (\boldsymbol{\xi} + \mathbf{n}_R)\|^2\}} \right) \quad (4.7)$$

$$R_{\text{RD},k} = \log_2 \left(1 + \frac{|\mathbb{E}_{\{\mathbf{g}\}} \{\mathbf{g}_{\text{RD},k}^T \mathbf{v}_k\}|^2}{\text{Var}_{\{\mathbf{g}\}} (\mathbf{g}_{\text{RD},k}^T \mathbf{v}_k) + \sum_{j=1, j \neq k}^K \mathbb{E}_{\{\mathbf{g}\}} \{|\mathbf{g}_{\text{RD},k}^T \mathbf{v}_j|^2\} + \sigma_w^2} \right) \quad (4.8)$$

where we denoted with some abuse of notation $\boldsymbol{\xi}$ and \mathbf{n}_R for CSCG vectors that are independent of the desired signal and have i.i.d. elements of variance (provided later) LI and σ_R^2 , respectively. Since the ergodic achievable rate depends on the weaker link, the lower bound of E2E achievable rate of the k th terminal pair reads [12]

$$R_k = \min\{R_{\text{SR},k}, R_{\text{RD},k}\}. \quad (4.9)$$

We analyze here matched-filtering (MF) and zero-forcing (ZF) processing for both detection and precoding, that is,

$$\mathbf{W}_{\text{MF}}^H = \hat{\mathbf{G}}_{\text{SR}}^H, \quad (4.10)$$

$$\mathbf{V}_{\text{MF}} = \hat{\mathbf{G}}_{\text{RD}}^* \mathbf{P}_{\text{MF}}, \quad (4.11)$$

$$\mathbf{W}_{\text{ZF}}^H = (\hat{\mathbf{G}}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}})^{-1} \hat{\mathbf{G}}_{\text{SR}}^H, \quad (4.12)$$

$$\mathbf{V}_{\text{ZF}} = \mathbf{B} \mathbf{P}_{\text{ZF}} = \hat{\mathbf{G}}_{\text{RD}}^* (\hat{\mathbf{G}}_{\text{RD}}^T \hat{\mathbf{G}}_{\text{RD}}^*)^{-1} \mathbf{P}_{\text{ZF}}, \quad (4.13)$$

where $\mathbf{P}_{\text{MF}} \in \mathbb{C}^{K \times K}$ and $\mathbf{P}_{\text{ZF}} \in \mathbb{C}^{K \times K}$ are power allocation matrices to be designed in the next subsection. The k th diagonals of \mathbf{P}_{MF} and \mathbf{P}_{ZF} are given by $p_{\text{MF},k} = \sqrt{\frac{m_{\text{MF},k}}{\mathbb{E}\{\|\hat{\mathbf{g}}_{\text{RD},k}\|^2\}}} = \sqrt{\frac{m_{\text{MF},k}}{M_{\text{Tx}} \hat{\beta}_{\text{RD},k}}}$ and $p_{\text{ZF},k} = \sqrt{\frac{m_{\text{ZF},k}}{\mathbb{E}\{\|\hat{\mathbf{b}}_k^T\|^2\}}} = \sqrt{(M_{\text{Tx}} - K) \hat{\beta}_{\text{RD},k} m_{\text{ZF},k}}$, where

³Although the relay station has the instantaneous channel estimate, i.e. $\hat{\mathbf{g}}_{\text{RD},k}$, we assume it always uses statistical channel estimates $\mathbb{E}_{\{\mathbf{g}\}} \{\mathbf{w}_k^H \hat{\mathbf{g}}_{\text{SR},k}\}$ for detection. The destination terminals use statistical channel estimates for detection, as they do not know instantaneous CSI.

\mathbf{b}_k is the k th column of matrix \mathbf{B} and $m_{*,k}$ denotes the relay's transmit power for the k th link.

Proposition 4.1. *The lower bound of E2E achievable rates of k th terminal pair with MF and ZF processing in (4.14) and (4.15),*

$$R_k^{\text{MF}} = \log_2 \left(1 + \min \left(\frac{\rho_{S,k} M_{\text{Rx}} \hat{\beta}_{\text{SR},k}}{\sum_{j=1}^K \rho_{S,j} \beta_{\text{SR},j} + \rho_{\xi_{\text{MF}}} + \sigma_{\text{MF},\text{R}}^2}, \frac{m_{\text{MF},k} \hat{\beta}_{\text{RD},k} M_{\text{Tx}}}{\beta_{\text{RD},k} m_{\text{MF},\text{tot}} + \sigma_{\text{w}}^2} \right) \right) \quad (4.14)$$

$$R_k^{\text{ZF}} = \log_2 \left(1 + \min \left(\frac{\rho_{S,k} (M_{\text{Rx}} - K) \hat{\beta}_{\text{SR},k}}{\sum_{j=1}^K \rho_{S,j} \tilde{\beta}_{\text{SR},j} + \rho_{\xi_{\text{ZF}}} + \sigma_{\text{ZF},\text{R}}^2}, \frac{(M_{\text{Tx}} - K) \hat{\beta}_{\text{RD},k} m_{\text{ZF},k}}{\tilde{\beta}_{\text{RD},k} m_{\text{ZF},\text{tot}} + \sigma_{\text{w}}^2} \right) \right) \quad (4.15)$$

where $m_{*,\text{tot}} = \sum_k m_{*,k}$ is the total transmit power of the relay. The powers of the residual LI for MF and ZF processing are given by $\rho_{\xi_{\text{MF}}} = \tilde{\beta}_{\text{R}} m_{\text{MF},\text{tot}}$ and $\rho_{\xi_{\text{ZF}}} = \tilde{\beta}_{\text{R}} (1 - \frac{K}{M_{\text{Tx}}}) m_{\text{ZF},\text{tot}}$, respectively. The power of noise + distortion are

$$\sigma_{*,\text{R}}^2 = (\mu_{\text{Tx}} + \mu_{\text{Rx}}) \beta_{\text{R}} m_{*,\text{tot}} + \mu_{\text{Rx}} \left(\sum_{k=1}^K \rho_{S,k} \beta_{\text{SR},k} + \mu_{\text{Tx}} \beta_{\text{R}} m_{*,\text{tot}} \right) + \sigma_{\text{w}}^2 (1 + \mu_{\text{Rx}}).$$

Proof. See Appendix C. □

4.3.2 Novel Power Allocation Scheme

For the k th terminal pair, increasing transmit power at the relay station yields a higher rate for the R→D link but increases the LI and hence decreases the rate of the S→R link. As the achievable rate depends on the weaker link, we propose adjusting $\mathbf{m}_* = (m_{*,1}, \dots, m_{*,K})$ so that the achievable rates of these two links are equal, i.e., $R_{\text{SR},k} = R_{\text{RD},k}, \forall k$. This can be achieved via simple iterative algorithms that are

given for MF and ZF processing respectively by

$$m_{\text{MF},k}^{(l)} = \frac{\rho_{\text{S},k} \hat{\beta}_{\text{SR},k} M_{\text{Rx}} (\beta_{\text{RD},k} m_{\text{MF,tot}}^{(l-1)} + \sigma_{\text{w}}^2)}{\hat{\beta}_{\text{RD},k} M_{\text{Tx}} (\sum_{j=1}^K \rho_{\text{S},j} \beta_{\text{SR},j} K + \tilde{\beta}_{\text{R}} m_{\text{MF,tot}}^{(l-1)} + \sigma_{\text{MF,R}}^2)}, \quad (4.16)$$

$$m_{\text{ZF},k}^{(l)} = \frac{\frac{M_{\text{Rx}} - K}{M_{\text{Tx}} - K} \rho_{\text{S},k} \hat{\beta}_{\text{SR},k} (\tilde{\beta}_{\text{RD},k} m_{\text{ZF,tot}}^{(l-1)} + \sigma_{\text{w}}^2)}{\hat{\beta}_{\text{RD},k} (\sum_{j=1}^K \rho_{\text{S},j} \tilde{\beta}_{\text{SR},j} + \tilde{\beta}_{\text{R}} (1 - \frac{K}{M_{\text{Tx}}}) m_{\text{ZF,tot}}^{(l-1)} + \sigma_{\text{ZF,R}}^2)}, \quad (4.17)$$

where l is the iteration index and $m_{*,\text{tot}}^{(l-1)}$ is the total transmit power of the relay station in the $l - 1$ th iteration. The initial point of the iteration can be found by treating large scale fading factors of all terminals the same and assuming $m_{*,1}^{(0)} = m_{*,2}^{(0)} = \dots = m_{*,K}^{(0)}$. Then (4.16) and (4.17) become quadratic equations that have only one real positive solution respectively, which can be used as the initial points of the iteration. Note that instantaneous CSI is not required for the proposed power allocation scheme, which makes the complexity very low. The iterative algorithm is summarized below.

Algorithm 1 Find $m_{*,k}, k = 1, 2, \dots, K$ in (4.16) and (4.17) to maximize E2E achievable rate in (4.14) and (4.15)

- 1. Initialization $l = 0$:** Initialize all the entries with $\mathbf{m}_*^{(0)}$, where $\mathbf{m}_*^{(0)}$ is calculated by assuming $m_{*,1}^{(0)} = m_{*,2}^{(0)} = \dots = m_{*,K}^{(0)}$,
 - 2. Iteration l :** Set $l = l + 1$ and calculate $\mathbf{m}_*^{(l)}$ as in (4.16) and (4.17),
 - 3. Decision:** Stop the iteration as $\frac{\|\mathbf{m}_*^{(l)} - \mathbf{m}_*^{(l-1)}\|_1}{\|\mathbf{m}_*^{(l-1)}\|_1} < \Delta$ or $l = 10000$, otherwise, return to Step 2.
-

The same power allocation scheme at the relay station can also be used for the case when LI is not canceled, by simply replacing $\tilde{\beta}_{\text{R}}$ by β_{R} in (4.16) and (4.17).

4.3.3 Spectral Efficiency and Energy Efficiency

Given the E2E achievable rate for k th terminal pair in (4.15), the average sum spectral efficiency is given by

$$\text{SE} = \begin{cases} \frac{aT - ST - 2K(a-S)}{aT} \mathbb{E}\{\sum_{k=1}^K R_k\}, & ST - M_{\text{Tx}} \leq 2K \\ \frac{aT - M_{\text{Tx}} - 2K(a-S+1)}{aT} \mathbb{E}\{\sum_{k=1}^K R_k\}, & ST - M_{\text{Tx}} > 2K \end{cases} \quad (4.18)$$

where the expectation is over the terminal locations and shadow fading. Energy efficiency is defined as

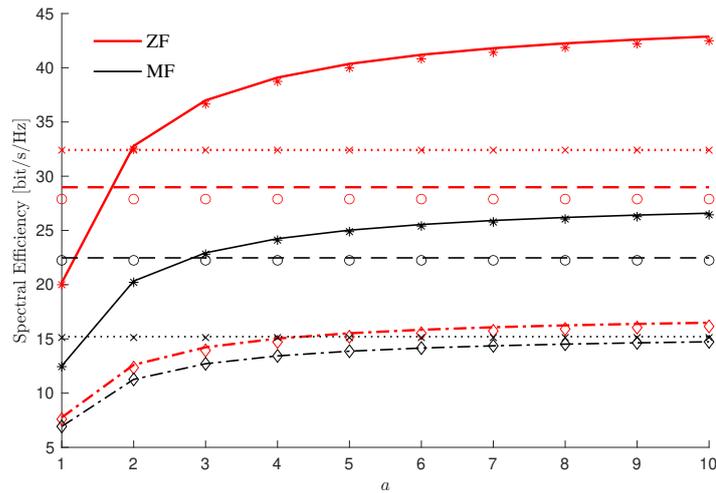
$$\text{EE} = \frac{\text{SE}}{E_{\text{tot}}} aT, \quad (4.19)$$

where E_{tot} denotes the average total energy consumption of the whole system during data and pilot transmission (a blocks). Note that in numerical examples, the energy consumption is based only on the power consumed by the amplifiers and baseband circuit power consumption is omitted from the analysis.

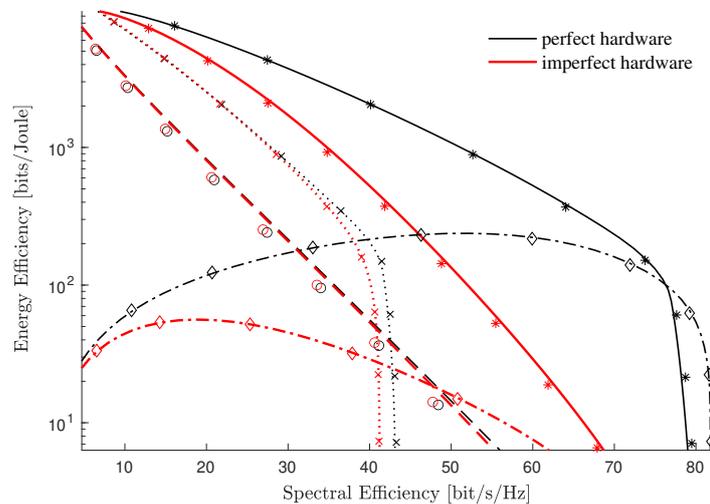
4.4 Numerical Results

Unless otherwise specified, the system parameters used in the numerical results are $T = 200$, $K = 10$, $M_{\text{Rx}} = M_{\text{Tx}} = 100$, $\sigma_w^2 = -101$ dBm and $\beta_{\text{R}} = -90$ dB, corresponding to a slightly optimistic but realistic level of passive and analog LI mitigation [36, 75]. The parameters of the geometric model are $d_{\text{SR}} = d_{\text{RD}} = 400$ m and $r_{\text{SR}} = r_{\text{RD}} = 100$ m with path loss exponent $\alpha = 4$ and log-normal shadowing with zero mean and 6 dB variance. Strict power constraint of 23 dBm is enforced at the relay station. Pilots used for estimating the LI channel at the relay are transmitted at the maximum power, i.e. $\rho_{\xi} = 23$ dBm. With hardware impairments, the transmit-side distortion coefficient is chosen as $\mu_{\text{Tx}} = 0.1^2$ that corresponds to $\text{EVM} = 0.1$, and is within the EVM range [0.08, 0.175] of the LTE standard. We apply statistics-based power control at the

source terminals $\rho_{S,k} = \frac{\gamma}{\beta_{SR,k}}$, where γ is a design parameter, and denote the average transmit power (over β 's) of the source terminals $\bar{\rho}_S$. In pilot transmission phase, we set $\rho_{pS,k} = \gamma_p/\beta_{SR,k}$ and $\rho_{pD,k} = \gamma_p/\beta_{RD,k}$, where γ_p is a design parameter. In numerical results $\gamma = \gamma_p$ is assumed.



(a) Spectral efficiency vs. the ratio of coherence time difference of the system with imperfect hardware.



(b) Energy efficiency of the whole system with ZF processing vs. sum spectral efficiency.

Fig. 4.3 Performance of the proposed pilot protocol and power allocation scheme, analytical results of FD w/ LI cancellation (solid), FD w/o LI cancellation (dashed), FD w/ LI cancellation and fixed relay power $m_{\text{tot}} = 23$ dBm (dash-dotted) and HD (dotted) presented with curves and Monte Carlo simulations with markers (stars, circles, diamonds and crosses, respectively).

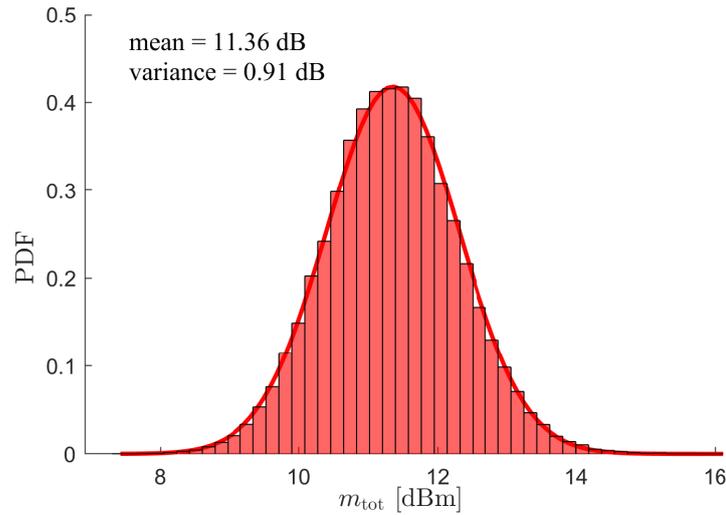
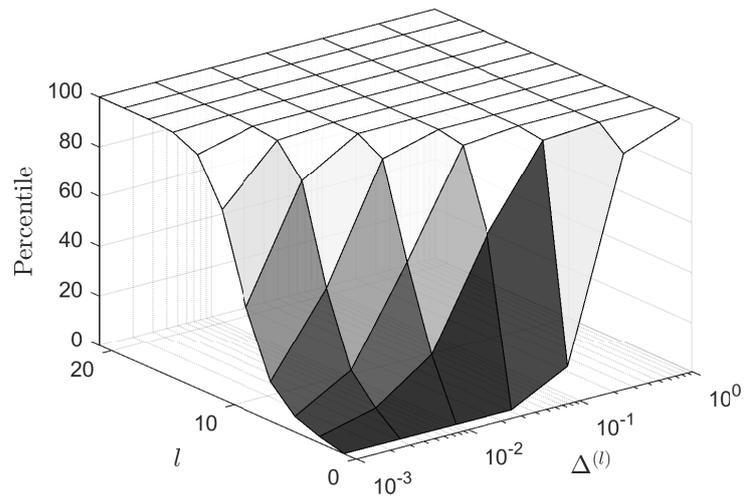


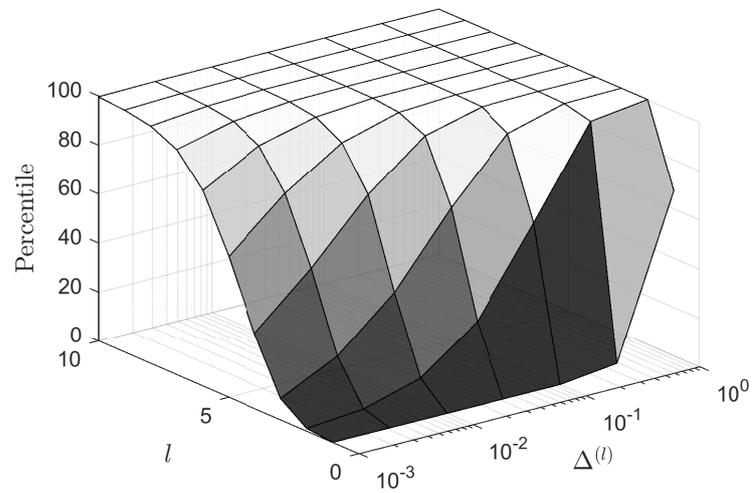
Fig. 4.4 Empirical PDF of the total power consumption of the proposed power allocation scheme at the relay station with ZF processing.

Fig. 4.3(a) plots the SE versus the coherence time ratio a for $\bar{\rho}_S = 11.57$ dBm ($\gamma = -95$ dBm) with ZF and MF processing, respectively. The curves for HD relaying (with relay power optimized so that $R_{\text{SR},k} = R_{\text{RD},k}$) and FD relaying without LI cancellation (as in [12, 63]) are horizontal since the relay does not transmit any pilots. Due to severe pilot overhead, the SE of the proposed pilot protocol is relatively low when the coherence times are equal. However, already for $a = 2$, ZF processing with LI cancellation and the proposed power allocation strategy offers the highest SE. Since MF processing is inter-pair interference limited, it does not benefit from LI channel estimation as much as ZF.

In Fig. 4.3(b) SE and EE tradeoff is investigated with ZF processing for $a = 8$, and performance of the considered schemes with perfect and imperfect hardware are demonstrated. The FD system with LI cancellation and the proposed power allocation scheme outperforms the fixed FD relay power case and HD relaying with significant margin. This is because that transmitting at the fixed relay power 23 dBm, which corresponds to the dash-dotted curves, can obviously increase $R_{\text{RD},k}$ to be significantly larger than $R_{\text{SR},k}$. In this case, the E2E rate for k th user-pair can only be $R_k = R_{\text{SR},k}$, since the E2E rate depends on the weaker link, as demonstrated in (4.9). It means



(a) MF processing.



(b) ZF processing.

Fig. 4.5 Convergence of the proposed power allocation scheme at the relay station as in (4.17).

the power of k th source terminal or the relay station is partly wasted. According to the definition of EE in (4.19), the relay station of FD fixed relay power case has been wasting more energy during the same amount of time, thus the EE of which is significantly lower. The similar explanation can be applied to HD case. With our novel power allocation scheme demonstrated in solid lines, the power of the relay station is wisely adjusted so that for each user-pair k , $R_{SR,k} = R_{RD,k}$. It leads to no energy waste. The relay power variation in the proposed scheme for ZF processing is illustrated in Fig. 4.4 that plots the empirical probability density function (PDF) of m_{tot} for the case $\bar{\rho}_S = 11.57$ dBm. The PDF is well approximated by a Gaussian distribution (the solid line) implying that the variations around the mean decay exponentially. It is also clear that the proposed algorithm does not cause violation of the relay power constraint 23 dBm.

Fig. 4.5 shows the percentile of power allocation instances that converge to a normalized difference $\Delta^{(l)} = \|\mathbf{q}^{(l)} - \mathbf{q}^{(l-1)}\|_1 / \|\mathbf{q}^{(l-1)}\|_1$ after l iterations when $\bar{\rho}_S = 11.57$ dBm. Clearly the proposed power allocation scheme converges very fast, typically within 20 iterations for all considered cases.

4.5 Conclusion

In this chapter, we proposed pilot-based LI cancellation and power allocation schemes for multipair FD relaying systems with hardware impairments and large but finite number of antennas. The low-complexity iterative power allocation scheme requires only channel statistic and converges very fast. The combination of both schemes improves the SE and EE of the FD relaying with fixed relay power or the HD relaying significantly under practical system parameters.

Chapter 5

Full-duplex Massive MIMO Relay Systems with Low Resolution ADCs

5.1 Introduction

In this chapter, the throughput and energy efficiency of a multipair full-duplex DF relaying system (see Fig. 4.1) that suffers from realistic levels of residual LI after passive and analog cancellation is investigated. Based on the mathematical analysis, we develop an iterative power allocation scheme that controls the link-wise transmit powers at the relay by taking into account the resolution of the ADCs and the transmit-side hardware impairments. The proposed scheme is based on an extension of last chapter, where full resolution ADCs were employed at the relay. By reducing the effects of quantization noise and residual LI in the system, the power allocation algorithm aims at maximizing the pairwise end-to-end (E2E) achievable rate. The proposed scheme uses statistical information, so that instantaneous channel state information (CSI) is not needed and the power allocation update frequency is low. In addition, we carry out detailed EE investigation and in the spirit of ultra-reliable and low-latency communications (URLLC) [99], examine the effects of imposing strict

delay constraint, i.e., finite code word length, on the system. The numerical results show that compared to HD relaying, the proposed FD transmission scheme provides significant gains in system throughput and energy efficiency in both delay sensitive and insensitive cases.

5.2 System Model

We continue to study the multipair relaying system in last chapter depicted in Fig. 4.1. As demonstrated in last chapter, the coherence time of the LI channel is typically much longer than that of the S→R and R→D channels, in this chapter, we assume that the relay station can estimate the LI channel accurately, with negligible loss in throughput.

The received signals at the relay and destination terminals are given by [12, 63, 64]

$$\mathbf{y}_R = \mathbf{G}_{SR}\mathbf{x}_S[i] + \mathbf{G}_{RR}\mathbf{x}_R[i] + \mathbf{G}_{RR}\mathbf{u}_{Tx} + \mathbf{n}'_R, \quad (5.1)$$

$$\mathbf{y}_D = \mathbf{G}_{RD}^T(\mathbf{x}_R[i] + \mathbf{u}_{Tx}) + \mathbf{n}_D, \quad (5.2)$$

where \mathbf{y}_R denotes the signal after analog circuit domain cancellation before ADCs. The signals transmitted by the sources and the relay at time instant i are $\mathbf{x}_S[i] \in \mathbb{C}^K$ and $\mathbf{x}_R[i] \in \mathbb{C}^{M_{Tx}}$, respectively. The elements of the thermal noise vectors \mathbf{n}'_R and \mathbf{n}_D are i.i.d. $\mathcal{CN}(0, \sigma_w^2)$. The transmit-side hardware impairment vector $\mathbf{u}_{Tx} \sim \mathcal{CN}(\mathbf{0}, \mu_{Tx} \text{diag}(\mathbb{E}\{\mathbf{x}_R\mathbf{x}_R^H\}))$ takes into account the combined effects of hardware impairments at the transmit-side of the relay, where the coefficient $\mu_{Tx} > 0$ is related to the error vector magnitude (EVM) requirements of the system [64, 63]. In the following, we omit \mathbf{u}_{Tx} from the R→D link since it has negligible impact on the rate. To focus on the impact of quantization noise due to low resolution ADCs, we do not consider other receive-side imperfection (such as I/Q imbalance) in this chapter. Such effects can be incorporated in the system model e.g. as done in [100] if desired. Since

Table 5.1 θ for different resolutions of ADCs

N	1	2	3	4	5	≥ 6
θ	0.6366	0.8825	0.96546	0.990503	0.997501	$1 - \frac{\pi\sqrt{3}}{2}2^{-2N}$

the source and destination terminals have only one antenna, they are assumed to be equipped with ideal hardware, including full resolution ADCs.

5.2.1 Quantization with low-resolution ADCs

At the receive-side of the relay station, low-resolution ADCs are assumed to be used for energy saving purposes. We model the quantized received signal as [101, 74]

$$\begin{aligned} \mathbf{r}_R &= \theta \mathbf{y}_R + \mathbf{n}_q \\ &= \theta \mathbf{G}_{SR} \mathbf{x}_S[i] + \theta \mathbf{G}_{RR} \mathbf{x}_R[i] + \underbrace{\theta \mathbf{G}_{RR} \mathbf{u}_{Tx} + \theta \mathbf{n}'_R}_{\triangleq \mathbf{n}_R} + \mathbf{n}_q. \end{aligned} \quad (5.3)$$

The quantization noise is modeled as a CSCG vector

$$\mathbf{n}_q \sim \mathcal{C}\mathcal{N}(\mathbf{0}, \theta(1 - \theta) \text{diag}(\mathbb{E}\{\mathbf{y}_R \mathbf{y}_R^H\})),$$

that is assumed to be uncorrelated with the received signal (5.1). Note that since the covariance matrix of \mathbf{n}_q depends on the statistics of the received signal after analog circuit domain cancellation, digital baseband LI cancellation cannot be used to reduce the level of quantization noise in the system. The coefficient $\theta > 0$ is related to the resolution N of the ADCs and Table 5.1 lists the values of θ for different choices of N [76]. For notational convenience, we combine the terms related to transmit-side hardware impairments, quantization noise and thermal noise as one term \mathbf{n}_R as shown in (5.3). Note that \mathbf{n}_R neither is Gaussian distributed nor has i.i.d. elements in general.

5.2.2 Channel Estimation

Block fading with a coherence time of T symbols is assumed for the channels between the terminals and the relay. To facilitate channel estimation, the source and destination terminals transmit pilot matrices Φ_S and Φ_D , respectively, at the beginning of each coherence block. To satisfy pilot orthogonality, we let Φ_S and Φ_D be $K \times K$ diagonal matrices with $\sqrt{K\rho_{pS,k}}$ and $\sqrt{K\rho_{pD,k}}$ as the k th diagonal entry, respectively. The received pilots at the relay station after ADCs are given by To satisfy pilot orthogonality and notational simplicity, we design Φ_S and Φ_D to be diagonal matrices with $\sqrt{K\rho_{pS,k}}$ and $\sqrt{K\rho_{pD,k}}$ on the k th diagonal, respectively. The quantized received pilots at the relay station are

$$\begin{aligned}\mathbf{R}_{SR} &= \theta \mathbf{G}_{SR} \Phi_S + \mathbf{N}_{SR}, \\ \mathbf{R}_{DR} &= \theta \mathbf{G}_{RD} \Phi_D + \mathbf{N}_{DR},\end{aligned}$$

where the additive noise matrices model the combination of quantization and thermal noise. The entries of \mathbf{N}_{SR} and \mathbf{N}_{DR} are assumed to be independent CSCG and the variance of the elements on the k th column are

$$\begin{aligned}\sigma_{SR,k}^2 &= \theta(1-\theta)K\rho_{pS,k}\beta_{SR,k} + \theta\sigma_w^2, \\ \sigma_{RD,k}^2 &= \theta(1-\theta)K\rho_{pD,k}\beta_{RD,k} + \theta\sigma_w^2,\end{aligned}$$

respectively.

After receiving the pilots from the terminals, the relay calculates the channel estimates $\hat{\mathbf{G}}_{SR}$ and $\hat{\mathbf{G}}_{RD}$. We denote $\tilde{\mathbf{G}}_* = \mathbf{G}_* - \hat{\mathbf{G}}_*$ for the error matrix and assume linear minimum mean squared error (LMMSE) estimator is used to obtain the CSI. As a result, the estimates and the errors are uncorrelated [29]. For the system under investigation, the entries in the k th column of $\tilde{\mathbf{G}}_{SR}$ and $\tilde{\mathbf{G}}_{RD}$ with LMMSE estimation

are independent CSCG with variance [76, 100]

$$\begin{aligned}\tilde{\beta}_{\text{SR},k} &= \frac{(1-\theta)K\rho_{\text{pS},k}\beta_{\text{SR},k}^2 + \sigma_{\text{w}}^2\beta_{\text{SR},k}}{K\rho_{\text{pS},k}\beta_{\text{SR},k} + \sigma_{\text{w}}^2}, \\ \tilde{\beta}_{\text{RD},k} &= \frac{(1-\theta)K\rho_{\text{pD},k}\beta_{\text{SR},k}^2 + \sigma_{\text{w}}^2\beta_{\text{SR},k}}{K\rho_{\text{pD},k}\beta_{\text{RD},k} + \sigma_{\text{w}}^2},\end{aligned}$$

respectively. The properties of the estimator also guarantee that $\hat{\beta}_* = \beta_* - \tilde{\beta}_*$ holds for all channels.

As mentioned earlier in this chapter, the coherence time of the LI channel is typically much longer than that of the S→R and R→D channels [98, Table I]. Therefore, we assume that the relay has perfect knowledge of the LI channel \mathbf{G}_{RR} .

5.2.3 Data Transmission

For notational convenience, we write the transmit signals from the source terminals at time instant i as

$$\mathbf{x}_{\text{S}}[i] = \text{diag}(\sqrt{\rho_{\text{S},1}}, \sqrt{\rho_{\text{S},2}}, \dots, \sqrt{\rho_{\text{S},K}})\bar{\mathbf{x}}[i],$$

where the entries of $\bar{\mathbf{x}}[i]$ are i.i.d. standard CSCG. After subtracting the known part of the LI from the quantized signal (5.3) by using the knowledge of θ , \mathbf{G}_{RR} and $\mathbf{x}_{\text{R}}[i]$, the receive-side signal at the relay is given by

$$\theta\mathbf{G}_{\text{SR}}\mathbf{x}_{\text{S}}[i] + \mathbf{n}_{\text{R}} = \mathbf{r}_{\text{R}}[i] - \theta\mathbf{G}_{\text{RR}}\mathbf{x}_{\text{R}}[i].$$

Note that although the data-dependent part of the LI is perfectly cancelled in baseband, the transmit-side hardware impairments received through the LI channel as well as the quantization noise can still have severe impact on the performance of the system via \mathbf{n}_{R} . The power of the noise term \mathbf{n}_{R} , defined in (5.3), will be calculated in the next section.

After ADCs, the relay uses linear estimator \mathbf{W} to separate the data streams transmitted by the source terminals. The k th estimated signal stream at the relay reads then

$$r_{R,k}[i] = \theta \mathbf{w}_k^H \mathbf{g}_{\text{SR},k} x_{\text{S},k}[i] + \theta \sum_{j \neq k} \mathbf{w}_k^H \mathbf{g}_{\text{SR},j} x_{\text{S},j}[i] + \mathbf{w}_k^H \mathbf{n}_R, \quad (5.4)$$

where $\mathbf{g}_{\text{SR},k}$, $\mathbf{g}_{\text{RR},k}$ and \mathbf{w}_k are the k th columns of \mathbf{G}_{SR} , \mathbf{G}_{RR} and \mathbf{W} , respectively, and $x_{\text{S},k}[i]$ is the k th element of $\mathbf{x}_{\text{S}}[i]$. Similarly, linear precoding with matrix \mathbf{V} is used at the relay to transmit the data to the destination terminals, where the received signal at the k th terminal reads

$$y_{\text{D},k}[i] = \mathbf{g}_{\text{RD},k}^T \mathbf{v}_k \bar{x}_k[i - \tau] + \sum_{j \neq k} \mathbf{g}_{\text{RD},k}^T \mathbf{v}_j \bar{x}_j[i - \tau] + n_{\text{D},k}.$$

Here $\mathbf{g}_{\text{RD},k}$, \mathbf{v}_k , $\bar{x}_k[i - \tau]$ and $n_{\text{D},k}$ denote the k th columns (or elements) of \mathbf{G}_{RD} , \mathbf{V} , $\bar{\mathbf{x}}[i - \tau]$ and \mathbf{n}_{D} , respectively. We assume that in DF relaying the decoding delay is $\tau \geq 1$ symbols so that $\mathbf{x}_{\text{R}}[i] = \mathbf{V} \bar{\mathbf{x}}[i - \tau]$ and, thus, the transmitted signal is uncorrelated with the received signal at the relay [95]. Finally, note that the estimation and precoding matrices \mathbf{W} and \mathbf{V} are functions of $\hat{\mathbf{G}}_{\text{SR}}$ and $\hat{\mathbf{G}}_{\text{RD}}$, respectively.

5.3 Achievable Rate Analysis

Since \mathbf{n}_R is not Gaussian distributed, we resort to calculating a lower bound on the achievable rate by replacing the combination of inter-pair interference and noise in (5.4) by a CSCG term that has the same variance but is independent of the desired signal [96]. For the k th S \rightarrow R and R \rightarrow D links, the lower bounds¹ on achievable rates are given by

$$R_{*,k} = \log_2(1 + \text{SINR}_{*,k}). \quad (5.5)$$

¹Although the relay has instantaneous channel estimates $\hat{\mathbf{g}}_{\text{SR},k}$, as in [12, 63], we assume that statistical channel estimates $\mathbb{E}_{\{\mathbf{g}\}}\{\mathbf{w}_k^H \hat{\mathbf{g}}_{\text{SR},k}\}$ are used for decoding. This provides a lower bound on the SINR of the S \rightarrow R link.

The signal-to-interference-plus-noise ratios associated with (5.5) are given by

$$\begin{aligned} \text{SINR}_{\text{SR},k} &= \frac{\rho_{\text{S},k} |\mathbb{E}_{\{\mathbf{g}\}} \{\mathbf{w}_k^H \mathbf{g}_{\text{SR},k}\}|^2}{\rho_{\text{S},k} \text{Var}_{\{\mathbf{g}\}} (\mathbf{w}_k^H \mathbf{g}_{\text{SR},k}) + \sum_{j=1, j \neq k}^K \rho_{\text{S},j} \mathbb{E}_{\{\mathbf{g}\}} \{|\mathbf{w}_k^H \mathbf{g}_{\text{SR},j}|^2\} + \mathbb{E}_{\{\mathbf{g}, \mathbf{n}_R\}} \{\|\mathbf{w}_k^H \mathbf{n}_R\|^2\}} \\ \text{SINR}_{\text{RD},k} &= \frac{|\mathbb{E}_{\{\mathbf{g}\}} \{\mathbf{g}_{\text{RD},k}^T \mathbf{v}_k\}|^2}{\text{Var}_{\{\mathbf{g}\}} (\mathbf{g}_{\text{RD},k}^T \mathbf{v}_k) + \sum_{j=1, j \neq k}^K \mathbb{E}_{\{\mathbf{g}\}} \{|\mathbf{g}_{\text{RD},k}^T \mathbf{v}_j|^2\} + \sigma_w^2}, \end{aligned}$$

where \mathbf{n}_R denotes—with some abuse of notation—a CSCG vector that is independent of the desired signal and has i.i.d. entries of variance (provided later in this subsection) σ_R^2 .

To calculate a lower bound on the ergodic E2E rate R_k for the k th terminal pair, we follow [12] and let

$$\begin{aligned} R_k &= \min\{R_{\text{SR},k}, R_{\text{RD},k}\} \\ &= \log_2 (1 + \min\{\text{SINR}_{\text{SR},k}, \text{SINR}_{\text{RD},k}\}). \end{aligned} \quad (5.6)$$

Since matched-filtering (MF) have been studied in previous chapters, we limit our analysis to zero-forcing (ZF) processing, which is a more commonly used method in real systems, for both estimation and precoding,

$$\begin{aligned} \mathbf{W}^H &= (\hat{\mathbf{G}}_{\text{SR}}^H \hat{\mathbf{G}}_{\text{SR}})^{-1} \hat{\mathbf{G}}_{\text{SR}}^H, \\ \mathbf{V} = \mathbf{B}\mathbf{P} &= \hat{\mathbf{G}}_{\text{RD}}^* (\hat{\mathbf{G}}_{\text{RD}}^T \hat{\mathbf{G}}_{\text{RD}}^*)^{-1} \mathbf{P}, \end{aligned}$$

where $\mathbf{P} \in \mathbb{C}^{K \times K}$ is a diagonal power allocation matrix. The k th diagonal entry of \mathbf{P} is given by

$$p_k = \sqrt{\frac{m_k}{\mathbb{E}\{\|\mathbf{b}_k\|^2\}}} = \sqrt{(M_{\text{Tx}} - K) \hat{\beta}_{\text{RD},k} m_k},$$

where \mathbf{b}_k is the k th column of \mathbf{B} and m_k denotes the relay's transmit power for the k th link. The m_k 's that maximize the pairwise E2E rate are designed in the next section.

Using similar techniques as in [100], a lower bound for the E2E achievable rate of the k th terminal pair reads

$$R_k = \log_2 \left(1 + \min \left(\frac{\theta^2 \rho_{S,k} (M_{\text{Rx}} - K) \hat{\beta}_{\text{SR},k}}{\theta^2 \sum_{j=1}^K \rho_{S,j} \tilde{\beta}_{\text{SR},j} + \sigma_{\text{R}}^2}, \frac{(M_{\text{Tx}} - K) \hat{\beta}_{\text{RD},k} m_k}{\tilde{\beta}_{\text{RD},k} m_{\text{tot}} + \sigma_{\text{w}}^2} \right) \right), \quad (5.7)$$

where $m_{\text{tot}} = \sum_k m_k$ is the total transmit power of the relay. The variance of the elements of \mathbf{n}_{R} is given by

$$\sigma_{\text{R}}^2 = \theta \mu_{\text{Tx}} \beta_{\text{R}} m_{\text{tot}} + \theta (1 - \theta) \left(\sum_{k=1}^K \rho_{S,k} \beta_{\text{SR},k} + \beta_{\text{R}} m_{\text{tot}} \right) + \theta \sigma_{\text{w}}^2.$$

Given the above achievable rate analysis, we define the average E2E throughput as

$$C = f_{\text{B}} \cdot \frac{T - 2K}{T} \cdot \mathbb{E} \left\{ \sum_{k=1}^K R_k \right\}, \quad (5.8)$$

where f_{B} denotes the system bandwidth. Note that the expectation in (5.8) is over the terminal locations and shadowing.

5.3.1 Delay Constrained Achievable Rate

To achieve reliable communication at a rate given in (5.7), the code word length needs to grow without bound. In principle, this implies infinite delay in decoding. While some applications can afford long delays so that (5.7) is a reasonable approximation, e.g. in URLLC [99] the delay is very short so that the previous rate calculations yield inaccurate results.

With the above in mind, for delay sensitive applications we replace the $R_{*,k}$ in (5.6) by the finite block length rate [102]

$$R_{*,k}^{\tau} = R_{*,k} - \sqrt{\frac{V}{\tau}} Q^{-1}(\epsilon) + \frac{\log_2 \tau}{2\tau},$$

where $Q^{-1}(x)$ denotes the inverse of the Q-function. The variable V is so-called channel dispersion and given for the system under consideration by [102]

$$V = \frac{\text{SINR}_{*,k}(2 + \text{SINR}_{*,k})}{(1 + \text{SINR}_{*,k})^2} (\log_2 e)^2,$$

since (5.5) represents the capacity of an AWGN channel. The code word length is the same as decoding delay and constrained to be $\tau < T - 2K$, that is, the decoding has to be carried out within one coherence block. For simplicity, in the numerical results we set the target error probability ε to be equal for the $S \rightarrow R$ and $R \rightarrow D$ links. Given the desired E2E decoding error is ε_{SD} , the code word error probability over one channel is simply $\varepsilon = 1 - \sqrt{1 - \varepsilon_{\text{SD}}}$.

5.4 Novel Power Allocation Scheme

Since the power of quantization noise depends on the level of residual LI before digital cancellation is carried out, the relay needs to adjust the transmit power by taking into account the effects of the transmit-side hardware impairments and quantization at the receive-side. This can be achieved by taking advantage of the analysis carried out in the previous section.

As shown in (5.6), the E2E rate depends on the weaker link. Therefore, we propose adjusting the link-wise powers $\{m_1, m_2, \dots, m_k\}$ so that $\text{SINR}_{\text{SR},k} = \text{SINR}_{\text{RD},k}$ for all $k = 1, 2, \dots, K$. This can be achieved via a simple iterative algorithm that allocates power

$$m_k^{(l)} = \frac{\frac{M_{\text{Rx}} - K}{M_{\text{Tx}} - K} \rho_{\text{S},k} \hat{\beta}_{\text{SR},k} (\tilde{\beta}_{\text{RD},k} m_{\text{tot}}^{(l-1)} + \sigma_{\text{w}}^2)}{\hat{\beta}_{\text{RD},k} (\sum_{j=1}^K \rho_{\text{S},j} \tilde{\beta}_{\text{SR},j} + \frac{1}{\theta^2} \sigma_{\text{R}}^2)}, \quad (5.9)$$

to the k th link at the l th iteration. The variable $m_{\text{tot}}^{(l-1)} = \sum_k m_k^{(l-1)}$ is the total transmit power of the relay station in the $(l-1)$ th iteration. The initial point for the iterations (5.9) can be obtained by assuming that the large scale fading factors of all terminals

are the same and setting $m_1^{(0)} = m_2^{(0)} = \dots = m_k^{(0)}$. Note that the proposed power allocation scheme uses only statistical information so that the power allocation needs to be updated only when the channel statistics change. The algorithm also converges fast, typically within 30 iterations.

5.5 Energy Efficiency

In this chapter, the energy consumed by the source and destination terminals is neglected and the energy efficiency for FD relaying is defined as

$$EE = \frac{C}{E_{\text{tot}}} T,$$

where E_{tot} denotes the average total energy consumption of the relay during one coherence block and is given by

$$E_{\text{tot}} = K\rho_{\text{Rx,tot}} + \frac{KM_{\text{Tx}}}{M_{\text{Rx}}}\rho_{\text{Rx,tot}} + (T - 2K)(\rho_{\text{Tx,tot}} + \rho_{\text{Rx,tot}}). \quad (5.10)$$

The first and second term include the energy consumed by pilot reception on the S→R and the R→D channel, respectively, while the last term denotes the energy consumed by data reception and transmission at the relay.

To evaluate (5.10), we need to obtain the total power consumption at the receive-side, $\rho_{\text{Rx,tot}}$, and transmit-side, $\rho_{\text{Tx,tot}}$, of the relay. At the receive-side, we single out the power consumed by one ADC, ρ_{ADC} , as in [71] and define the power consumed by the rest of the components in terms of the power $P_{\text{ADC,ref}}$ consumed by a reference N_{ref} -bit ADC

$$\rho_{\text{Rx,tot}} = M_{\text{Rx}}(2\rho_{\text{ADC}} + \eta_{\text{Rx}}P_{\text{ADC,ref}}).$$

The scalar $\eta_{\text{Rx}} > 0$ is a constant that depends on the system architecture. The transmit-side power excluding the power amplifiers (PAs), ρ_{PA} , is defined similarly

$$\rho_{\text{Tx,tot}} = M_{\text{Tx}} \eta_{\text{Tx}} \rho_{\text{ADC,ref}} + \rho_{\text{PA}}.$$

The parameter $\eta_{\text{Tx}} > 0$ depends on the architecture and $\rho_{\text{PA}} = m_{\text{tot}}/\delta$, where δ denotes the efficiency of the PA.

Finally, for $N > 1$ the power consumed by one ADC is assumed to grow exponentially with N according to [66]

$$\rho_{\text{ADC}} = E_{\text{con}} 2^N f_{\text{samp}}, \quad N \geq 2,$$

where f_{samp} denotes sampling rate and E_{con} depends on the type of ADC, the average of which is around 5 pJ. The power consumed by one-bit ADC is considered to be negligible [103].

For a relay that operates in HD mode, the active transmit- and receive-side power consumption is the same as in the FD mode. When one of the RF chains is not active, it is assumed to be in a sleep mode that consumes half of the corresponding power when active [104].

5.6 Numerical Results

The default parameters in the numerical examples are $T = 100$, $K = 10$, $M_{\text{Rx}} = M_{\text{Tx}} = 100$, $f_{\text{B}} = 20$ MHz, $N_{\text{ref}} = 2$, $\delta = 10\%$, $E_{\text{con}} = 5$ pJ, and $\sigma_{\text{w}}^2 = -101$ dBm. For the geometric model we choose $r_{\text{SR}} = r_{\text{RD}} = 100$ m, $d_{\text{SR}} = d_{\text{RD}} = 400$ m and the radio propagation model includes path loss with exponent $\alpha = 4$ and log-normal shadowing with zero-mean and 6 dB variance.

The power of residual LI channel after passive and analog mitigation is assumed to be $\beta_R = -90$ dB [36, 75]. Nyquist-rate sampling is used so that $f_{\text{samp}} = 2f_B$. Since transmitters consume typically more power than receivers, we let $\eta_{\text{Tx}} = 3\eta_{\text{Rx}}$. Power constraint $m_{\text{tot}} \leq 23$ dBm is set for the relay. According to the LTE standard, the required EVM range is [0.08, 0.175], so we model the transmit-side hardware imperfections using $\mu_{\text{Tx}} = \text{EVM}^2$ and select $\text{EVM} = 0.1$.

To avoid near-far problem at the relay station, simple power control based on long-term statistics is applied at the source terminals, i.e., $\rho_{\text{S},k} = \frac{\gamma}{\beta_{\text{SR},k}}$, where γ is a design parameter. We also denote $\bar{\rho}_{\text{S}} = \text{E}\{\rho_{\text{S},k}\}$ for notational simplicity, where the expectation is over the source terminals, locations and shadow fading. Similarly, in channel estimation phase, we set $\rho_{\text{pS},k} = \gamma_p/\beta_{\text{SR},k}$ and $\rho_{\text{pD},k} = \gamma_p/\beta_{\text{RD},k}$, where $\gamma = \gamma_p$ is assumed in the numerical results.

Fig. 5.1 demonstrates the throughput of the proposed power allocation scheme for two parameter values $\gamma = -100$ dBm ($\bar{\rho}_{\text{S}} = 6.92$ dBm) and $\gamma = -90$ dBm ($\bar{\rho}_{\text{S}} = 16.85$ dBm), that correspond to received signal-to-noise ratio (SNR) of roughly 0 dB and 10 dB, respectively. Clearly, as the relay station operates in FD mode, the proposed power allocation scheme gives the highest throughput. The curves for HD relaying saturate around $N = 4$ (with relay power optimized so that $R_{\text{SR},k} = R_{\text{RD},k}$). For FD relaying with or without proposed power allocation, it is not necessary to use more than 6-bit ADCs for near optimal performance.

The cumulative density functions (CDFs) of m_{tot} for the proposed power allocation scheme are illustrated in Fig. 5.2 for different resolutions of ADCs. Only the case $\gamma = -100$ dBm is shown ($\gamma = -90$ dBm yields similar results). The CDFs show how the proposed algorithm reduces the transmit power in case of low resolution ADCs to control quantization noise. The relay power constraint is also satisfied in all cases.

Fig. 5.3 plots the EE vs. the resolution of the ADCs when $\gamma = -100$ dBm. The optimum resolution of ADCs for maximizing the EE for typical range of η_{Rx} is from

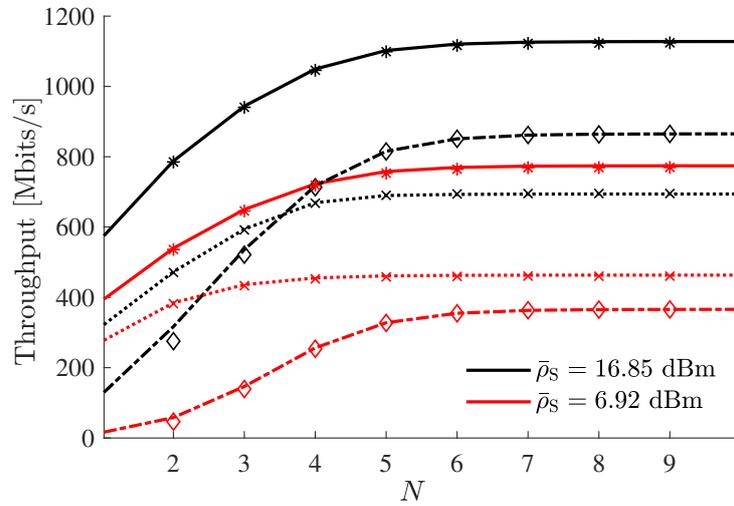


Fig. 5.1 Throughput of the system vs. the resolution of ADC, analytical results of the proposed power allocation scheme in FD mode (solid), the proposed power allocation scheme in HD mode (dotted) and fixed relay power $m_{\text{tot}} = 23$ dBm in FD mode (dash-dotted) presented with curves and Monte Carlo simulations with markers (stars, crosses and diamonds, respectively).

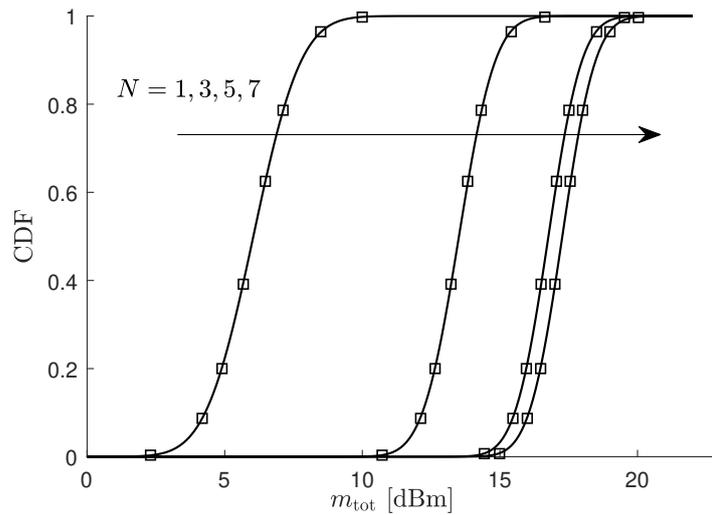


Fig. 5.2 Empirical CDF of total transmit power at the FD relay station with proposed power allocation scheme.

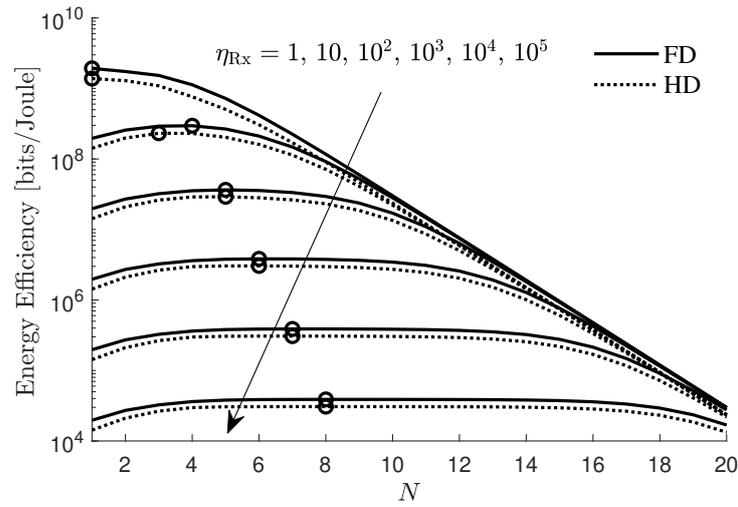


Fig. 5.3 Energy efficiency of the proposed power allocation scheme vs. ADC resolution N at the relay station with a circle at the peak of each curve.

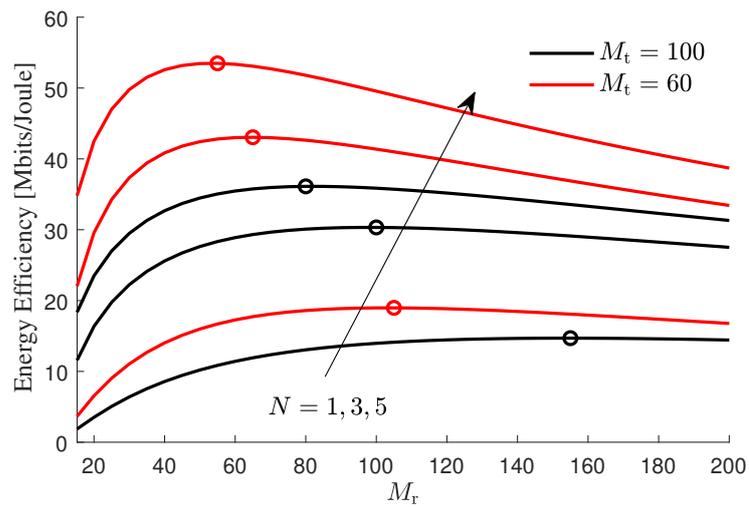


Fig. 5.4 Energy efficiency vs. the number of receive antennas at the FD relay station of the proposed power allocation scheme with a circle at the peak of each curve.

4 to 7 bits for both FD and HD. As η_{Rx} increases, the relative energy consumed by components other than the ADCs grows. As a result, the optimum resolution of the ADCs also increases. In all cases, the proposed FD scheme outperforms HD in energy efficiency.

The optimum number of receive antennas at the relay station in terms of EE is investigated in Fig. 5.4 for $\gamma = -100$ dBm and $\eta_{Rx} = 10^2$. For low resolution ADCs,

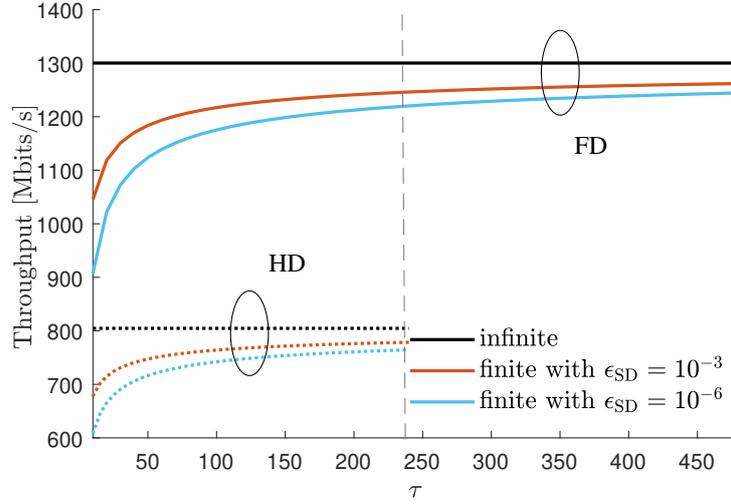


Fig. 5.5 Throughput vs. code length τ with $T = 500$.

the system should employ higher number of receive antennas than transmit antennas ($M_{R_x} > M_{T_x}$) to maximize EE. This is explained by the relatively low power consumption of the 1 and 3 bit ADCs. Higher resolution $N = 5$ ADCs, on the other hand, consume more power and the opposite ($M_{R_x} < M_{T_x}$) is true.

Fig. 5.5 illustrates how fast the system throughput with finite code word length approaches that of the infinite code word length. Since the pilot transmission consumes $2K$ symbols, the code word length with HD and FD relaying cannot exceed $(T - 2K)/2 = 240$ symbols and $T - 2K = 480$ symbols, respectively. It is slightly surprising that the throughput of FD relaying with very strict delay constraint outperforms HD relaying without delay constraint under the given system parameters. This indicates that the proposed FD relaying scheme is a suitable technique for URLLC applications.

5.7 Conclusion

In this chapter, we have analyzed the throughput and energy efficiency of a massive MIMO multipair decode-and-forward FD relaying system with low-resolution ADCs.

Based on the analysis, we have proposed a novel power allocation scheme that is aware of the transmit-side hardware imperfections and quantization noise. The proposed scheme aims at maximizing the achievable rate by adjusting the transmit power at the relay on link-by-link basis. The numerical examples demonstrate the benefits of the proposed scheme in improving both the throughput and energy efficiency compared to other considered methods.

Chapter 6

Carrier Frequency Offset Estimation

6.1 Introduction

Since applying large number of antennas at a node in wireless networks is circuit power consuming, implementing one-bit ADCs at the receivers has been proposed in such systems for its nearly zero power consumption and small area occupation (see [80, 79] for example). The study of synchronization problems in such systems are limited. To achieve better performance, synchronization can be carried out in time and frequency domain by estimating and compensating the time offset (TO) and carrier frequency offset (CFO) between user devices and BS/relaying with large number of antennas. Here we mainly consider the frequency synchronization problems.

In this chapter, rather than using thousands of pilot symbols for CFO estimation as in [89], we use only two pilot symbols for individual users to perform CFO estimations targeting on three different scenarios. The three scenarios are: a large but finite number of antennas at the receiver, infinite number of antennas at the receiver, and very small CFO. These estimation methods are extended to multiple pilot symbol case. These novel CFO estimation methods are low-complexity and can avoid pilot overhead.

6.2 System Model

We consider a system with massive MIMO receiver, which can be the uplink of a massive MIMO BS or the source-to-relay link in the FD relaying system mentioned in previous chapters. For notational simplicity, we assume K user devices are conducting uplink transmission to a base station (BS) with M antennas. We assume that each antenna is adapted with a individual RF chain, and all the RF chains at the BS share the same oscillator. Since the user devices and BS are hardly perfectly synchronized due to Doppler shift [13], the received signals from different user devices are affected by from different phase shifts. At time index t , the received signal at the base station is given by

$$\mathbf{y}[t] = \sum_{k=1}^K e^{j\omega_k t} \mathbf{g}_k x_k[t] + \mathbf{n}[t], \quad (6.1)$$

where $e^{j\omega_k t}$ is the CFO of the k th link, denoting the phase shift of the received signal. Note that CFO is a function of t . Even small ω_k can affect the detection with the growth of time t . Normalized additive noise $\mathbf{n}[t] \in \mathbb{C}^{M \times 1}$ is circularly symmetric complex Gaussian (CSCG) distributed. Each entry of $\mathbf{n}[t]$ has zero mean and variance σ_n^2 . Channel vector is decomposed as $\mathbf{g}_k = \beta_k \bar{\mathbf{g}}_k, \forall k$, where β_k denotes large-scale fading of the k th link. $\bar{\mathbf{g}}_k \in \mathbb{C}^{M \times 1}$ denotes small-scale fading, entries of which are i.i.d. CSCG random variables (RVs) with zero means and unit variances. For any time instance t , the Gaussian signalling is used at user devices. The transmit signals from the any user k has constant power $\mathbb{E}\{\|x_k[t]\|^2\} = \rho_k$.

To reduce the circuit power consumption, BS equips with one-bit ADCs, which gives binary outputs of real and imaginary parts of received signals, i.e. $\mathbf{r}[t] \in \frac{1}{\sqrt{2}}\{1 + j, -1 + j, 1 - j, -1 - j\}$. We denote quantization process as

$$\mathbf{r}[t] = \mathcal{Q}(\mathbf{y}[t]) = \frac{1}{\sqrt{2}} (\text{sign}(\Re(\mathbf{y}[t])) + j \text{sign}(\Im(\mathbf{y}[t]))), \quad (6.2)$$

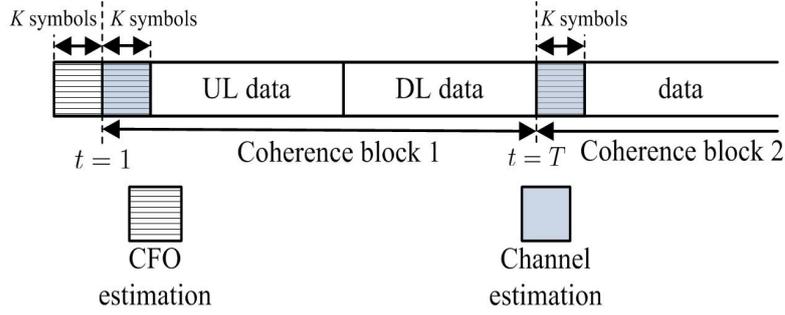


Fig. 6.1 Transmission Protocol.

where $\text{sign}(\cdot)$ denotes the sign function. $\Re(\cdot)$ and $\Im(\cdot)$ take the real and imaginary part of the signal, respectively.

6.3 CFO Estimation

As shown in Fig. 6.1, $2K$ symbols are allocated for CFO estimation. Here block fading is assumed so that channel stays unchanged within T symbols and $T > 2K$. Each user transmit two pilot symbols to the BS and the pilot symbols have equal power $K\rho$. Note that at any time instant t , only one user device is allowed to transmit pilot symbol in order to completely avoid pilot contamination caused by CFO. We denote the pilot symbols from k th user device as ϕ_{t_1} and ϕ_{t_2} , which are transmitted at time indices $t_1 \in [-K+1, 0]$ and $t_2 \in [1, K]$ respectively. The received pilots from k th user device at the BS are denoted by

$$\mathbf{Y}_p = \mathbf{g}_k \boldsymbol{\phi}_p^T + \mathbf{N}_p, \quad (6.3)$$

where $\mathbf{N}_p = (\mathbf{n}[t_1], \mathbf{n}[t_2])$. For notation simplicity, we combine CFO with pilot symbols as $\boldsymbol{\phi}_p^T = (e^{j\omega_k t_1} \phi_{t_1}, e^{j2\omega_k t_2} \phi_{t_2})$. With some abuse of subscript, the vectorized received pilot vector is given by

$$\mathbf{y}_p = \text{vec}(\mathbf{Y}_p) = \boldsymbol{\Phi}_p \mathbf{g}_k + \mathbf{n}_p, \quad (6.4)$$

where $\mathbf{n}_p = \text{vec}(\mathbf{N}_p)$ and $\Phi_p = \phi_p \otimes \mathbf{I}_M$. The direct analysis of one-bit quantization model is difficult, we herein introduce Bussgang decomposition [78, 79] to analyze the statistics of one-bit quantization. The quantized vectorized received pilots can be decomposed as

$$\begin{aligned} \mathcal{Q}(\mathbf{y}_p) &= \text{vec}(\mathbf{r}_p[t_1], \mathbf{r}_p[t_2]) \\ &= \mathbf{A}_p \mathbf{y}_p + \mathbf{q} \\ &= \mathbf{A}_p \Phi_p \mathbf{g}_k + \mathbf{A}_p \mathbf{n}_p + \mathbf{q}, \end{aligned} \quad (6.5)$$

where \mathbf{q} denotes the quantization noise, the statistics of which will be analyzed later. $\mathbf{r}_p[t_1]$ and $\mathbf{r}_p[t_2]$ are the quantized received pilots at time t_1 and t_2 respectively. Matrix \mathbf{A}_p can be obtained by minimizing the quantization noise,

$$\mathbf{A}_p = \mathbf{C}_{\mathbf{y}_p \mathbf{r}_p}^H \mathbf{C}_{\mathbf{y}_p \mathbf{y}_p}^{-1}, \quad (6.6)$$

where $\mathbf{C}_{\mathbf{y}_p \mathbf{y}_p}$ denotes auto-correlation of \mathbf{y}_p which is calculated by

$$\mathbf{C}_{\mathbf{y}_p \mathbf{y}_p} = \text{E}\{\mathbf{y}_p \mathbf{y}_p^H\} = \beta_k \Phi_p \Phi_p^H + \sigma_n^2 \mathbf{I}_{2M}. \quad (6.7)$$

According to Bussgang theorem, for one-bit quantizer and Gaussian inputs, we have the cross-correlation between \mathbf{r}_p and \mathbf{y}_p [78],

$$\mathbf{C}_{\mathbf{y}_p \mathbf{r}_p} = \sqrt{\frac{2}{\pi}} \mathbf{C}_{\mathbf{y}_p \mathbf{y}_p} \text{diag}(\mathbf{C}_{\mathbf{y}_p \mathbf{y}_p})^{-\frac{1}{2}}, \quad (6.8)$$

yielding the linear operator being

$$\begin{aligned} \mathbf{A}_p &= \sqrt{\frac{2}{\pi}} \text{diag}(\mathbf{C}_{\mathbf{y}_p \mathbf{y}_p})^{-\frac{1}{2}}, \\ &= \begin{bmatrix} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\beta_k |\phi_{t_1}|^2 + \sigma_n^2}} & 0 \\ 0 & \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\beta_k |\phi_{t_2}|^2 + \sigma_n^2}} \end{bmatrix} \otimes \mathbf{I}_M \end{aligned} \quad (6.9)$$

Here for notational simplicity we denote

$$a_{pt} = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\beta_k |\phi_t|^2 + \sigma_n^2}}. \quad (6.10)$$

The auto-correlation on quantization noise is given in [79]

$$\mathbf{C}_q = \frac{2}{\pi} (\arcsin(\mathbf{P}_p) + j \arcsin(\mathbf{Q}_p)) - \frac{2}{\pi} (\mathbf{P}_p + j\mathbf{Q}_p), \quad (6.11)$$

where \mathbf{P}_p and \mathbf{Q}_p are given by

$$\begin{aligned} \mathbf{P}_p &= \text{diag}(\mathbf{C}_{y_p, y_p})^{-\frac{1}{2}} \Re(\mathbf{C}_{y_p, y_p}) \text{diag}(\mathbf{C}_{y_p, y_p})^{-\frac{1}{2}} \\ &= \begin{bmatrix} 1 & \frac{\beta_k \Re(\phi_1 \phi_2^* e^{-j\omega_k(t_2-t_1)})}{\sqrt{(\beta_k |\phi_1|^2 + \sigma_n^2)(\beta_k |\phi_2|^2 + \sigma_n^2)}} \\ \frac{\beta_k \Re(\phi_2 \phi_1^* e^{j\omega_k(t_2-t_1)})}{\sqrt{(\beta_k |\phi_2|^2 + \sigma_n^2)(\beta_k |\phi_1|^2 + \sigma_n^2)}} & 1 \end{bmatrix} \otimes \mathbf{I}_M \end{aligned} \quad (6.12)$$

$$\begin{aligned} \mathbf{Q}_p &= \text{diag}(\mathbf{C}_{y_p, y_p})^{-\frac{1}{2}} \Im(\mathbf{C}_{y_p, y_p}) \text{diag}(\mathbf{C}_{y_p, y_p})^{-\frac{1}{2}} \\ &= \begin{bmatrix} 0 & \frac{\beta_k \Im(\phi_1 \phi_2^* e^{-j\omega_k(t_2-t_1)})}{\sqrt{(\beta_k |\phi_1|^2 + \sigma_n^2)(\beta_k |\phi_2|^2 + \sigma_n^2)}} \\ \frac{\beta_k \Im(\phi_2 \phi_1^* e^{j\omega_k(t_2-t_1)})}{\sqrt{(\beta_k |\phi_2|^2 + \sigma_n^2)(\beta_k |\phi_1|^2 + \sigma_n^2)}} & 0 \end{bmatrix} \otimes \mathbf{I}_M \end{aligned} \quad (6.13)$$

Note that (6.11) can be represented by using Kronecker product \otimes as

$$\mathbf{C}_q = \mathbf{C} \otimes \mathbf{I}_M \quad (6.14)$$

where the $\mathbf{C} \in \mathbb{C}^{2 \times 2}$ has the element on its second row and first column denoting $C_{t_2 t_1}$.

Note \mathbf{C} has identical diagonal elements $C_{\text{diag}} = 1 - \frac{2}{\pi}$.

We define the ‘‘unbiased’’ average correlation between two quantized received vectors

as

$$\frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2] = \phi_{t_1}^* \phi_{t_2} \zeta_1 e^{j\omega_k(t_2-t_1)} + \sum_{i=2}^9 \zeta_i, \quad (6.15)$$

where

$$\zeta_1 = \frac{1}{M} a_{pt_1} a_{pt_2} \mathbf{g}_k^H \mathbf{g}_k, \quad (6.16)$$

$$\zeta_2 = \frac{1}{M} a_{pt_1} a_{pt_2} \mathbf{n}_{t_1}^H \mathbf{n}_{t_2}, \quad (6.17)$$

$$\zeta_3 = \frac{1}{M} \phi_{t_1}^* e^{-j\omega_k t_1} a_{pt_1} a_{pt_2} \mathbf{g}_k^H \mathbf{n}_{pt_2}, \quad (6.18)$$

$$\zeta_4 = \frac{1}{M} \phi_{t_2} e^{j\omega_k t_2} a_{pt_1} a_{pt_2} \mathbf{n}_{pt_1}^H \mathbf{g}_k, \quad (6.19)$$

$$\zeta_5 = \frac{1}{M} \phi_{t_1}^* e^{-j\omega_k t_1} a_{pt_1} \mathbf{g}_k^H \mathbf{q}_{t_2}, \quad (6.20)$$

$$\zeta_6 = \frac{1}{M} \phi_{t_2} e^{j\omega_k t_2} a_{pt_2} \mathbf{q}_{t_1}^H \mathbf{g}_k, \quad (6.21)$$

$$\zeta_7 = \frac{1}{M} a_{pt_1} \mathbf{n}_{t_1}^H \mathbf{q}_{t_2}, \quad (6.22)$$

$$\zeta_8 = \frac{1}{M} a_{pt_2} \mathbf{q}_{t_1}^H \mathbf{n}_{t_2}, \quad (6.23)$$

$$\zeta_9 = \frac{1}{M} \mathbf{q}_{t_1}^H \mathbf{q}_{t_2}. \quad (6.24)$$

By analyzing the statistics of the “unbiased” average correlation in (6.15), we will propose the following three methods for CFO estimation. In method 1, we consider a practical scenario where large but finite number of antennas are deployed at the BS. Since ζ_1 depends on the channel gain $\|\mathbf{g}_k\|^2$ and usually far bigger than ζ_2, \dots, ζ_9 , which will be proved later, we treat the ω_k in the first term on the right hand side of (6.15) as desired part. We will then convert ζ_2, \dots, ζ_9 to phase noise¹ terms $\zeta'_{2,Q}, \dots, \zeta'_{9,Q}$, and analyse the statistics them. We proposed the method 1 based on the mean of $\zeta'_{2,Q}, \dots, \zeta'_{9,Q}$. The derived closed-form variances of these noise terms indicate the accuracy of the estimation. Method 2 and 3 utilize extreme cases of the infinite antenna M and small CFO assumptions, respectively. Due to the extreme assumptions, the performance analysis of method 2 and 3 is not feasible. We are only able to partially analyze the performance of method 1.

¹The phase noise is an ambiguous definition, since we can extract useful information about ω_k from the noise terms.

6.3.1 Method 1: Phase Noise Analysis

The first term on the left hand side of (6.15) can be written as

$$\phi_{t_1}^* \phi_{t_2} \zeta_1 e^{j\omega_k(t_2-t_1)} = |\phi_{t_1}^* \phi_{t_2}| \zeta_1 e^{j(\omega_k(t_2-t_1) + \angle \phi_{t_1}^* \phi_{t_2})}, \quad (6.25)$$

which is a shifted, scaled and rotated version of ζ_1 . As pilot sequence is known to the BS, $\angle \phi_{t_1}^* \phi_{t_2}$ in (6.27) can always be compensated easily. We therefore make a design choice of $\phi_{t_1} = \phi_{t_2} = \phi$, so that $a_{pt_1} = a_{pt_2} = a_p$. Note that ζ_1 in (6.25) is real. It does not bring phase noise. We can simplify (6.25) as

$$|\phi_{t_1}^* \phi_{t_2}| \zeta_1 e^{j(\omega_k(t_2-t_1) + \angle \phi_{t_1}^* \phi_{t_2})} = |\phi|^2 \zeta_1 e^{j\omega_k(t_2-t_1)}, \quad (6.26)$$

A simple strategy of obtaining the CFO ω_k is to take the angle of the term $\frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2]$. We assume the in-phase and quadrature of $\sum_{i=2}^9 \zeta'_i$, denoted by $\sum_{i=2}^9 \zeta'_{i,I}$ and $\sum_{i=2}^9 \zeta'_{i,Q}$ respectively.

Proposition 6.1. *If $\sum_{i=2}^9 \zeta'_i$ is far less than 1, the following formula holds for any ω_k ,*

$$\angle \frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2] = \omega_k(t_2 - t_1) + \sum_{i=2}^9 \zeta'_{i,Q}, \quad (6.27)$$

where

$$\zeta'_i = \frac{1}{|\phi|^2 \zeta_1} \zeta_i e^{-j\omega_k(t_2-t_1)}. \quad (6.28)$$

Proof. By substituting (6.26) into (6.15), we obtain

$$\frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2] = |\phi|^2 \zeta_1 e^{j\omega_k(t_2-t_1)} + \sum_{i=2}^9 \zeta_i \quad (6.29)$$

$$= |\phi|^2 \zeta_1 \left(1 + \frac{\sum_{i=2}^9 \zeta_i e^{-j\omega_k(t_2-t_1)}}{|\phi|^2 \zeta_1} \right) e^{j\omega_k(t_2-t_1)} \quad (6.30)$$

$$= \left(1 + \sum_{i=2}^9 \zeta'_i \right) |\phi|^2 \zeta_1 e^{j\omega_k(t_2-t_1)} \quad (6.31)$$

According to [105, eq. 6],

$$1 + \sum_{i=2}^9 \zeta'_i = \left(\left(1 + \sum_{i=2}^9 \zeta'_{i,I} \right)^2 + \left(\sum_{i=2}^9 \zeta'_{i,Q} \right)^2 \right)^{\frac{1}{2}} e^{\left(\arctan \left(\frac{\sum_{i=2}^9 \zeta'_{i,Q}}{1 + \sum_{i=2}^9 \zeta'_{i,I}} \right) \right)}, \quad (6.32)$$

Taylor series of $\frac{1}{1+x}$ up to second order at $x = 0$ is $\frac{1}{1+x} = 1 - x$. For $\sum_{i=2}^9 \zeta'_{i,I} \approx 0$, we have the following approximation

$$\arctan \left(\frac{\sum_{i=2}^9 \zeta'_{i,Q}}{1 + \sum_{i=2}^9 \zeta'_{i,I}} \right) \quad (6.33)$$

$$= \arctan \left(\sum_{i=2}^9 \zeta'_{i,Q} \left(1 - \sum_{i=2}^9 \zeta'_{i,I} \right) \right) \quad (6.34)$$

$$\approx \arctan \left(\sum_{i=2}^9 \zeta'_{i,Q} \right) \quad (6.35)$$

$$\approx \sum_{i=2}^9 \zeta'_{i,Q}, \quad (6.36)$$

where the last approximation is obtained from $\sum_{i=2}^9 \zeta'_{i,Q} < 1$.

By substituting (6.32) and (6.36) into (6.31) and taking the angle of (6.31), we conclude the proof. \square

Proposition 6.2. *The distribution of $\zeta'_{2,Q}, \zeta'_{3,Q}, \dots, \zeta'_{9,Q}$ are*

$$\zeta'_{2,Q} \sim \mathcal{N} \left(0, \frac{\sigma_n^4}{2|\phi|^4 \|\mathbf{g}_k\|^4} \right), \quad (6.37)$$

$$\zeta'_{3,Q}, \zeta'_{4,Q} \sim \mathcal{N} \left(0, \frac{\sigma_n^2}{2|\phi|^2 \|\mathbf{g}_k\|^2} \right), \quad (6.38)$$

$$\zeta'_{5,Q}, \zeta'_{6,Q} \sim \mathcal{N} \left(0, \frac{C_{\text{diag}}}{2a_p^2 |\phi|^2 \|\mathbf{g}_k\|^2} \right), \quad (6.39)$$

$$\zeta'_{7,Q}, \zeta'_{8,Q} \sim \mathcal{N} \left(0, \frac{\sigma_n^2 C_{\text{diag}}}{2a_p^2 |\phi|^4 \|\mathbf{g}_k\|^4} \right), \quad (6.40)$$

and $\zeta'_{9,Q}$ is normal distributed with mean and variance

$$\begin{aligned} \mathbb{E}\{\zeta'_{9,Q} | \mathbf{g}_k\} &= \frac{M}{|\phi|^2 a_p^2 \|\mathbf{g}_k\|^2} \left(\Im(C_{t_2 t_1}) \cos(\omega_k(t_2 - t_1)) \right. \\ &\quad \left. - \Re(C_{t_2 t_1}) \sin(\omega_k(t_2 - t_1)) \right) \end{aligned} \quad (6.41)$$

$$\begin{aligned} \text{Var}\{\zeta'_{9,Q} | \mathbf{g}_k\} &= \frac{M}{2|\phi|^4 a_p^4 \|\mathbf{g}_k\|^4} \left(C_{\text{diag}}^2 \right. \\ &\quad \left. - (\Im(C_{t_2 t_1})^2 - \Re(C_{t_2 t_1})^2) \cos 2\omega_k(t_2 - t_1) \right). \end{aligned} \quad (6.42)$$

All the random variables are mutually uncorrelated except for $\zeta'_{5,Q}$ and $\zeta'_{6,Q}$, of which the covariance is conditioned on \mathbf{g}_k as

$$\begin{aligned} \text{cov}(\zeta'_{5,Q}, \zeta'_{6,Q} | \mathbf{g}_k) &= \frac{\Re(\text{cov}(\zeta'_5, \zeta'_6 | \mathbf{g}_k))}{2} \\ &\quad - \frac{1}{2|\phi|^2 a_p^2 \|\mathbf{g}_k\|^2} (\Re(C_{21}) \cos \omega_k - \Im(C_{21}) \sin \omega_k), \end{aligned} \quad (6.43)$$

where

$$\begin{aligned} &\text{cov}(\zeta'_5, \zeta'_6 | \mathbf{g}_k) \\ &= \mathbb{E}\left\{ \frac{\phi^2}{|\phi|^4 a_p^2 \|\mathbf{g}_k\|^4} \mathbf{q}_2^H \mathbf{g}_k \mathbf{q}_1^H \mathbf{g}_k e^{j\omega_k(t_1+t_2)} | \mathbf{g}_k \right\} \end{aligned} \quad (6.44)$$

has no closed form. It can only be obtained numerically.

Proof. See appendix D. □

From proposition 6.1 and 6.2, we can see that $\angle \frac{1}{M} \mathbf{r}_{pt_1}^H \mathbf{r}_{pt_2}$ contains two parts, desired signal $\omega_k(t_2 - t_1)$ and random phase noise $\sum_{i=2}^9 \zeta'_{i,Q}$. In order to design a CFO estimator, we first eliminate the randomness in the phase noise part by taking the mean over the phase noise. Since the phase noise has a biased mean $\mathbb{E}\{\sum_{i=2}^9 \zeta'_{i,Q}\} = \mathbb{E}\{\zeta'_{9,Q}\}$ which is a non-linear function of ω_k . We define the the mean of $\angle \frac{1}{M} \mathbf{r}_{pt_1}^H \mathbf{r}_{pt_2}$ as the following

function

$$f(\boldsymbol{\omega}_k) = \boldsymbol{\omega}_k(t_2 - t_1) + \mathbb{E}\left\{\sum_{i=2}^9 \zeta'_{i,Q}\right\}. \quad (6.45)$$

At arbitrary time indices t_1, t_2 , we propose a CFO estimator for any user k as

$$\hat{\boldsymbol{\omega}}_k = \frac{f^{-1}\left(\angle \frac{1}{M} \mathbf{r}_{pt_1}^H \mathbf{r}_{pt_2}\right)}{t_2 - t_1}, \quad (6.46)$$

where f^{-1} denotes the inverse of function $f(\boldsymbol{\omega}_k)$.

The proposed estimator will work and perform well under following conditions:

- 1) Channel vector \mathbf{g}_k is fixed within a coherence block and $\|\mathbf{g}_k\|^2$ is a real number. However, at this stage, \mathbf{g}_k has not been estimated. We have to make a design choice to take the expectation over the channel.

Proposition 6.3. *The mean and variance of the phase noise term $\sum_{i=2}^9 \zeta'_{i,Q}$ are respectively*

$$\begin{aligned} \mathbb{E}\left\{\sum_{i=2}^9 \zeta'_{i,Q}\right\} &= \mathbb{E}\{\zeta'_{9,Q}\} \\ &= \frac{1}{|\phi|^2 a_p^2 \beta_k} \left(\cos(\boldsymbol{\omega}_k(t_2 - t_1)) \Im(C_{t_2 t_1}) \right. \\ &\quad \left. - \sin(\boldsymbol{\omega}_k(t_2 - t_1)) \Re(C_{t_2 t_1}) \right), \end{aligned} \quad (6.47)$$

$$\begin{aligned} \text{Var}\left\{\sum_{i=2}^9 \zeta'_{i,Q}\right\} &= \frac{1}{M} \left(\frac{\sigma_n^2}{|\phi|^2 \beta_k} + \frac{C_{\text{diag}}}{a_p^2 |\phi|^2 \beta_k} \right) \\ &+ \frac{1}{M^2} \left(\frac{\sigma_n^4}{2|\phi|^4 \beta_k^2} + \frac{\sigma_n^2 C_{\text{diag}}}{a_p^2 |\phi|^4 \beta_k^2} \right. \\ &\quad \left. + \frac{M}{2|\phi|^4 a_p^4 \beta_k^2} \left(C_{\text{diag}}^2 - (\Im(C_{21}))^2 - \Re(C_{21})^2 \right) \cos 2\boldsymbol{\omega}_k \right) \\ &- \frac{1}{2|\phi|^2 a_p^2 \beta_k M} \left(\Re(C_{21}) \cos \boldsymbol{\omega}_k - \Im(C_{21}) \sin \boldsymbol{\omega}_k \right), \end{aligned} \quad (6.48)$$

where the expectation is taken over \mathbf{g}_k , \mathbf{n} and \mathbf{q} .

Proof. See appendix F. □

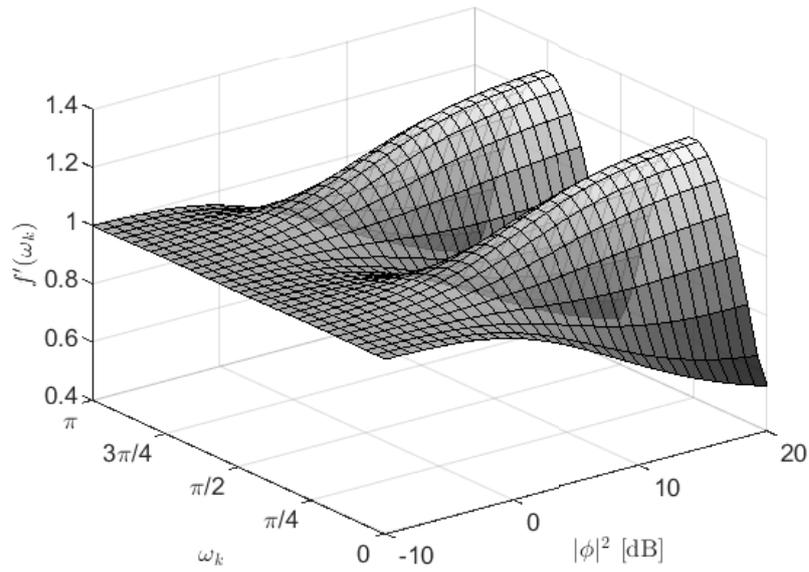


Fig. 6.2 Derivative of $f'(\omega_k)$ for different pilot power.

When the number of antennas approaches infinity, the variance of the phase noise can be ignored. However, as the number of antennas is large but finite, the variance of the phase noise will affect the mean square error (MSE) of the CFO estimation.

- 2) To show the inverse function $f^{-1}(\omega_k)$ exists, we numerically find the first derivative of $f(\omega_k)$, which is always positive within certain parameter range, as shown in Fig. 6.2. It means that $f(\omega_k)$ is monotonic and has an inverse function $f^{-1}(\omega_k)$. Due to the complex structure of $f(\omega_k)$, finding the closed form of the inverse function is no easy task. We can always find the inverse function numerically.

6.3.2 Method 2: Infinite Number of Antennas Assumption

We now consider a special case when the number of antennas at the BS is infinite. Different than the lemma D.1 in appendix D, channel vector \mathbf{g} is considered random due to the fact that M is assumed infinite. It is independent with the white and

quantization noise. According to law of large numbers [106], from (6.15) we can obtain

$$\begin{aligned}
\lim_{M \rightarrow \infty} \frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2] &= \phi_{t_2} \phi_{t_1}^* e^{j\omega_k(t_2-t_1)} \mathbf{E}\{\zeta_1\} + \mathbf{E}\left\{\sum_{i=2}^9 \zeta_i\right\} \\
&= \frac{2}{\pi} \frac{\beta \phi_1^* \phi_2 e^{j\omega_k(t_2-t_1)}}{\sqrt{(\beta|\phi_2|^2 + \sigma_n^2)(\beta|\phi_1|^2 + \sigma_n^2)}} \\
&\quad + \frac{2}{\pi} \arcsin\left(\frac{\beta \Re(\phi_{t_2} \phi_{t_1}^* e^{j\omega_k(t_2-t_1)})}{\sqrt{(\beta|\phi_{t_2}|^2 + \sigma_n^2)(\beta|\phi_{t_1}|^2 + \sigma_n^2)}}\right) \\
&\quad + j \frac{2}{\pi} \arcsin\left(\frac{\beta \Im(\phi_{t_2} \phi_{t_1}^* e^{j\omega_k(t_2-t_1)})}{\sqrt{(\beta|\phi_{t_2}|^2 + \sigma_n^2)(\beta|\phi_{t_1}|^2 + \sigma_n^2)}}\right) \\
&\quad - \frac{2}{\pi} \frac{\beta \Re(\phi_{t_2} \phi_{t_1}^* e^{j\omega_k(t_2-t_1)})}{\sqrt{(\beta|\phi_{t_2}|^2 + \sigma_n^2)(\beta|\phi_{t_1}|^2 + \sigma_n^2)}} \\
&\quad - j \frac{2}{\pi} \frac{\beta \Im(\phi_{t_2} \phi_{t_1}^* e^{j\omega_k(t_2-t_1)})}{\sqrt{(\beta|\phi_{t_2}|^2 + \sigma_n^2)(\beta|\phi_{t_1}|^2 + \sigma_n^2)}}. \tag{6.49}
\end{aligned}$$

Similar to method 1, we make a design choice that pilot symbols for user k are equivalent and adjacent in time, i.e., $\phi_1 = \phi_2 = \phi$, $t_2 = t_1 + 1$, yielding

$$\begin{aligned}
\lim_{M \rightarrow \infty} \frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2] \\
= \frac{2}{\pi} \arcsin\left(\frac{\beta|\phi|^2}{\beta|\phi|^2 + \sigma_n^2} \cos \omega_k\right) + j \frac{2}{\pi} \arcsin\left(\frac{\beta|\phi|^2}{\beta|\phi|^2 + \sigma_n^2} \sin \omega_k\right). \tag{6.50}
\end{aligned}$$

For infinite M , the inverse function of (6.50) would thus be

$$\begin{aligned}
\omega_k = \angle \left(\frac{\beta|\phi|^2 + \sigma_n^2}{\beta|\phi|^2} \left(\sin\left(\frac{\pi}{2} \Re\left(\lim_{M \rightarrow \infty} \frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2]\right)\right) \right. \right. \\
\left. \left. + j \sin\left(\frac{\pi}{2} \Im\left(\lim_{M \rightarrow \infty} \frac{1}{M} \mathbf{r}_p^H[t_1] \mathbf{r}_p[t_2]\right)\right) \right) \right). \tag{6.51}
\end{aligned}$$

Although infinite number of antennas would not be realistic, above equation is relatively accurate as long as M is large enough. A straight forward CFO estimator for

finite M is proposed as

$$\hat{\omega}_k = \angle \left(\frac{\beta|\phi|^2 + \sigma_n^2}{\beta|\phi|^2} \left(\sin \left(\frac{\pi}{2} \Re \left(\frac{1}{M} \mathbf{r}_p^H [t_1] \mathbf{r}_p [t_2] \right) \right) + j \sin \left(\frac{\pi}{2} \Im \left(\frac{1}{M} \mathbf{r}_p^H [t_1] \mathbf{r}_p [t_2] \right) \right) \right) \right). \quad (6.52)$$

Theoretically, unlike method 1, the only requirement of method 2 is that number of antennas M is large. However, due to the non-linear property introduced by the inverse trigonometric functions, performance analysis of method 2 is non-trivial. We will show some numerical results in section 6.5.

6.3.3 Method 3: Small CFO Assumption

We now explore the extreme case where ω_k is near zero. Since $\lim_{\omega_k \rightarrow 0} E\{\zeta'_{9,Q}\} = 0$, we can neglect the second term on the right hand side of (6.45). A simple estimator for very small CFO is then designed by

$$\hat{\omega}_k = \frac{\angle \frac{1}{M} \mathbf{r}_{pt_1}^H \mathbf{r}_{pt_2}}{t_2 - t_1}. \quad (6.53)$$

6.3.4 CFO Update

In principle, above methods are for arbitrary time indices t_1 and t_2 , which enables the CFO update at the beginning of next coherence block. To be specific, as shown in figure 6.1, CFO can be updated using $\mathbf{r}_p [t'_1]$ and $\mathbf{r}_p [t'_2]$ where $t'_1 \in [1, K]$ and $t'_1 \in [T + 1, T + K]$, respectively. Since the CFO update and channel estimation share the same pilot sequences, it does not generate redundancy.

6.3.5 Multiple Pilots Extension

If a user is assigned multiple pilot symbols for CFO estimation, we can still apply the proposed methods for adjacent pilot symbols to get multiple CFO estimation samples. Since Bussgang decomposition guarantees only the first two moments of statistics of the quantization noise, it is not possible to effectively combine these samples via analytically exploring the correlations among them as in [105], which requires the third moment of statistics of the quantization noise. We can, however, always simply average over the samples.

Specifically, we assume the length of pilots for estimating the CFO of user k is τ_p and with some abuse of the time instances, the received pilots from user k are denoted by $\mathbf{r}_p[1], \mathbf{r}_p[2], \dots, \mathbf{r}_p[\tau_p]$. We take “unbiased” average correlation of each adjacent pilot pairs $\frac{1}{M}\mathbf{r}_p^H[1]\mathbf{r}_p[2], \frac{1}{M}\mathbf{r}_p^H[2]\mathbf{r}_p[3], \dots, \frac{1}{M}\mathbf{r}_p^H[\tau_p-1]\mathbf{r}_p[\tau_p]$ as $\tau_p - 1$ samples. For t th sample, the estimate $\hat{\omega}_k^t$ can be obtained using above mentioned methods. Finally, the CFO estimates $\hat{\omega}_k^t$ from τ_p pilot symbols can be obtained by averaging over all the samples, i.e.,

$$\hat{\omega}_k = \frac{\sum_{t=1}^{\tau_p-1} \hat{\omega}_k^t}{\tau_p - 1}. \quad (6.54)$$

6.4 CFO Compensation and Channel Estimation

First K pilot symbols at the beginning of each block are re-used for channel estimation. Although the k th user can occupy any time indice to transmit its pilot symbol for CFO estimation, we assume it always occupies the k th one for notational simplicity. The pilot sequences of all users are thus a diagonal matrix $\Phi_c = \sqrt{K}\phi\mathbf{I}_K$, yielding the received signal with CFO at the BS

$$\mathbf{Y}_c = \mathbf{G}\Phi_c\mathbf{\Omega}_c + \mathbf{N}, \quad (6.55)$$

where the channel matrix is $\mathbf{G} = (\mathbf{g}_1, \dots, \mathbf{g}_K) \in \mathbb{C}^{M \times K}$ and the CFO is denoted by $\mathbf{\Omega}_c = \text{diag}(e^{j\omega_1}, e^{j2\omega_2}, \dots, e^{j\omega_K K})$. The vectorized received signal is given by,

$$\mathbf{y}_c = \text{vec}(\mathbf{Y}_c) = \mathbf{\Omega}_c \otimes \mathbf{I}_M \bar{\mathbf{\Phi}}_c \mathbf{g}_{\text{vec}} + \mathbf{n}_{\text{vec}},$$

where $\mathbf{n}_{\text{vec}} = \text{vec}(\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_K)$, $\mathbf{g}_{\text{vec}} = \text{vec}(\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_K)$ and $\bar{\mathbf{\Phi}}_c = \mathbf{\Phi}_c \otimes \mathbf{I}_M$. In this case the covariance matrix of \mathbf{y}_c is

$$\mathbf{C}_{\mathbf{y}_c} = \text{diag}(\beta_1 K |\phi|^2 + \sigma_n^2, \beta_2 K |\phi|^2 + \sigma_n^2, \dots, \beta_K K |\phi|^2 + \sigma_n^2) \otimes \mathbf{I}_M.$$

Digital CFO compensation is performed on the quantized pilots, by using the estimated CFO $\hat{\mathbf{\Omega}}_c = \text{diag}(e^{-j\hat{\omega}_1}, e^{-j2\hat{\omega}_2}, \dots, e^{-j\hat{\omega}_K K})$, yielding

$$\mathbf{r}_c = \hat{\mathbf{\Omega}}_c \otimes \mathbf{I}_M \mathcal{Q}(\mathbf{y}_c) \quad (6.56)$$

By using the Bussgang decomposition [78], we obtain the quantized vector as

$$\mathbf{r}_c = \mathbf{A}_c \bar{\mathbf{\Phi}}_c \mathbf{h}_{\text{vec}} + \mathbf{A}_c \mathbf{n}_{\text{vec}} + \mathbf{q}_c. \quad (6.57)$$

According to Bussgang decomposition, $\mathbf{A}_c = \mathbf{a}_c \otimes \mathbf{I}_M$, where \mathbf{a}_c is a $K \times K$ diagonal matrix with $\sqrt{\frac{2}{\pi(\beta_k K |\phi|^2 + \sigma_n^2)}}$ on k th diagonal. The covariance matrix of \mathbf{q}_c is $\mathbf{C}_{\mathbf{q}_c} = (1 - \frac{2}{\pi}) \mathbf{I}_{MK}$. We denote the residual CFO contaminated vectorized channel as

$$\begin{aligned} \mathbf{h}_{\text{vec}} &= \text{vec}(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K) \\ &= \text{vec}(e^{j\hat{\omega}_1} \mathbf{g}_1, e^{j2\hat{\omega}_2} \mathbf{g}_2, \dots, e^{j\hat{\omega}_K K} \mathbf{g}_K). \end{aligned} \quad (6.58)$$

Note that given the knowledge of pilot sequences at BS, we can only estimate the contaminated channel vector \mathbf{h}_{vec} . Since \mathbf{q}_c is uncorrelated with \mathbf{h}_{vec} [79], we can

obtain the covariance matrix $\mathbf{C}_{r_c} = \mathbf{I}_{MK}$. The Busgang LMMSE channel estimator for \mathbf{h}_{vec} is given by [79]

$$\hat{\mathbf{h}}_{\text{vec}} = \text{vec}(\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \dots, \hat{\mathbf{h}}_K) = \mathbf{C}_{\mathbf{h}_{\text{vec}}} \mathbf{A}_c \bar{\Phi}_c^H \mathbf{C}_{r_c}^{-1} \mathbf{r}_c, \quad (6.59)$$

where $\mathbf{C}_{\mathbf{h}_{\text{vec}}} = \text{diag}(\beta_1, \beta_2, \dots, \beta_K) \otimes \mathbf{I}_M$ denotes the covariance matrix of the vectorized channel matrix \mathbf{h}_{vec} . The property of LMMSE estimator guarantees that estimation error $\tilde{\mathbf{h}}_{\text{vec}} = \mathbf{h}_{\text{vec}} - \hat{\mathbf{h}}_{\text{vec}}$ is uncorrelated with $\hat{\mathbf{h}}$ and has zero mean. We obtain the covariance matrix of the estimation error as

$$\mathbb{E}\{\tilde{\mathbf{h}}_{\text{vec}} \tilde{\mathbf{h}}_{\text{vec}}^H\} = \mathbb{E}\{(\mathbf{h}_{\text{vec}} - \hat{\mathbf{h}}_{\text{vec}})(\mathbf{h}_{\text{vec}} - \hat{\mathbf{h}}_{\text{vec}})^H\} \quad (6.60)$$

$$= \mathbf{C}_{\mathbf{h}_{\text{vec}}} - \mathbf{C}_{\mathbf{h}_{\text{vec}}} \mathbf{A}_c \bar{\Phi}_c \bar{\Phi}_c^H \mathbf{A}_c \mathbf{C}_{\mathbf{h}_{\text{vec}}}, \quad (6.61)$$

which is a diagonal matrix, indicating no cross-correlation among antennas. Therefore the co-variance matrix of estimates and the MSE of the estimation for k th user-to-BS channel are respectively

$$\mathbb{E}\{\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H\} = \frac{2\beta_k^2 K |\phi|^2}{\pi(\beta_k K |\phi|^2 + \sigma_n^2)} \mathbf{I}_M = \hat{\beta}_k \mathbf{I}_M, \quad (6.62)$$

$$\mathbb{E}\{\tilde{\mathbf{h}}_k \tilde{\mathbf{h}}_k^H\} = \frac{(\pi - 2)\beta_k^2 K |\phi|^2 + \pi\sigma_n^2 \beta_k}{\pi(\beta_k K |\phi|^2 + \sigma_n^2)} \mathbf{I}_M = \tilde{\beta}_k \mathbf{I}_M. \quad (6.63)$$

It indicates that the power of both channel estimates and estimation error do not depend on residual CFO, since CFO only causes a phase rotation of the signal and does not affect the power.

6.5 Numerical Results

In this section, the properties of proposed three CFO estimation methods are shown. Unless stated otherwise, we assume that $M = 256$, $K = 10$ and the coherence time

$T = 100$. For any user k , the large scale fading factor is normalized, i.e. $\beta_k = 1$. The transmit power of pilot symbols is K times bigger than the transmit power of data symbols and equal power for all users are assumed $\rho = \rho_k = \frac{|\phi_k|^2}{K}, \forall k$. The power of additive Gaussian white noise is normalized to $\sigma_n^2 = 1$. Note that although the proposed methods are designed from Bussgang model, the simulations are based on one-bit quantization in (6.2) for precise results.

Fig. 6.3 depicts the mean-square-error (MSE) of the CFO estimation method 1 as a function of the CFO ω_k , based on two pilot symbols. The dotted, solid and dashed curves represent $\rho = 0$ dB, 5 dB and 10 dB, respectively. Using high transmit power ρ results in better overall performance and enlarged variation of MSE. The better MSE is observed in small ω_k region. It is obvious that the MSE is a cyclic function with respect to CFO ω_k , which has the period $\frac{\pi}{2}$. We are not able to obtain the closed form of MSE since method 1 is built on the inverse function (6.46). We can, however, get an implication from the periodic property of (6.48), which is caused by the trigonometric functions. It also proves that the inversion in (6.46) does not affect the periodicity of the function. Besides MSE, we also investigate the distribution of the estimation error in Fig. 6.4, which is from a point in Fig. 6.3, where $\omega_k = \frac{\pi}{8}$ and $\rho = 10$ dB. The PDF of $\tilde{\omega}_k$ is well approximated by Gaussian distribution, which is depicted in solid line. It implies that our the inverse function does not change the distribution of the noise terms. Albeit the mean of $\tilde{\omega}_k$ is not zero, it is very small and the variance is decaying exponentially.

In figure 6.5, we compare the MSE of three proposed methods as function of CFO ω_k . The cross, square and circle markers notate method 1, 2 and 3 respectively. The three methods are showing periodicity as the statistics of ζ_9 in (6.24) are trigonometric. Method 3 performs best among three methods in very low ω_k region, while the performance decays rapidly with the growth of ω_k . Method 2 outperforms method 1 in low to moderate ω_k region ($\omega \in (0, \frac{\pi}{4})$). This is due to that method 1 requires a strong

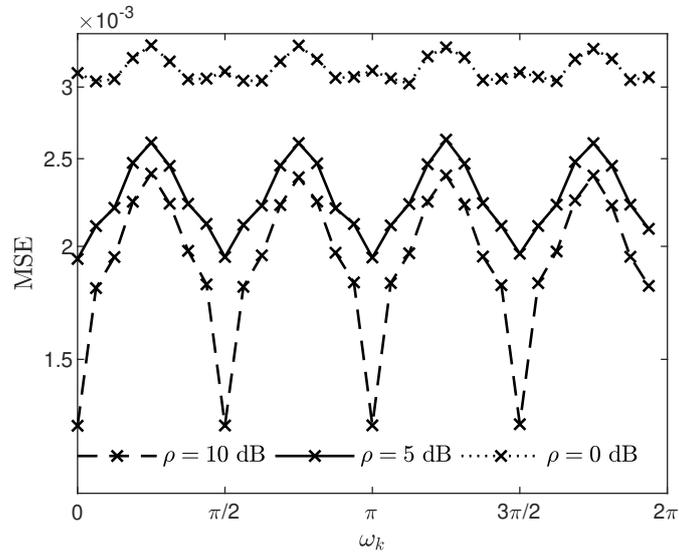


Fig. 6.3 Mean square error (MSE) of method 1 vs. ω_k for different pilot power ($\tau_p = 2$).

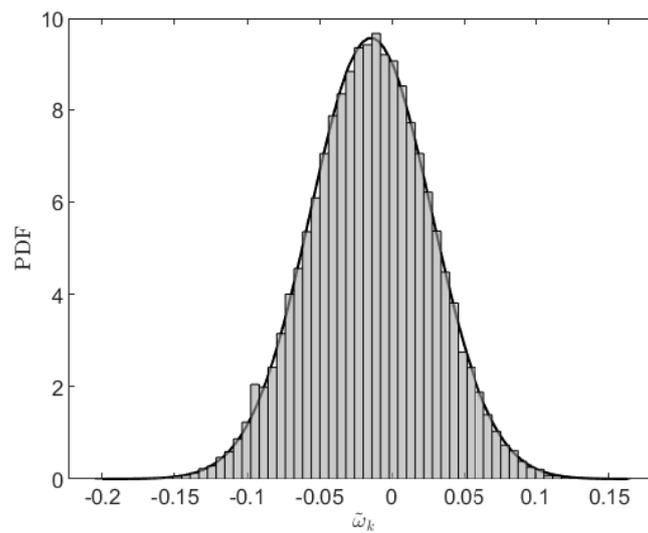


Fig. 6.4 Empirical PDF of $\tilde{\omega}_k$ in method 1 ($\omega_k = \frac{\pi}{8}$, $\rho = 10$ dB, $\tau_p = 2$).

assumption that $\zeta'_{2,Q}, \dots, \zeta'_{2,Q}$ should be small enough to be accurate, which does not always hold in this region.

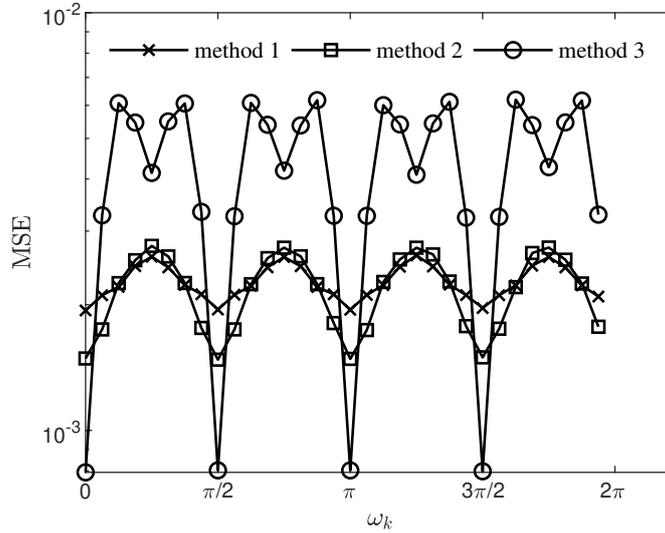


Fig. 6.5 MSE of CFO estimation vs. ω_k for different CFO estimation methods ($\tau_p = 2$, $\rho = 5$ dB).

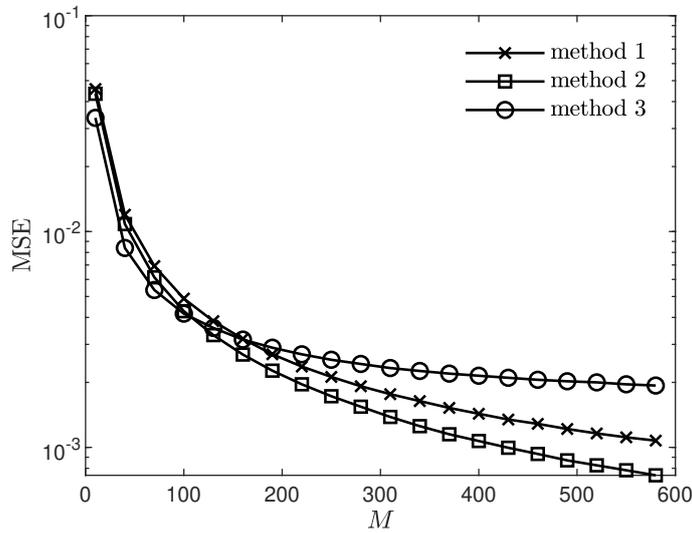


Fig. 6.6 MSE of CFO estimation vs. M for different CFO estimation methods ($\tau_p = 2$, $\rho = 5$ dB, $\omega_k = \frac{\pi}{20}$).

Fig. 6.6 describes the MSE three methods with increasing number of antennas M equipped at the BS, where the method 1, 2 and 3, depicted in cross, square and circle, use two pilot symbols to estimating CFO $\omega_k = \frac{\pi}{20}$. The MSE of all methods decrease

with the growth of M , since larger M provides more samples for the “unbiased” average correlation in (6.15) and reduce the variance of the phase noise terms. For small to moderate number of antennas ranges $M < 80$, method 3 has the best performance among the three, since it does not require large M assumption. Method 1 and 2, however, decrease faster in large M ranges.

Fig. 6.7 and Fig. 6.8 compare the MSE of three CFO estimation methods with 2 and 24 pilot symbols respectively. All of the curves start to flatten out at moderate to high ρ region. In Fig. 6.8, the curves for method 1 and 3 even show an increasing trend. It is because the power of the quantization noise is proportional to the received pilot power. Method 3 has the worst performance at high ρ region since it starts to break the $\lim_{\omega_k \rightarrow 0} E\{\zeta'_{9,Q}\} = 0$ assumption. In both figures, method 2 outperforms method 1 since the conversion of the noise terms ζ_2, \dots, ζ_9 to phase noise terms in proposition 6.1 brings inaccuracy.

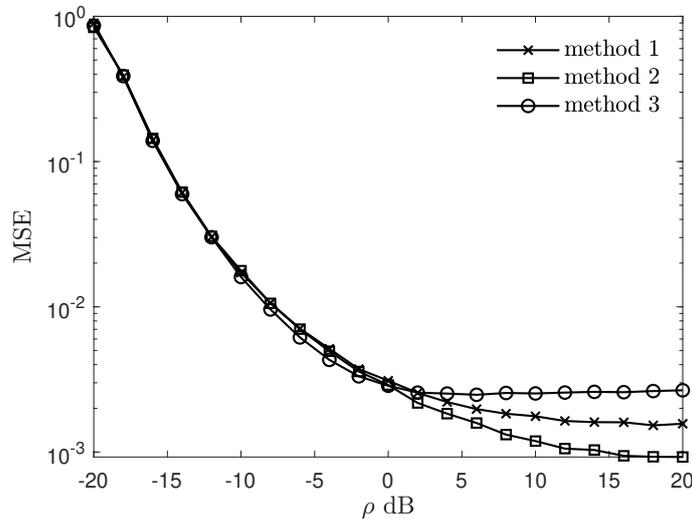


Fig. 6.7 MSE of CFO estimation vs. pilot power for different CFO estimation methods ($\tau_p = 2$, $\omega_k = \frac{\pi}{20}$).

Optimal pilot length for estimating CFO $\omega_k = \frac{\pi}{20}$ for $\rho = 0$ dB is investigated in 6.9. All the curves are showing the flatten-out trend, since there are correlations among all received pilot symbols due to the correlated quantization noise through time. The

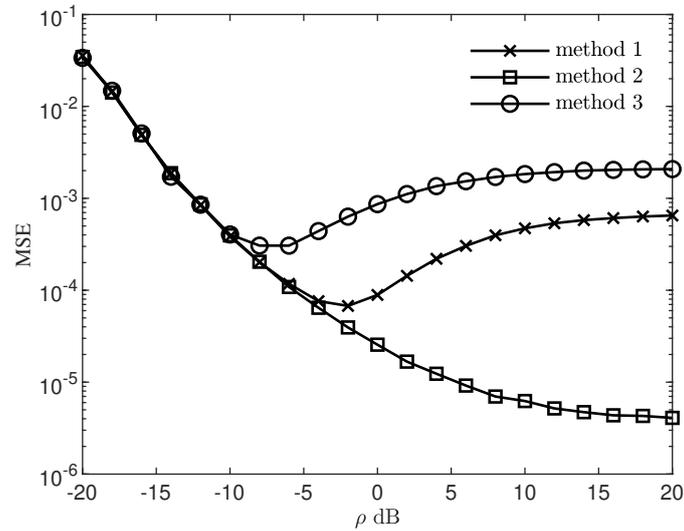


Fig. 6.8 MSE of CFO estimation vs. pilot power for different multiple-pilot CFO estimation methods ($\tau_p = 24$, $\omega_k = \frac{\pi}{20}$).

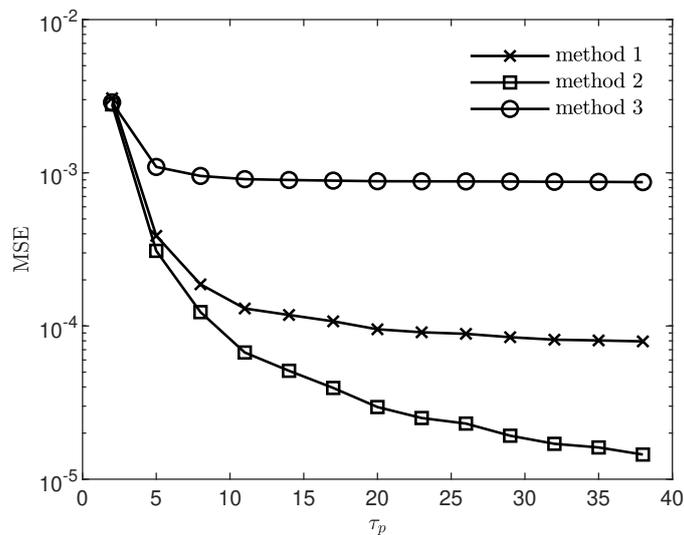


Fig. 6.9 MSE of CFO estimation vs. pilot length τ_p for different CFO estimation methods ($\rho = 0$ dB, $\omega_k = \frac{\pi}{20}$).

correlations, however, cannot be obtained from Bussgang model nor any existing model, which result in the bottleneck for the CFO estimation methods. Around 20 to 25 pilots should be used for method 1. As many as possible pilot symbols should be sent in method 2 for high accuracy estimation, although it might not be spectral efficient. Since method 3 neglect the mean of the phase noise terms, which is a function of the CFO, using more than 10 pilot symbols does not benefit the performance.

6.6 Conclusion

In this chapter, we proposed three CFO estimation methods using only two pilot symbols for an individual user in an uplink massive MIMO system with one-bit ADCs. Method 1 and 2 utilize the statistics of the phase noise terms for scenarios: a large but finite number of antennas and infinite number of antennas at the receivers respectively, while method 3 estimates the very small CFO. These estimation methods are extended to multiple pilot symbol scenario. Channel estimation is conducted using the same pilot sequences with CFO estimation and proved not depending on the residual CFO. Numerical results show that MSE is a cyclic function of the CFO for all three methods, and increasing the power of the pilot enlarges the variation of the MSE. Method 3 outperforms method 1 and 2 while small number of antennas are equipped at the BS, however, it does not benefit from large pilot power and long pilot sequences. While using multiple pilots, large pilot power is recommended for method 2 and medium pilot power benefits method 1.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

The thesis starts with a literature review of physical layer techniques in massive MIMO and FD relaying systems. The importance of studying interference mitigation and frequency synchronization problems in such systems especially with hardware impairments are emphasised.

A large number of antennas can be implement at any node in a FD relaying network. We investigate two FD relaying models, where the first one has large number of antennas deployed at both source and destination and several relaying nodes have just two antennas each, while the second model is a multi-pair system with large number of antennas implemented at one FD relaying node. LI cancellation schemes are proposed and achievable rate analysis for ZF and MF processing is discussed in details for two models.

Specifically, in the first model, co-located, distributed cooperative and distributed non-cooperative FD relaying schemes are investigated. The asymptotic analysis is carried out, showing the system is interference free as the number of antennas at

the source and destination grows without bound for three schemes with both MF and ZF processing. For large but finite number of antennas deployed at the source and destination, pilot-based LI cancellation for three relaying schemes are proposed. The simulation results demonstrate that FD operation at the relaying can provide significant achievable rate gain over HD with a smaller antenna configuration. Co-located and distributed cooperative relaying schemes benefit from instant inter-node information exchange, which would lead to better LI cancellation. Since the distributed non-cooperative FD relaying nodes cannot exchange information, only the LI from a relaying node itself can be cancelled. It implies that the distributed FD relaying nodes should be spaced far apart or blocked by large obstacles to obtain better performance.

In the second model, the FD relaying are deployed with large but finite number of antennas and is used to assist multi-pair user signal transmission. A pilot protocol enabling the LI channel estimation are proposed by exploiting the coherence time difference between static LI channel and mobile user-relaying channels. A power allocation scheme that adjust the per-link power at the FD relay is proposed to maximize the overall E2E achievable rate. It is solved by a low-complexity iterative algorithm that uses only channel statistics. The algorithm converges fast, typically within 10 to 20 iterations for ZF and MF processing, respectively. The power allocation scheme results in a Gaussian distributed total transmit power at the FD relay that does not violate the power constraint. The numerical results show that the proposed power allocation scheme improves the SE and EE of the FD relaying system. It outperforms the FD relaying with fixed relay power or the HD relaying significantly under practical system parameters.

The second model is then expanded to a massive MIMO FD relaying that adapts low-resolution ADCs at the receiver for circuit power saving. We use the additive CSCG quantization noise model for channel estimation, data transmission modelling and achievable rate analysis. Similar to the FD relaying with full-resolution ADCs,

the power allocation scheme that considers the quantization noise and other hardware imperfection is proposed. It can also be solved using iterative algorithm which is low-complexity and converges fast. EE of the system is analyzed considering the circuit energy consumption. It is shown that the optimum resolution of ADCs are 4 to 7 bits for maximizing the EE for both HD and FD relaying. The EE of FD relaying outperforms HD relaying no matter whether the power allocation scheme is used. The delay constrained achievable rate for the system is calculated adhering to the ultra-reliable and low-latency communication (URLLC) principle. The results show that the FD relaying with very strict delay constraint is superior than HD relaying without delay constraint, making the proposed FD relaying suitable for URLLC applications.

The frequency synchronization problem is finally studied for massive MIMO system with one-bit ADCs. We proposed three CFO estimation methods for individual users by using just two pilot symbols. In method 1 of CFO estimation, we start from the “unbiased” correlation of two received pilot symbols. By taking the angle of the “unbiased” correlation, a linear model consists of CFO and eight additive phase noise terms is obtained. We analyze the noise terms with the assumption that the number of antenna is large but finite and the in-phase of these noise terms are negligible. Based on the statistics of the noise terms, we developed the CFO estimation method 1. The statistics of the noise terms give us a hint on the parameters that would affect the mean-square-error of the estimation. Method 2 and 3 are formulated by assuming that the number of antennas at the BS is infinite and the CFO is around zero, respectively. The three methods are expanded to the scenario where multiple pilot symbols are used for individual users. Linear-minimum-mean-square-error channel estimation is conducted using the same pilot sequences with CFO estimation. It is proved not affected by CFO. Using the pilot sequences for channel estimation from two different coherence blocks, CFO can be updated without unnecessary redundancy. Based on the estimated CFO, we conduct CFO cancellation during data transmission phase.

Numerical results show that MSE is a cyclic function of the CFO for all three methods, and increasing the power of the pilot enlarges the variation of the MSE. With as little as 24 pilots, the proposed CFO estimator can reach 10^{-5} in terms of MSE.

7.2 Future Work

Some topics in the thesis could be extended in the future work.

- **Joint source-and-relay power allocation** In a system as described in Chapter 3, a joint power allocation scheme could be proposed to maximize the E2E achievable rate via adjusting the power of the source and the FD relay nodes. This could be formed as an optimization problem given the power constraint.
- **Millimeter wave extension** As mmWave is a potential solution to solve the inter-cell interference in pico, femto-cellular systems, the research on massive MIMO FD relaying systems could be extended to adapt mmWave frequency communication systems. The main difference will be: 1) Fading model is changed. 2) Precoding techniques in mmWave MIMO system are usually hybrid, taking LOS propagation into account, since the degree of freedom is often smaller. A novel hybrid precoding scheme could be proposed for aforementioned FD relaying systems.
- **OFDM extension** OFDM systems require the subcarriers to be strictly orthogonal. It is very sensitive to CFO. The proposed CFO estimators in Chapter 6 could be extended to adapt OFDM systems.

References

- [1] C. V. N. Index, “Cisco annual internet report (2018–2023) white paper,” 2020.
- [2] H. Song, X. Fang, L. Yan, and Y. Fang, “Control/user plane decoupled architecture utilizing unlicensed bands in lte systems,” *IEEE Wireless Communications*, vol. 24, no. 5, pp. 132–142, 2017.
- [3] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb 2005.
- [4] H. Song, X. Fang, and Y. Fang, “Unlicensed spectra fusion and interference coordination for lte systems,” *IEEE Transactions on Mobile Computing*, vol. 15, no. 12, pp. 3171–3184, 2016.
- [5] J. Rodriguez, *Fundamentals of 5G mobile networks*. John Wiley & Sons, 2015.
- [6] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, “What will 5G be?” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [7] S. Hong, J. Brand, J. I. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, “Applications of self-interference cancellation in 5G and beyond,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, February 2014.
- [8] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MultiUser MIMO*. John Wiley & Sons, 2010, pp. 395–417.
- [9] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.
- [11] 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, Jan. 2013.
- [12] H. Q. Ngo *et al.*, “Multipair full-duplex relaying with massive arrays and linear processing,” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1721–1737, 2014.

- [13] W. Zhang, F. Gao, S. Jin, and H. Lin, "Frequency synchronization for uplink massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 235–249, 2018.
- [14] P. Sabeti, A. Farhang, N. Marchetti, and L. Doyle, "Frequency synchronization for OFDM-based massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 2973–2986, 2019.
- [15] 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.
- [16] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 303–313, February 2013.
- [17] 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.
- [18] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [19] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part II: perturbation," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 537–544, March 2005.
- [20] C. Masouros, M. Sellahurai, and T. Ratnarajah, "Bridging the gap between linear and non-linear precoding in small- and large-scale MIMO downlinks," in *IEEE International Conference on Communications (ICC)*, June 2014, pp. 4483–4487.
- [21] 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.
- [22] B. Li, C. Z. Wu, H. H. Dam, A. Cantoni, and K. L. Teo, "A parallel low complexity zero-forcing beamformer design for multiuser MIMO systems via a regularized dual decomposition method," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4179–4190, Aug. 2015.
- [23] H. H. Dam, A. Cantoni, and B. Li, "A fast low complexity method for optimal zero-forcing beamformer MU-MIMO system," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1443–1447, Sept. 2015.
- [24] C. S. Park, Y. S. Byun, A. M. Bokiye, and Y. H. Lee, "Complexity reduced zero-forcing beamforming in massive MIMO systems," in *Information Theory and Applications Workshop (ITA)*, Feb. 2014, pp. 1–5.
- [25] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for ofdm signals," *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 257–268, 2008.

- [26] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 1059–1071, March 2013.
- [27] S. Mukherjee and S. K. Mohammed, "Constant-envelope precoding with time-variation constraint on the transmitted phase angles," *IEEE Wireless Communications Letters*, vol. 4, no. 2, pp. 221–224, April 2015.
- [28] S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity precoder for large multiuser MISO systems," in *VTC Spring 2008 - IEEE Vehicular Technology Conference*, May 2008, pp. 797–801.
- [29] S. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. PTR Prentice-Hall, 1993.
- [30] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, February 2016.
- [31] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.
- [32] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, Sept. 2014.
- [33] T. Zahir, K. Arshad, A. Nakata, and K. Moessner, "Interference management in femtocells," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 293–311, 2013.
- [34] A. Al-Dulaimi, X. Wang, and C. L. I, *Full-Duplex System Design for 5G Access*, 2018, pp. 77–134.
- [35] F. L. Luo and C. Zhang, *Design and Implementation of Full-duplex Transceivers*, 2016, pp. 402–428.
- [36] D. Bharadia, E. McMillin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*. ACM, 2013, pp. 375–386.
- [37] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Conference on Signals, Systems and Computers*, Nov. 2010, pp. 1558–1562.
- [38] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 680–694, Feb. 2014.
- [39] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, X. Zhang, and S. Barghi, "The case for antenna cancellation for scalable full-duplex wireless communications," in *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*. ACM, 2011, p. 17.
- [40] S. C. Cripps, "RF power amplifiers for wireless communications," 2006.

- [41] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3519–3532, 2015.
- [42] S. Jacobsson, C. Lindquist, G. Durisi, T. Eriksson, and C. Studer, "Timing and frequency synchronization for 1-bit massive MU-MIMO-OFDM downlink," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, 2019, pp. 1–5.
- [43] E. Antonio-Rodríguez, R. López-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Sinr optimization in wideband full-duplex MIMO relays under limited dynamic range," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, June 2014, pp. 177–180.
- [44] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Beamforming design for full-duplex mimo interference channels—QoS and energy-efficiency considerations," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4635–4651, 2016.
- [45] G. Zheng, "Joint beamforming optimization and power control for full-duplex MIMO two-way relay channel," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 555–566, Feb 2015.
- [46] D. G. Wilson-Nunn, A. Chaaban, A. Sezgin, and M. S. Alouini, "Antenna selection for full-duplex MIMO two-way communication systems," *IEEE Communications Letters*, vol. 21, no. 6, pp. 1373–1376, June 2017.
- [47] A. Shojaeifard, K. K. Wong, M. D. Renzo, K. A. Hamdi, and J. Tang, "Design and analysis of full-duplex massive MIMO cellular networks," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2016, pp. 1–6.
- [48] J. Lee, K. J. Choi, and K. S. Kim, "Massive MIMO full-duplex for high-efficiency next generation WLAN systems," in *International Conference on Information and Communication Technology Convergence (ICTC)*, Oct 2016, pp. 1152–1154.
- [49] A. F. Molisch, *Relaying, MultiHop, and Cooperative Communications*. Wiley-IEEE Press, 2011, pp. 521–563.
- [50] V. R. Cadambe and S. A. Jafar, "Can feedback, cooperation, relays and full duplex operation increase the degrees of freedom of wireless networks?" in *IEEE International Symposium on Information Theory*, July. 2008, pp. 1263–1267.
- [51] S. Simoens, O. Munoz-Medina, J. Vidal, and A. del Coso, "On the Gaussian MIMO relay channel with full channel state information," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3588–3599, Sept. 2009.
- [52] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," in *Conference on Signals, Systems and Computers*, Nov. 2012, pp. 1290–1294.

- [53] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3074–3085, Sept. 2011.
- [54] X. Xu, X. Chen, M. Zhao, S. Zhou, C. Y. Chi, and J. Wang, "Power-efficient distributed beamforming for full-duplex MIMO relaying networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1087–1103, Feb. 2017.
- [55] M. Wu, W. Xue, D. Wübben, A. Dekorsy, and S. Paul, "An improved inter-relay cooperation scheme for distributed relaying networks," in *IEEE International Workshop on Smart Antennas*, Mar. 2012, pp. 62–69.
- [56] L. Ahumada, R. Feick, R. A. Valenzuela, and C. Morales, "Measurement and characterization of the temporal behavior of fixed wireless links," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 6, pp. 1913–1922, Nov. 2005.
- [57] X. Xiong, X. Wang, T. Riihonen, and X. You, "Channel estimation for full-duplex relay systems with large-scale antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6925–6938, Oct 2016.
- [58] W. Liu, A. Zappone, C. Yang, and E. Jorswieck, "Global EE optimization of massive MIMO systems," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, June 2015, pp. 221–225.
- [59] M. Vehkaperä, T. Riihonen, R. Wichman, and B. Xu, "Power allocation for balancing the effects of channel estimation error and pilot overhead in full-duplex decode-and-forward relaying," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, July 2016, pp. 1–5.
- [60] H. Song, X. Fang, and C. Wang, "Cost-reliability tradeoff in licensed and unlicensed spectra interoperable networks with guaranteed user data rate requirements," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 200–214, 2017.
- [61] A. Papazafeiropoulos and T. Ratnarajah, "Rate-splitting robustness in multi-pair massive MIMO relay systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5623–5636, 2018.
- [62] A. Papazafeiropoulos, B. Clerckx, and T. Ratnarajah, "Rate-splitting to mitigate residual transceiver hardware impairments in massive mimo systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8196–8211, 2017.
- [63] X. Xia *et al.*, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6565–6580, 2015.
- [64] E. Björnson *et al.*, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

- [65] K. Roth, H. Pirzadeh, A. L. Swindlehurst, and J. A. Nossek, "A comparison of hybrid beamforming and digital beamforming with low-resolution ADCs for multiple users and imperfect CSI," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 3, pp. 484–498, 2018.
- [66] J. Jussila and K. Halonen, "Minimization of power dissipation of analog channel-select filter and nyquist-rate A/D converter in UTRA/FDD," in *IEEE International Symposium on Circuits and Systems*, vol. 4, May 2004, pp. IV–940.
- [67] J. Yoo, K. Choi, and D. Lee, "Comparator generation and selection for highly linear CMOS flash analog-to-digital converter," *Analog Integrated Circuits and Signal Processing*, vol. 35, no. 2-3, pp. 179–187, 2003.
- [68] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.
- [69] Q. Ding, Y. Deng, and X. Gao, "Spectral and energy efficiency of hybrid precoding for mmwave massive MIMO with low-resolution ADCs/DACs," *IEEE Access*, vol. 7, pp. 186 529–186 537, 2019.
- [70] M. Tang, M. Vehkaperä, X. Chu, and R. Wichman, "Power allocation for multipair massive MIMO FD relay systems with low resolution ADCs," in *IEEE International Symposium on Wireless Communication Systems*, 2019, pp. 505–510.
- [71] M. Sarajlić, L. Liu, and O. Edfors, "When are low resolution ADCs energy efficient in massive MIMO?" *IEEE Access*, vol. 5, pp. 14 837–14 853, 2017.
- [72] T. Liu, J. Tong, Q. Guo, J. Xi, Y. Yu, and Z. Xiao, "Energy efficiency of massive MIMO systems with low-resolution ADCs and successive interference cancellation," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 3987–4002, 2019.
- [73] Y. Shim, K. Lee, and H. Park, "Joint relay beamforming design for multilevel nondistributed and distributed amplify-and-forward relay networks," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2016.
- [74] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *IEEE International Symposium on Information Theory*, 2012.
- [75] A. Sabharwal *et al.*, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept. 2014.
- [76] C. Kong *et al.*, "Full-Duplex massive MIMO relaying systems with Low-Resolution ADCs," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5033–5047, Aug 2017.
- [77] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *IEEE International Symposium on Information Theory*, 2012, pp. 1–5.

- [78] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted gaussian signals," *Research Laboratory of Electronics Massachusetts Institute of Technology*, Mar 1952.
- [79] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 4075–4089, Aug 2017.
- [80] C. Kong, A. Mezghani, C. Zhong, A. L. Swindlehurst, and Z. Zhang, "Multipair massive mimo relaying systems with one-bit ADCs and DACs," *IEEE Transactions on Signal Processing*, vol. 66, no. 11, pp. 2984–2997, 2018.
- [81] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 191–193, 1995.
- [82] Y. Tsai, H. Huang, Y. Chen, and K. Yang, "Simultaneous multiple carrier frequency offsets estimation for coordinated multi-point transmission in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4558–4568, 2013.
- [83] J. Chen, Y. Wu, S. Ma, and T. Ng, "Joint CFO and channel estimation for multiuser MIMO-OFDM systems with optimal training sequences," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4008–4019, 2008.
- [84] Y. Wu, S. Attallah, and J. W. M. Bergmans, "Carrier frequency offset estimation for multi-user MIMO OFDM uplink using CAZAC sequences," in *IEEE Wireless Communications and Networking Conference*, 2009, pp. 1–5.
- [85] W. Zhang, Q. Yin, and F. Gao, "Computationally efficient blind estimation of carrier frequency offset for MIMO-OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7644–7656, 2016.
- [86] D. Wang, J. Wei, and X. Zhang, "A novel blind carrier synchronization method for MIMO OFDM system," in *IEEE Military Communications Conference*, 2007, pp. 1–4.
- [87] H. V. Cheng and E. G. Larsson, "Some fundamental limits on frequency synchronization in massive MIMO," in *Asilomar Conference on Signals, Systems and Computers*, 2013, pp. 1213–1217.
- [88] Y. Feng, W. Zhang, F. Gao, and Q. Sun, "Computationally efficient blind CFO estimation for massive MIMO uplink," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7795–7799, 2018.
- [89] N. J. Myers and R. W. Heath, "Message passing-based joint CFO and channel estimation in mmwave systems with one-bit ADCs," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3064–3077, 2019.
- [90] J. Zhu, H. Cao, and Z. Xu, "A two-stage approach to estimate CFO and channel with one-bit ADCs," *Signal Processing*, vol. 168, p. 107336, 2020.

- [91] 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, Aug. 2011.
- [92] H. T. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, 1946.
- [93] H. G. Schantz, A. H. Udden, M. Nikravan, and D.-H. Kwon, "Simple formulas for near-field transmission gain and fields," in *37th Antenna Applications Symposium, Allerton, IL*, 2013.
- [94] E. Bjornson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Communication Letters*, vol. 17, no. 1, pp. 91–94, Jan. 2013.
- [95] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3074–3085, 2011.
- [96] 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, Apr. 2003.
- [97] M. Vehkaperä, T. Riihonen, and R. Wichman, "Asymptotic analysis of full-duplex bidirectional MIMO link with transmitter noise," in *IEEE International Symposium on Personal and Indoor Mobile Radio Communications*, Sep. 2013.
- [98] O. Blume *et al.*, "Measurement and characterization of the temporal behavior of fixed massive MIMO links," in *IEEE International Workshop on Smart Antennas*, 2017, pp. 1–8.
- [99] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, 2016.
- [100] M. Tang, M. Vehkaperä, X. Chu, and R. Wichman, "LI cancellation and power allocation for multipair FD relay systems with massive antenna arrays," *IEEE Wireless Communication Letters*, vol. 8, no. 4, pp. 1077–1081, 2019.
- [101] A. K. Fletcher *et al.*, "Robust predictive quantization: Analysis and design via convex optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 618–632, Dec 2007.
- [102] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [103] K. Roth *et al.*, "A comparison of hybrid beamforming and digital beamforming with Low-Resolution ADCs for multiple users and imperfect CSI," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 3, pp. 484–498, June 2018.
- [104] G. Auer *et al.*, "How much energy is needed to run a wireless network?" *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 40–49, October 2011.

-
- [105] S. Tretter, “Estimating the frequency of a noisy sinusoid by linear regression,” *IEEE Trans. Info. Theory*, vol. 31, no. 6, pp. 832–835, November 1985.
 - [106] A. M. Tulino, S. Verdú *et al.*, “Random matrix theory and wireless communications,” *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.
 - [107] H. Cramér, *Random Variables and Probability Distributions*. Cambridge University Press, 2004, vol. 36.
 - [108] G. Grimmett, G. R. Grimmett, D. Stirzaker *et al.*, *Probability and random processes*. Oxford university press, 2001.

Appendix A

Proof of Proposition 3.1

The received signal $y_{R,k}^{\text{MF}}$ in (3.14) for MF reads

$$\begin{aligned} \frac{y_{R,k}^{\text{MF}}[i]}{\sqrt{M}} &= \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}} \frac{\mathbf{g}_{\text{SR},k}^T \hat{\mathbf{g}}_{\text{SR},k}^*}{M} \bar{x}_k[i] \\ &+ \sum_{j=1, j \neq k}^K \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}} \frac{\mathbf{g}_{\text{SR},k}^T \hat{\mathbf{g}}_{\text{SR},j}^*}{M} \bar{x}_j[i] + \frac{\sqrt{\rho_R} \mathbf{g}_{\text{RR},k}^T \mathbf{x}_R[i]}{\sqrt{M}} + \frac{n_{R,k}}{\sqrt{M}}. \end{aligned} \quad (\text{A.1})$$

Due to the properties of the MMSE estimates and the fact that the channels have CSCG elements, $\hat{\mathbf{g}}_{\text{SR},k}$ and $\tilde{\mathbf{g}}_{\text{SR},k}$ are independent. By applying the strong law of large numbers [107], the first term in (A.1) becomes

$$\begin{aligned} &\sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}} \frac{\mathbf{g}_{\text{SR},k}^T \hat{\mathbf{g}}_{\text{SR},k}^* + \tilde{\mathbf{g}}_{\text{SR},k}^T \hat{\mathbf{g}}_{\text{SR},k}^*}{M} \bar{x}_k[i] \\ &\xrightarrow{a.s.} \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}} \hat{\beta}_{\text{SR},k} \bar{x}_k[i], \quad \text{as } M \rightarrow \infty. \end{aligned} \quad (\text{A.2})$$

Similarly, since $\mathbf{g}_{\text{SR},k}$ and $\hat{\mathbf{g}}_{\text{SR},j}$ are independent for $k \neq j$, the inter-pair interference term in (A.1) converges almost surely to 0 as $M \rightarrow \infty$, i.e.

$$\sum_{j=1, j \neq k}^K \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}}} \frac{\mathbf{g}_{\text{SR},k}^T \hat{\mathbf{g}}_{\text{SR},j}^*}{M} \bar{x}_j[i] \xrightarrow{a.s.} 0, \quad (\text{A.3})$$

when $M \rightarrow \infty$. Moreover, the third term in (A.1) representing LI also approaches 0 almost surely, when $M \rightarrow \infty$, since $\mathbf{g}_{\text{RR},k}^T \mathbf{x}_R[i]$ is a summation of K random variables that are almost surely finite.

$$\frac{\sqrt{\rho_R} \mathbf{g}_{\text{RR},k}^T \mathbf{x}_R[i]}{\sqrt{M}} = \frac{\sqrt{\rho_R} \sum_{l=1}^K g_{\text{RR},kl} \bar{x}_l[i]}{\sqrt{M}} \xrightarrow{a.s.} 0, \quad (\text{A.4})$$

Here $g_{\text{RR},lk}$ and $x_{R,l}[i]$ denotes l th element of $\mathbf{g}_{\text{RR},k}$ and $\mathbf{x}_R[i]$, respectively. Finally, we note that the noise term is almost surely finite and converges to zero when divided by M that grows without bound. Therefore, even without LI cancellation the received signals at the relay are interference and noise free. Substituting (A.3) and (A.4) into (A.1) leads to (3.22), thus completes the proof.

For ZF, the received signals at relay nodes/antennas reads

$$\frac{\mathbf{y}_R^{\text{ZF}}}{\sqrt{M}} = \sqrt{\frac{\rho_S}{M}} \mathbf{G}_{\text{SR}}^T \mathbf{V}_{\text{ZF}} \bar{\mathbf{x}}[i] + \sqrt{\frac{\rho_R}{M}} \mathbf{G}_{\text{RR}} \mathbf{x}_R[i] + \frac{\mathbf{n}_R}{\sqrt{M}}. \quad (\text{A.5})$$

The first term in (A.5) can be written as

$$\begin{aligned} \sqrt{\frac{\rho_S}{M}} \mathbf{G}_{\text{SR}}^T \mathbf{V}_{\text{ZF}} \bar{\mathbf{x}}[i] &= \sqrt{\frac{\rho_S}{M}} \hat{\mathbf{G}}_{\text{SR}}^T \mathbf{V}_{\text{ZF}} \bar{\mathbf{x}}[i] + \sqrt{\frac{\rho_S}{M}} \tilde{\mathbf{G}}_{\text{SR}}^T \mathbf{V}_{\text{ZF}} \bar{\mathbf{x}}[i] \\ &= \sqrt{\frac{\rho_S}{M}} \lambda_{\text{ZF}} \bar{\mathbf{x}}[i] + \sqrt{\frac{\rho_S}{M}} \tilde{\mathbf{G}}_{\text{SR}}^T \mathbf{V}_{\text{ZF}} \bar{\mathbf{x}}[i], \end{aligned} \quad (\text{A.6})$$

where the second term in (A.6) is almost surely converges to $\mathbf{0}$ as $M \rightarrow \infty$ by applying law of large numbers, i.e.

$$\sqrt{\frac{\rho_S}{M}} \tilde{\mathbf{G}}_{\text{SR}}^T \mathbf{v}_{\text{ZF}} \bar{\mathbf{x}}[i] = \sqrt{\frac{\rho_S(M-K)K}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}^{-1}}} \frac{\tilde{\mathbf{G}}_{\text{SR}}^T \hat{\mathbf{G}}_{\text{SR}}^*}{M} \left(\frac{\hat{\mathbf{G}}_{\text{SR}}^T \hat{\mathbf{G}}_{\text{SR}}^*}{M} \right)^{-1} \bar{\mathbf{x}}[i] \xrightarrow{a.s.} 0. \quad (\text{A.7})$$

Thus, as $M \rightarrow \infty$, we have

$$\sqrt{\frac{\rho_S}{M}} \mathbf{g}_{\text{SR}}^T \mathbf{v}_{\text{ZF}} \bar{\mathbf{x}}_k[i] \xrightarrow{a.s.} \sqrt{\frac{\rho_S}{\frac{1}{K} \sum_{k=1}^K \hat{\beta}_{\text{SR},k}^{-1}}} \bar{\mathbf{x}}_k[i]. \quad (\text{A.8})$$

Other two terms in (A.5) almost surely approaches $\mathbf{0}$ as M grows large. The proof is omitted due to the similarity to MF case above. Equation (3.24) is therefore proved. The proofs of (3.23) and (3.25) are similar and omitted.

Appendix B

Proof of Proposition 3.2

i) Derivation of $R_{\text{SR},k}^{\text{MF}}$:

If linear MF precoding is used at source, by following similar derivation in [12], desired signal power in (3.28) can be calculated by

$$\begin{aligned} \frac{\rho_S}{K} |\mathbb{E}\{\mathbf{g}_{\text{SR},k}^H \mathbf{v}_k\}|^2 &= \frac{\rho_S}{K} \lambda_{\text{MF}}^2 |\mathbb{E}\{\mathbf{g}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^*\}|^2 \\ &= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} |\mathbb{E}\{\hat{\mathbf{g}}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^* + \tilde{\mathbf{g}}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^*\}|^2 \\ &= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} |\mathbb{E}\{\|\hat{\mathbf{g}}_{\text{SR},k}\|^2\}|^2 \\ &= \frac{\rho_S M \hat{\beta}_{\text{SR},k}^2}{\sum_{k=1}^K \hat{\beta}_{\text{SR},k}} \end{aligned} \tag{B.1}$$

By applying [106, Lemma 2.9], the term $\frac{\rho_S}{K} \text{Var}(\mathbf{g}_{\text{SR},k}^H \mathbf{v}_k)$ can be computed by

$$\begin{aligned}
& \frac{\rho_S}{K} \text{Var}(\mathbf{g}_{\text{SR},k}^H \mathbf{v}_k) \\
&= \frac{\rho_S}{K} \lambda_{\text{MF}}^2 \text{Var}(\mathbf{g}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^*) \\
&= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} (\mathbb{E}\{|\mathbf{g}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^*|^2\} - |\mathbb{E}\{\mathbf{g}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^*\}|^2) \\
&= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} \left(\mathbb{E}\{\|\hat{\mathbf{g}}_{\text{SR},k}\|^4\} + \mathbb{E}\{|\tilde{\mathbf{g}}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},k}^*|^2\} - M^2 \hat{\beta}_{\text{SR},k}^2 \right) \\
&= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} \left(M(M+1) \hat{\beta}_{\text{SR},k}^2 + M \hat{\beta}_{\text{SR},k} (\beta_{\text{SR},k} - \hat{\beta}_{\text{SR},k}) - M^2 \hat{\beta}_{\text{SR},k}^2 \right) \\
&= \frac{\rho_S \beta_{\text{SR},k} \hat{\beta}_{\text{SR},k}}{\sum_{k=1}^K \hat{\beta}_{\text{SR},k}}
\end{aligned} \tag{B.2}$$

Similarly, the inter-pair interference power is given by

$$\begin{aligned}
& \frac{\rho_S}{K} \sum_{j=1, j \neq k}^K \mathbb{E}\{|\mathbf{g}_{\text{SR},k}^H \mathbf{v}_j|^2\} \\
&= \frac{\lambda_{\text{MF}}^2 \rho_S}{K} \sum_{j=1, j \neq k}^K \mathbb{E}\{|\mathbf{g}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},j}^*|^2\} \\
&= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} \sum_{j=1, j \neq k}^K \mathbb{E}\{\hat{\mathbf{g}}_{\text{SR},j}^H \mathbf{g}_{\text{SR},k}^* \mathbf{g}_{\text{SR},k}^H \hat{\mathbf{g}}_{\text{SR},j}^*\} \\
&= \frac{\rho_S}{M \sum_{k=1}^K \hat{\beta}_{\text{SR},k}} \sum_{j=1, j \neq k}^K M \hat{\beta}_{\text{SR},j} \beta_{\text{SR},k} \\
&= \frac{\rho_S \beta_{\text{SR},k} \sum_{j=1, j \neq k}^K \hat{\beta}_{\text{SR},j}}{\sum_{k=1}^K \hat{\beta}_{\text{SR},k}}
\end{aligned} \tag{B.3}$$

By substituting (B.1)-(B.3) into (3.28), $R_{\text{SR},k}^{\text{MF}}$ is obtained.

ii) Derivation of $R_{RD,k}^{\text{MF}}$:

By applying Jensen's inequality to (3.29), we obtain that

$$\begin{aligned} & R_{RD,k} \\ &= \mathbb{E} \left\{ \log_2 \left(1 + \frac{\frac{\rho_R}{K} |\mathbf{w}_k^H \hat{\mathbf{g}}_{RD,k}|^2}{\sum_{j=1, j \neq k}^K \frac{\rho_R}{K} \mathbb{E}\{|\mathbf{w}_k^H \mathbf{g}_{RD,j}|^2\} + \frac{\rho_R}{K} \mathbb{E}\{|\mathbf{w}_k^H \tilde{\mathbf{g}}_{RD,k}|^2\} + \mathbb{E}\{\mathbf{n}_D^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{n}_D\}} \right) \right\} \\ &\geq \log_2 \left(1 + \frac{\mathbb{E} \left\{ \frac{1}{\frac{\rho_R}{K} |\mathbf{w}_k^H \hat{\mathbf{g}}_{RD,k}|^2} \right\}^{-1}}{\sum_{j=1, j \neq k}^K \frac{\rho_R}{K} \mathbb{E}\{|\mathbf{w}_k^H \mathbf{g}_{RD,j}|^2\} + \frac{\rho_R}{K} \mathbb{E}\{|\mathbf{w}_k^H \tilde{\mathbf{g}}_{RD,k}|^2\} + \mathbb{E}\{\mathbf{n}_D^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{n}_D\}} \right) \end{aligned}$$

where when MF receiver is applied, we calculate

$$\mathbb{E} \left\{ \frac{1}{\frac{\rho_R}{K} |\mathbf{w}_k^H \hat{\mathbf{g}}_{RD,k}|^2} \right\}^{-1} = \mathbb{E} \left\{ \frac{1}{\frac{\rho_R}{K} \|\hat{\mathbf{g}}_{RD,k}\|^4} \right\}^{-1} = \frac{\rho_R \hat{\beta}_{RD,k}^2}{K} \mathbb{E} \left\{ \text{tr}(\mathbf{L}^{-2}) \right\}^{-1} \quad (\text{B.4})$$

where we denote $\mathbf{L} \sim \mathcal{W}_1(M, \mathbf{I})$ is a Wishart matrix. The property of the Wishart matrix guarantees [106]

$$\mathbb{E} \left\{ \text{tr}\{\mathbf{L}^{-2}\} \right\} = \frac{M}{(M-1)^3 - (M-1)},$$

and for very large M ,

$$\mathbb{E} \left\{ \text{tr}\{\mathbf{L}^{-2}\} \right\} \approx \frac{1}{M^2}.$$

We can obtain that

$$\mathbb{E} \left\{ \frac{1}{\frac{\rho_R}{K} |\mathbf{w}_k^H \hat{\mathbf{g}}_{RD,k}|^2} \right\}^{-1} = \mathbb{E} \left\{ \frac{1}{\frac{\rho_R}{K} \|\hat{\mathbf{g}}_{RD,k}\|^4} \right\}^{-1} = \frac{\rho_R \hat{\beta}_{RD,k}^2 M^2}{K} \quad (\text{B.5})$$

Similarly to (B.3), the power of inter-pair interference in R→D link is given by

$$\begin{aligned} & \frac{\rho_R}{K} \sum_{j=1, j \neq k}^K \mathbb{E}\{|\mathbf{w}_k^H \mathbf{g}_{RD,j}|^2\} \\ &= \frac{\rho_R}{K} \sum_{j=1, j \neq k}^K \mathbb{E}\{|\hat{\mathbf{g}}_{RD,k}^H \mathbf{g}_{RD,j}|^2\} \\ &= \frac{\rho_R}{K} M \hat{\beta}_{RD,k} \sum_{j=1, j \neq k}^K \beta_{RD,j} \end{aligned} \quad (\text{B.6})$$

The estimate error power is given by

$$\begin{aligned} \frac{\rho_R}{K} \mathbb{E}\{|\mathbf{w}_k^H \tilde{\mathbf{g}}_{\text{RD},k}|^2\} &= \frac{\rho_R}{K} \mathbb{E}\{|\hat{\mathbf{g}}_{\text{RD},k}^H \tilde{\mathbf{g}}_{\text{RD},k}|^2\} \\ &= \frac{\rho_R}{K} M \hat{\beta}_{\text{RD},k} \left(\beta_{\text{RD},k} - \hat{\beta}_{\text{RD},k} \right) \end{aligned} \quad (\text{B.7})$$

Moreover, as channel estimate is independent with additive noise, noise term can be calculated by

$$\mathbb{E}\{\|\mathbf{w}_k^H\|^2\} = \mathbb{E}\{\hat{\mathbf{g}}_{\text{RD},k}^H \hat{\mathbf{g}}_{\text{RD},k}\} = M \hat{\beta}_{\text{RD},k} \quad (\text{B.8})$$

By substituting (B.5)-(B.8) to (3.29), we can obtain $R_{\text{RD},k}^{\text{MF}}$.

We substitute $R_{\text{SR},k}^{\text{MF}}$ and $R_{\text{RD},k}^{\text{MF}}$ to (3.30), yielding (3.31). The proof of (3.32) is similar, thus is omitted.

iii) Derivation of $\rho_{\xi_k}^{\varepsilon}$:

The residual power of loop interference term $\rho_{\xi_k}^{\varepsilon}$ after LI cancellation can be calculated as

$$\rho_{\xi_k}^{\varepsilon} = \begin{cases} \frac{\rho_R}{K} \mathbb{E}\{\|\tilde{\mathbf{g}}_{\text{RR},k}^H\|^2\} = \rho_R \tilde{\beta}, & \text{co-located;} \\ \frac{\rho_R}{K} \mathbb{E}\{|\tilde{g}_{\text{RR},kk} + \sum_{l=1, l \neq k}^K g_{\text{RR},kl}|^2\} = \frac{\rho_R}{K} \left(\tilde{\beta}_k + \sum_{l=1, l \neq k}^K \beta_{k,l} \right), & \text{cooperative;} \\ \frac{\rho_R}{K} \mathbb{E}\{\|\tilde{\mathbf{g}}_{\text{RR},k}^H\|^2\} = \frac{\rho_R}{K} \left(\tilde{\beta}_k + \sum_{l=1, l \neq k}^K \tilde{\beta}_{kl} \right), & \text{non-cooperative.} \end{cases} \quad (\text{B.9})$$

Appendix C

Proof of Proposition 4.1

Equation (4.14) and (4.15) are derived from (4.7) and (4.8). Derivations of majority terms are omitted due to the similarity of [12]. The different terms containing the power allocation matrix \mathbf{P}_{MF} and \mathbf{P}_{ZF} will be derived:

i) Derivation of R_k^{MF} :

As $\hat{\mathbf{G}}_{\text{RD}}$, $\tilde{\mathbf{G}}_{\text{RR}}$ and $\hat{\mathbf{g}}_{\text{SR},k}$ are independent. With MF precoding, the power of residual LI can be calculated by

$$\begin{aligned}
 E\{\|\mathbf{w}_{\text{MF},k}^H \boldsymbol{\xi}_{\text{MF}}\|^2\} &= E\{\|\hat{\mathbf{g}}_{\text{SR},k}^H \tilde{\mathbf{G}}_{\text{RR}} \mathbf{x}_{\text{R}}\|^2\} \\
 &= E\{\|\hat{\mathbf{g}}_{\text{SR},k}^H \tilde{\mathbf{G}}_{\text{RR}} \hat{\mathbf{G}}_{\text{RD}}^* \mathbf{P}_{\text{MF}} \bar{\mathbf{x}}\|^2\} \\
 &= E\{\hat{\mathbf{g}}_{\text{SR},k}^H \tilde{\mathbf{G}}_{\text{RR}} \hat{\mathbf{G}}_{\text{RD}}^* \mathbf{P}_{\text{MF}} \bar{\mathbf{x}} \bar{\mathbf{x}}^H \mathbf{P}_{\text{MF}}^H \hat{\mathbf{G}}_{\text{RD}}^T \tilde{\mathbf{G}}_{\text{RR}}^H \hat{\mathbf{g}}_{\text{SR},k}\} \\
 &= M_{\text{Tx}} M_{\text{Rx}} \tilde{\beta}_{\text{R}} \hat{\beta}_{\text{SR},k} \sum_{j=1}^K \frac{m_{\text{MF},j}}{M_{\text{Tx}} \hat{\beta}_{\text{RD},j}} \hat{\beta}_{\text{RD},j} \\
 &= M_{\text{Rx}} \tilde{\beta}_{\text{R}} \hat{\beta}_{\text{SR},k} m_{\text{MF},\text{tot}}.
 \end{aligned}$$

By similar calculation of other terms, one can obtain that $M\hat{\beta}_{\text{SR},k}$ is a common factor of both the upstairs and downstairs of (4.7). We can then obtain that

$$\rho_{\xi_{\text{MF}}} = \tilde{\beta}_{\text{R}} m_{\text{MF},\text{tot}}.$$

The power of inter-stream interference of R→D link is derived by

$$\begin{aligned} \sum_{j=1, j \neq k}^K \mathbb{E}\{|\mathbf{g}_{\text{RD},k}^T \mathbf{v}_{\text{MF},j}|^2\} &= \sum_{j=1, j \neq k}^K \frac{m_{\text{MF},j}}{M_{\text{Tx}} \hat{\beta}_{\text{RD},j}} \mathbb{E}\{|\mathbf{g}_{\text{RD},k}^T \hat{\mathbf{g}}_{\text{RD},j}^*|^2\} \\ &= \sum_{j=1, j \neq k}^K \frac{m_{\text{MF},j}}{M_{\text{Tx}} \hat{\beta}_{\text{RD},j}} M_{\text{Tx}} \beta_{\text{RD},k} \hat{\beta}_{\text{RD},j} \\ &= \beta_{\text{RD},k} \sum_{j=1, j \neq k}^K m_{\text{MF},j}. \end{aligned}$$

ii) Derivation of R_k^{ZF} :

With ZF precoding at the relay, the power of residual LI can be calculated by

$$\mathbb{E}\{\|\mathbf{w}_{\text{ZF},k}^H \boldsymbol{\xi}_{\text{ZF}}\|^2\} = \mathbb{E}\{\mathbf{w}_{\text{ZF},k}^H \tilde{\mathbf{G}}_{\text{RR}} \mathbf{V}_{\text{ZF}} \mathbf{V}_{\text{ZF}}^H \tilde{\mathbf{G}}_{\text{RR}}^H \mathbf{w}_{\text{ZF},k}\}. \quad (\text{C.1})$$

According to the law of large numbers, we have following approximation [12]

$$\mathbf{V}_{\text{ZF}} \mathbf{V}_{\text{ZF}}^H = \frac{1}{M_{\text{Tx}}^2} \mathbf{G}_{\text{RD}}^* \mathbf{D}_{\text{RD}}^{-2} \mathbf{P}_{\text{ZF}}^2 \mathbf{G}_{\text{RD}}^T. \quad (\text{C.2})$$

From (C.2), one can further calculate (C.1) as

$$\begin{aligned}
\mathbb{E}\{\|\mathbf{w}_{ZF,k}^H \boldsymbol{\xi}_{ZF}\|^2\} &= \mathbb{E}\{\mathbf{w}_k^H \tilde{\mathbf{G}}_{RR} \mathbf{V}_{ZF} \mathbf{V}_{ZF}^H \tilde{\mathbf{G}}_{RR}^H \mathbf{w}_k\} \\
&= \frac{\tilde{\beta}_R}{M} \sum_{j=1}^K \frac{(M_{Tx} - K) \hat{\beta}_{RD,j} m_{ZF,j}}{\hat{\beta}_{RD,j}} \mathbb{E}\{\|\mathbf{w}_k\|^2\} \\
&\stackrel{(a)}{=} \frac{(M_{Tx} - K) \tilde{\beta}_R}{M_{Tx} (M_{Rx} - K) \hat{\beta}_{RD,k}} \sum_{j=1}^K m_{ZF,j} \\
&= \frac{(M_{Tx} - K) \tilde{\beta}_R m_{ZF,tot}}{M_{Tx} (M_{Rx} - K) \hat{\beta}_{RD,k}}.
\end{aligned}$$

where (a) is obtained by using the property of Wishart matrix [106, 12]. Again,

one can easily find that $\frac{1}{(M_{Rx} - K) \hat{\beta}_{RD,k}}$ is a common factor and obtain

$$\rho_{\xi_{ZF}} = \tilde{\beta}_R \left(1 - \frac{K}{M_{Tx}}\right) m_{ZF,tot}.$$

Appendix D

Proof of Proposition 6.2

As ζ'_* is a scaled and rotated version of ζ_* , we first need to obtain the distribution of ζ_* .

Lemma D.1. *The distributions of $\zeta_2, \zeta_3, \dots, \zeta_8$ are given by*

$$\zeta_2 \sim \mathcal{CN}\left(0, \frac{a_p^4 \sigma_n^4}{M}\right), \quad (D.1)$$

$$\zeta_3, \zeta_4 \sim \mathcal{CN}\left(0, \frac{|\phi|^2 a_p^4 \sigma_n^2 \|\mathbf{g}\|^2}{M^2}\right), \quad (D.2)$$

$$\zeta_5, \zeta_6 \sim \mathcal{CN}\left(0, \frac{|\phi|^2 a_p^2 C_{\text{diag}} \|\mathbf{g}\|^2}{M^2}\right) \quad (D.3)$$

$$\zeta_7, \zeta_8 \sim \mathcal{CN}\left(0, \frac{a_p^2 \sigma_n^2 C_{\text{diag}}}{M}\right), \quad (D.4)$$

and $\zeta_9 = \frac{1}{M} \mathbf{q}_{t_1}^H \mathbf{q}_{t_2}$ is improper Gaussian distributed with mean and variance

$$\mathbb{E}\{\zeta_9\} = C_{t_2 t_1}, \quad (D.5)$$

$$\text{Var}\{\zeta_{9, \text{I}}\} = \frac{1}{2M} \left(C_{\text{diag}}^2 - \Im(C_{t_2 t_1})^2 + \Re(C_{t_2 t_1})^2 \right), \quad (D.6)$$

$$\text{Var}\{\zeta_{9, \text{Q}}\} = \frac{1}{2M} \left(C_{\text{diag}}^2 + \Im(C_{t_2 t_1})^2 - \Re(C_{t_2 t_1})^2 \right), \quad (D.7)$$

where I and Q denote the in-phase and quadrature of the random variable.

Proof. See appendix E. □

For notational simplicity, we assume that $t_2 = 2$ and $t_1 = 1$ in the derivation, so that $C_{t_2 t_1} = C_{21}$. We also omit the subscript k , so that $\omega_k = \omega$.

i) Finding the mean of $\zeta'_{9,Q}$:

The mean of ζ'_9 can be calculated by

$$\begin{aligned} \mathbb{E}\{\zeta'_9 | \mathbf{g}\} &= \frac{1}{|\phi|^2 \zeta_1} \mathbb{E}\left\{ \zeta_9 e^{-j\omega} \right\} \\ &= \frac{1}{|\phi|^2 \zeta_1} \mathbb{E}\left\{ (\Re(C_{21}) + j\Im(C_{21})) (\cos \omega - j \sin \omega) \right\} \\ &= \frac{1}{|\phi|^2 \zeta_1} (\cos(\omega) \Im(C_{21}) + \sin(\omega) \Re(C_{21})) \\ &\quad + j \frac{1}{|\phi|^2 \zeta_1} (\cos(\omega) \Re(C_{21}) - \sin(\omega) \Im(C_{21})). \end{aligned} \quad (\text{D.8})$$

By taking the quadrature part of the above equation, we conclude the derivation of (6.42).

ii) Finding the variance of $\zeta'_{9,Q}$:

$$\begin{aligned} \zeta'_9 &= \frac{1}{|\phi|^2 \zeta_1} (\zeta_{9,I} + j\zeta_{9,Q}) (\cos \omega - j \sin \omega) \\ &= \frac{1}{|\phi|^2 \zeta_1} \left((\zeta_{9,I} \cos \omega + \zeta_{9,Q} \sin \omega) \right. \\ &\quad \left. + j (\zeta_{9,Q} \cos \omega - \zeta_{9,I} \sin \omega) \right). \end{aligned} \quad (\text{D.9})$$

After some simple derivation, the mean of the real and imaginary parts of ζ'_9 are

$$\begin{aligned} \mathbb{E}\{\zeta'_{9,I} | \mathbf{g}\} &= \frac{1}{|\phi|^2 \zeta_1} (\Re(C_{21}) \cos \omega + \Im(C_{21}) \sin \omega), \\ \mathbb{E}\{\zeta'_{9,Q} | \mathbf{g}\} &= \frac{1}{|\phi|^2 \zeta_1} (\Im(C_{21}) \cos \omega - \Re(C_{21}) \sin \omega). \end{aligned}$$

The variance of $\zeta'_{9,Q}$ is then derived by

$$\begin{aligned}
& \text{Var}\{\zeta'_{9,Q} \mid \mathbf{g}\} \\
&= \frac{1}{|\phi|^4 \zeta_1^2} \mathbb{E}\left\{ \left| \Im(C_{21}) \cos \omega - \Re(C_{21}) \sin \omega - \mathbb{E}\{\zeta'_{9,Q}\} \right|^2 \right\} \\
&= \frac{1}{|\phi|^4 \zeta_1^2} \mathbb{E}\left\{ \left(\zeta_{9,Q}^2 + \Im(C_{21})^2 - 2\Im(C_{21})\zeta_{9,Q} \right) \cos^2 \omega \right. \\
&\quad \left. + \left(\zeta_{9,I}^2 + \Re(C_{21})^2 - 2\Re(C_{21})\zeta_{9,I} \right) \sin^2 \omega \right. \\
&\quad \left. + 2 \left(\zeta_{9,Q}\Re(C_{21}) + \zeta_{9,I}\Im(C_{21}) - \Im(C_{21})\Re(C_{21}) \right. \right. \\
&\quad \left. \left. - \zeta_{9,I}\zeta_{9,Q} \right) \cos \omega \sin \omega \right\} \\
&= \frac{1}{|\phi|^4 \zeta_1^2} \left(\text{Var}\{\Im(\zeta_9)\} \cos^2 \omega + \text{Var}\{\Re(\zeta_9)\} \sin^2 \omega \right. \\
&\quad \left. - 2\mathbb{E}\{\zeta_{9,I}\zeta_{9,Q}\} \cos \omega \sin \omega \right. \\
&\quad \left. + 2\Im(C_{21})\Re(C_{21}) \cos \omega \sin \omega \right), \tag{D.10}
\end{aligned}$$

where $\mathbb{E}\{\zeta_{9,I}\zeta_{9,Q}\}$ is derived by using the same model as in Appendix E,

$$\begin{aligned}
\mathbb{E}\{\zeta_{9,I}, \zeta_{9,Q}\} &= \mathbb{E}\left\{ \Im\left(\frac{1}{M}\mathbf{q}_1^H \mathbf{q}_2\right) \Re\left(\frac{1}{M}\mathbf{q}_1^H \mathbf{q}_2\right) \right\} \\
&= \frac{1}{M^2} \mathbb{E}\left\{ \sum_{m=1}^M (q_{m1,I}q_{m2,I} + q_{m1,Q}q_{m2,Q}) \right. \\
&\quad \left. \sum_{m=1}^M (q_{m1,I}q_{m2,Q} - q_{m1,Q}q_{m2,I}) \right\}. \tag{D.11}
\end{aligned}$$

According to the structure of correlation matrix \mathbf{C}_q , there is no cross-correlation between different antennas, and \mathbf{q}_* has i.i.d. elements. Using the model in

(E.6)-(E.11), we can obtain

$$\begin{aligned}
& \mathbb{E}\{\zeta_{9,I}, \zeta_{9,Q}\} \\
&= \frac{1}{M^2} \mathbb{E}\left\{ \sum_{m=1}^M (q_{m1,I}q_{m2,I} + q_{m1,Q}q_{m2,Q}) \right. \\
&\quad \left. (q_{m1,I}q_{m2,Q} - q_{m1,Q}q_{m2,I}) \right\} \\
&= \mathbb{E}\left\{ (q_{m1,I}q_{m2,I} + q_{m1,Q}q_{m2,Q}) \right. \\
&\quad \left. (q_{m1,I}q_{m2,Q} - q_{m1,Q}q_{m2,I}) \right\} \\
&= \Im(C_{21})\Re(C_{21}). \tag{D.12}
\end{aligned}$$

Thus, we conclude the derivation as

$$\begin{aligned}
& \text{Var}\{\zeta'_{9,Q} \mid \mathbf{g}\} \\
&= \frac{1}{|\phi|^4 \zeta_1^2} (\text{Var}\{\Im(\zeta_9)\} \cos^2 \omega + \text{Var}\{\Re(\zeta_9)\} \sin^2 \omega) \\
&= \frac{M}{2|\phi|^4 a_p^4 \|\mathbf{g}\|^4} (C_{\text{diag}}^2 - (\Im(C_{21}))^2 - \Re(C_{21})^2) \cos^2 \omega \\
&\quad + (\Im(C_{21}))^2 - \Re(C_{21})^2) \sin^2 \omega \\
&= \frac{M}{2|\phi|^4 a_p^4 \|\mathbf{g}\|^4} (C_{\text{diag}}^2 - (\Im(C_{21}))^2 - \Re(C_{21})^2) \cos 2\omega).
\end{aligned}$$

iii) Calculate $\text{cov}(\zeta'_{5,Q}, \zeta'_{6,Q} \mid \mathbf{g})$:

$$\begin{aligned}
& \text{cov}(\zeta'_5, \zeta'_6) = \mathbb{E}\{\zeta_5'^* \zeta_6' \mid \mathbf{g}\} \\
&= \mathbb{E}\left\{ \frac{\phi^2}{|\phi|^4 a_p^2 \|\mathbf{g}\|^4} \mathbf{q}_2^H \mathbf{g} \mathbf{q}_1^H \mathbf{g} e^{j3\omega} \mid \mathbf{g} \right\} \\
&\triangleq b_I + j b_Q, \tag{D.13}
\end{aligned}$$

where we denote the real and imaginary part of $\text{cov}(\zeta'_5, \zeta'_6)$ as b_I and b_Q respectively. Although the expectation does not have a closed-form expression, we can

always calculate it numerically. And since

$$\begin{aligned} \mathbb{E}\{\zeta_5'^* \zeta_6' | \mathbf{g}\} &= \mathbb{E}\{\zeta_{5,I}' \zeta_{6,I}' + \zeta_{5,Q}' \zeta_{6,Q}' | \mathbf{g}\} \\ &\quad + j \mathbb{E}\{\zeta_{5,I}' \zeta_{6,Q}' - \zeta_{5,Q}' \zeta_{6,I}' | \mathbf{g}\}, \end{aligned} \quad (\text{D.14})$$

we can obtain

$$\mathbb{E}\{\zeta_{5,Q}' \zeta_{6,Q}' | \mathbf{g}\} = b_I - \mathbb{E}\{\zeta_{5,I}' \zeta_{6,I}' | \mathbf{g}\}. \quad (\text{D.15})$$

The real part of pseudo-correlation between $\zeta_{5,Q}'$ and $\zeta_{6,Q}'$ is

$$\Re(\mathbb{E}\{\zeta_5' \zeta_6' | \mathbf{g}\}) = \mathbb{E}\{\zeta_{5,I}' \zeta_{6,I}' - \zeta_{5,Q}' \zeta_{6,Q}' | \mathbf{g}\} \quad (\text{D.16})$$

$$= b_I - 2\mathbb{E}\{\zeta_{5,Q}' \zeta_{6,Q}' | \mathbf{g}\}. \quad (\text{D.17})$$

Thus, the co-variance is calculated by

$$\begin{aligned} &\text{cov}(\zeta_{5,Q}', \zeta_{6,Q}' | \mathbf{g}) \\ &= \frac{b_I}{2} - \frac{1}{2} \Re(\mathbb{E}\{\zeta_5' \zeta_6' | \mathbf{g}\}) \\ &= \frac{b_I}{2} - \frac{1}{2} \Re\left(\mathbb{E}\left\{\frac{1}{|\phi|^2 a_p^2 \|\mathbf{g}\|^4} \mathbf{g}^H \mathbf{q}_2 \mathbf{q}_1^H \mathbf{g} e^{j\omega} | \mathbf{g}\right\}\right) \\ &= \frac{b_I}{2} - \frac{1}{2|\phi|^2 a_p^2 \|\mathbf{g}\|^2} \Re\left(\text{tr}\left\{\mathbf{g} \mathbf{g}^H \mathbb{E}\{\mathbf{q}_2 \mathbf{q}_1^H\}\right\}\right) \\ &= \frac{b_I}{2} - \frac{1}{2|\phi|^2 a_p^2 \|\mathbf{g}\|^2} \Re(C_{21} e^{-j\omega}) \\ &= \frac{b_I}{2} - \frac{1}{2|\phi|^2 a_p^2 \|\mathbf{g}\|^2} (\Re(C_{21}) \cos \omega - \Im(C_{21}) \sin \omega). \end{aligned} \quad (\text{D.18})$$

Appendix E

Proof of Lemma D.1

i) Finding the distribution of ζ_2 :

As \mathbf{n}_{p1} and \mathbf{n}_{p2} are independent with zero mean, the mean of ζ_2 is zero.

For notational simplicity, we denote the m th elements of vectors \mathbf{n}_{p1} and \mathbf{n}_{p2} as $n_{1m} = n_{1m,I} + jn_{1m,Q}$ and $n_{2m} = n_{2m,I} + jn_{2m,Q}$, where zero mean RVs $n_{*,I}$ and $n_{*,Q}$ denote the real and imaginary parts of n_{*m} respectively. They are mutually independent and has the same variance $\sigma_n^2/2$. We can write ζ_2 as

$$\zeta_2 = \frac{1}{M} a_p^2 \sum_{m=1}^M (n_{1m,I}n_{2m,I} + n_{1m,Q}n_{2m,Q} - jn_{1m,Q}n_{2m,I} + jn_{2m,Q}n_{1m,I}), \quad (\text{E.1})$$

where $n_{1m,*}$ and $n_{2m,*}$ are independent, so that

$$\begin{aligned} & \text{Var}\{n_{1m,*}n_{2m,*}\} \\ &= \text{E}\{n_{1m,*}^2 n_{2m,*}^2\} - \text{E}\{n_{1m,*}n_{2m,*}\}^2 = \frac{\sigma_n^4}{4}. \end{aligned}$$

According to the central limit theorem [108], the summation of M i.i.d. RVs with variance σ^2 can be approximated as a Gaussian distributed random variable,

with the same mean and variance σ^2/M . Thus we obtain

$$\zeta_2 \sim \mathcal{N}\left(0, \frac{a_p^4 \sigma_n^4}{M}\right). \quad (\text{E.2})$$

ii) Finding the distribution of ζ_3 :

For a certain block, \mathbf{g} is fixed. Since both \mathbf{n}_{p2} has zero mean, we can simply obtain that $\text{E}\{\zeta_3 | \mathbf{g}\} = 0$. We denote the m th elements of vector \mathbf{g} and \mathbf{n}_{p2} as g_m and n_{2m} respectively. The variance of ζ_3 can be calculated by

$$\begin{aligned} \text{Var}\{\zeta_3 | \mathbf{g}\} &= \frac{1}{M^2} |\phi|^2 a_p^4 \text{E}\{\mathbf{n}_{p2}^H \mathbf{g} \mathbf{g}^H \mathbf{n}_{p2} | \mathbf{g}\} \\ &= \frac{1}{M^2} |\phi|^2 a_p^4 \text{tr}\left\{\mathbf{g} \mathbf{g}^H \text{E}\{\mathbf{n}_{p2} \mathbf{n}_{p2}^H\}\right\} \\ &= \frac{1}{M^2} |\phi|^2 a_p^4 \left\{\sum_{m=1}^M |g_m|^2 \text{E}\{|n_{2m}|^2\}\right\} \\ &= \frac{1}{M^2} |\phi|^2 a_p^4 \sigma_n^2 \|\mathbf{g}\|^2. \end{aligned} \quad (\text{E.3})$$

iii) Finding the distribution of ζ_9 :

According to (6.14), we denote the m th elements of vectors \mathbf{q}_{p*} as $q_{m*} = q_{m*,I} + jq_{m*,Q}$, where $q_{m*,I}$ and $q_{m*,Q}$ are mutually independent and has the same variance $R/2$ and $\text{cov}(q_{mt_2}, q_{mt_1}) = \text{E}\{q_{mt_2} q_{mt_1}^*\} = C_{t_2 t_1}$, yielding the mean of $q_{mt_1}^* q_{mt_2}$ being

$$\text{E}\{q_{mt_1}^* q_{mt_2}\} = C_{t_2 t_1}. \quad (\text{E.4})$$

As there is no cross-correlation between antennas, the mean of ζ_9 is calculated by

$$\text{E}\{\zeta_9\} = \text{E}\left\{\frac{1}{M} \sum_{m=1}^M q_{mt_1}^* q_{mt_2}\right\} = C_{t_2 t_1}. \quad (\text{E.5})$$

Due to the correlation between q_{mt_1} and q_{mt_2} , ζ_9 is not CSCG. We then find the variance of the real and imaginary part of the term $q_{mt_1}^* q_{mt_2}$ respectively. By

taking account the cross-correlation $C_{t_2t_1}$ and the variances C_{diag} of q_{mt_1} and q_{mt_2} , the real and imaginary parts of q_{mt_1} and q_{mt_2} can be modeled as

$$q_{m1,I} = \sqrt{C_{\text{diag}}}X_I, \quad (\text{E.6})$$

$$q_{m1,Q} = \sqrt{C_{\text{diag}}}X_Q, \quad (\text{E.7})$$

$$q_{m2,I} = \frac{1}{\sqrt{C_{\text{diag}}}} \left(\Re(C_{t_2t_1})X_I \right. \quad (\text{E.8})$$

$$\left. + \sqrt{C_{\text{diag}}^2 - |C_{t_2t_1}|^2}Z_I - \Im(C_{t_2t_1})X_Q \right), \quad (\text{E.9})$$

$$q_{m2,Q} = \frac{1}{\sqrt{C_{\text{diag}}}} \left(\Re(C_{t_2t_1})X_Q \right. \quad (\text{E.10})$$

$$\left. + \sqrt{C_{\text{diag}}^2 - |C_{t_2t_1}|^2}Z_Q + \Im(C_{t_2t_1})X_I \right). \quad (\text{E.11})$$

The cross-correlation between q_{mt_1} and q_{mt_2} is calculated by

$$\begin{aligned} & \text{Var}\{\Im(q_{mt_1}^* q_{mt_2})\} \\ &= \text{E}\{|q_{m1,I}q_{m2,Q} - q_{m1,Q}q_{m2,I} \\ & \quad - \text{E}\{q_{m1,I}q_{m2,Q} - q_{m1,Q}q_{m2,I}\}|^2\} \\ &= \text{Var}\{q_{m1,I}q_{m2,Q}\} + \text{Var}\{q_{m1,Q}q_{m2,I}\} \\ & \quad - 2\text{cov}(q_{m1,I}q_{m2,Q}, q_{m1,Q}q_{m2,I}) \end{aligned} \quad (\text{E.12})$$

By using equations (E.6) to (E.11), and after some simple derivations, we obtain

$$\begin{aligned} & \text{Var}\{\Im(q_{mt_1}^* q_{mt_2})\} \\ &= \frac{1}{2} \left(C_{\text{diag}}^2 + \Im(C_{t_2t_1})^2 - \Re(C_{t_2t_1})^2 \right). \end{aligned} \quad (\text{E.13})$$

Similarly, the variance of the real part is

$$\begin{aligned} & \text{Var}\{\Re(q_{mt_1}^* q_{mt_2})\} \\ &= \frac{1}{2} \left(C_{\text{diag}}^2 - \Im(C_{t_2t_1})^2 + \Re(C_{t_2t_1})^2 \right). \end{aligned} \quad (\text{E.14})$$

According to the structure of the auto-correlation of quantization noise, there is no cross antenna correlation. The Central limit theorem [citation] guarantees when M is large, $\frac{1}{M}\mathbf{q}_{p1}^H\mathbf{q}_{p2}$ is approximately Gaussian distributed with the variances

$$\begin{aligned} & \text{Var}\left\{\Im\left(\frac{1}{M}\mathbf{q}_{p1}^H\mathbf{q}_{p2}\right)\right\} \\ &= \frac{1}{2M}\left(C_{\text{diag}}^2 + \Im(C_{t_2t_1})^2 - \Re(C_{t_2t_1})^2\right), \end{aligned} \quad (\text{E.15})$$

$$\begin{aligned} & \text{Var}\left\{\Re\left(\frac{1}{M}\mathbf{q}_{p1}^H\mathbf{q}_{p2}\right)\right\} \\ &= \frac{1}{2M}\left(C_{\text{diag}}^2 - \Im(C_{t_2t_1})^2 + \Re(C_{t_2t_1})^2\right). \end{aligned} \quad (\text{E.16})$$

- iv) The rest of the terms are all CSCG. Following the similar derivation as in i) and ii), we can obtain the results.

Appendix F

Proof of Proposition 6.3

For notational simplicity, we assume that $t_2 = 2$ and $t_1 = 1$ in the derivation, so that $C_{t_2 t_1} = C_{21}$.

i) Finding the mean of $\sum_{i=2}^9 \zeta'_{i,Q}$:

As ζ_2, \dots, ζ_8 are CSCG with zero mean and $\zeta'_2, \dots, \zeta'_9$ are just scaled and rotated versions, the expectations of $\zeta'_2, \dots, \zeta'_8$ are still zero. We take the expectation of the result in (6.41) over channel realizations and obtain

$$\begin{aligned} \mathbb{E} \left\{ \sum_{i=2}^9 \zeta'_{i,Q} \right\} &= \mathbb{E} \{ \zeta'_{9,Q} \} \\ &= \frac{M}{|\phi|^2 a_p^2} (\Im(C_{21}) \cos \omega - \Re(C_{21}) \sin \omega) \mathbb{E} \left\{ \frac{1}{\|\mathbf{g}\|^2} \right\} \\ &= \frac{M}{|\phi|^2 a_p^2 \beta} (\Im(C_{21}) \cos \omega - \Re(C_{21}) \sin \omega) \mathbb{E} \{ \text{tr} \{ \mathbf{L}^{-1} \} \}, \end{aligned}$$

where we denote $\mathbf{L} \sim \mathcal{W}_1(M, \mathbf{I})$ is a Wishart matrix. The property of the Wishart matrix guarantees [106]

$$\mathbb{E}\left\{\text{tr}\{\mathbf{L}^{-1}\}\right\} = \frac{1}{M-1} \approx \frac{1}{M} \quad (\text{F.1})$$

$$\mathbb{E}\left\{\text{tr}\{\mathbf{L}^{-2}\}\right\} = \frac{M}{(M-1)^3 - (M-1)} \approx \frac{1}{M^2}, \quad (\text{F.2})$$

where the approximations hold as the number of antennas is large. Thus the mean of $\sum_{i=2}^9 \zeta'_{i,Q}$ is

$$\mathbb{E}\left\{\sum_{i=2}^9 \zeta'_{i,Q}\right\} = \frac{1}{|\phi|^2 a_p^2 \beta} (\Im(C_{21}) \cos \omega - \Re(C_{21}) \sin \omega).$$

ii) Finding the variance of $\sum_{i=2}^9 \zeta'_{i,Q}$:

Due to the correlation in quantization noise, ζ'_5 and ζ'_6 are correlated, while other terms are uncorrelated. We obtain that

$$\text{Var}\left\{\sum_{i=2}^9 \zeta'_{i,Q}\right\} = \sum_{i=2}^9 \text{Var}\{\zeta'_{i,Q}\} + 2\text{cov}(\zeta'_{5,Q}, \zeta'_{6,Q}) \quad (\text{F.3})$$

a) Derive $\sum_{i=2}^9 \text{Var}\{\zeta'_{i,Q}\}$:

We take the expectation over the channel realization to analyze the overall variance. Thus,

$$\begin{aligned}
& \sum_{i=2}^9 \text{Var}\{\zeta'_{i,Q}\} \\
&= \sum_{i=2}^8 \frac{1}{2} \text{Var}\{\zeta'_i\} + \text{Var}\{\zeta'_{9,Q}\} \\
&= \text{E} \left\{ \frac{\sigma_n^4}{2|\phi|^4 \|\mathbf{g}\|^4} + \frac{\sigma_n^2}{|\phi|^2 \|\mathbf{g}\|^2} + \frac{C_{\text{diag}}}{a_p^2 |\phi|^2 \|\mathbf{g}\|^2} + \frac{\sigma_n^2 C_{\text{diag}}}{a_p^2 |\phi|^4 \|\mathbf{g}\|^4} \right. \\
&\quad \left. + \frac{M}{2|\phi|^4 a_p^4 \|\mathbf{g}\|^4} (C_{\text{diag}}^2 - (\Im(C_{21})^2 - \Re(C_{21})^2) \cos 2\omega) \right\}. \\
&= \text{E}\{\text{tr}(\mathbf{L}^{-1})\} \left(\frac{\sigma_n^2}{|\phi|^2 \beta} + \frac{C_{\text{diag}}}{a_p^2 |\phi|^2 \beta} \right) + \text{E}\{\text{tr}(\mathbf{L}^{-2})\} \left(\frac{\sigma_n^4}{2|\phi|^4 \beta^2} + \frac{\sigma_n^2 C_{\text{diag}}}{a_p^2 |\phi|^4 \beta^2} \right. \\
&\quad \left. + \frac{M}{2|\phi|^4 a_p^4 \beta^2} (C_{\text{diag}}^2 - (\Im(C_{21})^2 - \Re(C_{21})^2) \cos 2\omega) \right).
\end{aligned}$$

By applying (F.1) and (F.2) to above equation, we obtain

$$\begin{aligned}
& \sum_{i=2}^9 \text{Var}\{\zeta'_{i,Q}\} \\
&= \frac{1}{M} \left(\frac{\sigma_n^2}{|\phi|^2 \beta} + \frac{C_{\text{diag}}}{a_p^2 |\phi|^2 \beta} \right) + \frac{1}{M^2} \left(\frac{\sigma_n^4}{2|\phi|^4 \beta^2} + \frac{\sigma_n^2 C_{\text{diag}}}{a_p^2 |\phi|^4 \beta^2} \right. \\
&\quad \left. + \frac{M}{2|\phi|^4 a_p^4 \beta^2} (C_{\text{diag}}^2 - (\Im(C_{21})^2 - \Re(C_{21})^2) \cos 2\omega) \right).
\end{aligned}$$

b) Calculate $\text{cov}(\zeta'_{5,Q}, \zeta'_{6,Q})$:

As ζ'_5 and ζ'_6 are zero mean, and \mathbf{g} is CSCG, we can obtain

$$\text{cov}(\zeta'_5, \zeta'_6) \text{E} \left\{ \frac{\phi^2}{|\phi|^4 a_p^2 \beta^2 M^2} \mathbf{q}_2^H \mathbf{g} \mathbf{q}_1^H \mathbf{g} e^{j3\omega} \right\} = 0, \quad (\text{F.4})$$

and

$$\text{E}\{\zeta'_{5,Q} \zeta'_{6,Q}\} = -\text{E}\{\zeta'_{5,I} \zeta'_{6,I}\}. \quad (\text{F.5})$$

The real part of pseudo-correlation between $\zeta'_{5,Q}$ and $\zeta'_{6,Q}$ is

$$\Re(\mathbb{E}\{\zeta'_5 \zeta'_6\}) = -2\mathbb{E}\{\zeta'_{5,Q} \zeta'_{6,Q}\}. \quad (\text{F.6})$$

Thus, we obtain the co-variance

$$\text{cov}(\zeta'_{5,Q}, \zeta'_{6,Q}) = -\frac{1}{2}\Re(\mathbb{E}\{\zeta'_5 \zeta'_6\}) \quad (\text{F.7})$$

$$= -\frac{1}{2|\phi|^2 a_p^2 \beta M} (\Re(C_{21}) \cos \omega - \Im(C_{21}) \sin \omega). \quad (\text{F.8})$$