TRANSCEIVER OPTIMIZATION FOR BROADBAND COOPERATIVE WIRELESS COMMUNICATION SYSTEMS

by

Peiran Wu

M. Phil., The City University of Hong Kong, 2010
B. Eng., Sun Yat-sen University, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

November 2015

Abstract

Cooperative broadband communications is a promising technique to improve the reliability, throughput, and coverage of the next generation wireless communication systems. Single-carrier transmission with frequency-domain equalization (SC-FDE) and orthogonal frequency-division multiplexing (OFDM) are the prevailing block based broadband schemes widely adopted in major wireless standards. Traditionally, these broadband schemes are deployed for point-to-point communications without the cooperation of any intermediate transmission nodes. However, as the communication systems evolve, both service providers and users are demanding higher data rates and non-seamless connection over large areas. As a result, it is necessary to design novel transceiver architectures that meet these stringent requirements. This dissertation proposes four such cooperative transceiver designs which are tailored for different communication scenarios.

Firstly, for single-user SC-FDE broadband systems with multiple multi-antenna amplify-and-forward (AF) relays, we optimize the relay beamforming (rBF) filters and destination equalization filter based on the minimum mean-square error (MSE) criterion under an aggregate relay transmit power constraint. We also propose suboptimal rBF schemes which perform close to the optimal rBF scheme. Subsequently, we investigate single-user SC-FDE broadband MIMO AF relay systems. By exploiting the properties of the block-circulant channel matrices and the majorization theory, the problem is transformed into an equivalent power optimization problem with

scalar variables. An alternating optimization algorithm is devised to obtain the optimal solution for the source and relay power allocation. Thirdly, we study a robust transceiver design for multiuser broadband systems with multiple single-antenna AF relays and in the presence of channel estimation errors. Our proposed design treats multiuser SC-FDE and OFDM based systems in a unified manner, where the goal is to maximize the network ABR subject to different types of relay power constraints. Lastly, we propose a robust transceiver design for single-user SC-FDE based multi-hop full-duplex decode-and-forward relay systems. The optimization problem is formulated as the minimization of the sum MSE or maximum MSE of different hops which takes into account the loopback interference of the full-duplex relays. We propose two algorithms to solve the resulting non-convex power allocation problems based on sequential geometric programming and alternating optimization, respectively.

Preface

Chapters 2–5 are based on works under the supervision of Professor Robert Schober and Professor Vijay K. Bhargava.

For all chapters, I conducted the paper survey on related topics, formulated the problems, proposed problem solutions, and performed the analysis and the simulations of the considered communication systems. I also wrote all paper drafts.

Two papers related to Chapter 2 have been published:

- P.Wu and R.Schober, "Cooperative beamforming for single-carrier frequency-domain equalization systems with multiple relays," *IEEE Trans. Wireless Commun.*, vol.11, pp.2276–2286, Jun. 2012.

- P.Wu and R.Schober, "Cooperative frequency-domain beamforming for broadband SC-FDE systems," *IEEE GlobeCom 2011*, Houston, TX, U.S.A, Dec. 2011.

Two papers related to Chapter 3 have been published:

- P.Wu, R.Schober, and V.K.Bhargava, "Transceiver design for SC-FDE based MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol.12, pp.4376–4391, Sept. 2013.

- P.Wu, R.Schober, and V.K.Bhargava, "Joint transceiver design for MIMO relay systems employing SC-FDE," *IEEE GlobeCom 2012*, Anaheim, CA, U.S.A, Dec. 2012.

Two papers related to Chapter 4 have been published:

- P.Wu, R.Schober, and V.K.Bhargava, "Robust transceiver design for broadband multiuser multi-relay networks," *IEEE Trans. Wireless Commun.*, vol.14, pp. 5643–5658, Oct. 2015.
- P.Wu, R.Schober, and V.K.Bhargava, "Robust cooperative beamforming for SC-FDMA based multi-relay networks," *IEEE ICC 2014*, Sydney, Australia, Jun. 2014.

Two papers related to Chapter 5 have been published:

- P.Wu, R.Schober, and V.K.Bhargava, "Robust transceiver design for SC-FDE multi-hop full-duplex decode-and-forward relaying systems," *IEEE Trans. Wireless Commun.*, vol.PP, no.99, Sept. 2015.
- P.Wu, R.Schober, and V.K.Bhargava, "Robust MMSE design for full-duplex decode-and-forward SC-FDE relay systems," *IEEE ICC 2015*, London, UK, Jun. 2015.

Other contributions not presented in this dissertation: Some research works not directly related to this dissertation have been published during my time as a Ph.D. student at UBC in the following three letters.

- P.Wu, R.Schober, and V.K.Bhargava, "Optimal power allocation for wideband cognitive radio networks employing SC-FDMA," *IEEE Commun. Lett.*, vol.17, pp.669–672, Apr. 2013.
- P.Wu, R.Schober, and V.K.Bhargava, "Optimal Tx-BF for MIMO SC-FDE systems," *IEEE Commun. Lett.*, vol.17, pp.1509–1512, Aug. 2013.

- P.Wu, R.Schober, and V.K.Bhargava, "Rate optimization for hybrid SC-FDE/ OFDM decode-and-forward MIMO relay systems," *IEEE Commun. Lett.*, vol.18, pp.1210–1213, Jul. 2014.

Table of Contents

Appendices

List of Tables

List of Figures

List of Abbreviations

3GPP	3rd Generation Partnership Project
ABR	Achievable Bit Rate
AF	Amplify–and–Forward
Agg-PC	Aggregate Power Constraint
AMSE	Arithmetic Mean–Square Error
AO	Alternating Optimization
AWGN	Additive White Gaussian Noise
BER	Bit–Error Rate
BF	Beamforming
BS	Base Station
CCDF	Complementary Cumulative Density Function
CF	Compress–and–Forward
CIR	Channel Impulse Response
CP	Cyclic Prefix
CSI	Channel State Information
dEQ	Destination Equalization
DF	Decode–and–Forward
DFE	Decision–Feedback Equalization
DFT	Discrete Fourier Transform
Direc-Tx	Direct Transmission

EPA	Equal Power Allocation
EPA-T	Equal Power Allocation Per–Tone
EPA-TR	Equal Power Allocation Per–Tone Per–Relay
EVD	Eigen–Value Decomposition
FBF	Feedback Filter
FD	Frequency Domain
FDD	Frequency Division Duplex
FDR	Full Duplex Relay
FFF	Feedforward Filter
FFT	Fast Fourier Transform
GMD	Geometric–Mean Decomposition
GMSE	Geometric Mean–Square Error
GP	Geometric Programming
HDR	Half Duplex Relay
Ind-PC	Individual Power Constraint
i.i.d.	Independent and Identically Distributed
ISI	Inter-Symbol Interference
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
KKT	Karush–Kuhn–Tucker
LE	Linear Equalization
LI	Loopback Interference
LTE-A	Long Term Evolution - Advanced
MIMO	Multiple–Input Multiple–Output
MISO	Multiple–Input Single–Output

maxMSE	Maximum Mean–Square Error
MMSE	Minimum Mean–Square Error
MF	Matched Filter
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MSE	Mean Square Error
N-AF	Naive AF
NSD	Negative Semidefinite
OFDM	Orthogonal Frequency–Division Multiplexing
OFDMA	Orthogonal Frequency–Division Multiple Access
OPA	Optimal Power Allocation
PAM	Pulse Amplitude Modulation
PAPR	Peak–to–Average–Power Ratio
PSD	Positive Semdefinite
QoS	Quality of Service
QPSK	Quaternary Phase-Shift Keying
RF	Radio Frequency
rBF	Relay Beamforming
rPB	Relay Power Budget
sBF	Source Beamforming
SC-FDE	Single–Carrier Transmission with Frequency–Domain Equalization
SC-FDMA	Single–Carrier Frequency–Division Multiple Access
SIC	Successive Interference Cancellation
SINR	Signal–to–Interference–plus–Noise Ratio
SNR	Signal–to–Noise Ratio

SVD	Singular–Value Decomposition
TD	Time Domain
TDD	Time Division Duplex

Notation

$\mathbf{A}^T$	Transpose of matrix $\mathbf{A}$
$\mathbf{A}^\dagger$	Hermitian transpose of matrix $\mathbf{A}$
$\mathbf{A}^*$	Complex conjugate of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	Inverse of a square matrix $\mathbf{A}$
$\mathbf{0}_X$	All–zero column vector of length X
$\mathbf{1}_X$	All–one column vector of length X
circ($\bullet$)	A circular matrix
diag($\bullet$)	A diagonal matrix
blkcirc($\bullet$)	A block circular matrix
blkdiag($\bullet$)	A block diagonal matrix
$\mathbb{E}_X\{\cdot\}$	Statistical expectation with respect to random variable X
det($\cdot$)	Matrix determinant
$\left[\,\cdot\,\right]^+$	$\max\{0, x\}, \quad x \in \mathbb{R}$
$x^\star$	The optimal value of x
$\otimes$	The Kronecker product
$\mathbb{C}^{N\times M}$	The space of all $N \times M$ matrices with complex entries
$\lvert\cdot\rvert$	Absolute value of a complex number
$\lVert\cdot\rVert^2$	Euclidean norm
$\lVert\cdot\rVert_F$	Frobenius norm
tr($\cdot$)	Trace of a matrix

$\text{vec}(\cdot)$	Vectorization of a matrix (stacking the columns on top of each other)
$\mathcal{CN}(\mu, \sigma^2)$	A complex Gaussian random variable with mean μ and variance σ^2

Acknowledgments

First and foremost I would like to express my deep and sincere gratitude to my supervisors, Professor Robert Schober and Professor Vijay K. Bhargava, for their support and invaluable advice during my Ph.D. study. Their engineering insights and technical depth in wireless communications, together with the critical ability to connect theory to practice, have been a source of inspiration. I am thankful for the opportunity to work with them and for the cultivating time during my PhD.

Also, I greatly thank the members of my doctoral committee, Prof. Lutz Lampe, Prof. Brian Marcus, and Prof. Shahriar Mirabbasi, and the external examiner, Prof. Lajos Hanzo, for the time and effort in evaluating my work and providing valuable feedback and suggestions.

I enjoyed the intellectual environment UBC harbors and appreciate the interactions with other students and researchers. To this aspect, I must thank all the members of Lab Kaiser 4090, for being dynamic colleagues and for creating a stimulating research environment.

Financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the University of British Columbia, the German Academic Exchange Service (DAAD), and the China Scholarship Council (CSC) is gratefully acknowledged.

Dedication

To My Family

Chapter 1

Introduction

The ever-increasing demands for higher data rates, more reliable communication, and seamless connection from everywhere have always been the major driving forces for the evolution of wireless communication systems [1, 2]. Cooperative communications [3, 4, 5], multiple-input-multiple-output (MIMO) systems [6, 7], broadband communication technologies such as orthogonal frequency division multiplexing (OFDM) [8, 9] and single-carrier transmission with frequency-domain equalization (SC-FDE) [10, 11, 12], are some of the key enabling techniques to accomplish the aforementioned objectives. Hence, we will provide a brief overview of these related techniques in this chapter.

This chapter is organized as follows. In Sections 1.1, we briefly review two major broadband communication schemes, SC-FDE and OFDM, as well as their multiple-access extensions, single-carrier frequency-division multiple access (SC-FDMA) and orthogonal frequency-division multiple access (OFDMA). In Sections 1.2, we discuss the existing cooperative beamforming for multiple relay networks for both single-user and multiuser scenarios. In Section 1.3, we introduce the concept of MIMO relay communication and discuss the associated transceiver optimization problems. In Section 1.4, we provide a brief review on full-duplex relay systems, which have recently received lots of research interest. The contributions made in this thesis are summarized in Section 1.5, and the thesis organization is provided in Section 1.6.

1.1 Broadband Transmission and Multiple-access Schemes

SC-FDE and OFDM, as well as their multiple access extensions, SC-FDMA and OFDMA, are the prevailing block based broadband schemes [14, 15] that have been widely adopted in the major wireless standards, such as the IEEE 802.16 family of standards [17, 18, 19] and 3rd Generation Partnership Project (3GPP) Long-Term Evolution-Advanced (LTE-A) [14, 15, 16]. OFDM/OFDMA is well known for its advantage of flexibility in radio resource allocation and resilience to inter-symbol interference in frequency-selective fading channels [20]-[28]. However, they suffer from a large transmit peak-to-average power ratio (PAPR) and a high sensitivity to carrier frequency offsets. Although many PAPR reduction and carrier frequency offset compensation techniques have been proposed in the literature [29, 30], these methods introduce additional complexities and usually do not guarantee optimal system performance. On the other hand, as an efficient solution to the aforementioned problems, SC-FDE/SC-FDMA has been recognized as an attractive alternative to OFDM/OFDMA for the uplink transmission, providing an improved uncoded bit-error rate (BER) and a similar achievable bit rate (ABR), while enjoying a comparable implementation complexity [11, 12, 13]. Therefore, this thesis will mainly focus on the discussion of SC-FDE/SC-FDMA based broadband systems and consider various transceiver schemes employing this type of broadband technology. It is worth mentioning that we do consider SC-FDMA and OFDMA based broadband systems in a unified manner in Chapter 4 .

1.2 Cooperative Beamforming with Multiple Relays

Next generation wireless systems are required to provide a certain quality-of-service (QoS) to their users, including those at the cell edge. In this context, relays have been recognized as an effective means to extend the coverage of wireless networks [3, 4, 5, 31]. The three most popular relaying techniques are amplify-and-forward (AF), compress-and-forward (CF), and decode-and-forward (DF). Compared to CF and DF relays, AF relays have the advantage of simpler signal processing and end-to-end transparency of data transmission [3, 31]. If the channel state information (CSI) is available at the relays, cooperative relay beamforming (rBF) [32, 33, 34, 35] across multiple parallel AF relays can be employed to create a virtual antenna array, which introduces additional spatial diversity and yields considerable performance gains.

1.2.1 Cooperative Relay Beamforming with Perfect CSI

With perfect CSI available at the relay nodes and the destination node, and for flat fading channels, the optimal rBF filters which maximize the received signal-to-noise ratio (SNR) and the ABR of the network under individual relay power constraints (Ind-PCs) were reported in [36] and [37], respectively. The rBF design minimizing the received mean-square error (MSE) under Ind-PCs was investigated in [38, 39]. In high-rate data communications, the channels become frequency-selective and significant inter-symbol interference (ISI) is present at the receiver. To reduce the performance degradation caused by ISI, the design of filter-and-forward rBF for a continuous SC transmission with and without time-domain (TD) equalization at the destination under an aggregate relay power constraint (Agg-PC) was investigated in [41] and [42], respectively. For frequency-selective fading channels, optimal rBF across multiple relays was also studied for OFDM with ergodic capacity as the opti-

mality criterion in [40, 43, 44]. As pointed out earlier in this chapter, OFDM is not a good option if low transmit PAPRs are desired. On the other hand, for channels with large length of channel impulse responses (CIRs), TD-BF and TD equalization introduce a high complexity in systems employing continuous SC transmission [10, 11]. In this regard, block-based SC-FDE is appealing as it enables both low PAPR and low-complexity frequency-domain processing for efficient mitigation of frequency-selective fadings [12, 13]. Nevertheless, the literature on cooperative BF for SC-FDE is very sparse, with [45] being the only notable exception, where the application of FD linear equalization (FD-LE) has been considered in a multi-user uplink network with a single-antenna relay station. However, [45] imposes a predetermined linear equalization structure on the relay and only optimizes the power allocation across different frequency tones. In addition, in future wireless systems, multiple multi-antenna relays may cooperate for improving the link quality and the extension of [45] to this scenario is not straightforward. This motivates us to consider a systematic transceiver design for a multiple multi-antenna AF relay network with both linear and nonlinear equalization receivers in Chapter 2.

1.2.2 Cooperative Relay Beamforming with Imperfect CSI

The rBF schemes mentioned so far are based on the assumption that perfect CSI is available at the relays and the destination. In practical wireless systems, CSI is usually imperfect due to channel estimation errors and/or feedback quantization errors [46]. Two different approaches are commonly used to model imperfect CSI, namely, statistical models and deterministic models. In the former case, the statistics of the CSI errors are assumed to follow some known distribution such as Gaussian, which makes this model suitable for modeling channel estimation errors [47]. In the latter

case, the CSI error is assumed to lie in an uncertainty region with known boundary, making this model suitable for the characterization of quantization errors [48]. There are several works on robust rBF design, which have adopted one of these two models. For example, the authors of [82] investigated the optimal rBF filter which maximizes the received SNR of a multi-relay network based on the deterministic CSI error model. In [50], the optimal minimum MSE relay precoding scheme was developed for a multiple-antenna single-relay system based on the statistical CSI error model. The relay precoding matrix design for a multi-relay network was studied in [51] for both of these error models. However, all these works are only applicable to flat fading channels, where the optimization is essentially performed on a per-subcarrier basis. The extension of robust rBF designs to multiuser multi-relay broadband systems in frequency-selective channels is more challenging as the dimensionality of the resulting problem is much higher. We note that for perfect CSI, the rBF optimization for SC-FDMA multi-relay systems with Ind-PCs has been addressed in [52], where the objective is the minimization of the maximum user-wise MSE in the network. However, for a multiuser network, the ultimate system performance depends on the aggregate ABR, which is a complex nonlinear function of the user-wise MSEs, and thus cannot be optimized using the formulation in [52] . To the best of our knowledge, a robust rBF filter design which directly maximizes the network ABR for SC-FDMA and OFDMA under either an Agg-PC or Ind-PCs has not been investigated in the literature yet. This motivates us to investigate a unified transceiver scheme, which includes the robust design of rBF filters and destination equalization (dEQ) filter, for SC-FDMA/OFDMA based multiuser broadband systems with multiple AF relays in Chapter 4.

1.3 MIMO Relay Systems

In the previous section, we have discussed various rBF schemes for both single-user and multi-user multi-relay systems. A common aspect of these schemes is that only a single spatial data stream is transmitted through the communication link [53]. For the multi-user case, each user only transmits a single data stream. However, future wireless communication systems are expected to support very high-speed data rates with limited available bandwidth resources. Spatial multiplexing [54, 55], which could transmit multiple data streams through multiple spatial channels created by the MIMO systems, is one of the key technique for achieving this goal. Therefore, MIMO relay systems [56, 57], where multiple antennas are equipped at source, relay and destination nodes, have received significant research interest for exploiting the benefits of both multiple-antenna and relay technologies. An important research problem for MIMO relay systems is the design of optimal node processing matrices to improve spectral efficiency and/or error performance through efficient utilization of transmit CSI [58, 59]. For example, assuming availability of perfect CSI at the source and relay nodes and linear processing at the destination, the source and relay processing matrices were optimized for maximization of the relay channel capacity and minimization of the MSE in [60, 61] and [62, 63], respectively. In [64], a general framework for linear transceiver optimization in MIMO AF relay systems was provided for a large family of objective functions, which includes the capacity maximizing and the MSE minimizing designs as special cases. The extension of the results in [64] to multi-hop MIMO AF relay systems with linear and decision-feedback equalization receivers was investigated in [66] and [67], respectively. More recently, the design of MIMO AF relay systems with partial or imperfect CSI at source and relay was considered in [68, 50].

Existing works on transceiver design for MIMO AF relay systems are based on the assumption of either frequency-nonselective (flat) channels [61, 62, 66, 67, 50] or frequency-selective channels in combination with OFDM [60, 64, 68, 65]. Since OFDM decomposes a frequency-selective channel into multiple parallel flat subchannels [70, 121], the transceiver designs developed for frequency-nonselective channels can be extended to OFDM based MIMO relay systems by solving an additional subcarrier power allocation problem across different subcarriers. However, the optimization of SC-FDE based MIMO relay systems has not been considered in the literature so far. The main difference of MIMO SC-FDE systems compared to MIMO OFDM systems lies in the special structure of its MSE matrix in the time domain, which is the average of the subcarrier-wise MSE matrices of the corresponding MIMO OFDM systems in the frequency domain [72, 73]. Therefore, the design problem for MIMO SC-FDE systems has a very different structure compared to that for MIMO OFDM systems even in point-to-point non-cooperative communication scenarios. This motivates us to study the joint transceiver design for SC-FDE based MIMO AF relay systems in Chapter 3.

1.4 Full-duplex Relay Systems

Conventionally, relay systems are operated in the half-duplex mode, where two orthogonal time slots are required to accomplish the signal transmission [31]. This would result in a low spectrum efficiency, especially when there is a large number of hops in the relay network. Recently, due to the promising advances in hardware technology for interference cancellation [74, 75, 76, 77], full-duplex relays (FDRs) that can transmit and receive signals simultaneously are drawing a growing interest from the research community. The key enabling technique for full-duplex radios is

the so-called self-interference cancellation. The authors of [78] investigated several signal cancellation techniques, including natural isolation, time-domain cancellation, and spatial-domain suppression, to alleviate the loopback interference (LI) inherent to FDRs. In [79], a spatial-domain nulling scheme was proposed for cancelling the echo FDR channel interference. Transmit beamforming and power allocation are also effective means to enhance the performance of FDR systems. In [80], a power allocation scheme was employed to maximize the throughput of hybrid half-duplex relay (HDR) and FDR systems. Power allocation for minimization of the outage probability of a cognitive radio FDR system was studied in [81]. Taking into account the effects of imperfect LI cancellation, the authors of [82] studied the primary-cognitive rate region of a cooperative cognitive system by deriving optimal and suboptimal beamforming schemes for the cognitive FDR. In terms of performance analysis, the diversity-multiplexing tradeoff and outage performance of FDR systems were studied in [83] and [84], respectively. To determine the optimal duplex mode, the authors of [85] studied the SNR outage probability for a DF relay system and derived the condition under which FDRs outperform HDRs. More recently, a low-complexity joint precoding and decoding strategy based on the zero-forcing criterion was proposed in [87] for a multi-antenna FDR system.

The aforementioned works on FDR systems are all based on the assumption of flat-fading channels. In the presence of frequency-selective fading, the design of FDR systems becomes even more challenging as the signal received at the relay suffers not only from LI, but also from inter-symbol interference [88]. Two prevailing solutions to overcome channel frequency selectivity are OFDM and SC-FDE, and both approaches have found successful application in HDR system design, e.g., [60, 89, 90]. Unfortunately, these HDR designs are not applicable to FDR systems due to the non-

negligible residual LI at the relay, which couples the transmit signals of consecutive hops and makes the design problem more difficult. For OFDM based FDR systems, some results for system optimization have been reported recently. For example, [91] investigated an adaptive interference cancellation scheme for OFDM FDR systems which aims to minimize the residual LI power assuming perfect knowledge of the LI channel at the relay. In [92], a resource allocation scheme was proposed to maximize the bit rate of a hybrid HDR/FDR system assuming imperfect LI channel estimation but perfect transmit channel estimation. Nevertheless, for SC-FDE based FDR systems, the corresponding transceiver design problem has not been investigated in the literature yet. This motivates us to study the robust transceiver design for SC-FDE based FDR systems assuming imperfect CSI at the transceivers in Chapter 5.

1.5 Contributions of the Thesis

This thesis considers the design of various transceiver schemes for performance enhancement that may find application in several current or upcoming wireless communication standards. The main contributions of this thesis are listed in the following.

1. We propose a novel frequency-domain approach to rBF design for SC-FDE based broadband systems with one single-antenna source, multiple multi-antenna AF relays, and one single-antenna destination. The proposed scheme could effectively reduce the inter-symbol interference originated from the multi-path fading and the inter-relay interference resulted from the concurrent transmissions of multiple relays. We also propose suboptimal rBF schemes which perform close to the optimal rBF scheme and show a remarkable robustness against imperfect CSI.

2. We propose a joint transceiver design for an MIMO SC-FDE based AF relay system with multi-antenna source, relay, and destination. Exploiting the mathemat-

ical tools from majorization theory, we find the optimal structure of the source and relay precoding matrices for minimizing a family of Schur-convex and Schur-concave objective functions. The remaining source and relay power allocation problems are solved by adopting an alternating optimization algorithm based on a high SNR approximation of the objective function. Furthermore, we propose low-complexity suboptimal precoding schemes with less signalling overhead at the cost of only a moderate performance degradation.

3. We investigate the robust transceiver design for a broadband multiuser network with multiple single-antenna source, multiple single-antenna AF relays, and one single-antenna destination, in the presence of channel estimation errors. The proposed unified transceiver scheme for both SC-FDMA and OFDMA maximizes the lower bound of the network ABR under different types of power constraints at the relay nodes, namely Ind-PCs and an Agg-PC. For the Agg-PC, we obtain closed-form solutions for the rBF filter coefficients and global optimal solution for the relay power allocation. For Ind-PCs, we find a local optimal solution of the rBF problem by reformulating the rBF problem into a reverse-convex problem and by applying the constrained convex-concave procedure (CCCP).

4. We study the robust transceiver design for a multi-hop full-duplex DF relay system. Two relevant objective functions are considered for the purpose of transceiver optimization: (1) minimization of the sum MSE of different hops. (2) minimization of the maximum MSE across different hops. For both cases, we optimize the transmit precoding coefficients and receive equalization coefficients under separate node transmit power constraints. We find that the the optimal equalization filters at the receiving nodes are identical for both considered optimization criteria and take the form of robust Wiener filters. We further propose two efficient algorithms to solve the

remaining non-convex optimization problems for the transmit precoding coefficients. The first scheme is based on the successive geometric programming algorithm, and the second scheme combines the different-of-convex programming algorithm and the alternating optimization algorithm to find a stationary point solution to the original problem.

1.6 Organization of the Thesis

In the following, we provide a brief overview of the remainder of this thesis.

In Chapter 2, we propose a novel single-data stream, frequency-domain rBF scheme for SC-FDE broadband systems. Assuming perfect CSI at the relays, the rBF filter coefficients are optimized for three different receiver structures at the destination, namely, FD-LE and FD decision feedback equalization (DFE) as well as an idealized matched filter (MF) receiver, which constitutes a performance upper bound for any receiver. Simulation results for a typical frequency-selective fading channels confirm the excellent performance of the proposed scheme and show that proposed FD rBF has a more favorable performance than Naive-rBF schemes and multi-antenna source BF scheme in conventional non-cooperative communication systems.

In Chapter 3, we consider the transceiver optimization for an SC-FDE based MIMO AF relay networks with one multi-antenna source, one multi-antenna AF relay, and one multi-antenna destination. Assuming perfect CSI at all nodes in the system, the transmit precoding matrices at the source and the relay are jointly optimized for the minimization of a general objective functions in terms of the stream-wise MSE. We solve the associated optimization problems in two stages. In the first stage, we find the optimal structure of the precoding matrices by exploiting the property of the MSE matrix and the tools from majorization thoery. For Schur-concave objective

function, the optimal precoding matrices diagonalize the end-to-end effective MIMO relay channel matrix. For Schur-convex objective function, the optimal precoding matrices diagonalize the effective channel up to a unitary rotation at the source precoder. In the second stage, an efficient numerical algorithm based on alternating optimization principle is proposed for finding the optimal power allocation at the source and the relay. Furthermore, we propose efficient low-complexity suboptimal methods for the source and relay precoding matrices which results in less signalling overhead at the cost of moderate performance degradation. Simulation results show the benefits of having multiple antennas at the source, relay and destination and illustrate the excellent BER and ABR performance of the proposed optimal/suboptimal precoding schemes.

In Chapter 4, we investigate robust transceiver schemes for multiuser multi-relay networks . In contrast to prior work, which concentrated on system design with perfect CSI, we consider here a more piratical scenario where the CSI available at the relays and the destination are imperfect. For the power constraints at the relays, we investigate two different cases: (1) the Agg-PC and (2) Ind-PCs. For both cases, we optimize the rBF filter coefficients and dEQ filter coefficients for maximization a lower bound on the ABR of the network. The optimal dEQ filters and the phases of the optimal rBF filter coefficients are derived for both types of constraints. Then, for the Agg-PC, a two-step solution for the optimal amplitudes of the rBF filter coefficients is proposed by employing primal decomposition. In the first step, for each subcarrier of the users, the optimal rBF allocates power across relays by taking into account the CSI error variances. In the second step, a convex power allocation procedure further distributes the power across subcarriers and among users to enhance the weighed ABR performance. For Ind-PCs, the optimal amplitudes of the rBF filter coefficients

are obtained by solving a sequence of convex optimization problems. Our simulation results reveal that the proposed robust rBF schemes outperform conventional non-robust scheme remarkably, especially with large values of CSI error variances, and also achieves a significant ABR performance gain over the naive-AF relaying scheme.

In Chapter 5, we consider robust transceiver design for a multi-hop SC-FDE based full-duplex DF relay networks with channel uncertainty. For the performance metrics, we consider (1) the sum MSE of different hops. (2) the maximum MSE of different hops. For both problems, we propose two efficient algorithms to find the solution of transmit precoding/power allocation coefficients. The first scheme is based on the successive geometric programming algorithm, and the second scheme combine the different-of-convex programming algorithm and alternating optimization algorithm to obtain a stationary point solution to the original problem. Numerical results show that both approaches yield practically identical MSE and ABR performances. In addition, the proposed robust transceiver design outperforms non-robust FDR systems and robust half-duplex relay systems by a considerable margin, especially in high SNR region and with large channel estimation errors.

Finally, Chapter 6 summarizes the contributions of this thesis and outlines areas of future research. Appendices contain the proofs of the propositions, lemmas, and theorems.

Chapter 2

Transceiver Design For SC-FDE Systems with Multiple Relays

2.1 Introduction

As pointed out in Chapter 1, next generation wireless systems are required to guarantee a certain level of QoS for the cell-edge users. In this context, relays have been recognized as an effective means to achieve this goal [3, 4, 5, 31]. If the channel state information is available at the relays, cooperative rBF [32, 33, 34, 35] across multiple parallel AF relays can be employed to create a virtual antenna array, which introduces additional spatial diversity and yields considerable performance gains.

In this chapter, we propose a systematic transceiver design for cooperative broadband SC-FDE networks with multiple multi-antenna AF relays. At the destination, we consider both FD-LE and FD-DFE as well as an idealized MF receiver, which constitutes a performance upper bound for any receiver. Assuming the equalizers at the destination are optimized based on the minimum MSE criterion, we develop a unified framework for optimization of the rBF matrices, which is applicable for all considered receivers. In particular, for an aggregate relay power constraint, we derive a closed-form solution for the structure of the optimal rBF matrices which is identical for all considered receivers. In contrast, the optimal power allocation across frequency tones depends on the adopted receiver. We develop a numerical algorithm

for optimal power allocation and two suboptimal power allocation schemes assigning equal powers to all relays and/or frequency tones. These suboptimal power allocation schemes are shown to achieve near-optimal performance for SC-FDE and constitute viable alternatives to more complex optimal power allocation.

The remainder of this chapter is organized as follows. In Section 2.2, the system model is presented. In Section 2.3, the optimal minimal MSE receivers are provided and the corresponding objective functions for rBF matrix optimization are derived. The structure of the optimal rBF matrices and power allocation schemes are developed in Section 2.4, and simulation results are provided in Section 2.5. Some conclusions are drawn in Section 2.6.

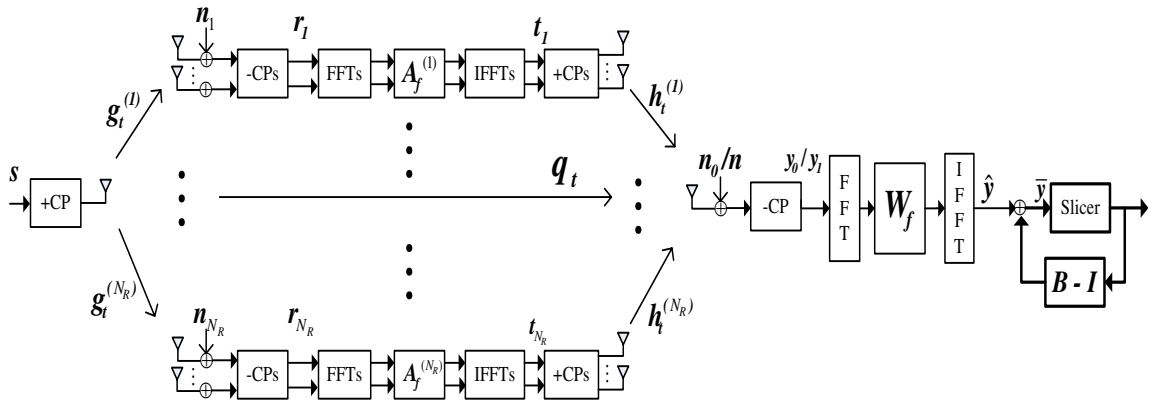

Figure 2.1: System model for SC-FDE block transmission with cooperative relay BF.

2.2 System Model

We consider a block-based SC-FDE system with one source node, S, N_R relays, R_i, $i = 1, \ldots, N_R$, and one destination node, D, as shown in Fig. 2.1. We assume that S and D are equipped with a single antenna, respectively, and the relays are equipped with N antennas. The CIR coefficients of the S-D, S-R_i, and R_i-D channels

are collected in vectors $\mathbf{q}_t = [q_0, q_1, \ldots, q_{L_q-1}]^T$, $\mathbf{g}_t^{(i)} = [\mathbf{g}_0^{(i)T}, \mathbf{g}_1^{(i)T}, \ldots, \mathbf{g}_{L_g-1}^{(i)T}]^T$, and $\mathbf{h}_t^{(i)} = [\mathbf{h}_0^{(i)}, \mathbf{h}_1^{(i)}, \ldots, \mathbf{h}_{L_h-1}^{(i)}]^T$, respectively, where L_q, L_g, and L_h are the corresponding CIR lengths. Here, $\mathbf{g}_l^{(i)} = [g_{i1l}, \cdots, g_{iNl}]^T \in \mathbb{C}^{N\times 1}$ and $\mathbf{h}_l^{(i)} = [h_{i1l}, \cdots, h_{iNl}] \in \mathbb{C}^{1\times N}$, where g_{inl} (h_{inl}) represents the lth tap of the CIR between the nth antenna of R_i and S (D). The transmission is organized in two phases.

2.2.1 First Phase of Transmission

During the first phase, the source node broadcasts its information to the relay nodes and the destination node. Specifically, at the source, N_c independent, identically distributed (i.i.d.) symbols s_n, $n = 1, \ldots, N_c$, which are drawn from a signal constellation with variance $\sigma_s^2 = E[|s_n|^2]$, are collected in a vector $\mathbf{s} = [s_1, \ldots, s_{N_c}]^T$.

A cyclic-prefix (CP) in form of the last $N_{g,s}$ ($N_{g,s} \geq \max\{L_q, L_g\}$) elements of $\mathbf{s}$ is added at the beginning of $\mathbf{s}$ and the resulting signal is transmitted by the source node. The CP transforms the linear convolution of the CIRs and the transmitted signal into a circular convolution. Thus, the received signals (after CP removal) at the destination, $\mathbf{y}_0 \in \mathbb{C}^{N_c\times 1}$, and at the ith relay, $\mathbf{r}_i = [\mathbf{r}_{i1}^T, \ldots \mathbf{r}_{iN_c}^T]^T \in \mathbb{C}^{NN_c\times 1}$ ($\mathbf{r}_{ik}$ denotes the received vector at the ith relay at time k), are given by

$$\mathbf{y}_0 = \mathbf{Q}\mathbf{s} + \mathbf{n}_0 \quad \text{and} \quad \mathbf{r}_i = \mathbf{G}_i\mathbf{s} + \mathbf{n}_i, \quad i = 1, \ldots, N_R, \tag{2.1}$$

respectively, where $\mathbf{Q} = \text{circ}\{[\mathbf{q}_t^T\, \mathbf{0}_{1\times N_c-L_q}]^T\} \in \mathbb{C}^{N_c\times N_c}$ and $\mathbf{G}_i = \text{blkcirc}\{\mathbf{g}_0^{(i)}, \ldots, \mathbf{g}_{L_g-1}^{(i)}, \mathbf{0}_{N\times 1}, \ldots, \mathbf{0}_{N\times 1}\} \in \mathbb{C}^{NN_c\times N_c}$, are column circular and block-circular matrices, respectively, and $\mathbf{n}_0 \sim \mathcal{CN}(\mathbf{0}, \sigma_{n_0}^2\mathbf{I}_{N_c})$ and $\mathbf{n}_i = [\mathbf{n}_{i1}^T, \ldots \mathbf{n}_{iN_c}^T]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_{n_1}^2\mathbf{I}_{NN_c})$ are the noise vectors at the destination in the first phase of transmission and at the ith relay, respectively. We assume that all noise vectors are independent of each other.

2.2.2 Second Phase of Transmission

In the second phase, the relays process the received signals before forwarding them to the destination. Specifically, the ith relay first converts $\mathbf{r}_i$ into the FD using N parallel fast Fourier transforms (FFTs). Subsequently, the FD vector $(\mathbf{F}\otimes\mathbf{I}_N)\mathbf{r}_i$ is multiplied by a block-diagonal FD-BF matrix $\mathbf{A}_f^{(i)} = \text{blkdiag}\{\mathbf{A}_{ik} \in \mathbb{C}^{N\times N},\ k = 1, ..., N_c\}$, where $\mathbf{F}$ is the DFT matrix of size N_c. The resulting FD vector is transformed back into the TD using N parallel inverse FFTs (IFFTs) leading to vector $\mathbf{t}_i = [\mathbf{t}_{i1}^T, \dots \mathbf{t}_{iN_c}^T]^T \in \mathbb{C}^{NN_c\times 1}$,

$$\mathbf{t}_i = \mathbf{A}_i\mathbf{r}_i, \quad i = 1, \dots, N_R, \tag{2.2}$$

where $\mathbf{A}_i = (\mathbf{F}\otimes\mathbf{I}_N)^\dagger\mathbf{A}_f^{(i)}(\mathbf{F}\otimes\mathbf{I}_N)$.[1] Subsequently, a CP of length $N_{g,r}$ ($N_{g,r} \geq L_h$) is added to the vector formed by the nth element of $\mathbf{t}_{ik}$, $k = 1, \dots, N_c$, and the resulting signal is transmitted over the nth antenna of relay R_i. After CP removal, the received signal at the destination in the second phase of transmission can be written as

$$\mathbf{y}_1 = \sum_{i=1}^{N_R} \mathbf{H}_i\mathbf{t}_i + \mathbf{n} = \bar{\mathbf{H}}\mathbf{s} + \bar{\mathbf{n}}, \tag{2.3}$$

where $\mathbf{H}_i = \text{blkcirc}\{\mathbf{h}_0^{(i)}, \dots, \mathbf{h}_{L_h-1}^{(i)}, \mathbf{0}_{1\times N}, \dots, \mathbf{0}_{1\times N}\} \in \mathbb{C}^{N_c\times NN_c}$, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_{n_2}^2\mathbf{I}_{N_c})$, $\bar{\mathbf{H}} = \sum_{i=1}^{N_R} \mathbf{H}_i\mathbf{A}_i\mathbf{G}_i$ represents the effective channel matrix between S and D, and $\bar{\mathbf{n}} = \sum_{i=1}^{N_R} \mathbf{H}_i\mathbf{A}_i\mathbf{n}_i + \mathbf{n}$ is the effective noise vector. Stacking the signals received at

[1] We note that the proposed FD based processing at the relays bears a complexity advantage compared to an equivalent TD based processing. In particular, each FFT and IFFT operation requires $\log_2(N_c)N_c/2$ complex multiplications [93] and the multiplication with block-diagonal matrix $\mathbf{A}_f^{(i)}$ entails N^2N_c complex multiplications. Thus, for the proposed FD based processing, each relay has to perform $\log_2(N_c)NN_c/2 + N^2N_c$ complex multiplications. On the other hand, the direct (TD) multiplication of $\mathbf{A}_i$ and $\mathbf{r}_i$ in (2.2) requires $N^2N_c^2$ complex multiplications. Thus, FD based processing enables considerable computational savings, especially if the number of subcarriers N_c is large.

the destination in both transmission phases into one long vector, we obtain

$$\mathbf{y} = [\mathbf{y}_0^T \quad \mathbf{y}_1^T]^T = \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{n}}, \tag{2.4}$$

where $\tilde{\mathbf{H}} = [\mathbf{Q}^T \quad \bar{\mathbf{H}}^T]^T$ and $\tilde{\mathbf{n}} = [\mathbf{n}_0^T \quad \bar{\mathbf{n}}^T]^T$. It is useful to note that since $\mathbf{Q}$, $\mathbf{G}_i$, and $\mathbf{H}_i$ are circular/block-circular matrices, they admit the following diagonal/block-diagonal decompositions:

$$\mathbf{Q} = \mathbf{F}^\dagger \mathbf{Q}_f \mathbf{F}, \quad \mathbf{G}_i = (\mathbf{F} \otimes \mathbf{I}_N)^\dagger \mathbf{G}_f^{(i)} \mathbf{F}, \quad \mathbf{H}_i = \mathbf{F}^\dagger \mathbf{H}_f^{(i)} (\mathbf{F} \otimes \mathbf{I}_N)^\dagger, \tag{2.5}$$

where $\mathbf{Q}_f$, $\mathbf{G}_f^{(i)}$, and $\mathbf{H}_f^{(i)}$ are diagonal/block-diagonal matrices with the kth diagonal/block-diagonal entry equal to q_k, $\mathbf{g}_{ik} \in \mathbb{C}^{N\times 1}$, and $\mathbf{h}_{ik} \in \mathbb{C}^{1\times N}$, respectively. Here, q_k, $\mathbf{g}_{ik}$, and $\mathbf{h}_{ik}$ represent the frequency responses on the kth frequency tone of the respective channels. Based on these observations, we can decompose $\bar{\mathbf{H}}$ and $\tilde{\mathbf{H}}$ as

$$\bar{\mathbf{H}} = \mathbf{F}^\dagger \bar{\mathbf{H}}_f \mathbf{F} \quad \text{and} \quad \tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{F}^\dagger \mathbf{Q}_f \mathbf{F} \\ \mathbf{F}^\dagger \bar{\mathbf{H}}_f \mathbf{F} \end{bmatrix} = (\mathbf{I}_2 \otimes \mathbf{F})^\dagger \tilde{\mathbf{H}}_f \mathbf{F}, \tag{2.6}$$

where $\bar{\mathbf{H}}_f = \sum_{i=1}^{N_R} \mathbf{H}_f^{(i)} \mathbf{A}_f^{(i)} \mathbf{G}_f^{(i)}$ and $\tilde{\mathbf{H}}_f = [\mathbf{Q}_f^T \quad \bar{\mathbf{H}}_f^T]^T$ are the FD representations of $\bar{\mathbf{H}}$ and $\tilde{\mathbf{H}}$, respectively.

2.2.3 Processing at the Destination

At the destination, the two vector elements of the received signal, $\mathbf{y}$, namely the signal directly received from the source, $\mathbf{y}_0$, and the signal received from the relays, $\mathbf{y}_1$, are converted to the FD domain via FFTs. For this purpose, the same FFT module can be used twice, of course. Subsequently, the FD signal components are filtered by an FD feedforward filter (FD-FFF) $\mathbf{W}_f$ and transformed back to the TD

using an FFT, as shown in Fig. 2.1. The overall TD filter $\mathbf{W}$ can be expressed as

$$\mathbf{W} = \mathbf{F}^{\dagger}\mathbf{W}_f(\mathbf{I}_2 \otimes \mathbf{F}), \tag{2.7}$$

where $\mathbf{W}_f = [\mathbf{W}_f^{(1)} \quad \mathbf{W}_f^{(2)}]$ with $\{\mathbf{W}_f^{(1)}, \mathbf{W}_f^{(2)}\} \in \mathbb{C}^{N_c \times N_c}$ being two diagonal matrices. The diagonal elements of $\mathbf{W}_f^{(1)}$ and $\mathbf{W}_f^{(2)}$ represent the FDE coefficients for the signal received at the destination during the first and second phase of transmission, respectively.

For FD-DFE, the output signal of the FD-FFF after conversion to the TD, $\hat{\mathbf{y}} = \mathbf{W}\mathbf{y}$, is further fed into a symbol-by-symbol decision feedback module, which can be described by a column circular matrix $\mathbf{B} = \text{circ}\{[\mathbf{b}\ \mathbf{0}_{1\times(N_c-N_{fb}-1)}]^T\} \in \mathbb{C}^{N_c \times N_c}$, where $\mathbf{b} = [b_0, b_1, \ldots, b_{N_{fb}}]$ with b_l denoting the feedback coefficient at the lth tap, and N_{fb} is the number of taps of the feedback filter (FBF). Note that b_0 is set to 1 to enable causal ISI cancellation. The circular matrix $\mathbf{B}$ allows the decomposition

$$\mathbf{B} = \mathbf{F}^{\dagger}\mathbf{B}_f\mathbf{F}, \tag{2.8}$$

where the kth diagonal entry of $\mathbf{B}_f$ is given by $b_{f,k}$. As usual, for equalizer design, we assume the decision feedback process to be error free [106, 96], and write the output of the FBF as $\bar{\mathbf{y}} = \hat{\mathbf{y}} - (\mathbf{B} - \mathbf{I}_{N_c})\mathbf{s}$. Note that when $N_{fb} = 0$, $\mathbf{B}$ becomes an identity matrix and FD-DFE degenerates to FD-LE. Therefore, for both FD-LE and FD-DFE, the error vector between the filtered received signal and the desired signal can be expressed as

$$\mathbf{e} = \bar{\mathbf{y}} - \mathbf{s} = \mathbf{W}\mathbf{y} - \mathbf{B}\mathbf{s}. \tag{2.9}$$

2.2.4 Required CSI and Feedback

CSI: In this paper, we assume that the system operates in the time division duplex (TDD) mode such that all channels are reciprocal. Relay R_i, $i = 1, \ldots, N_R$, estimates its own source-relay and relay-destination channels, i.e., g_{ink} and h_{ink}, $\forall n, k$, based on training symbols emitted by the source and the destination, respectively. One node, usually the source or the destination, assumes the role of the central node. The central node acquires the CSI of all links in the network and computes the optimal rBF matrices. Assuming the destination (source) is the central node, for CSI acquisition, the source (destination) emits training symbols and the relays forward the corresponding received signals to the destination (source). This enables the destination (source) to estimate the overall channel coefficients $\bar{h}_{ink} = g_{ink}h_{ink}$, $\forall i, n, k$. Subsequently, each relay emits training symbols and the destination (source) estimates h_{ink} (g_{ink}), $\forall i, n, k$, and calculates $g_{ink} = \bar{h}_{ink}/h_{ink}$ ($h_{ink} = \bar{h}_{ink}/g_{ink}$), $\forall i, n, k$. Throughput this paper, we assume perfect CSI acquisition. The effect of imperfect CSI acquisition will be investigated in Section 2.5.

Feedback: As will be seen in Section 2.4, relay R_i can compute its optimal rBF matrix based on its own local CSI, i.e., g_{ink} and h_{ink}, $\forall n, k$, and $2N_c$ real numbers that have to be fed back by the central node and are identical for all relays. For this purpose, we assume a zero-delay error-free feedback broadcast channel. Similar assumptions are commonly made in the related literature [32, 42].

2.3 Receiver Structures and Objective Functions

In this section, we derive the optimal receive filters for FD-LE, FD-DFE, and the idealized MF receiver. In addition, we define for each considered receiver an objective

function for rBF matrix optimization.

2.3.1 FD-LE and FD-DFE

We first examine the MSE of the considered system, which is related to the diagonal entries of the error covariance matrix $\mathbf{E} = E[\mathbf{e}\mathbf{e}^\dagger]$. Using the definition of $\mathbf{E}$, along with (2.4) and (2.9), we obtain

$$\mathbf{E} = \mathbf{W}\left(\sigma_s^2\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger + \tilde{\mathbf{C}}\right)\mathbf{W}^\dagger - \sigma_s^2\mathbf{W}\tilde{\mathbf{H}}\mathbf{B}^\dagger - \sigma_s^2\mathbf{B}\tilde{\mathbf{H}}^\dagger\mathbf{W}^\dagger + \sigma_s^2\mathbf{B}\mathbf{B}^\dagger, \tag{2.10}$$

where the covariance matrix of the effective noise at D is given by

$$\tilde{\mathbf{C}} \triangleq E[\tilde{\mathbf{n}}\tilde{\mathbf{n}}^\dagger] = \begin{bmatrix} \sigma_{n_0}^2\mathbf{I}_{N_c} & \mathbf{0}_{N_cN_c} \\ \mathbf{0}_{N_cN_c} & \sigma_{n_1}^2\left(\sum_{i=1}^{N_R}\mathbf{H}_i\mathbf{A}_i\mathbf{A}_i^\dagger\mathbf{H}_i^\dagger\right) + \sigma_{n_2}^2\mathbf{I}_{N_c} \end{bmatrix}, \tag{2.11}$$

which can be further decomposed as $\tilde{\mathbf{C}} = (\mathbf{I}_2 \otimes \mathbf{F})^\dagger\tilde{\mathbf{C}}_f(\mathbf{I}_2 \otimes \mathbf{F})$ with

$$\tilde{\mathbf{C}}_f = \begin{bmatrix} \sigma_{n_0}^2\mathbf{I}_{N_c} & \mathbf{0}_{N_cN_c} \\ \mathbf{0}_{N_cN_c} & \sigma_{n_1}^2\left(\sum_{i=1}^{N_R}\mathbf{H}_f^{(i)}\mathbf{A}_f^{(i)}\mathbf{A}_f^{(i)\dagger}\mathbf{H}_f^{(i)\dagger}\right) + \sigma_{n_2}^2\mathbf{I}_{N_c} \end{bmatrix}. \tag{2.12}$$

Inserting (2.7) into (2.10) and differentiating $\mathrm{tr}\{\mathbf{E}\}$ with respect to $\mathbf{W}_f$ and setting the result to zero, we obtain for the optimal FFF

$$\mathbf{W} = \mathbf{F}\mathbf{B}_f\tilde{\mathbf{H}}_f^\dagger\left(\tilde{\mathbf{H}}_f\tilde{\mathbf{H}}_f^\dagger + \tilde{\mathbf{C}}_f\right)^{-1}(\mathbf{I}_2 \otimes \mathbf{F}). \tag{2.13}$$

Using this result in (2.10) and after some matrix manipulations, the MSE matrix is obtained as

$$\mathbf{E} = \sigma_s^2 \mathbf{F} \mathbf{B}_f \mathbf{\Psi}_f^{-1} \mathbf{B}_f^\dagger \mathbf{F}^\dagger, \tag{2.14}$$

where $\mathbf{\Psi}_f$ is a diagonal matrix given by

$$\mathbf{\Psi}_f = \sigma_s^2 \tilde{\mathbf{H}}_f^\dagger \tilde{\mathbf{C}}_f^{-1} \tilde{\mathbf{H}}_f + \mathbf{I}_{N_c}. \tag{2.15}$$

We observe from (2.14) that $\mathbf{E}$ is a circular matrix, i.e., all its diagonal elements are identical. Since the kth diagonal element of $\mathbf{E}$ represents the MSE for the kth transmitted symbol, the MSEs for all transmitted symbols are identical. Furthermore, by exploiting the expression for $\bar{\mathbf{H}}_f$ in (2.6) and $\tilde{\mathbf{C}}_f$ in (2.12), we can show that the diagonal entries of $\mathbf{\Psi}_f$ are given by

$$\Psi_k = \frac{\sigma_s^2 \sum_{i=1}^{N_R} (\mathbf{h}_{ik} \mathbf{A}_{ik} \mathbf{g}_{ik})^2}{\sigma_{n_1}^2 \sum_{i=1}^{N_R} \mathbf{h}_{ik} \mathbf{A}_{ik} \mathbf{A}_{ik}^\dagger \mathbf{h}_{ik}^\dagger + \sigma_{n_2}^2} + \frac{\sigma_s^2}{\sigma_{n_0}^2} q_k^2 + 1, \quad \forall k. \tag{2.16}$$

For FD-LE, we have $\mathbf{B}_f = \mathbf{I}_{N_c}$, which can be used in (2.13) and (2.14) to obtain the LE filter $\mathbf{W}$ and the MSE matrix $\mathbf{E}$. As the MSEs for all symbols are identical, we obtain from (2.14)

$$\begin{aligned} \mathrm{MSE}_{k,\mathrm{FD-LE}} &= \frac{\sigma_s^2}{N_c} \mathrm{tr} \left\{ \mathbf{F} \mathbf{\Psi}_f^{-1} \mathbf{F}^\dagger \right\} \\ &\overset{(a)}{=} \frac{\sigma_s^2}{N_c} \mathrm{tr} \left\{ \mathbf{\Psi}_f^{-1} \right\} = \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \Psi_k^{-1}, \quad \forall k, \end{aligned} \tag{2.17}$$

where in (a) the shift property of the trace operator and the unitary property of $\mathbf{F}$ were used.

Using the same reasoning as for FD-LE, for FD-DFE the MSE for the kth symbol is given by

$$\mathrm{MSE}_{k,\mathrm{FD-DFE}} = \frac{\sigma_s^2}{N_c}\mathrm{tr}\left\{\mathbf{B}_f \mathbf{\Psi}_f^{-1}\mathbf{B}_f^{\dagger}\right\} = \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1} b_{f,k}\Psi_k^{-1}{b_{f,k}}^*, \quad \forall k,$$

which depends on the FD-FBF coefficients $b_{f,k}$. Therefore, we first need to find the optimal FBF. Since the FBF is implemented in the TD, we rewrite the $b_{f,k}$ in terms of the TD coefficients, b_n, resulting in

$$\begin{aligned}\mathrm{MSE}_{k,\mathrm{FD-DFE}} =& \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\left(\sum_{n=0}^{N_{fb}} b_n e^{-j\frac{2\pi}{N_c}nk}\right)\Psi_k^{-1}\left(\sum_{m=0}^{N_{fb}} b_m^* e^{j\frac{2\pi}{N_c}mk}\right)\\ =& \frac{\sigma_s^2}{N_c}\sum_{n=0}^{N_{fb}}\sum_{m=0}^{N_{fb}} b_n\left(\sum_{k=0}^{N_c-1}\Psi_k^{-1}e^{-j\frac{2\pi}{N_c}(n-m)k}\right)b_m^*\\ =& \frac{\sigma_s^2}{N_c}\mathbf{b}\mathbf{V}\mathbf{b}^{\dagger}, \quad \forall k, \end{aligned} \tag{2.18}$$

where $\mathbf{V}$ is a Hermitian matrix whose (i,j)th entry is given by $\mathbf{V}_{i,j} = \sum_{k=0}^{N_c-1}\Psi_k^{-1} e^{-j\frac{2\pi}{N_c}(i-j)k}$. Recall that the first coefficient of the TD-FBF is set to 1 to ensure causality. Thus, the optimal FBF coefficients are obtained from

$$\min_{\mathrm{s.t.}\mathbf{b}\mathbf{\Phi}=1} \quad \mathbf{b}\mathbf{V}\mathbf{b}^{\dagger}, \tag{2.19}$$

where $\mathbf{\Phi} = [1, \quad \mathbf{0}_{N_{fb}-1}]^T$. Using the standard Lagrange multiplier method, the solution to (2.19) can be derived as [106]

$$\mathbf{b} = (\mathbf{\Phi}^{\dagger}\mathbf{V}^{-1}\mathbf{\Phi})^{-1}\mathbf{\Phi}^{\dagger}\mathbf{V}^{-1}. \tag{2.20}$$

Substituting the optimal $\mathbf{b}$ in (2.20) into (2.18), we obtain the corresponding ex-

pression for the minimum MSE. Unfortunately, this expression depends on the rBF coefficients in a complicated way, which renders the corresponding optimization problem intractable. To find a tractable objective function, we exploit the equivalence between FD-DFE and TD-DFE with infinite length filters. In the asymptotic scenario where both N_c and N_{fb} tend to infinity, the MSE becomes the geometric mean of the main diagonal elements of the error covariance matrix [96, 97]

$$\mathrm{MSE}_{k,\mathrm{FD-DFE}} = \sigma_s^2 \det(\mathbf{\Psi}_f)^{-1/N_c} = \sigma_s^2 \prod_{k=1}^{N_c} \Psi_k^{-1/N_c}, \quad \forall k. \tag{2.21}$$

In the following, for FD-DFE rBF matrix optimization, we use (2.21) as an approximation for the MSE of FD-DFE with finite N_{fb} and N_c. We found through extensive simulations that this choice achieves a high performance even for relatively small values of N_{fb} and N_c.

2.3.2 Idealized Matched Filter Receiver

For the idealized MF receiver, single-symbol transmission is assumed, i.e., ISI is not present, which results in a performance upper bound for any realizable receiver structure. Taking into account that the effective noise at the destination is colored and defining the equivalent channel matrix after pre-whitening, $\mathbf{H}_{\mathrm{eq}} = \tilde{\mathbf{C}}^{-1/2}\tilde{\mathbf{H}}$, the signal-to-interference-plus-noise ratio (SINR) at the output of the MF is obtained as

$$\mathrm{SINR}_{k,\mathrm{MF}} = \frac{\sigma_s^2}{N_c} \mathrm{tr}\left\{\mathbf{H}_{\mathrm{eq}}\mathbf{H}_{\mathrm{eq}}^{\dagger}\right\} = \frac{\sigma_s^2}{N_c} \mathrm{tr}\left\{\mathbf{\Psi}_f - \mathbf{I}_{N_c}\right\} = \frac{\sigma_s^2}{N_c} \sum_{k=1}^{N_c} \left(\Psi_k - 1\right), \quad \forall k. \tag{2.22}$$

2.3.3 Objective Function for rBF Matrix Optimization

Since the performance of SC-FDE is directly influenced by the MSE at the destination, a reasonable objective for rBF matrix optimization is the minimization of the MSE (or equivalently the maximization of the SINR). Consequently, based on (2.17), (2.21), and (2.22) the considered objective functions to be minimized can be compactly expressed as

$$f_{\mathrm{X}}(\Psi_k) = \begin{cases} \sum_{k=1}^{N_c} \Psi_k^{-1}, & \mathrm{X} = \text{FD-LE} \\ -\sum_{k=1}^{N_c} \log\Psi_k, & \mathrm{X} = \text{FD-DFE} \\ -\sum_{k=1}^{N_c} (\Psi_k - 1), & \mathrm{X} = \text{MF} \end{cases}, \tag{2.23}$$

where for FD-DFE we consider the logarithm of the MSE in (2.21) to facilitate the subsequent optimization. Because of the monotonicity of the logarithm, this has no effect on the optimal solution. We note that all objective functions in (2.23) are monotonically decreasing function of Ψ_k. This property facilitates a unified treatment of rBF matrix optimization for all considered SC receiver structures.

2.4 Problem Formulation and Solution

In this section, we derive the structure of the optimal rBF matrices, $\mathbf{A}_{ik}$, $i = 1, \ldots, N_R$, $k = 1, \ldots, N_c$, and the optimal subcarrier power allocation for the objective functions introduced in the previous section. Thereby, an aggregate relay power constraint is adopted. Other power constraints such as individual relay power constraints and a joint source and relay power constraint can be considered as well. Furthermore, the extension to multiuser systems is possible by adopting SC-FDMA. However, in this chapter, we focus on the joint relay power constraint and single-user

systems, and leave the consideration of other power constraints and multiple users to Chapter 4.

2.4.1 Problem Formulation

The transmit signal at the ith relay is given by $\mathbf{t}_i = \mathbf{A}_i\mathbf{G}_i\mathbf{s} + \mathbf{A}_i\mathbf{n}_i$ and the aggregate power of the relays is $P_R = \sum_{i=1}^{N_R} \mathrm{tr}\{\mathbf{t}_i\mathbf{t}_i^\dagger\}$. P_R can be rewritten as

$$\begin{aligned} P_R &= \sum_{i=1}^{N_R} \mathrm{tr}\left\{\mathbf{A}_i\left(\sigma_s^2\mathbf{G}_i\mathbf{G}_i^\dagger + \sigma_{n_1}^2\mathbf{I}_{NN_c}\right)\mathbf{A}_i^\dagger\right\} \\ &= \sum_{i=1}^{N_R} \mathrm{tr}\left\{\mathbf{A}_f^{(i)}\left(\sigma_s^2\mathbf{G}_f^{(i)}\mathbf{G}_f^{(i)^\dagger} + \sigma_{n_1}^2\mathbf{I}_{NN_c}\right)\mathbf{A}_f^{(i)^\dagger}\right\} \\ &= \sum_{i=1}^{N_R}\sum_{k=1}^{N_c} \mathrm{tr}\left\{\mathbf{A}_{ik}\left(\sigma_s^2\mathbf{g}_{ik}\mathbf{g}_{ik}^\dagger + \sigma_{n_1}^2\mathbf{I}_N\right)\mathbf{A}_{ik}^\dagger\right\}. \end{aligned} \tag{2.24}$$

Hence, the optimal rBF filters are given by

$$\mathbf{A}_{ik}^{\mathrm{opt}} = \underset{\mathbf{A}_{ik},\, \sum_{k=1}^{N_c} P_k \le P_{R,\max}}{\arg\min} f_{\mathrm{X}}\left(\Psi_k\right), \tag{2.25}$$

where $P_{R,\max}$ is the maximum relay power available and P_k is the power allocated to the kth subchannel, i.e.,

$$P_k = \sum_{i=1}^{N_R} \mathrm{tr}\left\{\mathbf{A}_{ik}\left(\sigma_s^2\mathbf{g}_{ik}\mathbf{g}_{ik}^\dagger + \sigma_{n_1}^2\mathbf{I}_N\right)\mathbf{A}_{ik}^\dagger\right\}. \tag{2.26}$$

In the next section, we show that for given P_k, the structure of the optimal rBF matrices can be obtained in closed form.

2.4.2 Structure of Optimal rBF Matrices

Since the considered family of objective functions, $f_{\mathrm{X}}(\Psi_k)$, are monotonically decreasing functions in Ψ_k, without loss of optimality, we can first maximize Ψ_k for a given subchannel power P_k, $k = 1, \cdots, N_c$, before optimizing the power allocation across subchannels. To this end, it is convenient to rewrite Ψ_k as $\Psi_k = \hat{\Psi}_k + \frac{\sigma_s^2}{\sigma_{n_0}^2} q_k^2 + 1$, where

$$\hat{\Psi}_k = \frac{\sigma_s^2 \sum_{i=1}^{N_R} (\mathbf{h}_{ik}\mathbf{A}_{ik}\mathbf{g}_{ik})^2}{\sigma_{n_1}^2 \sum_{i=1}^{N_R} \mathbf{h}_{ik}\mathbf{A}_{ik}\mathbf{A}_{ik}^{\dagger}\mathbf{h}_{ik}^{\dagger} + \sigma_{n_2}^2}. \tag{2.27}$$

Clearly, maximizing Ψ_k with respect to $\mathbf{A}_{ik}$ is equivalent to maximizing $\hat{\Psi}_k$. Furthermore, by incorporating (2.26) into $\hat{\Psi}_k$, we obtain

$$\hat{\Psi}_k = \frac{P_k \sigma_s^2 (\sum_{i=1}^{N_R} \mathbf{h}_{ik}\mathbf{A}_{ik}\mathbf{g}_{ik})^2}{P_k \sigma_{n_1}^2 \sum_{i=1}^{N_R} \mathrm{tr}\{\mathbf{A}_{ik}^{\dagger}\mathbf{h}_{ik}^{\dagger}\mathbf{h}_{ik}\mathbf{A}_{ik}\} + \sum_{i=1}^{N_R} \mathrm{tr}\{\mathbf{A}_{ik}(\sigma_s^2 \mathbf{g}_{ik}\mathbf{g}_{ik}^{\dagger} + \sigma_{n_1}^2 \mathbf{I}_N)\mathbf{A}_{ik}^{\dagger}\}\sigma_{n_2}^2}. \tag{2.28}$$

For the following, it is convenient to introduce $\bar{\mathbf{a}}_k = [\mathbf{a}_{1k}^T, ..., \mathbf{a}_{N_R k}^T]^T$ with $\mathbf{a}_{ik} = \mathrm{vec}(\mathbf{A}_{ik})$ and to rewrite (2.28) as

$$\hat{\Psi}_k = \frac{P_k \bar{\mathbf{a}}_k^{\dagger} \mathbf{U}_k \bar{\mathbf{a}}_k}{\bar{\mathbf{a}}_k^{\dagger}(P_k \sigma_{n_1}^2 \mathbf{Z}_k + \sigma_{n_2}^2 \mathbf{W}_k)\bar{\mathbf{a}}_k} = \frac{P_k \bar{\mathbf{a}}_k^{\dagger} \mathbf{U}_k \bar{\mathbf{a}}_k}{\bar{\mathbf{a}}_k^{\dagger} \mathbf{\Delta}_k \bar{\mathbf{a}}_k}, \tag{2.29}$$

where $\mathbf{U}_k = \bar{\mathbf{u}}_k^{\dagger}\bar{\mathbf{u}}_k$, $\bar{\mathbf{u}}_k = [\mathbf{u}_{1k}, ..., \mathbf{u}_{N_R k}]$, $\mathbf{u}_{ik} = (\mathrm{vec}((\mathbf{g}_{ik}\mathbf{h}_{ik})^T))^T$, $\mathbf{Z}_k = \mathrm{blkdiag}\{(\mathbf{I}_N \otimes \mathbf{h}_{ik}^{\dagger}\mathbf{h}_{ik}), i = 1, \cdots, N_R\}$, $\mathbf{W}_k = \mathrm{blkdiag}\{(\sigma_s^2 \mathbf{g}_{ik}^* \mathbf{g}_{ik}^T + \sigma_{n_1}^2 \mathbf{I}_N) \otimes \mathbf{I}_N, i = 1, \cdots, N_R\}$, and $\mathbf{\Delta}_k = P_k \sigma_{n_1}^2 \mathbf{Z}_k + \sigma_{n_2}^2 \mathbf{W}_k$. Here, we used the identities $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{vec}\left(\mathbf{A}^T\right)^T \mathrm{vec}(\mathbf{B})$, $\mathrm{tr}\{\mathbf{ABC}\} = \mathrm{vec}\left(\mathbf{A}^T\right)^T (\mathbf{I} \otimes \mathbf{B})\mathrm{vec}(\mathbf{C})$, and $\mathrm{tr}\{\mathbf{ABC}\} = \mathrm{vec}\left(\mathbf{C}^T\right)^T \left(\mathbf{B}^T \otimes \mathbf{I}\right)\mathrm{vec}(\mathbf{A})$ [99] to transform (2.28) into (2.29).

Since $\mathbf{U}_k$ is a Hermitian matrix and $\mathbf{\Delta}_k$ is a positive definite Hermitian matrix,

(2.29) is a generalized Rayleigh quotient [100], whose maximum value can be derived as

$$\begin{aligned}[\hat{\Psi}_k]_{\max} &= P_k \lambda_{\max}(\mathbf{\Delta}_k^{-\dagger/2}\mathbf{U}_k\mathbf{\Delta}_k^{-1/2}) \\ &\overset{(a)}{=} P_k \mathrm{tr}\{\mathbf{\Delta}_k^{-\dagger/2}\mathbf{U}_k\mathbf{\Delta}_k^{-1/2}\} \\ &\overset{(b)}{=} P_k \mathrm{tr}\{\mathbf{\Delta}_k^{-1}\bar{\mathbf{u}}_k^{\dagger}\bar{\mathbf{u}}_k\} \overset{(c)}{=} P_k \sum_{i=1}^{N_R} \mathrm{tr}\{\mathbf{\Delta}_{ik}^{-1}\mathbf{u}_{ik}^{\dagger}\mathbf{u}_{ik}\},\end{aligned} \tag{2.30}$$

where $\lambda_{\max}(\mathbf{X})$ denotes the maximum eigenvalue of matrix $\mathbf{X}$, step (a) follows from the fact that $\mathbf{\Delta}_k^{-\dagger/2}\mathbf{U}_k\mathbf{\Delta}_k^{-1/2}$ is a rank-one matrix, step (b) exploits the shift property of the trace operator, and step (c) uses the block diagonal structure of $\mathbf{\Delta}_k$ and the definition of $\bar{\mathbf{u}}_k$. The maximum value in (2.30) is achieved when $\bar{\mathbf{a}}_k = c_k\mathbf{\Delta}_k^{-1}\bar{\mathbf{u}}_k^{\dagger}$ or, equivalently

$$\mathbf{a}_{ik} = c_k\mathbf{\Delta}_{ik}^{-1}\mathbf{u}_{ik}^{\dagger}, \tag{2.31}$$

where c_k is a normalization factor that guarantees the power constraint (2.26).

We observe that both $[\hat{\Psi}_k]_{\max}$ in (2.30) and $\mathbf{a}_{ik}$ in (2.31) contain the term $\mathbf{\Delta}_{ik}^{-1}\mathbf{u}_{ik}^{\dagger}$. In order to gain some insight into the structure of $\mathbf{a}_{ik}$ and to obtain a tractable expression for $[\hat{\Psi}_k]_{\max}$ for subsequent optimization, we introduce the following proposition to simplify this term.

Proposition 2.1. *: $\mathbf{u}_{ik}^{\dagger}$ is an eigenvector of matrix $\mathbf{\Delta}_{ik}^{-1}$ and corresponds to the eigenvalue $1/(P_k\sigma_{n_1}^2||\mathbf{h}_{ik}||^2 + \sigma_{n_2}^2(\sigma_s^2||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2))$, i.e.,*

$$\mathbf{\Delta}_{ik}^{-1}\mathbf{u}_{ik}^{\dagger} = \frac{\mathbf{u}_{ik}^{\dagger}}{P_k\sigma_{n_1}^2||\mathbf{h}_{ik}||^2 + \sigma_{n_2}^2(\sigma_s^2||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)}. \tag{2.32}$$

Proof. Please refer to Appendix-A. □

From (2.30) and (2.32), we obtain

$$[\hat{\Psi}_k]_{\max} = \sum_{i=1}^{N_R} \frac{P_k ||\mathbf{h}_{ik}\mathbf{g}_{ik}||^2}{P_k \sigma_{n_1}^2 ||\mathbf{h}_{ik}||^2 + \sigma_{n_2}^2(\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)} \tag{2.33}$$

and $\mathbf{a}_{ik}$ can be expressed as

$$\mathbf{a}_{ik} = \frac{c_k \text{vec}(\mathbf{h}_{ik}^T \mathbf{g}_{ik}^T)^*}{P_k \sigma_{n_1}^2 ||\mathbf{h}_{ik}||^2 + \sigma_{n_2}^2(\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)}. \tag{2.34}$$

Upon de-vectorization of $\mathbf{a}_{ik}$, we obtain the final expression of the FD rBF matrix

$$\mathbf{A}_{ik} = \frac{c_k ||\mathbf{h}_{ik}|| \, ||\mathbf{g}_{ik}||}{P_k \sigma_{n_1}^2 ||\mathbf{h}_{ik}||^2 + \sigma_{n_2}^2(\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)} \frac{\mathbf{h}_{ik}^\dagger \mathbf{g}_{ik}^\dagger}{||\mathbf{h}_{ik}|| ||\mathbf{g}_{ik}||}. \tag{2.35}$$

Combining (2.35) and (2.26), the normalization factor c_k can be found explicitly as

$$c_k = \sqrt{P_k} \left[\sum_{i=1}^{N_R} \frac{||\mathbf{h}_{ik}\mathbf{g}_{ik}||^2 (\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)}{[\sigma_{n_1}^2 ||\mathbf{h}_{ik}||^2 P_k + (\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)\sigma_{n_2}^2]^2} \right]^{-\frac{1}{2}}. \tag{2.36}$$

Interestingly, all objective functions $f_{\text{X}}(\mathbf{\Psi})$ that are decreasing functions of Ψ_k, $k = 1, \ldots, N_c$, lead to the same structure for the optimal $\mathbf{A}_{ik}$ and only the scalar c_k is affected by the particular objective function via the allocated power P_k. Furthermore, the optimal structure of $\mathbf{A}_{ik}$ in (2.35) can be interpreted as follows. The first term, $c_k ||\mathbf{h}_{ik}|| \, ||\mathbf{g}_{ik}|| / (P_k \sigma_{n_1}^2 ||\mathbf{h}_{ik}||^2 + \sigma_{n_2}^2(\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2))$ is a positive scalar and depends on the fraction of power allocated to the kth frequency tone. The vector terms $\frac{\mathbf{g}_{ik}^\dagger}{||\mathbf{g}_{ik}||}$ and $\frac{\mathbf{h}_{ik}^\dagger}{||\mathbf{h}_{ik}||}$ have unit norm and represent a maximum ratio combining (MRC) filter for the S-R_i channel, $\mathbf{g}_{ik}$, and a maximum ratio transmission (MRT) filter for the R_i-D channel, $\mathbf{h}_{ik}$, respectively. Eq. (2.35) reveals that relay R_i, $i = 1, \ldots, N_R$, can compute its optimal rBF matrix based on local CSI, i.e., its own source-relay and

relay-destination channels, and the $2N_c$ real scalars c_k and P_k, $k = 1, \ldots, N_c$. Since c_k and P_k depend on the CSI of all channels in the system, they have to be computed at the central node and fed back to the relays, cf. Section 2.2.4.

2.4.3 Suboptimal Power Allocation Schemes

As will be shown in the next section, the optimal values for P_k can only be found with a numerical method, which may not be desirable in delay-sensitive or complexity-constrained systems. One alternative suboptimal solution is to adopt equal power allocation (EPA) across all tones, i.e., we set $P_k = P_{R,\max}/N_c$. A further complexity reduction is possible by performing EPA across tones and relays, which leads to

$$\mathbf{A}_{ik} = c'_k \frac{\mathbf{h}_{ik}^{\dagger} \mathbf{g}_{ik}^{\dagger}}{||\mathbf{h}_{ik}||\ ||\mathbf{g}_{ik}||}, \tag{2.37}$$

where $c'_k = \sqrt{P_{R,\max}/N_c}(\sum_{i=1}^{N_R}(\sigma_s^2||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2))^{-1/2}$. These two suboptimal schemes are referred to as EPA-T-rBF and EPA-TR-rBF, respectively. EPA-TR-rBF has the additional advantage that the central node has to know only the source-relay channels but not the relay-destination channels. This is advantageous if the source is the central node, and thus, can directly estimate the source-relay channel based on training symbols emitted by the relays.

Both EPA-T-rBF and EPA-TR-rBF reduce the feedback overhead compared to optimal power allocation as the central node has to broadcast only N_c real scalars c'_k, $k = 1, \ldots, N_c$.

2.4.4 Optimal Power Allocation (OPA)

Exploiting the optimal structure of the rBF matrices, Ψ_k can be expressed as

$$\Psi_k = \sum_{i=1}^{N_R} \frac{P_k \alpha_{ik}}{P_k \beta_{ik} + \gamma_{ik}} + \frac{\sigma_s^2}{\sigma_{n_0}^2} q_k^2 + 1, \tag{2.38}$$

where $\alpha_{ik} = ||\mathbf{h}_{ik}\mathbf{g}_{ik}||^2$, $\beta_{ik} = \sigma_{n_1}^2 ||\mathbf{h}_{ik}||^2$, and $\gamma_{ik} = \sigma_{n_2}^2(\sigma_s^2 ||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)$. Since Ψ_k is now only a function of P_k, we redefine the objective function and express it as a function of $\mathbf{P} = [P_1, \ldots, P_{N_c}]$, $f_{\mathrm{X}}(\mathbf{P})$, in the following. The optimal power allocation vector is given by

$$\mathbf{P} = \arg \min_{\sum_{k=1}^{N_c} P_k = P_{R,\max},\, P_k \geq 0} f_{\mathrm{X}}(\mathbf{P}), \tag{2.39}$$

where we have exploited the fact that the optimal power allocation meets the power constraint $P_{R,\max}$ with equality.

We first verify the convexity of the considered objective functions. To this end, we have tabulated the first and second derivatives of the considered $f_{\mathrm{X}}(\mathbf{P})$ in Table 2.1, where the first and second derivatives of Ψ_k are given by

$$\frac{\partial \Psi_k}{\partial P_k} = \sum_{i=1}^{N_R} \frac{\alpha_{ik}\gamma_{ik}}{(\beta_{ik}P_k + \gamma_{ik})^2} > 0, \quad \frac{\partial^2 \Psi_k}{\partial P_k^2} = \sum_{i=1}^{N_R} \frac{-2\alpha_{ik}\beta_{ik}\gamma_{ik}}{(\beta_{ik}P_k + \gamma_{ik})^3} < 0, \tag{2.40}$$

which shows that Ψ_k is an increasing concave function of P_k. We observe from Table 2.1 that $\partial^2 f_{\mathrm{X}}(\mathbf{P})/\partial P_k^2 > 0$ for all considered objective functions, i.e., the power allocation problem in (2.39) is a convex optimization problem.

The Lagrangian L of the power allocation problem is given by

$$L = f_{\mathrm{X}}(\mathbf{P}) + \mu(\sum_{k=1}^{N_c} P_k - P_{R,\max}) - \sum_{k=1}^{N_c} \nu_k P_k, \tag{2.41}$$

where μ and ν_k are the Lagrange multipliers associated with the total power constraint and the non-negativity constraint for the subchannel powers, respectively. The optimal power allocation policy across the different subchannels can be obtained from the following KKT conditions, which are sufficient and necessary for convex problems [101],

$$\begin{aligned} \frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k} + \mu - \nu_k = 0, \\ \sum_{k=1}^{N_c} P_k - P_{R,\max} = 0, \\ P_k \geq 0, \quad \nu_k \geq 0, \quad \nu_k P_k = 0. \end{aligned} \tag{2.42}$$

The last three conditions are the primal constraints, the dual constraints, and the complementary slackness conditions, respectively, and imply that if $\nu_k > 0$, then $P_k = 0$ and if $P_k > 0$, then $\nu_k = 0$. Combining the first condition, $\nu_k = \frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k} + \mu$, and the last three conditions yields

$$\begin{aligned} \frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k} + \mu > 0 \Rightarrow P_k = 0, \\ P_k > 0 \Rightarrow \frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k} + \mu = 0. \end{aligned} \tag{2.43}$$

Because of the convexity of $f_{\mathrm{X}}(\mathbf{P})$, $\frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k}$ is monotonically increasing. Thus, for a given value of μ, condition (2.43) can be verified by examining $[\frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k}]_{P_k=0} + \mu$. If this value is greater than 0, we set $P_k = 0$; otherwise, we find P_k as the unique positive solution of $\frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k} + \mu = 0$, which can be efficiently obtained by using e.g. Newton's method.

An algorithm for finding the optimum $\mathbf{P}$ is summarized in Table 2.2, where the required first and second derivatives of the objective function are given in Table 2.1.

Table 2.1: First and second derivatives of objective functions $f_{\mathrm{X}}(\mathbf{P})$.

Receiver	$\frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k}$	$\frac{\partial^2 f_{\mathrm{X}}(\mathbf{P})}{\partial P_k^2}$
FD-LE	$-\Psi_k^{-2}\frac{\partial \Psi_k}{\partial P_k}$	$2\Psi_k^{-3}[\frac{\partial \Psi_k}{\partial P_k}]^2 - \Psi_k^{-2}\frac{\partial^2 \Psi_k}{\partial P_k^2}$
FD-DFE	$-\Psi_k^{-1}\frac{\partial \Psi_k}{\partial P_k}$	$\Psi_k^{-2}[\frac{\partial \Psi_k}{\partial P_k}]^2 - \Psi_k^{-1}\frac{\partial^2 \Psi_k}{\partial P_k^2}$
MF	$-\frac{\partial \Psi_k}{\partial P_k}$	$-\frac{\partial^2 \Psi_k}{\partial P_k^2}$

We note that this algorithm is not limited to the objective functions given in (2.23) but is applicable to all objective functions which are convex in $\mathbf{P}$.

2.5 Simulation Results

In this section, we evaluate the performance of the proposed rBF schemes using simulations. Throughout this section, we assume quaternary phase shift keying (QPSK) modulation and define the transmit SNR as E_b/N_0, where E_b and N_0 denote the transmit energy per bit and the single-sided power spectral density of the underlying continuous-time noise process, respectively. We assume equal noise variances at the relays and the destination in both phases of transmission, i.e., $\sigma_{n_0}^2 = \sigma_{n_1}^2 = \sigma_{n_2}^2$. Furthermore, we assume that each data block contains $N_c = 64$ symbols. S and D are located on a horizontal line and have a normalized distance of $d_{S-D} = 2$ from each other. The relays are located on a vertical line which crosses the horizontal line connecting S and D at a distance of 1 from both S and D. The vertical distances between the crossing point and the relays are multiples of $1/4$. We assume a path loss factor of 4 with a reference distance of $d_{\mathrm{ref}} = 1$. Taking into account the pathloss, the channel vectors are modeled as $\mathbf{q} = \sqrt{E_b/(N_0(d_{S-D})^4)}\,\bar{\mathbf{q}}$, $\mathbf{g}_i = \sqrt{E_b/(N_0(d_{S-R_i})^4)}\,\bar{\mathbf{g}}_i$, and $\mathbf{h}_i = \sqrt{E_b/(N_0(d_{R_i-D})^4)}\,\bar{\mathbf{h}}_i$, where $\bar{\mathbf{s}} \in \{\bar{\mathbf{q}}, \bar{\mathbf{g}}_i, \bar{\mathbf{h}}_i\}$ is the normalized channel vector with unit variance, and d_{S-R_i} and d_{R_i-D} are the $S-R_i$ and the R_i-D distances,

Table 2.2: Algorithm 2.1 for finding the optimal power allocation. ϵ_1 and ϵ_2 are small constants, e.g. $\epsilon_1 = \epsilon_2 = 10^{-4}$.

1	Initialize the upper bound, μ_{ub}, and the lower bound, μ_{lb}, for μ with 0 and $\max_k\{-\frac{\partial f_{\mathrm{X}}(\mathbf{P})}{\partial P_k}\|_{P_k=0}\}$, respectively.
2	Initialize $P_k^{(1)}$, $\forall k$, such that $\sum_{k=1}^{N_c} P_k^{(1)} = P_{R,\max}$.

3	**Set** $\mu = \frac{1}{2}(\mu_{lb} + \mu_{ub})$.
	for $k = 1$ to N_c
	Initialize iteration number $m = 1$.
	Repeat
	Set $\Gamma_k^{(m)} = \frac{\partial f_{\mathrm{X}}(\mathbf{P})^{(m)}}{\partial P_k^{(m)}} + \mu$.
	If $\Gamma_k^{(m)} > 0$, set $P_k^{(m+1)} = 0$,
	otherwise, $P_k^{(m+1)} = P_k^{(m)} - \left(\frac{\partial^2 f_{\mathrm{X}}(\mathbf{P})^{(m)}}{\partial {P_k^{(m)}}^2}\right)^{-1} \Gamma_k^{(m)}$.
	$m = m + 1$.
	until $\|P_k^{(m)} - P_k^{(m-1)}\| < \epsilon_1$.
	end for

4	If $\sum_{k=1}^{N_c} P_k^{(m)} < P_{R,\max}$, set $\mu_{lb} = \mu$, otherwise, $\mu_{ub} = \mu$.
5	If $\|\sum_{k=1}^{N_c} P_k^{(m)} - P_{R,\max}\| < \epsilon_2$, then $P_k^{(m)}$, $\forall k$, are the desired results, otherwise, goto Step 2.

respectively. The normalized channel vectors are modeled as Rayleigh block fading channels. For simplicity, we assume that all relay antennas are uncorrelated but note that a non-zero correlation will generally degrade performance. The power delay profile of the channels is given by [102]

$$p[n] = \frac{1}{\sigma_t} \sum_{l=0}^{L_x-1} e^{-n/\sigma_t} \delta[n-l], \tag{2.44}$$

where $L_x \in \{L_g, L_h, L_q\}$ and σ_t is a factor, which influences the delay spread of the channels. For convenience, we assume L_q, L_g, L_h, $N_{g,s}$, and $N_{g,r}$ are all equal to 16. For FD-DFE, the number of FBF taps is set to $N_{fb} = 15$.

All simulations are averaged over at least 10,000 independent channel realizations. We compare the performance of the proposed rBF scheme with the following reference schemes: 1) Direct transmission (Direc-Tx) without relays; 2) Direct trans-

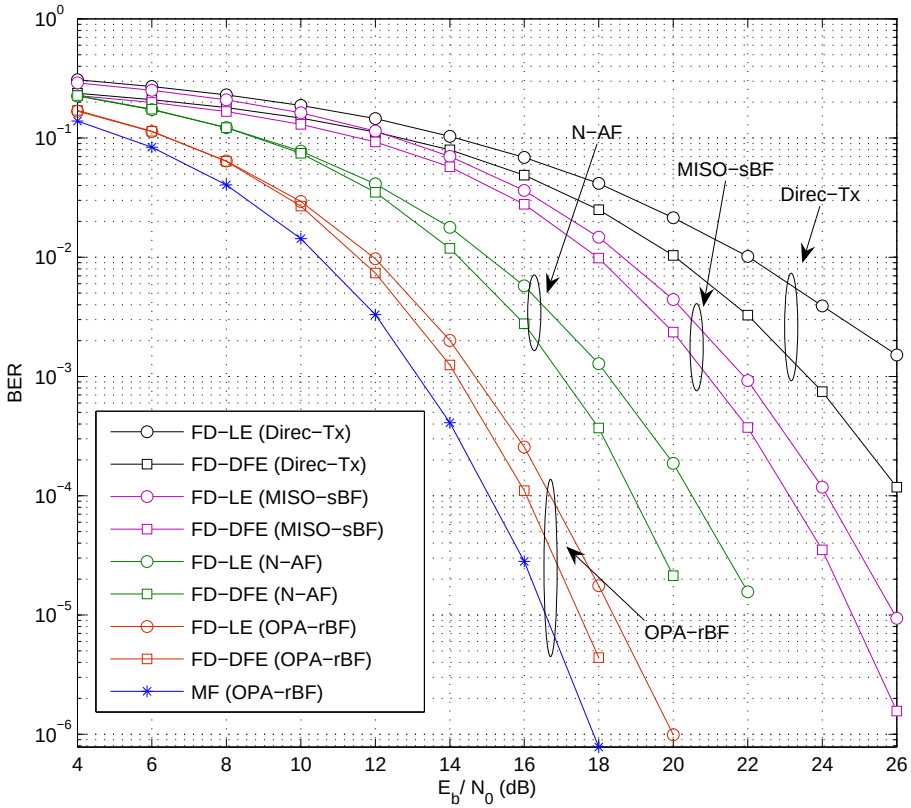

Figure 2.2: BERs of SC-FDE with OPA-rBF and N-AF. $N_R = 2$, $N = 1$, and $\sigma_t = 2$. For comparison, the BERs of SC-FDE with direct transmission using a single transmit antenna and MISO-sBF with two transmit antennas are also shown.

mission with conventional multiple antenna BF at the source (MISO-sBF) [103] ; and 3) Naive-AF relaying (N-AF), where the relay stations only scale the power of the received signal before forwarding it to the destination. For MISO-sBF we assume NN_R transmit antennas at the source to enable a fair comparison with the proposed rBF scheme. All schemes are compared for the same total system power budget.

2.5.1 Single-Antenna Relays: Impact of Receiver Structures, Power Allocation, and Frequency Selectivity

In Fig. 2.2, we depict the BERs for FD-LE and FD-DFE at the destination with OPA-rBF and N-AF relaying using two single-antenna relays (i.e., $N = 1$, $N_R = 2$). The performance of the idealized MF receiver with OPA-rBF is included as a benchmark. Furthermore, for comparison, the BERs of direct transmission using a single transmit antenna and MISO-sBF with two transmit antennas are also shown. We choose the delay spread factor as $\sigma_t = 2$, which corresponds to moderately frequency-selective fading. Fig. 2.2 shows that even N-AF achieves a considerable performance gain compared to direct transmission and MISO-sBF because of the pathloss gain induced by the relays. As expected, FD-DFE outperforms FD-LE since nonlinear equalization can better exploit the frequency diversity of the channel and is more robust to spectral nulls of the channel. However, for both MISO-sBF and rBF, the performance gap between FD-LE and FD-DFE is smaller than for direct transmission with a single transmit antenna since multiple antennas can alleviate the negative effect of spectral nulls. Furthermore, it is interesting to note that the BER curve of FD-DFE is parallel to that of the idealized MF receiver when OPA-rBF is adopted, with only a 1 dB SNR gap, while FD-LE suffers from a small loss in diversity gain compared to the idealized MF receiver.

In Fig. 2.3, we investigate the BER of the proposed OPA- and EPA-rBF schemes for the same system parameters as in Fig. 2.2. We observe that the BER curves of the proposed rBF schemes with OPA, EPA-T, and EPA-TR are parallel, i.e., the diversity gain is not negatively affected by the suboptimal power allocation. Furthermore, the loss in array gain due to the suboptimal power allocation, i.e., the horizontal shift of the SNR required to achieve a certain BER, is small. Therefore, EPA is an

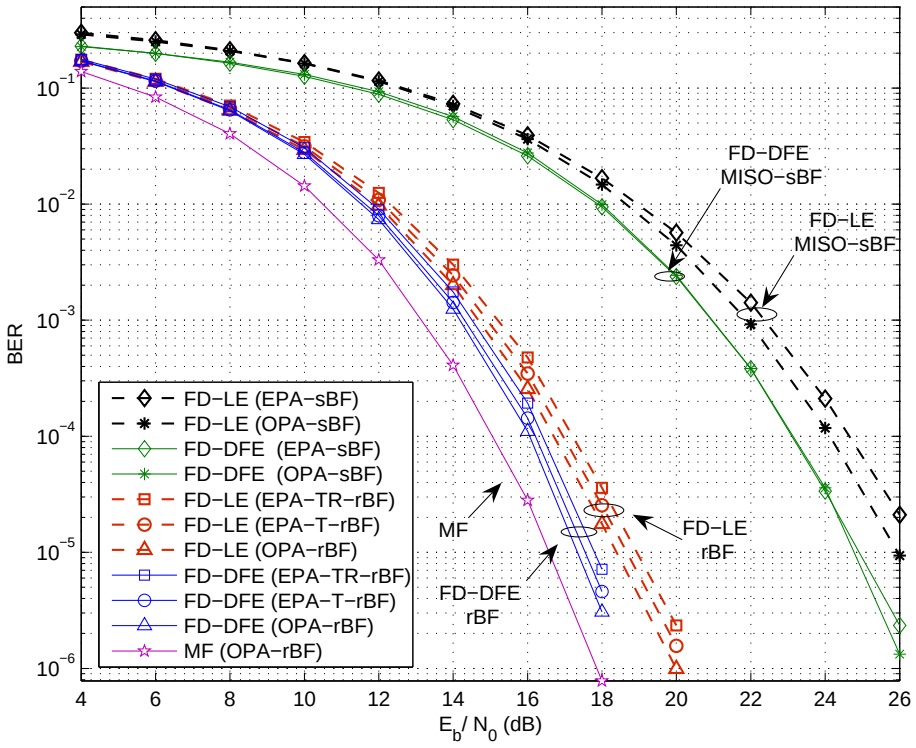

Figure 2.3: BERs of SC-FDE with OPA-rBF, EPA-T-rBF, and EPA-TR-rBF. $N_R = 2$, $N = 1$, and $\sigma_t = 2$. For comparison, the BERs of SC-FDE with direct transmission using a single transmit antenna and MISO-sBF with two transmit antennas are also shown.

attractive alternative to OPA because of the reduced feedback overhead and the low computational complexity. In addition, we note that for sufficiently high SNR, the BER curves for MISO-sBF have similar slopes as those for rBF, implying that both schemes achieve a comparable diversity gain. However, there is a large SNR gap between the BER curves for MISO-sBF and rBF since MISO-sBF cannot effectively combat the pathloss.

In Fig. 2.4, the effect of σ_t on the BER performance of rBF, N-AF, and direct transmission with and without MISO-sBF is investigated for FD-LE, FD-DFE, and an idealized MF receiver at the destination. Except for the value of σ_t, the system

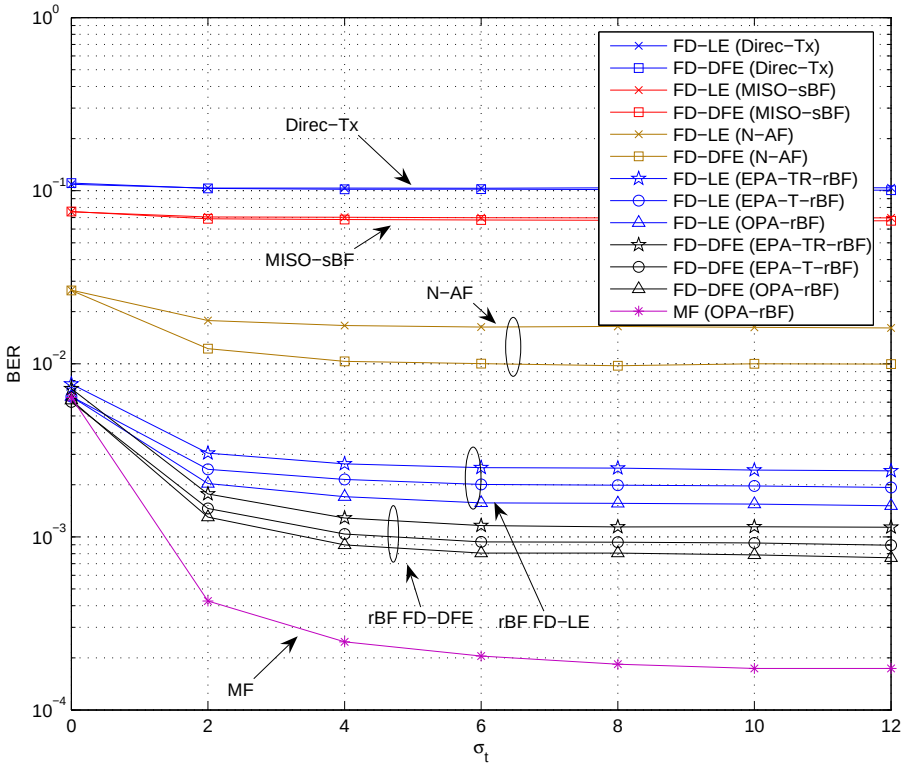

Figure 2.4: BERs of SC-FDE with rBF and N-AF vs. delay spread factor σ_t. $N_R = 2$, $N = 1$, and $E_b/N_0 = 14$ dB. For comparison, the BERs of SC-FDE with direct transmission using a single transmit antenna and MISO-sBF with two transmit antennas are also shown.

parameters are the same as in Fig. 2.2. The results show that for SC-FDE with rBF, increasing the channel frequency selectivity, i.e., increasing σ_t, leads to a considerably lower BER for both OPA and EPA due to the diversity gain extracted by rBF and FDE. In contrast, the performance of SC-FDE systems with MISO-sBF is almost invariant to the value of σ_t. This can be attributed to the fact that MISO-sBF cannot effectively compensate for the pathloss and, as a result, for the considered transmit SNR, the received SNR is quite low. Hence, for MISO-sBF, the noise dominates the performance and a diversity gain is not observed for increasing σ_t. Furthermore,

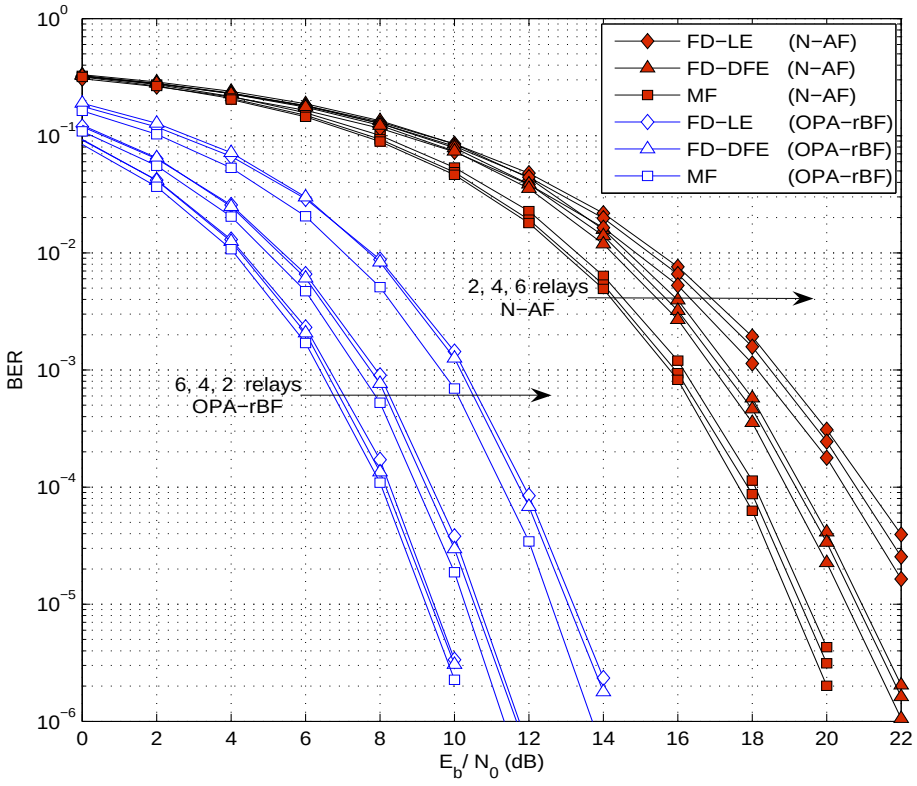

Figure 2.5: BERs of SC-FDE with rBF and N-AF for various numbers of relays. $N = 2$, $N_R = 2, 4, 6$, and $\sigma_t = 2$.

for rBF, the performance gain achieved by the idealized MF when σ_t increases is considerably larger than the gains for FD-LE and FD-DFE since the idealized MF is not negatively affected by the increased level of ISI and can take full advantage of the increased frequency diversity. We note that more complex trellis-based equalizers may be able to close the gap between FD-DFE and the idealized MF receiver at the expense of a considerably higher receiver complexity.

2.5.2 Multiple-Antenna Relays: Impact of the Number of Relays, Antennas, and Imperfect CSI

In Fig. 2.5, we investigate the effect of the number of relays on the BER of SC-FDE with OPA-rBF in a system with $N = 2$ relay antennas. As can be observed, for OPA-rBF, increasing N_R improves the BER performance substantially and allows FD-LE and FD-DFE to closely approach the performance of the idealized MF receiver. Thus, for relay networks with a sufficiently large numbers of relays, SC-FDE with MMSE based receivers is close to optimal and little improvement can be expected from more complex receiver structures. In contrast, for N-AF, performance slightly deteriorates with increasing number of relays since N-AF cannot exploit the increased diversity introduced by the additional relays.

In Fig. 2.6, we illustrate the effect of the number of relay antennas on the BER performance of SC-FDE with OPA-rBF for a system with $N_R = 2$ relays. As expected, increasing the number of relay antennas greatly improves performance of both rBF and MISO-sBF since additional spatial diversity is introduced. A comparison with Fig. 2.5 reveals that adding more relay antennas is more beneficial in rBF systems than increasing the number of relays assuming the total number of antennas is identical in both cases. This behavior is expected as co-located antennas enable a joint processing of all received signals, which is not possible for distributed antennas.

In Fig. 2.7, we examine the impact of imperfect CSI at the relays on the BER of a system with $N_R = 2$, $N = 2$, and $E_b/N_0 = 14$ dB. We assume that the CSI employed for rBF matrix computation is imperfect. We model the CSI estimation errors for all CIR coefficients as Gaussian random variables with variance σ_e^2. For example, the CSI estimate for g_{inl} is given by $\hat{g}_{inl} = g_{inl} + \Delta g_{inl}$, where the actual CSI g_{inl} and the CSI estimation error Δg_{inl} are mutually independent. We consider two

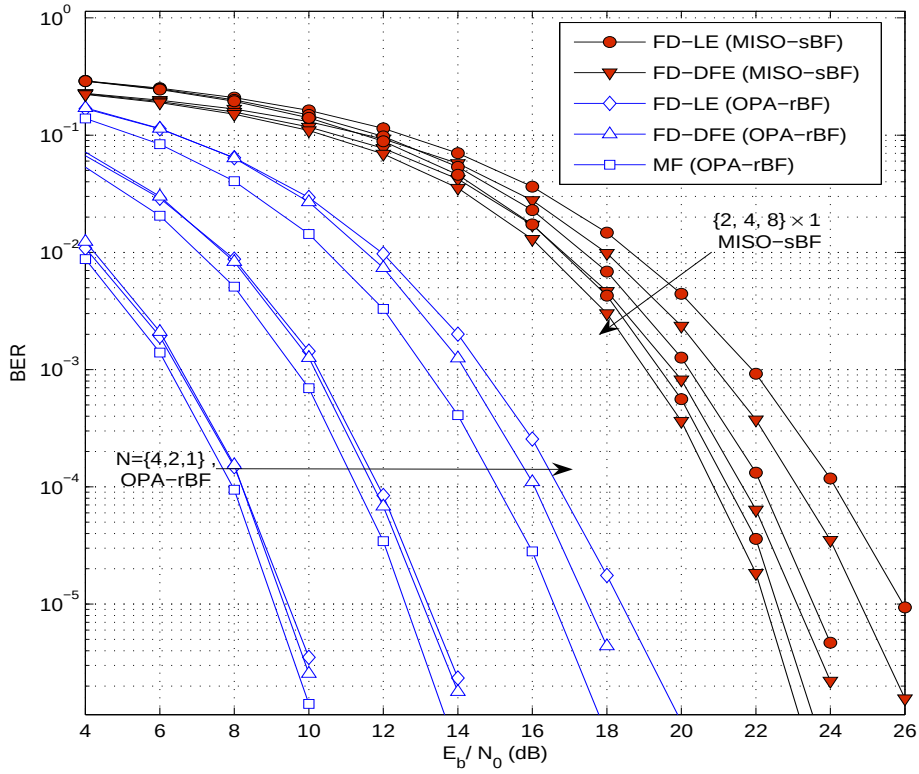

Figure 2.6: BERs of SC-FDE with rBF for various numbers of relay antennas. $N_R = 2$, $N = 1, 2, 4$, and $\sigma_t = 2$. For comparison, the BERs of SC-FDE with direct transmission using MISO-sBF with 2, 4, and 8 transmit antennas are also shown.

different cases for the imperfect CSI. For Case I (dotted lines), the CSI of all links is imperfect at the relays and the central node. In contrast, for Case II (solid lines), R_i has perfect CSI of its own links, i.e., $\mathbf{g}_t^{(i)}$ and $\mathbf{h}_t^{(i)}$, and the central node has imperfect CSI of all links (i.e., P_k and c_k, $k = 1, \ldots, N_c$, are computed based on imperfect CSI). The BERs of FD-LE employing rBF with OPA and EPA are shown as functions of σ_e^2 in Fig. 2.7. In both cases, the performance degrades with the increase of σ_e^2. For Case I, OPA, EPA-T, and EPA-TR yield a similar performance. In contrast for Case II, the suboptimal schemes are much more robust to imperfect CSI than OPA and yield significant performance gains for large estimation error variances of $\sigma_e^2 \geq 0.1$.

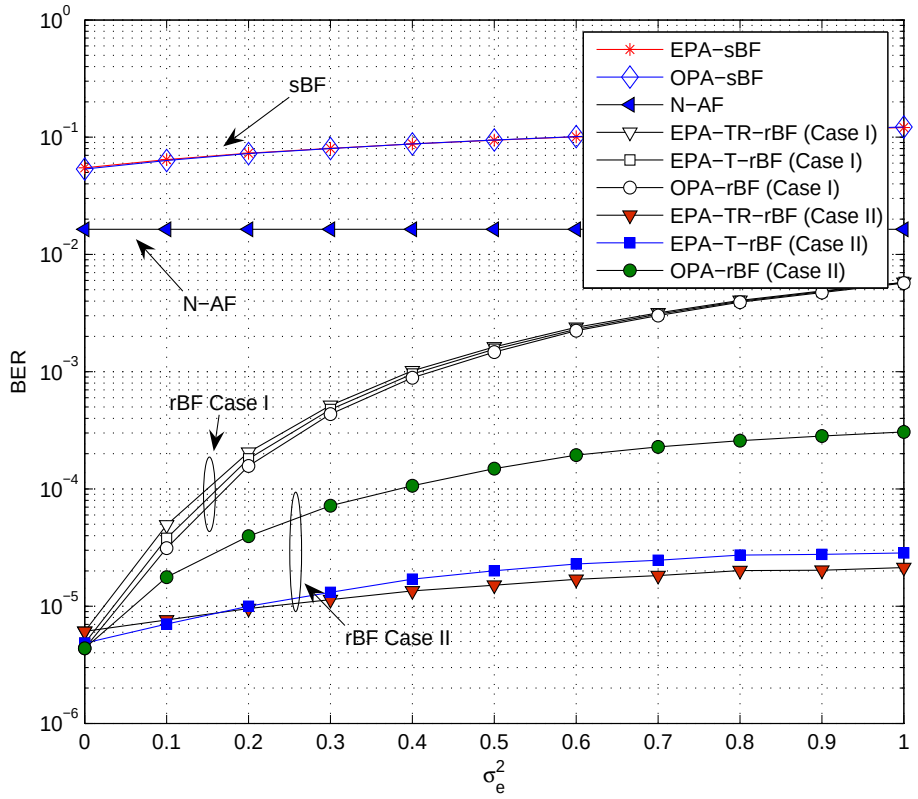

Figure 2.7: BERs of FD-LE with rBF and N-AF for imperfect CSI at the relays. $N_R = 2$, $N = 2$, $\sigma_t = 2$, and $E_b/N_0 = 14$ dB. For comparison, the BERs of SC-FDE with direct transmission using MISO-sBF with four transmit antennas are also shown.

This can be explained as follows. For OPA, both P_k and c_k are computed based on the imperfect CSI. In contrast, for EPA-T (EPA-TR) only c_k (c'_k) is affected by the imperfect CSI, which makes it more robust than OPA. EPA-TR has the additional benefit that c'_k depends only on the (imperfect) source-relay channels but not on the relay-destination channels. Thus, for $\sigma_e^2 \geq 0.1$, EPA-TR performs slightly better than EPA-T. As a general conclusion from Fig. 2.7 we note that for OPA to be beneficial, accurate CSI is required. Nevertheless, even with imperfect CSI, all considered rBF schemes outperform N-AF and MISO-sBF.

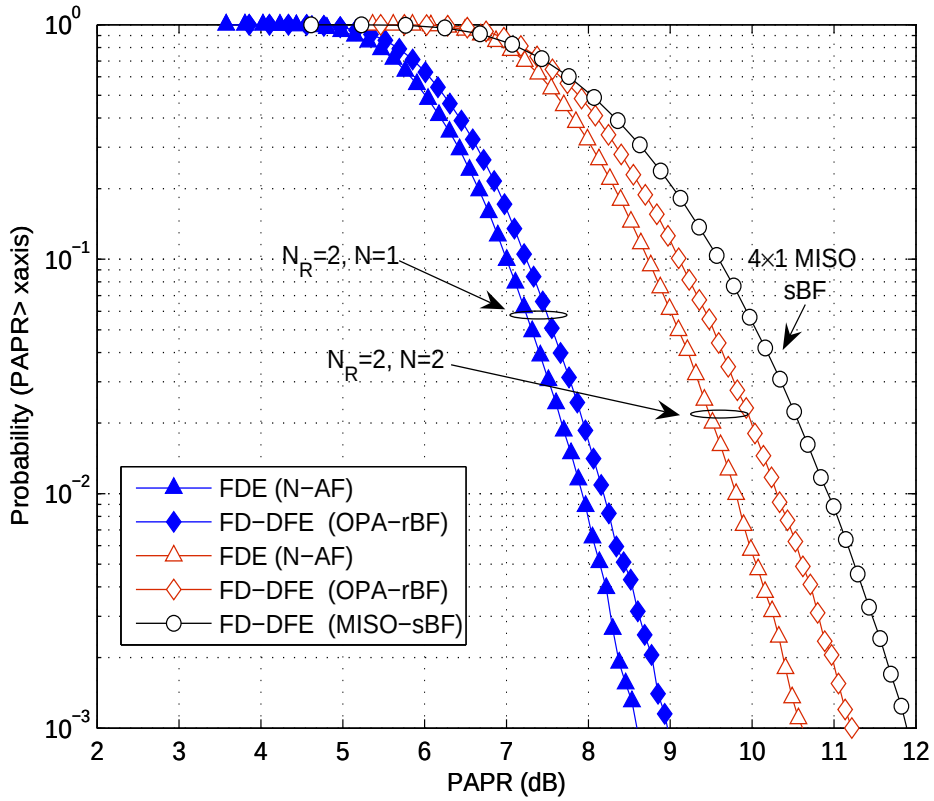

Figure 2.8: PAPR at relays for SC FD-DFE with N-AF and OPA-rBF. $N = \{1, 2\}$, $N_R = 2$, and $E_b/N_0 = 14$ dB. For comparison, the PAPRs of SC-FDE with direct transmission using MISO-sBF with four transmit antennas are also shown.

2.5.3 PAPR at Relays

Since one of the primary reasons for adopting SC-FDE is its low PAPR, we study in Fig. 2.8 the complementary cumulative density function (CCDF) of the PAPR at the output of the relays for a system with $N_R = 2$ relays employing one and two antennas, respectively. For comparison, we also include the PAPR at the source for MISO-sBF with four transmit antennas. From Fig. 2.8 we observe that N-AF has a lower PAPR than OPA-rBF. However, the differences between both schemes is quite small. Furthermore, Fig. 2.8 shows that additional relay antennas adversely affect

the PAPR of all considered schemes due to the increased level of signal fluctuation caused by the joint data processing across transmit antennas. For the same reason, MISO-sBF introduces a comparatively high PAPR at the source, which compromises the key advantage of SC transmission. In fact, the proposed rBF scheme shifts the PAPR problem from the source to the relays, which makes it attractive for application for uplink transmission.

2.6 Conclusion

In this chapter, we tackled the problem of optimal cooperative BF design for SC-FDE systems with multiple relays. We derived objective functions for rBF matrix optimization for FD-LE, FD-DFE, and idealized MF receivers and demonstrated that the structure of the optimal rBF matrices is identical for all considered receivers. Up to a power allocation factor, we obtained a closed-form expression for the optimal rBF matrices. The optimal power allocation across subchannels depends on the adopted receiver and we could show that the related optimization problem is convex, which allows for an efficient numerical solution. We also proposed suboptimal power allocation schemes, which perform close to OPA and show a remarkable robustness against imperfect CSI, hence constituting viable low-complexity alternatives to OPA in practice. Furthermore, for optimal rBF, simple FD-LE receivers approach the performance of the idealized MF receiver as the numbers of relays and/or relay antennas increase, making more complex nonlinear trellis-based receivers unnecessary.

Chapter 3

Transceiver Design For SC-FDE Based MIMO Relay Systems

3.1 Introduction

In Chapter 2, we have considered the transceiver optimization for a relay system with multi-antenna relays. However, the source and the destination are only equipped with a single transmit/receive antenna, and only a single spatial data stream is transmitted through the communication links, i.e., beamforming-based transmission is adopted [53]. In order to support the requirement of high-speed data rates for future wireless communication networks, multiple data stream transmission through spatial multiplexing [54] is highly desired [14, 17]. This requires the deployment of multiple antennas at all the transmit and receive terminals of the relay network, giving rise to the so-called MIMO relay systems [56, 57].

In this chapter, we shall investigate the joint transceiver design for MIMO broadband AF relay systems employing either FD-LE or FD-DFE at the destination. We optimize the source and relay precoding matrices for minimization of a general function of the MSEs of the spatial streams under separate power constraints for the source and the relay. Specifically, we adopt the arithmetic MSE (AMSE), the geometric MSE (GMSE), and the maximum MSE (maxMSE) [64, 70] as objective functions, which are closely related to the achievable bit rate and the bit-error rate performance. For

the case of FD-LE, we show that the optimal source and relay precoding matrices have a structure very similar to that of the optimal precoding matrices in MIMO-OFDM relay systems. However, the remaining power allocation problem is significantly different from the power allocation problem for MIMO-OFDM relay systems, especially for the GMSE and maxMSE criteria. For FD-DFE, the considered objective functions cannot be explicitly expressed in terms of the optimization variables and depend on the number of feedback filter taps, which makes a direct solution of the optimization problem challenging. However, we can show that for FD-DFE, the three considered objective functions are equivalent. Furthermore, we develop an upper bound for the objective function which is independent of the number of feedback filter taps and is a comparatively simple function of the optimization variables. Interestingly, this upper bound is shown to be identical to the GMSE objective function for the FD-LE receiver. Consequently, a unified solution for the power allocation problem for both FD-LE and FD-DFE can be obtained, which greatly simplifies the design procedure.

The remainder of this chapter is organized as follows. In Section 3.2, the system model is presented. In Section 3.3, the optimal minimum MSE FDE filters and the corresponding stream MSE matrices are derived. The optimal source and relay precoding matrices are presented in Section 3.4. Simulation results are given in Section 3.5, and some conclusions are drawn in Section 3.6.

3.2 System Model

We consider a block transmission system with one source node, S, one relay node, R, and one destination node, D, as shown in Fig. 3.1. The numbers of antennas at S, R, and D are denoted by N_s, N_r, and N_d, respectively. The number of spatial multiplexing data streams is $M \leq \min\{N_s, N_r, N_d\}$. The transmission is organized in

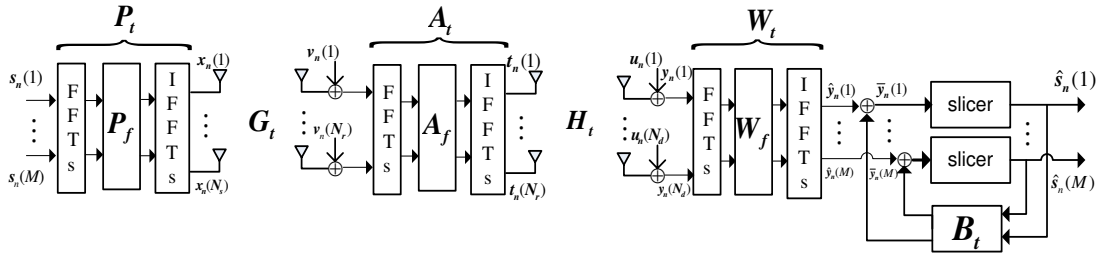

Figure 3.1: System model for a MIMO relay system with SC-FDE at the destination.

two phases. In the first phase, S processes the information symbols and sends them to R. In the second phase, R processes the received signal and retransmits it to D. We assume there is no direct link between S and D due to the large pathloss and/or shadowing.

The transmit signal of each source antenna is prepended by a CP, which comprises the last $N_{g,s} \geq L_g$ symbols of the transmitted source signal, where L_g denotes the largest CIR length between any S-R antenna pair[2]. Similarly, the transmit signal of each relay antenna is prepended by a CP, which comprises the last $N_{g,r} \geq L_h$ symbols of the transmitted relay signal, where L_h is the largest CIR length between any R-D antenna pair.

3.2.1 Precoding at Source and Relay

Let us denote the nth source data symbol vector as $\mathbf{s}_n = [s_n(1), s_n(2), \ldots, s_n(M)]^T$, $n = 0, \ldots, N_c - 1$, where N_c is the size of the data block, and $s_n(j)$ denotes the nth symbol of the jth data stream. Symbols $s_n(j)$ are independent and identically distributed random variables with zero mean and variance σ_s^2. By stacking all $\mathbf{s}_n$ into one vector, we obtain $\mathbf{s} = [\mathbf{s}_0^T, \ldots, \mathbf{s}_{N_c-1}^T]^T \in \mathbb{C}^{MN_c \times 1}$. The received signal at the

[2]For simplicity of presentation, the insertion and removal of the CPs are not shown in Fig. 3.1.

destination, $\mathbf{y}$, can be compactly written as

$$\mathbf{y} = \mathbf{H}_t\mathbf{A}_t\mathbf{G}_t\mathbf{P}_t\mathbf{s} + \mathbf{H}_t\mathbf{A}_t\mathbf{v} + \mathbf{u} \tag{3.1}$$

with block circular matrices

$$\begin{aligned}
\mathbf{P}_t &= \text{blkcirc}([\mathbf{P}_{t,0}^T, \cdots, \mathbf{P}_{t,N_c-1}^T]^T), \\
\mathbf{G}_t &= \text{blkcirc}([\mathbf{G}_{t,0}^T, \cdots, \mathbf{G}_{t,L_g-1}^T, \mathbf{0}_{N_s\times N_r(N_c-L_g)}]^T), \\
\mathbf{A}_t &= \text{blkcirc}([\mathbf{A}_{t,0}^T, \cdots, \mathbf{A}_{t,N_c-1}^T]^T), \\
\mathbf{H}_t &= \text{blkcirc}([\mathbf{H}_{t,0}^T, \cdots, \mathbf{H}_{t,L_h-1}^T, \mathbf{0}_{N_r\times N_d(N_c-L_h)}]^T),
\end{aligned}$$

where $\mathbf{P}_{t,l} \in \mathbb{C}^{N_s\times M}$, $\mathbf{G}_{t,l} \in \mathbb{C}^{N_r\times N_s}$, $\mathbf{A}_{t,l} \in \mathbb{C}^{N_r\times N_r}$, and $\mathbf{H}_{t,l} \in \mathbb{C}^{N_d\times N_r}$ denote the lth tap of the TD source precoding filter, the S-R CIR, the TD relay precoding filter, and the R-D CIR, respectively. The noise vectors at R and D are denoted by

$$\begin{aligned}
\mathbf{v} &= [\mathbf{v}_0^T, \ldots, \mathbf{v}_{N_c-1}^T]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I}_{N_rN_c}), \\
\mathbf{u} &= [\mathbf{u}_0^T, \ldots, \mathbf{u}_{N_c-1}^T]^T \sim \mathcal{CN}(\mathbf{0}, \sigma_u^2\mathbf{I}_{N_dN_c}),
\end{aligned}$$

where $\mathbf{v}_n = [v_n(1), v_n(2), \ldots, v_n(N_r)]^T$ and $\mathbf{u}_n = [u_n(1), u_n(2), \ldots, u_n(N_d)]^T$ denote the AWGN vectors at R and D at time n, respectively. It is noted that the block circular structure of the source and relay precoding matrices in (3.1) is imposed to facilitate the implementation of efficient FDE at the destination, cf. Section 3.3. The block circular matrices $\{\mathbf{P}_t, \mathbf{G}_t, \mathbf{A}_t, \mathbf{H}_t\}$ can be decomposed as

$$\begin{aligned}
\mathbf{P}_t &= \mathbf{F}_{N_s}^\dagger\mathbf{P}_f\mathbf{F}_M, \quad \mathbf{G}_t = \mathbf{F}_{N_r}^\dagger\mathbf{G}_f\mathbf{F}_{N_s}, \\
\mathbf{A}_t &= \mathbf{F}_{N_r}^\dagger\mathbf{A}_f\mathbf{F}_{N_r}, \quad \mathbf{H}_t = \mathbf{F}_{N_d}^\dagger\mathbf{H}_f\mathbf{F}_{N_r},
\end{aligned} \tag{3.2}$$

where $\mathbf{F}_\Upsilon = \mathbb{F}_{N_c}^\dagger \otimes \mathbf{I}_\Upsilon$, $\mathbb{F}_{N_c}$ is the DFT matrix of size N_c and $\Upsilon \in \{M, N_s, N_r, N_d\}$, $\mathbf{X}_f = \text{blkdiag}([\mathbf{X}_0^T, \cdots, \mathbf{X}_{N_c-1}^T]^T)$, and $\mathbf{X}_f \in \{\mathbf{P}_f, \mathbf{G}_f, \mathbf{A}_f, \mathbf{H}_f\}$. Here, $\mathbf{P}_k \in \mathbb{C}^{N_s \times M}$, $\mathbf{G}_k \in \mathbb{C}^{N_r \times N_s}$, $\mathbf{A}_k \in \mathbb{C}^{N_r \times N_r}$, and $\mathbf{H}_k \in \mathbb{C}^{N_d \times N_r}$ represent the FD source precoding, S-R channel, relay precoding, and R-D channel matrices for the kth frequency tone, respectively. For subsequent use, we introduce the transmitted signal vectors at the source and the relay explicitly as

$$\mathbf{x} = \mathbf{P}_t \mathbf{s} \quad \text{and} \quad \mathbf{t} = \mathbf{A}_t(\mathbf{G}_t \mathbf{x} + \mathbf{v}), \tag{3.3}$$

respectively. We further define the equivalent end-to-end channel matrix $\mathbf{Q}_t = \mathbf{H}_t \mathbf{A}_t \mathbf{G}_t \mathbf{P}_t$ and express it as $\mathbf{Q}_t = \mathbf{F}_{N_d}^\dagger \mathbf{Q}_f \mathbf{F}_M$, where $\mathbf{Q}_f = \text{blkdiag}([\mathbf{Q}_0^T, \cdots, \mathbf{Q}_{N_c-1}^T]^T)$ with

$$\mathbf{Q}_k = \mathbf{H}_k \mathbf{A}_k \mathbf{G}_k \mathbf{P}_k \in \mathbb{C}^{N_d \times M} \tag{3.4}$$

representing the equivalent S-D channel matrix on the kth frequency tone. Furthermore, the covariance matrix of the equivalent noise vector $\mathbf{n} = \mathbf{H}_t \mathbf{A}_t \mathbf{v} + \mathbf{u}$ can be obtained as

$$\mathbf{K} = E[\mathbf{n}\mathbf{n}^\dagger] = \mathbf{F}_{N_d}^\dagger \mathbf{K}_f \mathbf{F}_{N_d}, \tag{3.5}$$

where $\mathbf{K}_f = \sigma_v^2 \mathbf{H}_f \mathbf{A}_f \mathbf{A}_f^\dagger \mathbf{H}_f^\dagger + \sigma_u^2 \mathbf{I}_{N_d N_c}$.

In this chapter, we assume that relay and destination can perfectly estimate the S-R and R-D CIR coefficients using training symbols emitted by source and relay, respectively. These CIR coefficients are then fed back to the source and relay, respectively. We note that the overhead of feeding back time domain CSI is much lower than feeding back frequency domain CSI. For example, the feedback overhead

from the destination to the relay is $L_h N_d N_r$ and $N_c N_d N_r$ complex numbers for the time domain and the frequency domain CSI, respectively. Since we have in practice $L_h \ll N_c$, feeding back the time domain CSI incurs a much lower signalling overhead, especially for large values of N_c. In addition, with perfect CSI available at all nodes, adaptive modulation and coding across spatial data streams can be considered. This aspect will be partially investigated in Section 3.5, where the achievable bit rate of the considered system is optimized implying perfect channel loading with a continuous varying signal constellation size and optimal channel coding [11]. The system design for maximum bit rate performance with discrete constellations and practical coding schemes is not considered here but is an interesting topic for future work. Also, note that the source transmit symbols considered in this work are very general and include both discrete constellation symbols as well as Gaussian distributed symbols.

3.2.2 Equalization at the Destination

The received signal $\mathbf{y}$ is transformed into the FD using $\mathbf{F}_{N_d}$ and equalized by an FD-FFF $\mathbf{W}_f = \text{blkdiag}([\mathbf{W}_0^T, \cdots, \mathbf{W}_{N_c-1}^T]^T)$. The resulting signal is then transformed into the TD using $\mathbf{F}_M^\dagger$ resulting in

$$\hat{\mathbf{y}} \;=\; \mathbf{W}_t \mathbf{y}, \tag{3.6}$$

where $\mathbf{W}_t = \mathbf{F}_M^\dagger \mathbf{W}_f \mathbf{F}_{N_d}$ is the equivalent TD FFF and $\hat{\mathbf{y}} = [\hat{\mathbf{y}}_0^T, \ldots, \hat{\mathbf{y}}_{N_c-1}^T]^T$ with $\hat{\mathbf{y}}_n = [\hat{y}_n(1), \hat{y}_n(2), \ldots, \hat{y}_n(M)]^T$ denoting the nth signal vector at the output of the FFF. If FD-LE is employed, $\hat{\mathbf{y}}_n$ is the decision variable for the nth source symbol vector. On the other hand, for FD-DFE, $\hat{\mathbf{y}}_n$ is further processed using a TD-FBF to perform interference cancellation. Assuming correct feedback at the output of the

slicer[3], the signal corresponding to the mth data stream at time n at the input of the slicer is given by

$$\bar{y}_n(m) = \hat{y}_n(m) - \sum_{l=0}^{N_{fb}} [\mathbf{B}_{t,l}]_{(m,:)} \mathbf{s}_{(n-l) \bmod N_c}, \tag{3.7}$$

where $\mathbf{B}_{t,l}$ denotes the coefficient matrix of the lth tap of the FBF, $[\mathbf{X}]_{(m,:)}$ stands for the mth row of matrix $\mathbf{X}$, N_{fb} is the number of feedback taps, and $(\cdot) \bmod N$ denotes the modulo-N operation. From (3.7) we observe that at the initial stage of the feedback process, i.e., when $n = 0$, $[\mathbf{s}_{N_c - N_{fb}}, \cdots, \mathbf{s}_{N_c}]$ has to be known *a priori*, which can be accomplished by using known training symbols. Nevertheless, for detection of $s_n(m)$, $[s_n(1), \cdots, s_n(m-1)]$ is still unknown. Therefore, for causal detection, the 0th tap of the FBF, i.e., $\mathbf{B}_{t,0}$, has to be a lower triangular matrix with zero diagonal entries. By collecting all $\bar{y}_n(m)$ into a vector $\bar{\mathbf{y}} = [\bar{\mathbf{y}}_0^T, \ldots, \bar{\mathbf{y}}_{N_c-1}^T]^T$ with $\bar{\mathbf{y}}_n = [\bar{y}_n(1), \bar{y}_n(2), \ldots, \bar{y}_n(M)]^T$, we arrive at

$$\bar{\mathbf{y}} = \hat{\mathbf{y}} - \mathbf{B}_t \mathbf{s}, \tag{3.8}$$

where $\mathbf{B}_t = \text{blkcirc}([\mathbf{B}_{t,0}^T, \cdots, \mathbf{B}_{t,N_{fb}}^T, \mathbf{0}_{M \times M(N_c - N_{fb}-1)}]^T) \in \mathbb{C}^{MN_c \times MN_c}$ is the equivalent TD FBF. Thus, the error vector at the input of the slicer can be expressed as

$$\mathbf{e} = \bar{\mathbf{y}} - \mathbf{s} = \hat{\mathbf{y}} - \underbrace{(\mathbf{B}_t + \mathbf{I}_{MN_c})}_{\mathbf{C}_t} \mathbf{s} = \hat{\mathbf{y}} - \mathbf{C}_t \mathbf{s}, \tag{3.9}$$

where $\mathbf{C}_t = \text{blkcirc}([\mathbf{C}_{t,0}^T, \cdots, \mathbf{C}_{t,N_{fb}}^T, \mathbf{0}_{M \times M(N_c - N_{fb}-1)}]^T)$ with $\mathbf{C}_{t,n} = \mathbf{B}_{t,n}, \forall n \neq 0$ and $\mathbf{C}_{t,0} = \mathbf{B}_{t,0} + \mathbf{I}_M$. The block circular matrix $\mathbf{C}_t$ can be decomposed as $\mathbf{C}_t = \mathbf{F}_M^\dagger \mathbf{C}_f \mathbf{F}_M$,

[3] Correct feedback is a common assumption for the design of decision feedback equalizers [11, 106].

where $\mathbf{C}_f = \text{blkdiag}([\mathbf{C}_0^T, \cdots, \mathbf{C}_{N_c-1}^T]^T)$. We note that by setting $\mathbf{C}_t = \mathbf{I}_{MN_c}$, FD-DFE reduces to FD-LE.

3.3 Optimal Minimum MSE FDE Filter Design

In this section, we derive the optimal minimum MSE equalization filters at the destination and the corresponding MSE at the output of the equalizer as functions of the source and relay precoding matrices. Combining (3.1)-(3.6) and (3.9), the MSE matrix, $\mathbf{E} \triangleq E[\mathbf{e}\mathbf{e}^\dagger]$, can be expressed as

$$\mathbf{E} = \mathbf{F}_M^\dagger \Big(\mathbf{W}_f (\sigma_s^2 \mathbf{Q}_f \mathbf{Q}_f^\dagger + \mathbf{K}_f) \mathbf{W}_f^\dagger - \sigma_s^2 \mathbf{W}_f \mathbf{Q}_f \mathbf{C}_f^\dagger - \sigma_s^2 \mathbf{C}_f \mathbf{Q}_f^\dagger \mathbf{W}_f^\dagger + \sigma_s^2 \mathbf{C}_f \mathbf{C}_f^\dagger \Big) \mathbf{F}_M. \tag{3.10}$$

Following the conventional equalization design methodology, the optimum FD FFF is obtained by minimizing the sum of stream MSEs, $\text{tr}(\mathbf{E})$, which yields

$$\mathbf{W}_f^\star = \sigma_s^2 \mathbf{C}_f \mathbf{Q}_f^\dagger \left(\sigma_s^2 \mathbf{Q}_f \mathbf{Q}_f^\dagger + \mathbf{K}_f \right)^{-1}. \tag{3.11}$$

From (3.11), it is observed that in order to guarantee the block diagonal structure of the FDE filter matrix $\mathbf{W}_f^\star$ for efficient FD implementation, the matrices on the right hand side of (3.11) have to be block diagonal matrices. This justifies the special structure of the TD source and relay precoding matrices imposed in (3.1). Using (3.11) in (3.10) and simplifying the resulting expression, the MSE matrix becomes

$$\mathbf{E} = \sigma_s^2 \mathbf{F}_M^\dagger \mathbf{C}_f \mathbf{\Psi}_f^{-1} \mathbf{C}_f^\dagger \mathbf{F}_M, \tag{3.12}$$

where $\mathbf{\Psi}_f = \text{blkdiag}([\mathbf{\Psi}_0^T, \cdots, \mathbf{\Psi}_{N_c-1}^T]^T) \in \mathbb{C}^{MN_c \times MN_c}$ with

$$\mathbf{\Psi}_k = \sigma_s^2 \mathbf{Q}_k^\dagger \left(\sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d} \right)^{-1} \mathbf{Q}_k + \mathbf{I}_M. \tag{3.13}$$

From (3.12) we observe that $\mathbf{E}$ is a block circular matrix. Hence, its block diagonal entries, $\mathbf{E}_n \in \mathbb{C}^{M \times M}, \forall n$, are identical, i.e., $\mathbf{E}_n = \hat{\mathbf{E}}, \forall n$. Since the diagonal entries of $\mathbf{E}_n$ represent the MSEs of the different spatial streams at time n, symbols from the same stream experience identical MSEs. Exploiting the block circular structure of $\mathbf{E}$, we can obtain the MSE matrix $\hat{\mathbf{E}}$ for symbol vector $\mathbf{s}_n$ at each time $n = 0, \ldots, N_c - 1$ as

$$\hat{\mathbf{E}} = \frac{\sigma_s^2}{N_c} \sum_{k=0}^{N_c-1} \mathbf{C}_k^\dagger \mathbf{\Psi}_k^{-1} \mathbf{C}_k. \tag{3.14}$$

3.3.1 MSE Matrix and Filter Design for FD-LE

Eqs. (3.11) and (3.14) are valid for both FD-LE and FD-DFE. For the special case of FD-LE, we can set $\mathbf{C}_f = \mathbf{I}_{MN_c}$, which leads to

$$\mathbf{W}_f^\star = \sigma_s^2 \mathbf{Q}_f^\dagger \left(\sigma_s^2 \mathbf{Q}_f \mathbf{Q}_f^\dagger + \mathbf{K}_f \right)^{-1} \tag{3.15}$$

and MSE matrix

$$\hat{\mathbf{E}}_{\text{FD-LE}} = \frac{\sigma_s^2}{N_c} \sum_{k=0}^{N_c-1} \mathbf{\Psi}_k^{-1}. \tag{3.16}$$

Interestingly, $\hat{\mathbf{E}}_{\text{FD-LE}}$ is equal to the arithmetic mean of the subcarrier MSE matrices, $\mathbf{\Psi}_k^{-1}$, in MIMO-OFDM relay systems [64].

3.3.2 MSE Matrix and Filter Design for FD-DFE

The FD-DFE MSE matrix depends on the FD feedback filter matrices $\mathbf{C}_k$. Since the feedback filter has to be implemented in the TD, we express $\mathbf{C}_k$ in terms of the TD feedback filter coefficients $\mathbf{C}_{t,n}$ as $\mathbf{C}_k = \sum_{n=0}^{N_{fb}} \mathbf{C}_{t,n} e^{-j\frac{2\pi}{N_c}nk}$. Now, (3.14) can be rewritten as

$$\begin{aligned}\hat{\mathbf{E}}_{\text{FD-DFE}} &= \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\left[\Big(\sum_{n=0}^{N_{fb}}\mathbf{C}_{t,n}e^{-j\frac{2\pi nk}{N_c}}\Big)\mathbf{\Psi}_k^{-1}\Big(\sum_{m=0}^{N_{fb}}\mathbf{C}_{t,m}^{\dagger}e^{-j\frac{2\pi mk}{N_c}}\Big)\right] \\ &= \frac{\sigma_s^2}{N_c}\sum_{n=0}^{N_{fb}}\sum_{m=0}^{N_{fb}}\left(\mathbf{C}_{t,n}\sum_{k=0}^{N_c-1}\mathbf{\Psi}_k^{-1}e^{-j\frac{2\pi}{N_c}(n-m)k}\mathbf{C}_{t,m}^{\dagger}\right) = \frac{\sigma_s^2}{N_c}\hat{\mathbf{C}}\mathbf{Z}\hat{\mathbf{C}}^{\dagger}. \end{aligned} \tag{3.17}$$

To simplify the notation, we have used the definitions $\hat{\mathbf{C}} = [\mathbf{C}_{t,0}, ..., \mathbf{C}_{t,N_{fb}}]$ and

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_0 & \mathbf{z}_1 & \dots & \mathbf{z}_{N_{fb}} \\ \mathbf{z}_1^{\dagger} & \mathbf{z}_0 & \dots & \mathbf{z}_{N_{fb}-1} \\ \vdots & & \ddots & \vdots \\ \mathbf{z}_{N_{fb}}^{\dagger} & \mathbf{z}_{N_{fb}-1}^{\dagger} & \dots & \mathbf{z}_0 \end{bmatrix}, \tag{3.18}$$

where $\mathbf{z}_n = \sum_{k=0}^{N_c-1}\mathbf{\Psi}_k^{-1}e^{j\frac{2\pi}{N_c}kn}$. The optimal $\hat{\mathbf{C}}$ minimizing $\text{tr}\{\mathbf{E}_{\text{FD-DFE}}\}$ can be obtained as

$$\hat{\mathbf{C}}^{\star} = \arg\min_{\hat{\mathbf{C}}\mathbf{\Theta}=\mathbf{C}_{t,0}} \quad \text{tr}(\hat{\mathbf{C}}\mathbf{Z}\hat{\mathbf{C}}^{\dagger}), \tag{3.19}$$

where $\mathbf{\Theta} = [\mathbf{I}_M, \mathbf{0}_{M\times(N_{fb}-1)}]^{\dagger}$. Problem (3.19) can be solved using the standard Lagrange multiplier method, leading to [106]

$$\hat{\mathbf{C}}^{\star} = \mathbf{C}_{t,0}(\mathbf{\Theta}^{\dagger}\mathbf{Z}^{-1}\mathbf{\Theta})^{-1}\mathbf{\Theta}^{\dagger}\mathbf{Z}^{-1}. \tag{3.20}$$

By partitioning $\mathbf{Z}$ and $\mathbf{Z}^{-1}$ as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{12}^{\dagger} & \mathbf{Z}_{22} \end{bmatrix}, \mathbf{Z}^{-1} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{12}^{\dagger} & \mathbf{U}_{22} \end{bmatrix}, \tag{3.21}$$

where $\{\mathbf{Z}_{11}, \mathbf{U}_{11}\} \in \mathbb{C}^{M \times M}$, $\{\mathbf{Z}_{12}, \mathbf{U}_{12}\} \in \mathbb{C}^{M \times MN_{fb}}$, and $\{\mathbf{Z}_{22}, \mathbf{U}_{22}\} \in \mathbb{C}^{MN_{fb} \times MN_{fb}}$, we can exploit the formula for the inverse of partitioned matrices [100],

$$\mathbf{U}_{11} = (\mathbf{Z}_{11} - \mathbf{Z}_{12}\mathbf{Z}_{22}^{-1}\mathbf{Z}_{12}^{\dagger})^{-1}, \mathbf{U}_{12} = -\mathbf{U}_{11}^{-1}\mathbf{Z}_{12}\mathbf{Z}_{22}^{-1}, \tag{3.22}$$

to further express $\hat{\mathbf{C}}^{\star}$ as

$$\hat{\mathbf{C}}^{\star} = [\mathbf{C}_{t,0}, \quad -\mathbf{C}_{t,0}\mathbf{Z}_{12}\mathbf{Z}_{22}^{-1}]. \tag{3.23}$$

Substituting (3.23) into (3.17), the FD-DFE MSE matrix can be rewritten as

$$\mathbf{E}_{\mathrm{FD-DFE}} = \frac{\sigma_s^2}{N_c}\mathbf{C}_{t,0}\mathbf{U}_{11}^{-1}\mathbf{C}_{t,0}^{\dagger}. \tag{3.24}$$

To complete the feedback filter design, the optimal $\mathbf{C}_{t,0}$ has to be determined. To this end, we introduce the Cholesky decomposition of $\mathbf{U}_{11}^{-1}$ as

$$\mathbf{U}_{11}^{-1} = \mathbf{L}\mathbf{D}\mathbf{L}^{\dagger}, \tag{3.25}$$

where $\mathbf{L}$ is a unit-diagonal lower triangular matrix and $\mathbf{D}$ is a diagonal matrix with positive main diagonal entries. Now, it is easy to verify that the optimal $\mathbf{C}_{t,0}$ which minimizes $\mathrm{tr}(\mathbf{E}_{\mathrm{FD-DFE}})$ is given by $\mathbf{C}_{t,0}^{\star} = \mathbf{L}^{-1}$. Hence, the optimal $\hat{\mathbf{C}}$ is obtained as

$$\hat{\mathbf{C}}^{\star} = [\mathbf{L}^{-1}, \quad -\mathbf{L}^{-1}\mathbf{Z}_{12}\mathbf{Z}_{22}^{-1}]. \tag{3.26}$$

The structure of the optimal feedback filter can be interpreted as follows: $\mathbf{L}^{-1} \in \mathbb{C}^{M\times M}$ is a lower-triangular matrix which cancels the inter-stream interference in the current time slot, and the remaining feedback filter coefficients, $-\mathbf{L}^{-1}\mathbf{Z}_{12}\mathbf{Z}_{22}^{-1} \in \mathbb{C}^{M\times MN_{fb}}$, cancel both the inter-stream interference and the inter-symbol interference stemming from the previous $N_{fb}-1$ time slots. Inserting $\hat{\mathbf{C}}^{\star}$ into (3.24), the MSE matrix can be written as

$$\hat{\mathbf{E}}_{\mathrm{FD-DFE}} = \frac{\sigma_s^2}{N_c}\mathbf{D}. \tag{3.27}$$

Since $\mathbf{D}$ in (3.25) is a diagonal matrix, unlike for FD-LE, the MSE matrix for FD-DFE is a diagonal matrix, and also depends on the number of feedback filter taps N_{fb}.

3.4 Source and Relay Precoding Matrix Optimization

Exploiting the expressions for the MSE matrix obtained in the previous section, in this section, we minimize a general function $f(\mathrm{diag}[\hat{\mathbf{E}}])$ of the spatial stream MSEs at the output of the equalization filter under separate constraints on the transmit power consumed at the source and the relay, respectively[4]. Mathematically, the

[4]We note that our derivations can be extended to a joint source and relay transmit power constraint. While such a joint transmit power constraint offers more degrees of freedom for the system design, separate transmit power constraints appear more practical since usually the source node and the relay node have their own power supplies.

optimization problem is stated as

$$\begin{aligned} \min_{\{\mathbf{P}_k,\mathbf{A}_k\}} \quad & f(\mathrm{diag}[\hat{\mathbf{E}}]) \\ \text{s.t.} \quad & \mathrm{tr}\left(E[\mathbf{x}\mathbf{x}^\dagger]\right) \le P_S, \quad \mathrm{tr}\left(E[\mathbf{t}\mathbf{t}^\dagger]\right) \le P_R, \end{aligned} \tag{3.28}$$

where $\hat{\mathbf{E}} = \hat{\mathbf{E}}_{\mathrm{FD-LE}}$ and $\hat{\mathbf{E}} = \hat{\mathbf{E}}_{\mathrm{FD-DFE}}$ for FD-LE and FD-DFE, respectively, P_S and P_R are the transmit power limits for S and R, respectively, $\mathbf{x}$ and $\mathbf{t}$ are given in (3.3), and diag[$\mathbf{M}$] denotes a vector containing the main diagonal entries of matrix $\mathbf{M}$. The objective function $f(\mathrm{diag}[\hat{\mathbf{E}}])$ can be either a Schur-convex or a Schur-concave increasing function with respect to (w.r.t.) diag[$\hat{\mathbf{E}}$] [70]. For concreteness, in this chapter, we consider the three most important objective functions of this type, namely the arithmetic MSE (AMSE), the geometric MSE (GMSE), and the maximum MSE (maxMSE). We note that AMSE is an important performance metric for classical signal processing and communication systems, and the solution for AMSE minimization can be used as a building block for solving more complex problems such as maxMSE minimization [121]. On the other hand, GMSE minimization is appealing due to its equivalence to average bit rate maximization, see Section 3.4.2. Finally, maxMSE is closely related to the average system error rate performance since the worst-case MSE dominates the average uncoded symbol/bit-error rate [70]. Specifically, these objective functions can be written as

$$f(\mathrm{diag}[\hat{\mathbf{E}}]) = \begin{cases} \sum_{m=1}^{M} \hat{\mathbf{E}}_{mm}, & \text{AMSE} \\ \prod_{m=1}^{M} \hat{\mathbf{E}}_{mm}, & \text{GMSE} \\ \max_{m=1}^{M} \hat{\mathbf{E}}_{mm}, & \text{maxMSE} \end{cases}, \tag{3.29}$$

where $\hat{\mathbf{E}}_{mm}$ denotes the mth diagonal entry of $\hat{\mathbf{E}}$. The AMSE and GMSE are Schur-concave functions while the maxMSE is a Schur-convex function w.r.t. diag[$\hat{\mathbf{E}}$] [70]. We note that similar objective functions have been considered for MIMO-OFDM based relay systems in [64]. However, for MIMO-OFDM based relay systems, the AMSE, GMSE, and maxMSE are the sum, product, and maximum of the subcarrier MSEs of different spatial streams. In contrast, in (3.29), these three quantities are the sum, product and maximum of the stream MSEs of a single carrier.

The transmit power consumptions at source and relay are given by

$$\begin{aligned}
\mathrm{tr}\left(E[\mathbf{x}\mathbf{x}^\dagger]\right) &= \sigma_s^2 \sum_{k=0}^{N_c-1} \mathrm{tr}\left(\mathbf{P}_k\mathbf{P}_k^\dagger\right), \\
\mathrm{tr}\left(E[\mathbf{t}\mathbf{t}^\dagger]\right) &= \sum_{k=0}^{N_c-1} \mathrm{tr}\left(\mathbf{A}_k\left(\sigma_s^2\mathbf{G}_k\mathbf{P}_k\mathbf{P}_k^\dagger\mathbf{G}_k^\dagger + \sigma_v^2\mathbf{I}_{N_r}\right)\mathbf{A}_k^\dagger\right).
\end{aligned} \tag{3.30}$$

Since the optimization variables in (3.28) are matrices, solving the problem directly would incur high complexity. In the following, we will first derive the structure, i.e., the optimal form of the singular-value decompositions (SVDs), of the source and relay precoding matrices. Knowing this structure will allow us to transform the matrix optimization problem into an optimization problem with scalar variables.

3.4.1 Structure of the Optimal Precoding Matrices for FD-LE

We first derive the structure of the optimal source and relay precoding matrices for FD-LE. We begin by introducing the following SVDs of the FD channel matrices

$$\mathbf{G}_k = \mathbf{U}_G^{(k)}\mathbf{\Lambda}_G^{(k)}\mathbf{V}_G^{(k)\dagger}, \quad \mathbf{H}_k = \mathbf{U}_H^{(k)}\mathbf{\Lambda}_H^{(k)}\mathbf{V}_H^{(k)\dagger}, \quad \forall k, \tag{3.31}$$

where $\mathbf{U}_G^{(k)} \in \mathbb{C}^{N_r \times N_r}$, $\mathbf{V}_G^{(k)} \in \mathbb{C}^{N_s \times N_s}$ and $\mathbf{U}_H^{(k)} \in \mathbb{C}^{N_d \times N_d}$, $\mathbf{V}_H^{(k)} \in \mathbb{C}^{N_r \times N_r}$ are the singular-vector matrices of $\mathbf{G}_k$ and $\mathbf{H}_k$, respectively. Furthermore, $\mathbf{\Lambda}_G^{(k)} \in \mathbb{C}^{N_r \times N_s}$ and $\mathbf{\Lambda}_H^{(k)} \in \mathbb{C}^{N_d \times N_r}$ are the singular-value matrices of $\mathbf{G}_k$ and $\mathbf{H}_k$, which both have decreasing main diagonal elements.

Theorem 3.1. *For the optimization problem in (3.28), the following structures of* $\mathbf{P}_k$ *and* $\mathbf{A}_k$ *are optimal*

$$\mathbf{P}_k^\star = \bar{\mathbf{V}}_G^{(k)} \mathbf{\Lambda}_P^{(k)} \mathbf{V}_0, \quad \mathbf{A}_k^\star = \bar{\mathbf{V}}_H^{(k)} \mathbf{\Lambda}_A^{(k)} \bar{\mathbf{U}}_G^{(k)}, \quad \forall k, \tag{3.32}$$

where $\bar{\mathbf{V}}_G^{(k)}$, $\bar{\mathbf{U}}_G^{(k)}$, *and* $\bar{\mathbf{V}}_H^{(k)}$ *contain the* M *left-most columns of* $\mathbf{V}_G^{(k)}$, $\mathbf{U}_G^{(k)}$, *and* $\mathbf{V}_H^{(k)}$, *respectively.* $\mathbf{\Lambda}_P^{(k)}$ *and* $\mathbf{\Lambda}_A^{(k)}$ *are* $M \times M$ *diagonal matrices with the* m*th diagonal elements denoted by* p_{km} *and* a_{km}, *respectively. For Schur-concave functions,* $\mathbf{V}_0 = \mathbf{I}_M$. *For Schur-convex functions,* $\mathbf{V}_0$ *is a unitary matrix chosen in such a way that all main diagonal entries of* $\hat{\mathbf{E}}$ *are equal*[5].

Proof. Please refer to Appendix-B. □

Theorem 3.1 shows that the structures of the source and relay precoding matrices have to match those of the S-R and R-D channel matrices, respectively, such that the diagonal power allocation matrices can allocate the available source and relay transmit powers to the decomposed parallel channels in both the spatial domain and the frequency domain[6]. Moreover, for Schur-concave functions, the source and relay

[5] In practice, $\mathbf{V}_0$ can be chosen as a DFT matrix or a Hadamard matrix with appropriate dimensions.

[6] According to Theorem 3.1, the singular values of the S-R and R-D channels are sorted in the same order when deriving the optimal structure of the precoding matrices. This can be interpreted as optimal spatial subchannel pairing. As revealed in [60], employing only spatial subchannel pairing incurs a negligible performance loss compared to joint frequency and spatial subchannel pairing. Hence, to simplify the system design, we do not consider subcarrier pairing (frequency pairing) in this chapter. Nevertheless, the optimal subcarrier pairing for different optimization criteria is an interesting topic for future work.

precoding matrices jointly diagonalize the MIMO relay channels at each frequency tone, while for Schur-convex functions, the precoding matrices diagonalize the channels up to a unitary rotation at the source. Therefore, the original optimization problem involving matrix variables can be transformed into a scalar power allocation problem across different spatial beams and frequency tones.

3.4.2 Transformation of Optimization Problem for FD-LE

Since the maxMSE is a Schur-convex function, according to Theorem 3.1, the unitary matrix, $\mathbf{V}_0$, should be chosen to make all diagonal entries of $\hat{\mathbf{E}}$ equal. Recall from Section 3.3 that diag[$\hat{\mathbf{E}}$] represents the MSE of different spatial streams and all symbols of a particular stream have the same MSE. This means that for maxMSE, identical MSE is achieved for all symbols in the SC-FDE system. Hence, the remaining maxMSE power allocation problem is identical to that for the AMSE criterion. The only difference between the solutions for maxMSE and AMSE minimization lies in the choice of $\mathbf{V}_0$. We note that this is not true for MIMO-OFDM relay systems, where the unitary transformation at the source only achieves identical spatial MSEs on each subcarrier, while the MSEs across subcarriers are in general different. To balance these MSEs, multilevel waterfilling has to be carried out in such MIMO-OFDM relay systems, which entails a much higher complexity compared to the single-level waterfilling required for the AMSE criterion, cf. [64][7]. Additionally, for MIMO-OFDM relay systems, the unitary rotation matrices are in general different on each subcarrier as the number of transmitted data streams may vary from subcarrier to subcarrier. However, for SC-FDE, the rotation matrices are identical

[7]The maximum number of iterations for single-level and multilevel waterfilling algorithms is $\log_2(MN_c)$ and MN_c, respectively [104]. Therefore, for large values of MN_c, e.g., 128, multilevel waterfilling incurs a much higher complexity than single-level waterfilling.

for all frequency tones since the number of data streams is determined in the time domain.

Because of the equivalence of the power allocation problems for maxMSE and AMSE, in the following, we focus on the power allocation problem for the AMSE and GMSE criteria. From (3.13) and (3.32), we obtain $\mathbf{\Psi}_k = \mathbf{V}_0^\dagger \mathbf{\Phi}_k \mathbf{V}_0$ with

$$\mathbf{\Phi}_k = \sigma_s^2 \mathbf{\Lambda}_P^{(k)2} \bar{\mathbf{\Lambda}}_G^{(k)2} \mathbf{\Lambda}_A^{(k)2} \bar{\mathbf{\Lambda}}_H^{(k)2} \left(\sigma_v^2 \mathbf{\Lambda}_A^{(k)2} \bar{\mathbf{\Lambda}}_H^{(k)2} + \sigma_u^2 \mathbf{I}_M \right)^{-1} + \mathbf{I}_M, \tag{3.33}$$

where $\bar{\mathbf{\Lambda}}_G^{(k)}$ and $\bar{\mathbf{\Lambda}}_H^{(k)}$ are diagonal matrices whose diagonal entries contain the M largest singular values of $\mathbf{G}^{(k)}$ and $\mathbf{H}^{(k)}$, respectively. Now, using (3.16) and (3.33), we can rewrite the objective functions in (3.29) as

$$f_{\mathrm{X}}(\mathbf{\Phi}) = \begin{cases} \sum_{m=1}^{M} \left(\frac{1}{N_c} \sum_{k=0}^{N_c-1} \Phi_{km}^{-1} \right), & \mathrm{X{=}AMSE} \\ \sum_{m=1}^{M} \log_2 \left(\frac{1}{N_c} \sum_{k=0}^{N_c-1} \Phi_{km}^{-1} \right), & \mathrm{X{=}GMSE} \end{cases}, \tag{3.34}$$

where $\mathbf{\Phi} = \{\Phi_{km}, \forall k, m\}$ with

$$\Phi_{km} = \frac{\sigma_s^2 p_{km}^2 g_{km}^2 a_{km}^2 h_{km}^2}{\sigma_v^2 a_{km}^2 h_{km}^2 + \sigma_u^2} + 1. \tag{3.35}$$

Here, g_{km} and h_{km} denote the mth main diagonal elements of $\bar{\mathbf{\Lambda}}_G^{(k)}$ and $\bar{\mathbf{\Lambda}}_H^{(k)}$, respectively, and represent the corresponding channel gains of the mth spatial stream on the kth frequency tone. Note that, for the GMSE criterion, we have taken the logarithm of the original objective function in (3.29) to facilitate the subsequent optimization. Due to the monotonicity of the logarithm, the new objective function has the same

optimal solution as the original one. The new objective function can be rewritten as

$$-\sum_{m=1}^{M}\log_2\left(\mathrm{SINR}_m+1\right), \tag{3.36}$$

where $\mathrm{SINR}_m=(\frac{1}{N_c}\sum_{k=0}^{N_c-1}\Phi_{km}^{-1})^{-1}-1$ is the SINR of the mth data stream. This implies that (3.36) is essentially the negative sum of the channel capacity of different spatial streams, which can be approached with Gaussian signalling and ideal channel coding. Therefore, the minimization of the GMSE is equivalent to the maximization of the capacity of the considered MIMO SC-FDE relay system. By exploiting (3.32), the expression for the power consumption on the left hand side of the constraints in (3.28) can be expressed as

$$\begin{aligned}\mathrm{tr}\left(E[\mathbf{x}\mathbf{x}^\dagger]\right)&=\sigma_s^2\sum_{k=0}^{N_c-1}\mathrm{tr}\left(\mathbf{\Lambda}_P^{(k)2}\right)=\sum_{k=0}^{N_c-1}\sum_{m=1}^{M}P_{s,km}\\ \mathrm{tr}\left(E[\mathbf{t}\mathbf{t}^\dagger]\right)&=\sum_{k=0}^{N_c-1}\mathrm{tr}\left(\mathbf{\Lambda}_A^{(k)2}\left(\sigma_s^2\mathbf{\Lambda}_P^{(k)2}\bar{\mathbf{\Lambda}}_G^{(k)2}+\sigma_v^2\mathbf{I}_M\right)\right)=\sum_{k=0}^{N_c-1}\sum_{m=1}^{M}P_{r,km},\end{aligned} \tag{3.37}$$

where

$$P_{s,km}=\sigma_s^2p_{km}^2,\quad P_{r,km}=a_{km}^2(\sigma_s^2p_{km}^2g_{km}^2+\sigma_v^2) \tag{3.38}$$

can be interpreted as the power allocated to the kth frequency tone and the mth spatial stream at the source and the relay, respectively. By rewriting Φ_{km} in (3.35) in terms of the newly introduced variables $P_{s,km}$ and $P_{r,km}$ as

$$\Phi_{km}=\frac{P_{s,km}P_{r,km}g_{km}^2h_{km}^2}{\sigma_v^2P_{r,km}h_{km}^2+\sigma_u^2(P_{s,km}g_{km}^2+\sigma_v^2)}+1, \tag{3.39}$$

problem (3.34) can be reformulated as the following power allocation problem

$$\begin{aligned} \min_{\{P_{s,km},P_{r,km}\}} \; & f_{\mathrm{X}}\left(\mathbf{\Phi}\right) \\ \text{s.t.} \quad & \sum_{k=0}^{N_c-1}\sum_{m=1}^{M} P_{s,km} \le P_S, \quad \sum_{k=0}^{N_c-1}\sum_{m=1}^{M} P_{r,km} \le P_R \\ & P_{s,km} \ge 0, P_{r,km} \ge 0, \forall k, m, \end{aligned} \tag{3.40}$$

where the constraints $P_{s,km} \ge 0, P_{r,km} \ge 0, \forall k, m$, ensure that the allocated powers are not negative.

3.4.3 Structure of the Optimal Precoding Matrices for FD-DFE

For the FD-DFE receiver, we observe from (3.27) that $\mathbf{E}_{\mathrm{FD-DFE}}$ is not an explicit function of optimization variables $\mathbf{P}_k$ and $\mathbf{A}_k$, which renders the optimization a challenging task. In this section, we will show that by using proper transformations, an upper bound for the original objective function can be derived, which is equivalent to one of the objective functions considered for the FD-LE receiver. To this end, we will first show that for FD-DFE, the three considered objective functions are equivalent.

Equivalence of Objective Functions

Since $\mathbf{E}_{\mathrm{FD-DFE}}$ in (3.27) is a diagonal matrix, we invoke the following matrix arithmetic-geometric mean inequality

$$\frac{1}{M}\mathrm{tr}(\mathbf{D}) = \frac{1}{M}\sum_{m=1}^{M}\mathbf{D}_{mm} \ge (\prod_{m=1}^{M}\mathbf{D}_{mm})^{\frac{1}{M}} = \det(\mathbf{D})^{\frac{1}{M}}, \tag{3.41}$$

where the equality holds if and only if (i.f.f.) all main diagonal elements of $\mathbf{D}$ are equal. The inequality provides some important insights into the objective function for FD-DFE. First, it implies that the AMSE, i.e., tr($\mathbf{D}$), is lower bounded by the term involving the GMSE, i.e., det($\mathbf{D}$). Second, this lower bound is achieved i.f.f. the MSEs of all streams are identical. Therefore, making the diagonal entries of $\mathbf{D}$ identical will enable us to minimize the AMSE, GMSE, and maxMSE simultaneously. Consequently, for FD-DFE, the three considered objective functions become equivalent. In the sequel, we will show how this can be achieved by applying a suitable unitary matrix at the source precoder. From (3.25), we obtain

$$\mathbf{U}_{11}^{-1} \;=\; \mathbf{L}\mathbf{D}^{1/2}(\mathbf{L}\mathbf{D}^{1/2})^{\dagger} = (\mathbf{Q}\mathbf{R})^{\dagger}\mathbf{Q}\mathbf{R}, \tag{3.42}$$

where $\mathbf{Q}$ is an arbitrary unitary matrix of appropriate dimension and $\mathbf{R} = (\mathbf{L}\mathbf{D}^{1/2})^{\dagger}$ is a lower triangular matrix whose main diagonal elements are equal to the square root of the main diagonal elements of $\mathbf{D}$. Therefore, finding a diagonal matrix $\mathbf{D}$ with equal diagonal elements is equivalent to finding a triangular matrix $\mathbf{R}$ with equal diagonal elements. In the following, we provide an explicit construction for $\mathbf{R}$. By expressing $\mathbf{P}_k$ as the product of a unitary matrix $\mathbf{V}_1$ and a general matrix $\tilde{\mathbf{P}}_k$,

$$\mathbf{P}_k \;=\; \tilde{\mathbf{P}}_k\mathbf{V}_1^{\dagger}, \tag{3.43}$$

we can write $\mathbf{\Psi}_k$ in (3.13) as

$$\mathbf{\Psi}_k = \mathbf{V}_1\hat{\mathbf{\Psi}}_k\mathbf{V}_1^{\dagger}, \tag{3.44}$$

where

$$\hat{\mathbf{\Psi}}_k = \sigma_s^2 \tilde{\mathbf{P}}_k^\dagger \mathbf{G}_k^\dagger \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger \Big(\sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d} \Big)^{-1} \mathbf{H}_k \mathbf{A}_k \mathbf{G}_k \tilde{\mathbf{P}}_k + \mathbf{I}_M. \tag{3.45}$$

Note that $\hat{\mathbf{\Psi}}_k$ has the same form as $\mathbf{\Psi}_k$ in (3.13) but with $\mathbf{P}_k$ replaced by $\tilde{\mathbf{P}}_k$. Therefore, matrix $\mathbf{Z}$ in (3.18) can be written as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{V}_1 \bar{\mathbf{z}}_0 \mathbf{V}_1^\dagger & \mathbf{V}_1 \bar{\mathbf{z}}_1 \mathbf{V}_1^\dagger & \dots & \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}} \mathbf{V}_1^\dagger \\ \mathbf{V}_1 \bar{\mathbf{z}}_1^\dagger \mathbf{V}_1^\dagger & \mathbf{V}_1 \bar{\mathbf{z}}_0 \mathbf{V}_1^\dagger & \dots & \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}-1} \\ \vdots & & \ddots & \vdots \\ \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}}^\dagger \mathbf{V}_1^\dagger & \mathbf{V}_1 \bar{\mathbf{z}}_{N_{fb}-1}^\dagger \mathbf{V}_1^\dagger & \dots & \mathbf{V}_1 \bar{\mathbf{z}}_0 \mathbf{V}_1^\dagger \end{bmatrix} = (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1) \bar{\mathbf{Z}} (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1^\dagger), \tag{3.46}$$

where $\bar{\mathbf{Z}}$ has the same form as $\mathbf{Z}$ in (3.18) with $\mathbf{z}_n$ replaced by $\bar{\mathbf{z}}_n = \sum_{k=0}^{N_c-1} \hat{\mathbf{\Psi}}_k^{-1} e^{j\frac{2\pi}{N_c}kn}$. By noting that

$$\mathbf{Z}^{-1} = (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1) \bar{\mathbf{Z}}^{-1} (\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1^\dagger), \tag{3.47}$$

where we have used $(\mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1)^{-1} = \mathbf{I}_{N_{fb}} \otimes \mathbf{V}_1^\dagger$, we obtain from (3.21)

$$\mathbf{U}_{11} = \mathbf{V}_1 \bar{\mathbf{U}}_{11} \mathbf{V}_1^\dagger, \tag{3.48}$$

where $\bar{\mathbf{U}}_{11}$ is the first $M \times M$ submatrix of $\bar{\mathbf{Z}}^{-1}$. Using (3.48) in (3.42), we obtain that $\mathbf{J} \bar{\mathbf{U}}_{11}^{-1/2} \mathbf{V}_1^\dagger = \mathbf{Q}\mathbf{R}$, where $\mathbf{J}$ is an arbitrary unitary matrix. Therefore, we need the following decomposition for our purpose

$$\bar{\mathbf{U}}_{11}^{-1/2} = \tilde{\mathbf{Q}} \mathbf{R} \mathbf{V}_1 \tag{3.49}$$

where $\tilde{\mathbf{Q}} = \mathbf{J}^{\dagger}\mathbf{Q}$ is also a unitary matrix. Such a decomposition is referred to as equal-diagonal QR decomposition (E-QRD) or geometric-mean decomposition (GMD) and its efficient implementation can be found in the literature [107], [108]. Hence, for a given $\bar{\mathbf{U}}_{11}$, we can always find a unitary matrix $\mathbf{V}_1$ which achieves the MSE lower bound in (3.41).

As $\bar{\mathbf{U}}_{11}$ is a function of relay precoding matrix $\mathbf{A}_k$ as well as the remaining part of the source precoding matrix, i.e., $\tilde{\mathbf{P}}_k$, in the following, we need to determine these matrices. By noting that

$$\det(\mathbf{D}) = \det(\mathbf{L}\mathbf{D}\mathbf{L}^{\dagger}) = \det(\mathbf{U}_{11}^{-1}), \tag{3.50}$$

where we have exploited the properties $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$ and $\det(\mathbf{L}) = 1$ [100], we can further use (3.22) to express the objective function for FD-DFE as

$$\text{OBJ} = \det(\mathbf{U}_{11}^{-1}) = \det\left(\mathbf{Z}_{11} - \mathbf{Z}_{12}\mathbf{Z}_{22}^{-1}\mathbf{Z}_{12}^{\dagger}\right). \tag{3.51}$$

Upper Bound on Objective Function

Unfortunately, the expression for OBJ in (3.51) depends on the feedback filter length N_{fb}, cf. (3.46), which is not desirable in practice. Additionally, due to the presence of $\mathbf{Z}_{22}^{-1}$ in (3.51), it is also not straightforward to express the objective function in terms of $\mathbf{A}_k$ and $\tilde{\mathbf{P}}_k$. To avoid these problems, we derive an upper bound for OBJ, which is independent of N_{fb} and directly related to the optimization variables.

Since matrix $\mathbf{Z}$ in (3.18) is a positive semidefinite (PSD) matrix, $\mathbf{Z}^{-1}$, $\mathbf{Z}_{22}$, and $\mathbf{U}_{11}$ are PSD matrices as well. Thus, $\mathbf{U}_{11}^{-1}$ and $\mathbf{Z}_{12}\mathbf{Z}_{22}^{-1}\mathbf{Z}_{12}^{\dagger}$ are also PSD matrices. By exploiting the fact that $\det(\mathbf{A}+\mathbf{B}) \geq \det(\mathbf{A})$ if $\mathbf{A}$ and $\mathbf{B}$ are PSD matrices [100],

we obtain

$$\det(\mathbf{U}_{11}^{-1}) \leq \det\left(\mathbf{Z}_{11}\right), \tag{3.52}$$

where equality holds i.f.f. $N_{fb} = 0$. In other words, for the case of $N_{fb} = 0$, $\det\left(\mathbf{Z}_{11}\right)$ is the exact value of OBJ. In this case, the feedback filter matrix reduces to a lower triangular matrix, which only cancels the inter-stream interference in the current time slot, cf. (3.26). Otherwise, it is an upper bound for OBJ, which can be expressed as

$$\mathrm{OBJ}_{ub} = \det(\mathbf{Z}_{11}) \overset{(a)}{=} \det\left(\sum_{k=0}^{N_c-1} \mathbf{\Psi}_k^{-1}\right) \overset{(b)}{=} \det\left(\mathbf{V}_1 \left[\sum_{k=0}^{N_c-1} \hat{\mathbf{\Psi}}_k^{-1}\right] \mathbf{V}_1^{\dagger}\right) \overset{(c)}{=} \det\left(\sum_{k=0}^{N_c-1} \hat{\mathbf{\Psi}}_k^{-1}\right), \tag{3.53}$$

where (a) is due to the fact that $\mathbf{Z}_{11} = \mathbf{z}_0 = \sum_{k=0}^{N_c-1} \mathbf{\Psi}_k^{-1}$, cf. (3.18) and (3.21), (b) is due to (3.44), and to obtain (c) we have exploited the properties $\det(\mathbf{AB}) = \det(\mathbf{BA})$ and $\mathbf{V}_1^{\dagger}\mathbf{V}_1 = \mathbf{I}_M$. From (3.53) we observe that OBJ_{ub} is independent of $\mathbf{V}_1$.

Structures of Optimal Source and Relay Precoding Matrices

Since we can always choose $\tilde{\mathbf{P}}_k$ such that $\sum_{k=0}^{N_c-1} \hat{\mathbf{\Psi}}_k^{-1}$ is a diagonal matrix, cf. (3.45), the determinant in (3.53) is essentially the product of the diagonal entries of $\sum_{k=0}^{N_c-1} \hat{\mathbf{\Psi}}_k^{-1}$. Consequently, OBJ_{ub} is equivalent to the objective function of the FD-LE receiver under the GMSE criterion. From Theorem 3.1, we obtain the following optimal structures for $\tilde{\mathbf{P}}_k$ and $\mathbf{A}_k$

$$\tilde{\mathbf{P}}_k^{\star} = \bar{\mathbf{V}}_G^{(k)} \mathbf{\Lambda}_P^{(k)}, \quad \mathbf{A}_k^{\star} = \bar{\mathbf{V}}_H^{(k)} \mathbf{\Lambda}_A^{(k)} \bar{\mathbf{U}}_G^{(k)\dagger}, \tag{3.54}$$

and the optimal $\mathbf{P}_k^\star$ is thus given by $\tilde{\mathbf{P}}_k^\star \mathbf{V}_1^\dagger$. The remaining power allocation problem is identical to that of the GMSE criterion for FD-LE, cf. (3.40). It is worth mentioning that for $N_{fb} > 0$, the upper bound OBJ_{ub} constitutes a tight approximation of the objective function OBJ as is illustrated in Section V.

3.4.4 Asymptotically Optimal Power Allocation

From the previous two subsections, it can be concluded that only two different types of power allocation problems have to be solved, namely the problems for the AMSE and GMSE criteria for FD-LE. The solutions to these problems are also applicable for the maxMSE criterion for FD-LE and for all three criteria for FD-DFE. However, since the objective functions for the AMSE and GMSE criteria in (3.40) are not jointly convex w.r.t. the power allocation variables, the global optimal solution is difficult to obtain. Thus, in the following, we adopt a high SNR approximation for Φ_{km} [60], i.e., we assume $\sigma_u^2\sigma_v^2$ is sufficient small such that it can be ignored in the denominator of (3.39), which leads to

$$\Phi_{km} \approx \tilde{\Phi}_{km} = \frac{P_{s,km}P_{r,km}g_{km}^2 h_{km}^2}{\sigma_v^2 P_{r,km}h_{km}^2 + \sigma_u^2 P_{s,km}g_{km}^2} + 1. \tag{3.55}$$

Proposition 3.1. *The optimization problem (3.40) with Φ_{km} approximated by $\tilde{\Phi}_{km}$ is a convex optimization problem.*[8]

Proof. The proof is provided in Appendix C. □

We are now ready to derive an iterative power allocation algorithm. To this end,

[8] In [109], we have solved the power allocation problem without the high-SNR approximation. Although the solution in [109] is also not globally optimal, under simplifying assumptions, e.g., fixed power allocation at the source, the solution is globally optimal. We have compared the bit error rate of this scheme using and not using the high-SNR approximation. We found that the performance difference is negligible even at low SNR. Therefore, we expect the proposed asymptotically optimal power allocation to also work well for low-to-medium SNRs.

we introduce the Lagrangian of the considered power allocation problem

$$\mathcal{L} = f_{\mathrm{X}}(\tilde{\mathbf{\Phi}}) + \lambda[\sum_{k,m} P_{s,km} - P_S] + \mu[\sum_{k,m} P_{r,km} - P_R] - \sum_{k,m}[\beta_{km}P_{s,km} + \gamma_{km}P_{r,km}], \tag{3.56}$$

where λ and μ are the Lagrange multipliers for the sum power constraints for source and relay, respectively, and β_{km} and γ_{km} are the Lagrange multipliers for the individual power constraints for source and relay, respectively. Applying the KKT conditions to (3.56), which are sufficient and necessary conditions for convex optimization problems [101], we obtain

$$\begin{aligned}
\frac{\partial f_{\mathrm{X}}(\tilde{\mathbf{\Phi}})}{\partial \tilde{\Phi}_{km}} \frac{\partial \tilde{\Phi}_{km}}{\partial P_{s,km}} + \lambda - \beta_{km} = 0, \\
\frac{\partial f_{\mathrm{X}}(\tilde{\mathbf{\Phi}})}{\partial \tilde{\Phi}_{km}} \frac{\partial \tilde{\Phi}_{km}}{\partial P_{r,km}} + \lambda - \gamma_{km} = 0, \\
\beta_{km}P_{s,km} = 0, \quad \gamma_{km}P_{r,km} = 0, \\
\lambda \left[\sum_{k,m} P_{s,km} - P_S\right] = 0, \;\; \mu \left[\sum_{k,m} P_{r,km} - P_R\right] = 0,
\end{aligned} \tag{3.57}$$

where

$$\frac{\partial f_{\mathrm{X}}(\tilde{\mathbf{\Phi}})}{\partial \tilde{\Phi}_{km}} = \begin{cases} -\frac{\sigma_s^2}{N_c}\tilde{\Phi}_{km}^{-2}, & \text{X=AMSE} \\ \frac{-\tilde{\Phi}_{km}^{-2}}{\sum_{k=0}^{N_c-1}\tilde{\Phi}_{km}^{-1}}, & \text{X=GMSE} \end{cases}, \tag{3.58}$$

and

$$\frac{\partial \tilde{\Phi}_{km}}{\partial P_{s,km}} = \frac{\sigma_v^2 P_{r,km}^2 h_{km}^4 g_{km}^2}{\left[\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 P_{s,km} g_{km}^2\right]^2}, \;\; \frac{\partial \tilde{\Phi}_{km}}{\partial P_{r,km}} = \frac{\sigma_u^2 P_{s,km}^2 g_{km}^4 h_{km}^2}{\left[\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 P_{s,km} g_{km}^2\right]^2}. \tag{3.59}$$

For given λ and μ, we obtain from (3.57)-(3.59)

$$P_{s,km} = \frac{\sigma_v^2 P_{r,km} h_{km}^2}{g_{km}^2 (P_{r,km} h_{km}^2 + \sigma_u^2)} \left[\sqrt{\frac{g_{km}^2}{\lambda B_m \sigma_u^2 \sigma_v^2}} - 1 \right]^+, \tag{3.60}$$

$$P_{r,km} = \frac{\sigma_u^2 P_{s,km} g_{km}^2}{h_{km}^2 (P_{s,km} g_{km}^2 + \sigma_v^2)} \left[\sqrt{\frac{h_{km}^2}{\mu B_m \sigma_u^2 \sigma_v^2}} - 1 \right]^+, \tag{3.61}$$

where $B_m = N_c/\sigma_s^2$ and $B_m = (\ln 2) \sum_{k=0}^{N_c-1} (\tilde{\Phi}_{km} + 1)^{-1}$ for the AMSE and the GMSE criteria, respectively, and $[x]^+ = \max(0, x)$. The Lagrange multipliers λ and μ, which are chosen to satisfy the sum power constraint for source and relay, respectively, can be found with the following subgradient method [111, 112]

$$\lambda^{[i+1]} = \left[\lambda^{[i]} + \varepsilon_1 \left(\sum_{k=0}^{N_c-1} \sum_{m=1}^{M} P_{s,km} - P_S \right) \right]^+ \tag{3.62}$$

$$\mu^{[i+1]} = \left[\mu^{[i]} + \varepsilon_2 \left(\sum_{k=0}^{N_c-1} \sum_{m=1}^{M} P_{r,km} - P_R \right) \right]^+, \tag{3.63}$$

where i is the iteration index, and $\varepsilon_j, j = 1, 2$, are step sizes. From (3.60) and (3.61), we observe that the optimal $P_{s,km}$ depends on $P_{r,km}$ and vice versa. To tackle this problem, we propose Algorithm 3.1 in Table 3.1 to iteratively find the optimal power allocations. Convergence of this algorithm to the optimal solution is guaranteed because of the convexity of the consider optimization problem. Note that if either $P_{s,km}$ or $P_{r,km}$ is equal to 0, the other variable will also be 0. This result is intuitively pleasing since, if for example the (m, k)th subchannel is shut down in the S-R link, there is no need to waste power on this subchannel in the R-D link. It is also worth noting that for the GMSE criterion, $P_{s,km}$ and $P_{r,km}$ are functions of $\tilde{\Phi}_{km}$, which means the optimal $P_{s,km}$ and $P_{r,km}$ for the kth frequency tone depend on the power allocations in all other frequency tones. Therefore, finding the optimal solution

Table 3.1: Algorithm 3.1 for finding the optimal power allocation. ϵ_1 and ϵ_2 are small constants, e.g. $\epsilon_1 = \epsilon_2 = 10^{-4}$.

1	Initialize $\mu^{[1]}$ and $\lambda^{[1]}$
2	Initialize $P_{s,km}^{[1]}$, $P_{r,km}^{[1]}$. Set $P_{s,km}^{rec} = P_{s,km}^{[1]}$, $P_{r,km}^{rec} = P_{r,km}^{[1]}$, $\forall k, m$.

3	**Repeat**
	Set iteration number to $i = 2$.
	Repeat
	for $m = 1 : M$, $k = 0 : N_c - 1$
	Find $P_{s,km}^{[i]}$ from (3.60) using $P_{r,km}^{rec}$ and $\lambda^{[i-1]}$.
	end for
	Update $\lambda^{[i]}$ using (3.62). $i = i + 1$.
	until $\lvert\lambda^{[i+1]} - \lambda^{[i]}\rvert < \epsilon_1$, **set** $P_{s,km}^{rec} = P_{s,km}^{[i]}$.
	Set iteration number to $l = 2$.
	Repeat
	for $m = 1 : M$, $k = 0 : N_c - 1$
	Find $P_{r,km}^{[l]}$ from (3.61) using $P_{s,km}^{rec}$, and $\mu^{[l-1]}$.
	end for
	Update $\mu^{[l]}$ using (3.63). $l = l + 1$.
	until $\lvert\mu^{[l+1]} - \mu^{[l]}\rvert < \epsilon_2$, **set** $P_{r,km}^{rec} = P_{r,km}^{[l]}$.
	until $P_{r,km}^{rec}$ and $P_{s,km}^{rec}$ converge.

4	$P_{s,km}^{rec}$ and $P_{r,km}^{rec}$, $\forall k, m$, are the optimal solution.

requires a higher complexity for the GMSE criterion than for the AMSE criterion.

3.4.5 Suboptimal Power Allocation Schemes

Since the proposed precoding matrix optimization scheme involves an iterative power allocation algorithm and considerable feedback overhead from the relay and destination to the source[9], it is desirable to investigate suboptimal approaches with lower complexity and reduced feedback overhead. One option is to adopt equal power allocation at the source and to optimize only the power allocation at the relay. We refer to the corresponding scheme as EPA-S. EPA-S eliminates the iterative updating of the source power variables, hence guaranteeing faster convergence of the power allocation algorithm. However, the EPA-S scheme still requires CSI feedback of the S-R channel for computing the unitary part of the source precoding matrix. In order to completely avoid CSI feedback to the source, one can perform precoding at the

[9]The CSI of both the S-R and R-D channels is required at the source for computation of the source power allocation. This requires the feedback of $(L_g N_r N_s + L_h N_d N_r)$ complex numbers.

relay only, which we refer to as relay only precoding (ROP) scheme. For FD-DFE, we also introduce the unitary precoding at source (UPS) scheme, which applies only the unitary precoding matrix $\mathbf{V}_1$ at the source. This is motivated by the result in Section 3.43, where it is shown that this unitary matrix can balance the MSEs of the different spatial streams. Similar to ROP, the UPS scheme has the advantage of a reduced feedback overhead compared to optimal power allocation and the EPA-S scheme as the source only needs to acquire knowledge of the $M \times M$ unitary matrix $\mathbf{V}_1$.

3.5 Simulation Results

In this section, we evaluate the performance of the proposed source and relay precoding schemes using simulations. We assume that each data block contains $N_c = 64$ symbols. The channels are modeled as uncorrelated Rayleigh block fading channels with power delay profile $p[n] = \frac{1}{\sigma_t} \sum_{l=0}^{L_x-1} e^{-n/\sigma_t} \delta[n-l]$ [102], where $L_x \in \{L_g, L_h\} = 16$ and $\sigma_t = 2$, which corresponds to moderately frequency-selective fading. Unless stated otherwise, we set the values of $N_{g,s}$, $N_{g,r}$ and N_{fb} all equal to 15. We assume identical noise variances for both links, i.e., $\sigma_u^2 = \sigma_v^2$, and define the received SNRs at the relay and destination as $SNR_r \triangleq \frac{P_S}{N_s N_c \sigma_u^2}$ and $SNR_d \triangleq \frac{P_R}{N_r N_c \sigma_v^2}$, respectively. The corresponding energy per bit to noise power spectral density ratios are given by $(E_b/N_0)_r = \frac{SNR_r}{N_b}$ and $(E_b/N_0)_d = \frac{SNR_d}{N_b}$, where N_b is the number of bits per symbol. For the bit error rate (BER) simulation results, we set $(E_b/N_0)_d = 16$ dB and examine the performance as a function of $(E_b/N_0)_d$. For the achievable bit rate result in Fig. 3.6, we set $SNR_r = 20$ dB and examine the performance as a function of SNR_d. All simulations are averaged over at least 10,000 independent channel realizations and data blocks. In the following, the proposed joint source and relay precoding design

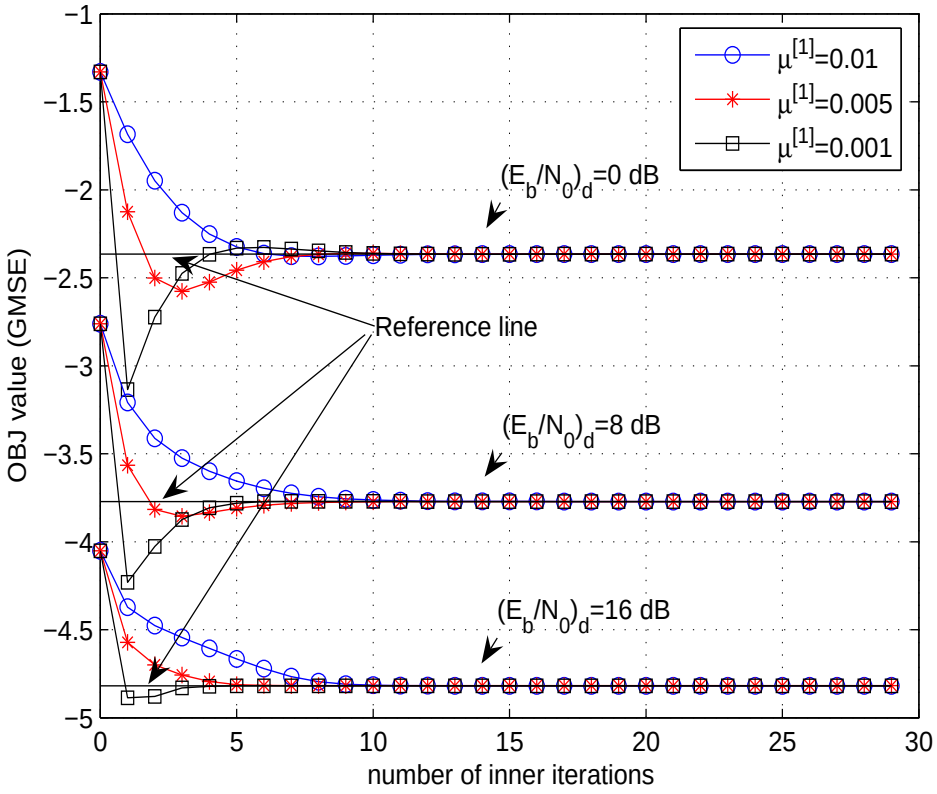

Figure 3.2: Objective function value (GMSE criterion) versus number of inner iterations.

is referred to as JSR, and the notation $\{M, N_s, N_r, N_d\}$ is used to specify a system with the parameters appearing in the brackets.

3.5.1 Convergence of the Algorithm and Tightness of OBJ_{ub} for FD-DFE

We first examine the convergence of the proposed power allocation algorithm in terms of the numbers of inner and outer iterations for a $\{2, 2, 2, 2\}$ MIMO relay system optimized for GMSE criterion[10]. We define an outer iteration as one optimization

[10]Similar results also hold for the AMSE criterion.

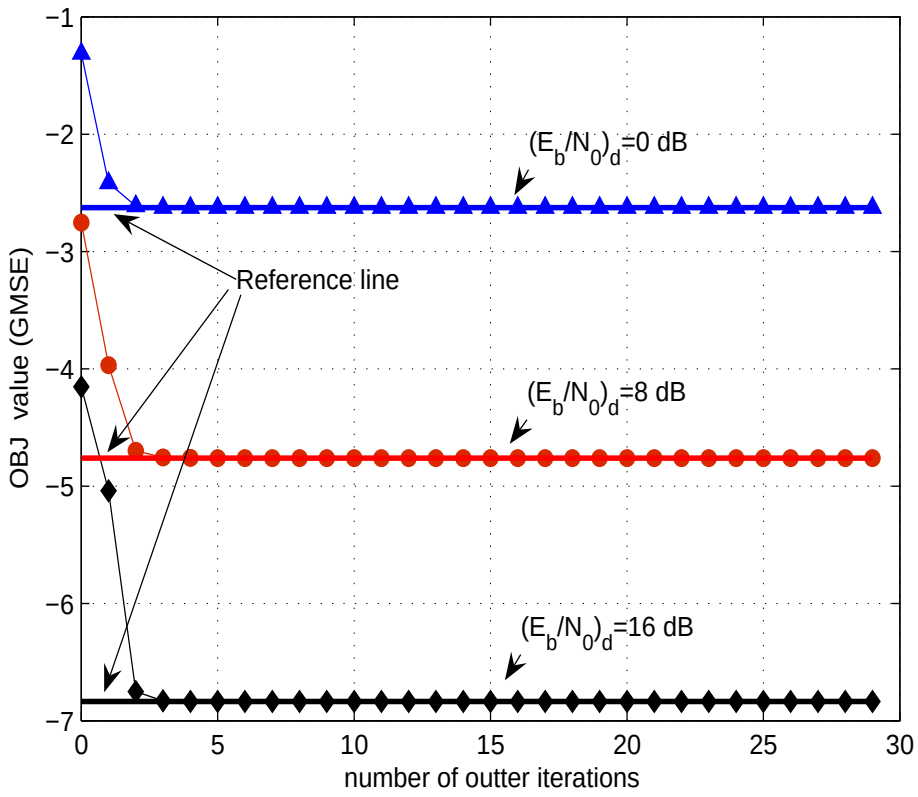

Figure 3.3: Objective function value (GMSE criterion) versus number of outer iterations.

of $\{P_{s,km}\}$ or $\{P_{r,km}\}$ in the algorithm shown in Table I, and the update of $\{P_{s,km}\}$ ($\{P_{r,km}\}$) in each outer iteration as one inner iteration. The reference lines indicate the optimal values of the objective function. In Fig. 3.2, we take the optimization of $\{P_{r,km}\}$ as an example, where we choose $0.01, 0.005$, and 0.001 as three specific initial values for the Lagrange multiplier μ, and randomly initialize $\{P_{s,km}\}$ and $\{P_{r,km}\}$. It is observed that the choices of the initial value of μ affect the convergence of the inner iterations. Nevertheless, for the considered three cases of initialization, the objective function values converge within 10 iterations. In practice, the initial values can be optimized offline for different SNRs. In Fig. 3.3, we investigate the convergence of

the algorithm in terms of the outer iterations, where both $\{P_{s,km}\}$ and $\{P_{r,km}\}$ are randomly initialized. Note that according to the definition of the outer iteration, the initial values of λ and μ do not affect the convergence of the outer iterations. From Fig. 3.3 we observe that it takes at most three outer iterations to obtain the final solutions for $\{P_{s,km}\}$ and $\{P_{r,km}\}$. Moreover, the largest improvement of the objective function value is obtained in the first and second outer iterations when $(E_b/N_0)_d$ is small and large, respectively. This suggests that for low SNR, optimizing the source or the relay power allocation is sufficient to realize most of the achievable performance gain, while for high SNR, a joint optimization of the source and relay power allocations is beneficial.

In Fig. 3.4, we show the values of the objective function, OBJ, for FD-DFE, cf. (3.51), for different values of N_{fb}. Note that OBJ for $N_{fb} = 0$ serves as the upper bound, OBJ_{ub}, for the general objective function. From the figure, we observe that the upper bound is very close to the objective function for all considered values of N_{fb}, especially for medium-to-high SNR. Therefore, OBJ_{ub} constitutes a good approximation for the objective function for the FD-DFE receiver.

3.5.2 Comparison of SC-FDE and OFDM for JSR Precoding

In Fig. 3.5, we show the BER of uncoded quaternary phase-shift keying (QPSK) as a function of $(E_b/N_0)_d$ for the proposed FD-LE based MIMO relay system for the three considered precoding matrix optimization criteria. For FD-DFE, only the GMSE criterion is considered as for FD-DFE all three criteria are equivalent. For comparison, the performance of a MIMO-OFDM relay system optimized under the same criteria is also included [64]. The figure shows that for the $\{2, 2, 2, 2\}$ system, the proposed MIMO relay system with an FD-LE receiver outperforms the corresponding

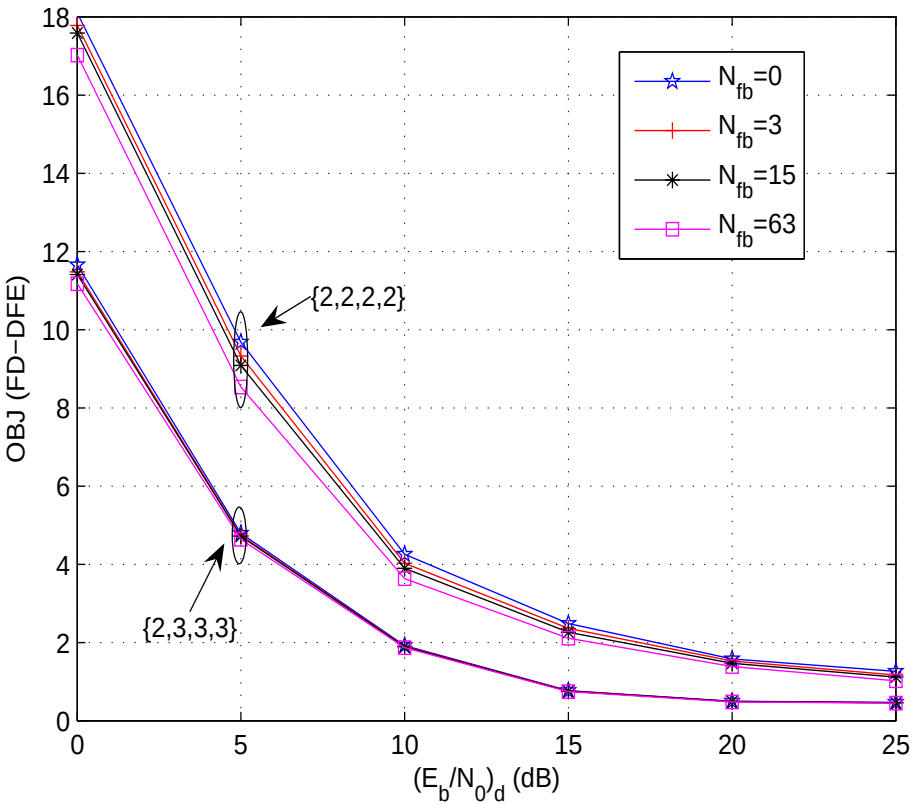

Figure 3.4: Objective function value of the FD-DFE receiver for different values of N_{fb}.

OFDM-based system by a large margin since, in contrast to uncoded OFDM, FD-LE is able to exploit the frequency diversity offered by the channel. In addition, for both FD-LE and OFDM, the system employing the maxMSE criterion offers the best performance since the worst-case MSE is minimized. In this case, FD-LE obtains a better error rate performance than OFDM, while enjoying the advantage of a single-level waterfilling solution. For FD-DFE, the performance improvement compared to FD-LE and OFDM is remarkable and a much higher diversity gain is observed. On the other hand, for the $\{2, 3, 3, 3\}$ system, we observe that the performance gaps between FD-DFE, FD-LE, and OFDM become smaller. Surprisingly, using the maxMSE

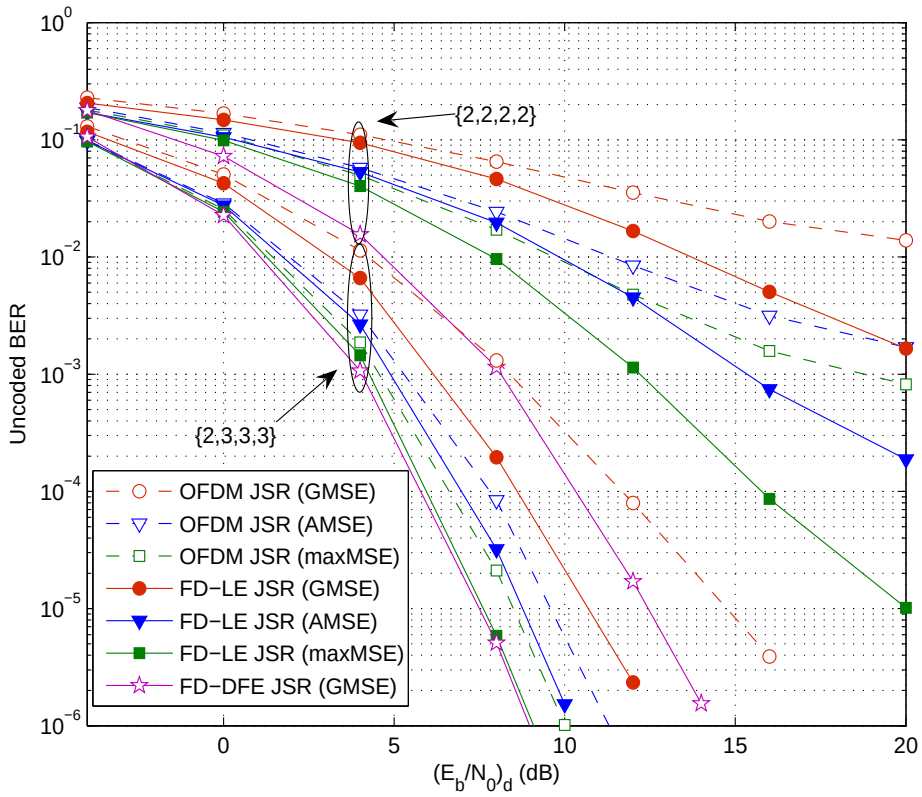

Figure 3.5: Uncoded BER of $\{2, 2, 2, 2\}$ and $\{2, 3, 3, 3\}$ MIMO relay systems for JSR precoding using different optimization criteria.

criterion, the optimized OFDM and FD-LE systems achieve a performance very close to that of FD-DFE. This is due to fact that the additional antennas offer additional spatial diversity which helps OFDM and FD-LE to effectively avoid the deep spectrum nulls that otherwise negatively affect their performance in frequency-selective fading.

In Fig. 3.6, we investigate the achievable bit rates (ABRs)[11, 113] of the OFDM and SC-FDE systems under different optimization criteria. The ABR is calculated

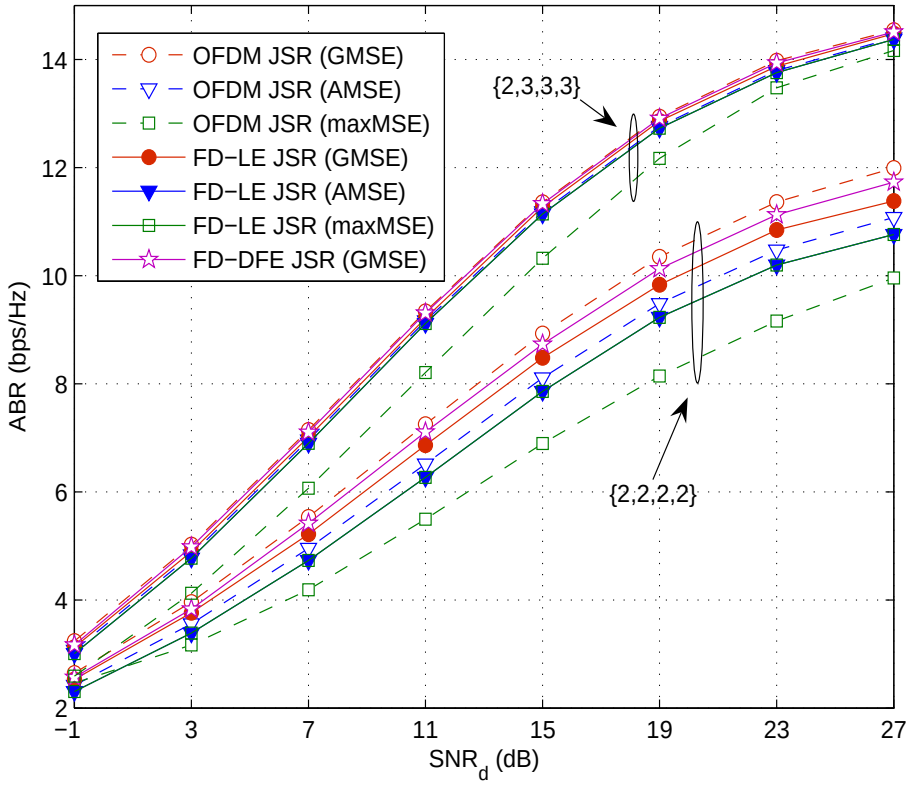

Figure 3.6: ABR of $\{2,2,2,2\}$ and $\{2,3,3,3\}$ MIMO relay systems with JSR precoding for different optimization criteria.

as[11]

$$\text{ABR} = \frac{1}{MN_c} \sum_{m=1}^{M} \sum_{k=0}^{N_c-1} \log_2 \left(\text{SINR}_{km,\text{X}} + 1\right) \tag{3.64}$$

where $\text{SINR}_{km,\text{X}} = \Phi_{km}$ for X=OFDM [64] and $\text{SINR}_{km,\text{X}} = ([\hat{\mathbf{E}}_{\text{X}}]_{mm})^{-1} - 1, \forall k$, for X={FD-LE, FD-DFE}. As expected, the systems optimized under the GMSE criterion have the best performance since minimizing the GMSE is equivalent to maximizing the ABR. In general, the ABR achieved by the considered MIMO-OFDM

[11]Note that for the ABR we implicitly assume Gaussian transmit symbols and ideal channel coding [11].

relay systems is higher than that of the corresponding FD-LE relay systems, except for the case when both systems are optimized based on the maxMSE criterion. Indeed, the OFDM system optimized under the maxMSE criterion suffers from the worst ABR performance among all the considered schemes since the available power is mainly used to improve the MSE of the subcarriers with bad channel conditions instead of taking advantage of the subcarriers with good channel conditions. In addition, Fig. 3.6 shows that for FD-LE, the AMSE and maxMSE criteria lead to exactly the same ABR, which implies that the unitary rotation of the source precoding matrix does not influence the ABR of the system. Furthermore, the ABR achieved with FD-DFE is larger than that achieved with any of the FD-LE schemes and very close to that of OFDM. This is due to the lower stream MSEs of FD-DFE compared to FD-LE, which translates into larger stream SINRs and larger system ABR. For the $\{2, 3, 3, 3\}$ system, we observe that SC-FDE and OFDM achieve almost the same performance for the AMSE and GMSE criteria, implying that with more source/relay/destination antennas, SC-FDE will approach the achievable ABR of the OFDM system.

3.5.3 Performance of Suboptimal Power Allocation Schemes

In Figs. 3.7 and 3.8, we plot the uncoded and coded BERs for the suboptimal power allocation schemes discussed in Section 3.45 using QPSK for a {2,2,2,2} system, respectively. For the coded case, the standard rate-1/2 convolution code with generator matrix $(133, 171)_{oct}$ is adopted. The OFDM and FD-LE systems are both optimized under the maxMSE criterion. The FD-DFE system is optimized under the GMSE criterion since for FD-DFE all three considered criteria are equivalent to the GMSE criterion. From Fig. 3.7 we observe that for uncoded transmission, the FD-LE system outperforms the OFDM system if both employ the same precoding

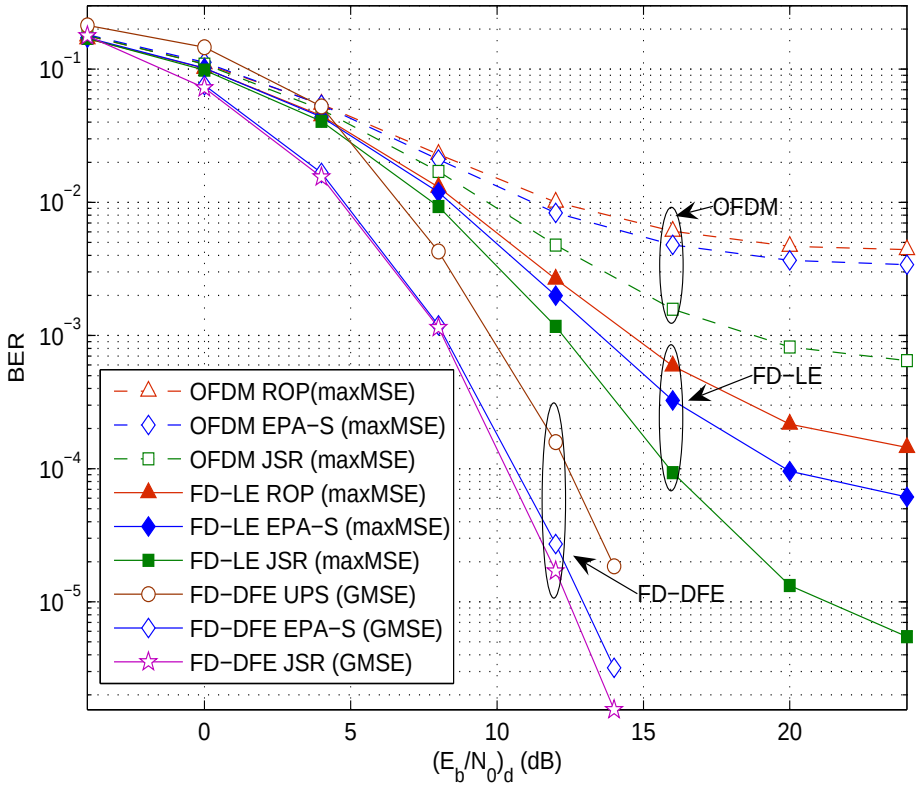

Figure 3.7: Uncoded BER of a $\{2, 2, 2, 2\}$ MIMO relay system with JSR and suboptimal precoding schemes.

technique. Fig. 3.7 also shows that for FD-LE and OFDM, EPA-S and ROP suffer from a considerable performance degradation compared to JSR, while for FD-DFE, the performance loss is relatively small for UPS and almost negligible for EPA-S. For coded systems, the channel coding helps to spread the information bits across different subcarriers, hence OFDM systems can also exploit the frequency diversity of the channel and significantly improve their BER performance, cf. Fig. 3.8. Nevertheless, the coded FD-LE system still outperforms the OFDM system if the same precoding technique is assumed in both cases. Also, Fig. 3.8 reveals that channel coding significantly reduces the performance loss caused by suboptimal precoding techniques for

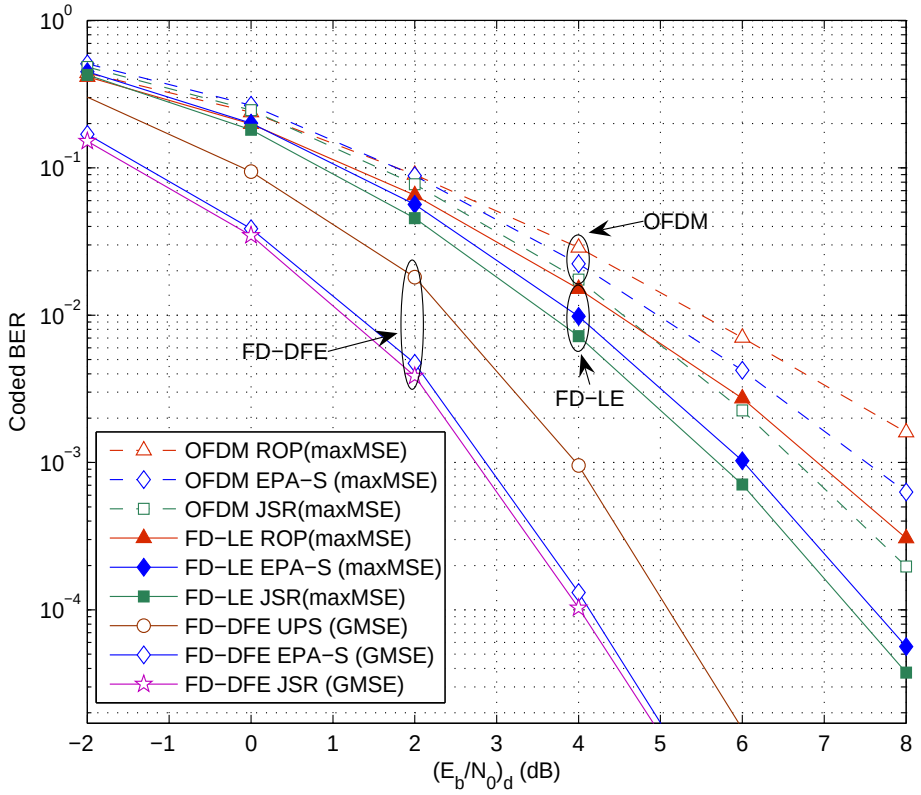

Figure 3.8: Coded BER of a $\{2, 2, 2, 2\}$ MIMO relay system with JSR and suboptimal precoding schemes.

both OFDM and FD-LE.

Since the performance of FD-DFE depends on the number of feedback filter taps, in Fig. 3.9, we investigate the influence of N_{fb} on the performance of a {2,2,2,2} QPSK system. The results show that while the value of N_{fb} has limited impact on the performance of EPA-S, it does play a critical role for UPS. The reason is that for EPA-S, the equivalent *S-R-D* channel is diagonalized into M parallel channels, thus eliminating the inter-stream interference at the receiver. However, for the case of UPS , the equivalent end-to-end channel is not fully diagonalized and the received symbols experience inter-stream interference. Consequently, a feedback filter with

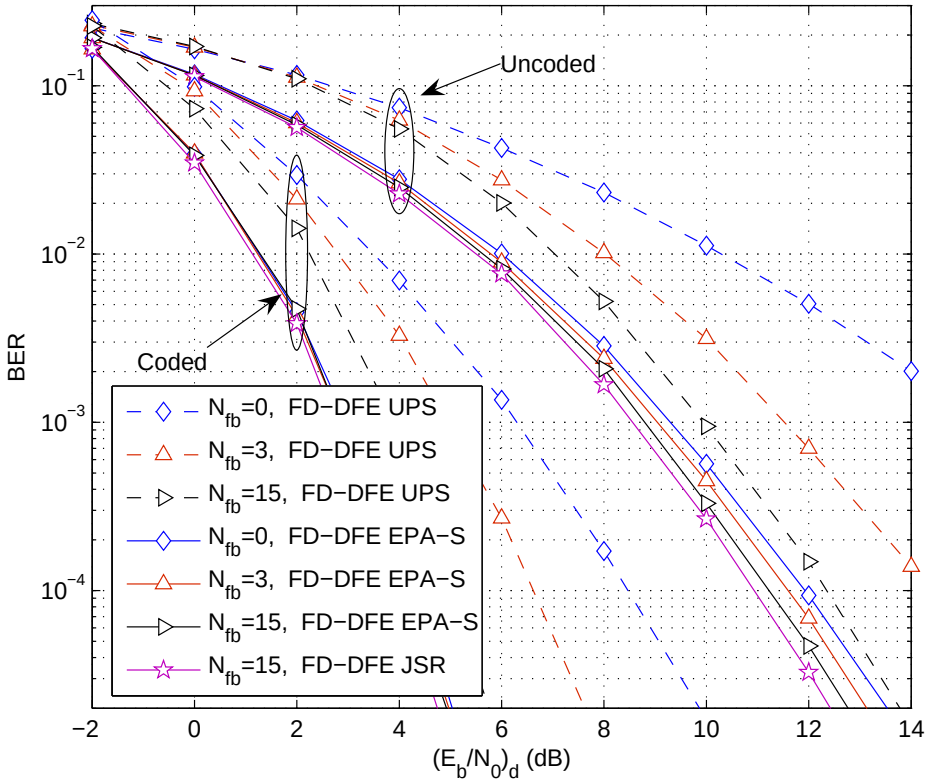

Figure 3.9: Uncoded and coded BER of a $\{2, 2, 2, 2\}$ MIMO relay system with FD-DFE using JSR and suboptimal precoding schemes with different numbers of feedback filter taps.

sufficiently large N_{fb} is required to cancel out this interference. As can be inferred from Fig. 3.9, there is a complexity tradeoff between the transmitter and the receiver for FD-DFE. For EPA-S, since a small number of feedback filter taps (e.g., $N_{fb} = 3$) is sufficient to achieve good performance, the receiver complexity is similar to that of FD-LE. However, comparatively complex FD signal processing has to be carried out at the transmitter. This characteristic makes EPA-S suitable for the downlink transmission. For the UPS scheme, on the other hand, the transmit processing is very simple since the single tap precoding matrix $\mathbf{V}_1$ can be directly implemented in the TD. In addition, the feedback overhead is low as $\mathbf{V}_1$ is identical for all frequency tones.

However, UPS requires a longer and thus more complex feedback filter to achieve a high performance. These characteristics make UPS a very promising scheme for uplink transmission.

3.6 Conclusion

In this chapter, we have tackled the problem of transceiver design for MIMO relay systems employing SC-FDE. The optimal minimum MSE FD-LE and FD-DFE filters at the destination were derived, and the optimal structures of the source and relay precoding matrices were obtained in closed form for a general family of objective functions. For systems employing an FD-DFE receiver, we first showed that the considered objective functions are all equivalent, and we derived an upper bound on the original objective functions, which was shown to be equal to the GMSE objective function for the FD-LE receiver. The remaining power allocation problem was solved globally by using a high SNR approximation of the objective function and efficient convex optimization methods. Our results show that the proposed SC-FDE relaying schemes outperform the corresponding OFDM schemes in terms of both coded and uncoded BER for fixed modulation and coding rates. However, the performance gap between SC-FDE and OFDM relay systems decreases when the number of source/relay/destination antennas is larger than the number of data streams. Assuming Gaussian signalling and ideal channel coding, SC-FDE and OFDM attain similar achievable bit rates. Furthermore, we have shown that the proposed suboptimal power allocation schemes can reduce the system complexity and feedback overhead at the expense of a moderate performance degradation, especially in case of coded transmission, making them promising candidates for practical relay systems.

Chapter 4

Robust Transceiver Design For Broadband Multiuser Multi-Relay Networks

4.1 Introduction

The transceiver designs discussed so far have assumed perfect CSI at relays and the destination. In practical wireless systems, CSI is usually imperfect due to channel estimation errors and/or feedback quantization errors. Two different approaches are commonly used to model imperfect CSI, namely, statistical models and deterministic models. In the former case, the statistics of the CSI errors are assumed to follow some known distribution such as Gaussian, which makes this model suitable for modeling channel estimation errors [47]. In the latter case, the CSI error is assumed to lie in an uncertainty region with known boundary, making this model suitable for the characterization of quantization errors [48]. Start from this chapter, we will focus on robust transceiver designs for cooperative relay communication systems with imperfect CSI.

In this chapter, we consider robust rBF and destination equalization (dEQ) filter design for multiuser SC-FDMA and OFDMA systems with multiple single-antenna AF relays under channel estimation errors. For the CSI errors, we adopt the statis-

tical Bayesian model due to its suitability for characterization of channel estimation errors. Our design goal is to maximize the weighted ABR of the network subject to either Ind-PCs or an Agg-PC[12]. Since in the presence of CSI errors, there is no analytically tractable expression for the network ABR, we first develop a closed-form lower bound on the ABR. Adopting this lower bound as the key component of the objective function, joint optimization problems for the robust rBF and dEQ filters are formulated and solved. Specifically, the optimal dEQ filters and the phases of the optimal rBF filter coefficients are independent of the relay power constraints and can be obtained in closed-form via solving unconstrained optimization problems. The optimal amplitudes of the rBF filter coefficients, on the other hand, depend on the relay power constraints. For the Agg-PC, we show that the optimization problem can be decomposed into two subproblems using primal decomposition. In the first subproblem, a closed-form solution is obtained for fixed, given power allocation variables across subcarriers and users. The second subproblem is shown to be a convex power allocation problem, for which the global optimal solution is obtained via the Karush-Kuhn-Tucker conditions. For Ind-PCs, we transform the optimization problem into a reverse-convex optimization problem with convex constraints, and obtain a local optimal solution using the constrained convex concave procedure.

This chapter is organized as follows. In Section 4.2, a unified system model for OFDMA and SC-FDMA multi-relay networks is introduced. In Section 4.3, the joint rBF and dEQ filter optimization problem is formulated, and the optimal dEQ filter and the phases of the optimal rBF filter coefficients are derived. In Section 4.4 and 4.5, the optimization problems for the amplitudes of the rBF filter coefficients under the Agg-PC and Ind-PCs are solved, respectively. Simulation results are provided in

[12]The Agg-PC is applicable if the relays belong to the network infrastructure and share a common power supply, e.g., the power grid. Ind-PCs, on the other hand, are more appropriate if the relays have their own power supplies, e.g., precharged batteries.

Section 4.6, and some conclusions are drawn in Section 4.7.

4.2 System Model

In this section, we introduce the OFDMA and SC-FDMA signal models as well as the channel estimation error model for the considered multiuser multi-relay systems.

4.2.1 Signal Model

We consider a network consisting of U source nodes, $S_u, u = 1, \ldots, U$, N_R relay nodes, $R_i, i = 1, \ldots, N_R$, and one destination node D, all equipped with a single antenna, as shown in Fig. 4.1[13]. We assume that direct links between the source nodes and the destination do not exist due to large path losses and heavy shadowing[14]. The information symbol vector of S_u is given by $\mathbf{s}^{(u)} = [s_1^{(u)}, \ldots, s_Q^{(u)}]^T$, where Q is the data block size, and symbols $s_n^{(u)}, n = 1, \ldots, Q$, are modeled as i.i.d. random variables with zero mean and unit variance. For SC-FDMA, $\mathbf{s}^{(u)}, \forall u$, is first converted into the FD using a Q-point FFT, mapped to the frequency band allocated to user u, zero-padded to length N_c, and converted back into the TD using an N_c-point IFFT. For OFDMA, $\mathbf{s}^{(u)}, \forall u$, is processed in the same fasion as for SC-FDMA, except that the conversion of the signal into the FD is not required. Hence, the transmit signal for both SC-FDMA and OFDMA can be written as

$$\mathbf{x}^{(u)} = \sqrt{\frac{P_S}{N_c}} \mathbf{F}_{N_c}^{\dagger} \mathbf{M}_f^{(u)} \mathbf{P}_Q \mathbf{s}^{(u)}, \tag{4.1}$$

[13]The insertion and removal of the cyclic prefixes are not shown in Fig. 4.1 for simplicity oft presentation.

[14]If direct links between the source nodes and the destination exist, they provide extra spatial diversity and thus enhance the system performance. In this case, the proposed optimization techniques are still applicable since the direct links only contribute constant gains to the end-to-end SINR but have no impact on the rBF filter coefficients.

where P_S is the transmit power limit for all users, N_c is the total number of subcarriers, $\mathbf{P}_Q = \mathbf{F}_Q$ for SC-FDMA, $\mathbf{P}_Q = \mathbf{I}_Q$ for OFDMA, and $\mathbf{M}_f^{(u)} \in \mathbb{C}^{N_c \times Q}$ is the subcarrier-mapping matrix which depends on the adopted subcarrier mapping scheme, e.g., localized or interleaved subcarrier mapping [114]. Here, we have assumed that all the symbols are transmitted with the same power $\frac{P_S}{N_c}$. In the following, we assume that $P_S = N_c$ to simplify the notation. Before transmission, $\mathbf{x}^{(u)}$ is prepended with a CP of length $N_{g,u}$, $N_{g,u} \geq \max_i\{L_{g,u,i}\}$, where $L_{g,u,i}$ is the length of the CIR between S_u and R_i. The CP converts the linear convolution of the CIR and the transmitted signal into a circular convolution. Thus, the TD S_u-R_i channel matrix, $\mathbf{G}_{t,i}^{(u)}, \forall u, i$, is a circular matrix, which can be decomposed as

$$\mathbf{G}_{t,i}^{(u)} = \mathbf{F}_{N_c}^{\dagger} \mathbf{G}_{f,i}^{(u)} \mathbf{F}_{N_c}, \tag{4.2}$$

where $\mathbf{G}_{f,i}^{(u)} \in \mathbb{C}^{N_c \times N_c}$ is the diagonal FD S_u-R_i channel matrix. The received signal at R_i after CP removal can be written as

$$\mathbf{r}_i \;=\; \sum_{u=1}^{U} \mathbf{G}_{t,i}^{(u)} \mathbf{x}^{(u)} + \mathbf{z}_{r,i}, \tag{4.3}$$

where $\mathbf{z}_{r,i}$ is the TD noise vector at R_i, whose entries are i.i.d. Gaussian random variables with zero mean and variance σ_r^2.

At $R_i, \forall i$, the received signal $\mathbf{r}_i$ is converted into the FD using an N_c-point FFT, demapped for user u using $\mathbf{M}_f^{(u)\dagger}$, processed by an FD rBF matrix $\mathbf{A}_{f,i}^{(u)}$, remapped to the frequency band allocated to user u using $\mathbf{M}_f^{(u)}$, and converted back into the TD using an N_c-point IFFT, resulting in

$$\mathbf{t}_i \;=\; \sum_{u=1}^{U} \left(\mathbf{F}_{N_c}^{\dagger} \mathbf{M}_f^{(u)} \mathbf{A}_{f,i}^{(u)} \mathbf{M}_f^{(u)\dagger} \mathbf{F}_{N_c} \right) \mathbf{r}_i = \sum_{u=1}^{U} \mathbf{F}_{N_c}^{\dagger} \mathbf{M}_f^{(u)} \mathbf{A}_{f,i}^{(u)} \bar{\mathbf{G}}_{f,i}^{(u)} \mathbf{P}_Q \mathbf{s}^{(u)} + \mathbf{z}'_{r,i}, \tag{4.4}$$

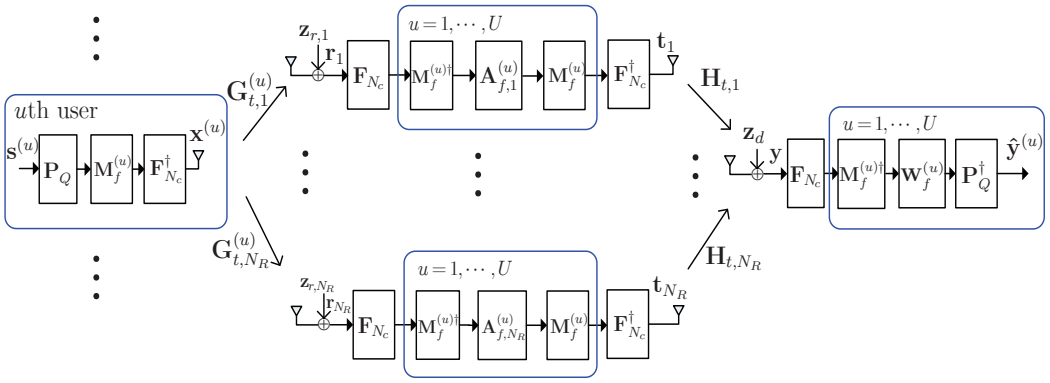

Figure 4.1: System model for OFDMA and SC-FDMA multiuser multi-relay networks.

where $\mathbf{A}_{f,i}^{(u)}$ is a diagonal matrix whose diagonal entries are given by $a_{ik}^{(u)}, k = 1, \ldots, Q$, $\mathbf{z}'_{r,i} = \sum_{u=1}^{U}(\mathbf{F}_{N_c}^\dagger \mathbf{M}_f^{(u)} \mathbf{A}_{f,i}^{(u)} \mathbf{M}_f^{(u)\dagger} \mathbf{F}_{N_c}) \mathbf{z}_{r,i}$ is the amplified TD noise vector at relay R_i, and to obtain (4.4) we have used the following property of the subcarrier mapping/demapping matrices: $\mathbf{M}_f^{(u)\dagger}(\sum_{u'=1}^{U} \mathbf{G}_{f,i}^{(u')} \mathbf{M}_f^{(u')}) = \bar{\mathbf{G}}_{f,i}^{(u)}$, with $\bar{\mathbf{G}}_{f,i}^{(u)} \in \mathbb{C}^{Q \times Q}$ denoting the FD S_u-R_i channel matrix for the non-dummy subcarriers[15]. Before transmission, $\mathbf{t}_i$ is prepended with a CP of length N_h, $N_h \geq \max_i\{L_{h,i}\}$, where $L_{h,i}$ is the length of the CIR between R_i and D. We denote the TD R_i-D channel matrix as $\mathbf{H}_{t,i}$. Similar to $\mathbf{G}_{t,i}^{(u)}$, $\mathbf{H}_{t,i}$ is also a circular matrix, which can be decomposed as

$$\mathbf{H}_{t,i} = \mathbf{F}_{N_c}^\dagger \mathbf{H}_{f,i} \mathbf{F}_{N_c}, \tag{4.5}$$

where $\mathbf{H}_{f,i} \in \mathbb{C}^{N_c \times N_c}$ is the diagonal FD R_i-D channel matrix. The received signal at D after CP removal is given by

$$\mathbf{y} = \sum_{i=1}^{N_R} \mathbf{H}_{t,i} \mathbf{t}_i + \mathbf{z}_d, \tag{4.6}$$

[15]For each user, only Q of the N_c subcarriers carry information. We refer to the information-carrying subcarriers as the non-dummy subcarriers.

where $\mathbf{z}_d$ is the TD noise vector at D, whose entries are i.i.d. Gaussian random variables with zero mean and variance σ_d^2. At D, $\mathbf{y}$ is converted into the FD using an N_c-point FFT, demapped for user u using $\mathbf{M}_f^{(u)\dagger}$, and equalized by a diagonal FD equalization matrix $\mathbf{W}_f^{(u)}$. For SC-FDMA, the signal is further converted back into the TD using an N_c-point IFFT before being fed into the decision device. For OFDMA, such conversion is not required. Therefore, the decision variable vector for both SC-FDMA and OFDMA can be written as

$$\begin{aligned}\hat{\mathbf{y}}^{(u)} =&(\mathbf{P}_Q^\dagger \mathbf{W}_f^{(u)} \mathbf{M}_f^{(u)\dagger} \mathbf{F}_{N_c})\mathbf{y} \\ =&\mathbf{P}_Q^\dagger \mathbf{W}_f^{(u)} \left(\mathbf{H}_{\text{eff}}^{[u]} \mathbf{P}_Q \mathbf{s}^{(u)} + \sum_{i=1}^{N_R} \bar{\mathbf{H}}_{f,i}^{(u)} \mathbf{A}_{f,i}^{(u)} \tilde{\mathbf{z}}_{r,i}^{(u)} + \tilde{\mathbf{z}}_d^{(u)} \right), \end{aligned} \tag{4.7}$$

where $\mathbf{H}_{\text{eff}}^{(u)} = \sum_{i=1}^{N_R} \bar{\mathbf{H}}_{f,i}^{(u)} \mathbf{A}_{f,i}^{(u)} \bar{\mathbf{G}}_{f,i}^{(u)}$ is the equivalent end-to-end FD channel matrix for user u, with $\bar{\mathbf{H}}_{f,i}^{(u)}$ denoting the FD R_i-D channel matrix for the non-dummy subcarriers. Furthermore, $\tilde{\mathbf{z}}_{r,i}^{(u)} = \mathbf{M}_f^{(u)\dagger} \mathbf{F}_{N_c} \mathbf{z}_{r,i}$ and $\tilde{\mathbf{z}}_d^{(u)} = \mathbf{M}_f^{(u)\dagger} \mathbf{F}_{N_c} \mathbf{z}_d$ are the effective FD noise vectors for user u at R_i and D, respectively, and have the same statistical properties as $\mathbf{z}_{r,i}$ and $\mathbf{z}_d$, respectively, since $\mathbf{M}_f^{(u)\dagger} \mathbf{F}_{N_c}$ is a unitary matrix. We note that to arrive at (4.7), the following properties of the subcarrier mapping/demapping matrices were used: $\mathbf{M}_f^{(u)\dagger} \mathbf{H}_{f,i} \mathbf{M}_f^{(u)} = \bar{\mathbf{H}}_{f,i}^{(u)}$ and $\mathbf{M}_f^{(u)\dagger} \mathbf{H}_{f,i} \mathbf{M}_f^{(u')} = \mathbf{0},\ \forall u' \neq u$.

4.2.2 Channel Estimation Error Model

In this work, we assume the system is operating in the TDD mode. The FD CSI matrices of the source-relay and the relay-destination channels, i.e., $\bar{\mathbf{G}}_{f,i}^{(u)}$ and $\bar{\mathbf{H}}_{f,i}^{(u)}$, are estimated at the relays via training symbols sent by the source and the destination, respectively. These estimates are then fed back to the destination using dedicated

zero-delay error-free feedback channels[16].

It is well known that if a linear MSE estimator is used, the CSI errors can be modelled as Gaussian random variables with known statistics. Specifically, the source-relay and relay-destination CSI matrices for the non-dummy subcarriers of user u can be written as

$$\bar{\mathbf{G}}_{f,i}^{(u)} = \hat{\mathbf{G}}_{f,i}^{(u)} + \Delta\mathbf{G}_{f,i}^{(u)}, \quad \bar{\mathbf{H}}_{f,i}^{(u)} = \hat{\mathbf{H}}_{f,i}^{(u)} + \Delta\mathbf{H}_{f,i}^{(u)}, \tag{4.8}$$

where $\hat{\mathbf{G}}_{f,i}^{(u)}$ and $\hat{\mathbf{H}}_{f,i}^{(u)}$ denote the estimated FD channel matrices, whose diagonal entries are given by $\hat{g}_{ik}^{(u)}$ and $\hat{h}_{ik}^{(u)}, \forall u, k, i$, respectively. $\Delta\mathbf{G}_{f,i}^{(u)}$ and $\Delta\mathbf{H}_{f,i}^{(u)}$ denote the CSI estimation error matrices, whose diagonal entries follow complex Gaussian distributions with zero means and variances σ_{e,g_i}^2 and σ_{e,h_i}^2, respectively.

4.2.3 Implementation Issues

In this work, we assume that the destination is the central node which has the estimated global CSI of $\{\hat{\mathbf{G}}_{f,i}^{(u)},\ \hat{\mathbf{H}}_{f,i}^{(u)},\ \forall i, u\}$, $\{\sigma_{e,g_i}^2,\ \sigma_{e,h_i}^2\ \forall i\}$, σ_r^2, and σ_d^2. Therefore, the required feedback overhead from the relays to the destination is $2N_cN_R$ complex numbers and $(2N_R + 1)$ real numbers in total. The equalization filters, $\mathbf{W}_f^{(u)}$, $\forall u$, are computed at the destination using the estimated global CSI. The phases of the rBF filter coefficients are obtained at each relay using the estimated local CSI, i.e., $\{\hat{\mathbf{G}}_{f,i}^{(u)},\ \hat{\mathbf{H}}_{f,i}^{(u)},\ \forall u\}$ for relay i. On the other hand, where the amplitudes of the rBF filter coefficients are computed depends on the adopted power constraint. In particular, for the Agg-PC, the amplitudes of the rBF filter coefficients are calculated at the

[16] Alternatively, the destination can also perform the channel estimation. This causes a mismatch between the CSI at the relays and the CSI at the destination if the estimation is not perfect. However, this is not a problem for transceiver optimization as the destination can design the equalizers based on the estimated aggregated end-to-end channels which include the effect of the rBF filters.

relays using the estimated local CSI and some global system information broadcasted by the destination[17]. For the Ind-PCs, the amplitudes of the rBF filter coefficients are calculated at the destination using the estimated global CSI and are then fed back to the relays, cf. Section 4.5 for details. Hence, for the Agg-PC and Ind-PCs, the feedback overhead from the destination to the relays is $2N_c$ and $N_R N_c$ real numbers, respectively. Therefore, for a small number of relays, e.g., two relays, the system signalling overhead is similar for the Agg-PC and Ind-PCs. However, if the number of relays is large, e.g., eight relays, a system with Ind-PCs requires a considerably higher feedback overhead than a system with the Agg-PC.

4.3 Problem Formulation and Optimal Equalization Filters

In this section, we formulate the joint optimization of the rBF and dEQ filters as a weighted ABR maximization problem with either Ind-PCs or an Agg-PC. To this end, we first need to find an expression for the ABR.

4.3.1 Lower Bound on the ABR

With the decision vector given in (4.7), we can calculate the error vector for user u as

$$\mathbf{e}^{(u)} = \hat{\mathbf{y}}^{(u)} - \mathbf{s}^{(u)} = \left(\mathbf{P}_Q^{\dagger}\mathbf{W}_f^{(u)}\mathbf{H}_{\text{eff}}^{(u)}\mathbf{P}_Q - \mathbf{I}_Q\right)\mathbf{s}^{(u)} + \mathbf{P}_Q^{\dagger}\mathbf{W}_f^{(u)}\left(\sum_{i=1}^{N_R}\bar{\mathbf{H}}_{f,i}^{(u)}\mathbf{A}_{f,i}^{(u)}\tilde{\mathbf{z}}_{r,i}^{(u)} + \tilde{\mathbf{z}}_d^{(u)}\right),$$

[17] The global system information mentioned here refers to the power allocation across users and subcarriers (i.e., $\{P_{ku}, \forall k, u\}$) and the power normalization factors (i.e., $\{c_k^{(u)}, \forall k, u\}$) on different subcarriers, see Section 4.4 for details.

which can be used to evaluate the MSE matrix as a function of the actual channel matrices, $\mathbf{E}^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)}) = \mathbb{E}_{\{\mathbf{s},\tilde{\mathbf{z}}_{r,i}^{(u)},\tilde{\mathbf{z}}_{d}^{(u)}\}}[\mathbf{e}^{(u)}\mathbf{e}^{(u)\dagger}]$, where $\mathbb{E}_{\{\mathbf{s},\tilde{\mathbf{z}}_{r,i}^{(u)},\tilde{\mathbf{z}}_{d}^{(u)}\}}(\cdot)$ is the expectation with respect to (w.r.t.) the statistics of the signals and noises. Since the noise vectors are independent of the transmit signal vectors, we obtain the MSE matrix conditioned on the actual channel matrices as

$$\mathbf{E}^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)}) = \mathbf{P}_Q^{\dagger}\mathbf{E}_f^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)})\mathbf{P}_Q, \tag{4.9}$$

where

$$\begin{aligned}\mathbf{E}_f^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)}) =& \mathbf{W}_f^{(u)}\mathbf{H}_{\text{eff}}^{(u)}\mathbf{H}_{\text{eff}}^{(u)\dagger}\mathbf{W}_f^{(u)\dagger} - \mathbf{W}_f^{(u)}\mathbf{H}_{\text{eff}}^{(u)} - \mathbf{H}_{\text{eff}}^{(u)\dagger}\mathbf{W}_f^{(u)\dagger} + \mathbf{I}_Q \\ &+ \mathbf{W}_f^{(u)}\left(\sigma_r^2\sum_{i=1}^{N_R}\bar{\mathbf{H}}_{f,i}^{(u)}\mathbf{A}_{f,i}^{(u)}\mathbf{A}_{f,i}^{(u)\dagger}\bar{\mathbf{H}}_{f,i}^{(u)\dagger} + \sigma_d^2\mathbf{I}_Q\right)\mathbf{W}_f^{(u)\dagger}.\end{aligned} \tag{4.10}$$

For OFDMA, $\mathbf{P}_Q = \mathbf{I}_Q$ and the MSEs of the symbols are given by the diagonal entries of $\mathbf{E}_f^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)})$, which are denoted as $E_k^{(u)}, \forall k, u$. For SC-FDMA, we have $\mathbf{P}_Q = \mathbf{F}_Q$, which implies that $\mathbf{E}^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)})$ is a circular matrix with identical diagonal entries; hence, all TD symbols for user u in SC-FDMA have identical MSEs given by $\frac{1}{Q}\sum_{k=1}^{Q}E_k^{(u)}$, $\forall u$, which is equal to the arithmetic mean of the subcarrier MSEs for user u in OFDMA. Exploiting the relation that $\text{SINR} = \text{MSE}^{-1} - 1$ [116],

we can obtain the ABR for user u conditioned on the actual channel matrices as

$$\begin{aligned}\mathrm{ABR}_u^{[\mathrm{OFDMA}]}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)}) &= -\sum_{k=1}^{Q} \log_2\left(E_k^{(u)}\right) \\ &= -\log_2\left(\det\left[\mathbf{E}_f^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)})\right]\right), \qquad (4.11)\\ \mathrm{ABR}_u^{[\mathrm{SC-FDMA}]}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)}) &= -Q\log_2\left(\frac{1}{Q}\sum_{k=1}^{Q} E_k^{(u)}\right) \\ &= -Q\log_2\left(\frac{1}{Q}\mathrm{tr}\left[\mathbf{E}_f^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)})\right]\right), \qquad (4.12)\end{aligned}$$

where we have implicitly assumed that the transmit symbols at the sources are Gaussian distributed[18]. By averaging over the random CSI estimation error matrices, the unconditional ABR for user u can be obtained as

$$\mathrm{ABR}_u^{[\mathrm{Y}]} = \mathbb{E}_{\{\Delta\mathbf{G}_{f,i}^{(u)}, \Delta\mathbf{H}_{f,i}^{(u)}\}}\mathrm{ABR}_u^{[\mathrm{Y}]}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)}), \qquad (4.13)$$

where Y={OFDMA, SC-FDMA}. Unfortunately, due to the nonlinearity of the logarithm, it is difficult to find an analytically tractable expression for $\mathrm{ABR}_u^{[\mathrm{Y}]}$. However, as $\log_2\det(\cdot)$ and $\log_2\mathrm{tr}(\cdot)$ are both concave functions, we can infer that $\mathrm{ABR}_u^{[\mathrm{Y}]}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i})$ is a convex function in $\mathbf{E}_f^{(u)}(\bar{\mathbf{G}}_{f,i}^{(u)}, \bar{\mathbf{H}}_{f,i}^{(u)})$. Then, by applying Jensen's

[18] We note that the optimal capacity-achieving input distribution for the multiple access relay channel with imperfect CSI is an open problem in the information theory literature. Therefore, a transceiver design which includes the optimization of the input distribution for such channels is a very challenging task. In this work, we are mostly concerned with the optimization of the rBF and dEQ filters for practical broadband systems, and not with the theoretical limits of such channels. Thus, we adopt the assumption of Gaussian input signals as is commonly done in the existing literature on (relay) transceiver optimization under imperfect CSI, e.g., [47] [50]. In addition, we note that by using the gap approximation for the ABR [115], our problem formulations and solutions are also applicable to practical modulation schemes such as quadrature amplitude modulation (QAM) and pulse amplitude modulation (PAM).

inequality, we can arrive at the following lower bound for the ABR for user u

$$\begin{aligned} \mathrm{ABR}_u^{[\mathrm{OFDMA}]} &\geq -\log_2\left(\det\left[\hat{\mathbf{E}}_f^{(u)}\right]\right) = \underline{\mathrm{ABR}}_u^{[\mathrm{OFDMA}]}, \\ \mathrm{ABR}_u^{[\mathrm{SC-FDMA}]} &\geq -Q\log_2\left(\frac{1}{Q}\mathrm{tr}\left[\hat{\mathbf{E}}_f^{(u)}\right]\right) = \underline{\mathrm{ABR}}_u^{[\mathrm{SC-FDMA}]}, \end{aligned} \tag{4.14}$$

where $\hat{\mathbf{E}}_f^{(u)} = \mathbb{E}_{\{\Delta\mathbf{G}_{f,i}^{(u)}, \Delta\mathbf{H}_{f,i}\}}[\mathbf{E}_f^{(u)}]$ is the unconditional FD MSE matrix. In the following, we adopt this ABR lower bound as the key component of the objective function in order to make the design problem tractable. In Section 4.6, we will investigate the tightness of these lower bounds.

4.3.2 Problem Formulation and Optimal Equalization Filters

Our design goal is to maximize the weighted ABR (lower bound) of the network, subject to either an average Agg-PC, i.e., $\mathbb{E}_{\{\mathbf{z}'_{r,i}, \Delta\mathbf{G}_{f,i}^{[u]}\}}[\sum_i \mathrm{tr}(\mathbf{t}_i\mathbf{t}_i^\dagger)] \leq P_R$, or average Ind-PCs, i.e., $\mathbb{E}_{\{\mathbf{z}'_{r,i}, \Delta\mathbf{G}_{f,i}^{[u]}\}}[\mathrm{tr}(\mathbf{t}_i\mathbf{t}_i^\dagger)] \leq P_R^{(i)}, \forall i$[19]. Mathematically, the optimization

[19] We note that in this chapter, the power constraints are imposed on a per-fading block basis. If long-term power constraints were imposed, i.e., if the power was averaged with respect to several channel fading blocks, the constraint set would be enlarged, leading to an improved performance due to the larger number of degrees of freedom for power allocation. However, finding closed-form expressions for the objective functions associated with the long-term power constraints is difficult and requires statistical information of the estimated channel coefficients in the FD. Therefore, to make the optimization problem tractable, we focus on the short-term power constraint.

problem for the rBF and dEQ filters can be formulated as

$$\begin{aligned}\max_{\{\mathbf{W}_f^{(u)},\mathbf{A}_f^{(u)}\}} \quad & \sum_{u=1}^{U}\tau_u \underline{\mathrm{ABR}}_u^{[\mathrm{Y}]} \\ \text{s.t.} \quad & \sum_{u=1}^{U}\mathrm{tr}\Big[\mathbf{A}_{f,i}^{(u)}(\hat{\mathbf{G}}_{f,i}^{(u)}\hat{\mathbf{G}}_{f,i}^{(u)\dagger}+\sigma_{e,g_i}^2\mathbf{I}_Q+\sigma_r^2\mathbf{I}_Q)\mathbf{A}_{f,i}^{(u)\dagger}\Big]\le P_R^{(i)},\ \forall i,\ \text{for Ind-PCs},\end{aligned} \tag{4.15}$$

$$\sum_{u=1}^{U}\sum_{i=1}^{N_R}\mathrm{tr}\Big[\mathbf{A}_{f,i}^{(u)}(\hat{\mathbf{G}}_{f,i}^{(u)}\hat{\mathbf{G}}_{f,i}^{(u)\dagger}+\sigma_{e,g_i}^2\mathbf{I}_Q+\sigma_r^2\mathbf{I}_Q)\mathbf{A}_{f,i}^{(u)\dagger}\Big]\le P_R,\quad \text{for Agg-PC}, \tag{4.16}$$

where $P_R^{(i)}$ is the transmit power limit for the ith relay, P_R is the total transmit power limit for the relays, and τ_u is a weighting factor for user u which specifies the priority of the user. From (4.15) and (4.16), we observe that the optimization of $\mathbf{W}_f^{(u)}$ is independent of the power constraints. Therefore, we can perform an unconstrained optimization of $\mathbf{W}_f^{(u)}$ for each user. To this end, we need to determine the expectation of the MSE matrix w.r.t. the statistic of the channel estimation errors, cf. (4.14). For the estimation error model given in (4.8), the unconditional MSE matrix, $\hat{\mathbf{E}}_f^{(u)}$, can be derived as

$$\hat{\mathbf{E}}_f^{(u)}=\mathbf{W}_f^{(u)}\left(\hat{\mathbf{H}}_{\text{eff}}^{(u)}\hat{\mathbf{H}}_{\text{eff}}^{(u)\dagger}+\mathbf{K}_{\text{eff}}^{(u)}\right)\mathbf{W}_f^{(u)\dagger}-\mathbf{W}_f^{(u)}\hat{\mathbf{H}}_{\text{eff}}-\hat{\mathbf{H}}_{\text{eff}}^{(u)\dagger}\mathbf{W}_f^{(u)\dagger}+\mathbf{I}_Q, \tag{4.17}$$

where $\hat{\mathbf{H}}_{\text{eff}}^{(u)}=\sum_{i=1}^{N_R}\hat{\mathbf{H}}_{f,i}^{(u)}\mathbf{A}_{f,i}^{(u)}\hat{\mathbf{G}}_{f,i}^{(u)}$, and

$$\mathbf{K}_{\text{eff}}^{(u)}=\sum_{i=1}^{N_R}\mathbf{A}_{f,i}^{(u)}\left[(\sigma_{e,g_i}^2+\sigma_r^2)\hat{\mathbf{H}}_{f,i}^{(u)}\hat{\mathbf{H}}_{f,i}^{(u)\dagger}+\sigma_{e,h_i}^2(\hat{\mathbf{G}}_{f,i}^{(u)}\hat{\mathbf{G}}_{f,i}^{(u)\dagger}+\sigma_{e,g_i}^2\mathbf{I}_Q+\sigma_r^2\mathbf{I}_Q)\right]\mathbf{A}_{f,i}^{(u)\dagger}+\sigma_d^2\mathbf{I}_Q$$

can be interpreted as the covariance matrix of the equivalent end-to-end noise vector.

Since the ABR lower bound for each user is a monotonic function w.r.t. the

term inside the logarithm, the optimal equalization matrix at the destination can be derived by minimizing either the trace (for SC-FDMA) or the determinant (for OFDMA) of $\hat{\mathbf{E}}_f^{(u)}$. After some algebraic manipulations, the optimal equalization filter matrix and the unconditional FD MSE matrix can be found as

$$\mathbf{W}_f^{(u)} = \hat{\mathbf{H}}_{\text{eff}}^{(u)\dagger}\left(\hat{\mathbf{H}}_{\text{eff}}^{(u)}\hat{\mathbf{H}}_{\text{eff}}^{(u)\dagger} + \mathbf{K}_{\text{eff}}^{(u)}\right)^{-1}, \quad \hat{\mathbf{E}}_f^{(u)} = \left(\hat{\mathbf{H}}_{\text{eff}}^{(u)\dagger}[\mathbf{K}_{\text{eff}}^{(u)}]^{-1}\hat{\mathbf{H}}_{\text{eff}}^{(u)} + \mathbf{I}_Q\right)^{-1}. \quad (4.18)$$

Explicitly, the kth diagonal entry of $\mathbf{W}_f^{(u)}$ is given by

$$w_k^{(u)} = \left[\sum_{i=1}^{N_R} \hat{h}_{ik}^{(u)} a_{ik}^{(u)} \hat{g}_{ik}^{(u)}\right]^* \left(\left[\sum_{i=1}^{N_R} \hat{h}_{ik}^{(u)} a_{ik}^{(u)} \hat{g}_{ik}^{(u)}\right]^2 + \sum_{i=1}^{N_R} |a_{ik}^{(u)}|^2 \beta_{ik}^{(u)} + \sigma_d^2\right)^{-1}. \quad (4.19)$$

where $\beta_{ik}^{(u)} = (\sigma_{e,g_i}^2 + \sigma_r^2)|\hat{h}_{ik}^{(u)}|^2 + \sigma_{e,h_i}^2(|\hat{g}_{ik}^{(u)}|^2 + \sigma_{e,g_i}^2 + \sigma_r^2)$. As can be observed from (4.19), the optimal equalization filter is in the form of a Wiener filter which takes the CSI estimation errors via $\beta_{ik}^{(u)}$ into consideration. The unconditional MSEs of user u for OFDMA and SC-FDMA are given by, respectively, the diagonal entries of $\hat{\mathbf{E}}_f^{(u)}$ and the arithmetic mean of these entries, namely

$$[\hat{E}_k^{(u)}]^{[\text{OFDMA}]} = \left(\Phi_k^{(u)} + 1\right)^{-1}, \quad [\hat{E}_k^{(u)}]^{[\text{SC-FDMA}]} = \frac{1}{Q}\sum_{j=1}^{Q}\left(\Phi_j^{(u)} + 1\right)^{-1}, \quad \forall k \quad (4.20)$$

where

$$\Phi_k^{(u)} = \frac{\left|\sum_{i=1}^{N_R} \hat{h}_{ik}^{(u)} a_{ik}^{(u)} \hat{g}_{ik}^{(u)}\right|^2}{\sum_{i=1}^{N_R} \left|a_{ik}^{(u)}\right|^2 \beta_{ik}^{(u)} + \sigma_d^2} \quad (4.21)$$

can be interpreted as the SINR for the information symbol on the kth subcarrier of user u for OFDMA and the corresponding virtual SINR[20] for SC-FDMA. Now, by

[20]Since for SC-FDMA, the information symbols are transmitted in the TD, we refer to the SINR

rewriting the ABR lower bound as a function of $\Phi_k^{(u)}$, we obtain

$$\begin{aligned}\underline{\mathrm{ABR}}_u^{[\mathrm{OFDMA}]}(\Phi_k^{(u)}) &= -\sum_{k=1}^{Q}\log_2\left(\Phi_k^{(u)}+1\right)^{-1},\\ \underline{\mathrm{ABR}}_u^{[\mathrm{SC-FDMA}]}(\Phi_k^{(u)}) &= -Q\log_2\left(\frac{1}{Q}\sum_{k=1}^{Q}(\Phi_k^{(u)}+1)^{-1}\right).\end{aligned} \tag{4.22}$$

Hence, after rewriting the relay power constraints (4.15) and (4.16) in scalar form, we arrive at the following optimization problem for the rBF filters

$$\max_{\{a_{ik}^{(u)}\}} \quad \sum_{u=1}^{U}\tau_u\underline{\mathrm{ABR}}_u^{[\mathrm{Y}]}(\Phi_k^{(u)}) \tag{4.23a}$$

$$\text{s.t.} \quad \sum_{u=1}^{U}\sum_{k=1}^{Q}|a_{ik}^{(u)}|^2\gamma_{ik}^{(u)} \le P_R^{(i)}, \quad \forall i \quad \text{for Ind-PCs}, \tag{4.23b}$$

$$\sum_{i=1}^{N_R}\sum_{u=1}^{U}\sum_{k=1}^{Q}|a_{ik}^{(u)}|^2\gamma_{ik}^{(u)} \le P_R, \quad \text{for Agg-PC}, \tag{4.23c}$$

where $\gamma_{ik}^{(u)} = |\hat{g}_{ik}^{(u)}|^2 + \sigma_{e,g_i}^2 + \sigma_r^2$.

4.3.3 Optimal Phase of the rBF Filters

It is observed from (4.23) that the power constraints are independent of the phases of the rBF filter coefficients. Furthermore, we can rewrite the rBF coefficients as $a_{ik}^{(u)} = \left|a_{ik}^{(u)}\right| e^{j\angle\arg(a_{ik}^{(u)})}$, where $\left|a_{ik}^{(u)}\right|$ and $\arg(a_{ik}^{(u)})$ are the amplitude and the phase of $a_{ik}^{(u)}$, respectively. Since the power constraints are not affected by $\arg(a_{ik}^{(u)})$ and $\underline{\mathrm{ABR}}_u^{[Y]}$ is monotonically increasing in $\Phi_k^{(u)}$, we can find the optimal phases on a per-subcarrier basis by maximizing $\Phi_k^{(u)}$. From (4.21), we observe that $\arg(a_{ik}^{(u)})$ appears only in the numerator of $\Phi_k^{(u)}$. Thus, by using the inequality $\left|\sum_{i=1}^{N_R}\hat{h}_{ik}^{(u)}a_{ik}^{(u)}\hat{g}_{ik}^{(u)}\right|^2 \le$

for the signal representations in the FD as "virtual" SINR.

$\left(\sum_{i=1}^{N_R}\left|\hat{g}_{ik}^{(u)}\hat{h}_{ik}^{(u)}\right|\left|a_{ik}^{(u)}\right|\right)^2$, where equality holds when $a_{ik}^{(u)} = |a_{ik}^{(u)}|\frac{\hat{g}_{ik}^{(u)*}\hat{h}_{ik}^{(u)*}}{|\hat{g}_{ik}^{(u)}||\hat{h}_{ik}^{(u)}|}$, it follows that $\Phi_k^{(u)}$ is maximized when $\arg(a_{ik}^{(u)}) = -\arg(\hat{g}_{ik}^{(u)}\hat{h}_{ik}^{(u)})$, i.e., when the phases of the rBF filter coefficients align with those of the corresponding estimated end-to-end FD channel coefficients. Equipped with this result, we can rewrite the maximized $\Phi_k^{(u)}$, denoted as $\tilde{\Phi}_k^{(u)}$, as a function of $|a_{ik}^{(u)}|$ as

$$\tilde{\Phi}_k^{(u)} = \frac{\left(\sum_{i=1}^{N_R}\left|\hat{g}_{ik}^{(u)}\hat{h}_{ik}^{(u)}\right|\left|a_{ik}^{(u)}\right|\right)^2}{\sum_{i=1}^{N_R}\left|a_{ik}^{(u)}\right|^2\beta_{ik}^{(u)} + \sigma_d^2}. \tag{4.24}$$

Replacing $\Phi_k^{(u)}$ by $\tilde{\Phi}_k^{(u)}$ in (4.23), we can formulate the optimization problem for the amplitudes of the rBF filter coefficients as

$$\max_{\{|a_{ik}^{(u)}|\}} \quad \sum_{u=1}^{U}\tau_u\underline{\mathrm{ABR}}_u^{[\mathrm{Y}]}(\tilde{\Phi}_k^{(u)}) \quad \text{s.t.} \quad \text{(4.23b), (4.23c).} \tag{4.25}$$

Since the the objective function in (4.25) is not convex in the optimization variables $\{|a_{ik}^{(u)}|\}$, solving the problem for both constraints directly is challenging, even though the constraints can be shown to be convex in $\{|a_{ik}^{(u)}|\}$.

4.4 Optimal rBF Filters For The Agg-PC

In this section, we shall first tackle the problem in (4.25) for the Agg-PC. We shall derive the optimal structure of the rBF filter coefficient amplitudes and the optimal power allocation among subcarriers and users, respectively.

4.4.1 Structure of the Optimal rBF Filter Coefficient Amplitudes

In the following, we introduce additional real-valued optimization variables, $\{P_{ku}\}$, which can be regarded as the power limit of the kth subcarrier for user u. This allows us to rewrite problem (4.25) as

$$\max_{\{|a_{ik}^{(u)}|,P_{ku}\}} \quad \sum_{u=1}^{U} \tau_u \underline{\mathrm{ABR}}_u^{[\mathrm{Y}]}(\tilde{\Phi}_k^{(u)}) \quad \text{s.t.} \quad \sum_{i=1}^{N_R} |a_{ik}^{(u)}|^2 \gamma_{ik}^{(u)} \leq P_{ku}, \forall k, u$$
$$\sum_{u=1}^{U}\sum_{k=1}^{Q} P_{ku} \leq P_R, \quad P_{ku} \geq 0, \forall k, u. \quad (4.26)$$

We observe that the constraints in (4.26) are decoupled in $\{|a_{ik}^{(u)}|\}$ and $\{P_{ku}\}$. This motivates us to adopt the primal decomposition technique [121] to solve this problem by decomposing it into two problems, i.e., a subproblem

$$\max_{\{|a_{ik}^{(u)}|\}} \quad \sum_{u=1}^{U} \tau_u \underline{\mathrm{ABR}}_u^{[\mathrm{Y}]}(\tilde{\Phi}_k^{(u)}) \quad \text{s.t.} \quad \sum_{i=1}^{N_R} |a_{ik}^{(u)}|^2 \gamma_{ik}^{(u)} \leq P_{ku}, \forall k, u, \quad (4.27)$$

and an upper-level master problem

$$\max_{\{P_{ku}\}} \quad \sum_{u=1}^{U} \tau_u \underline{\mathrm{ABR}}_u^{[\mathrm{Y}]}(\tilde{\Phi}_k^{(u)}) \quad \text{s.t.} \quad \sum_{u=1}^{U}\sum_{k=1}^{Q} P_{ku} \leq P_R, \quad P_{ku} \geq 0, \forall k, u. \quad (4.28)$$

The physical interpretation of (4.27) and (4.28) is as follows: the master problem (4.28) is responsible for finding the optimal distribution of the system transmit power among the users and the subcarriers. Once these power variables are optimized, the subproblem in (4.27) will further determine the optimal amplitudes of the rBF filter coefficients. For subproblem (4.27) with fixed P_{ku}, the optimization of $\{|a_{ik}^{(u)}|\}, \forall k, u$ is decoupled among users and subcarriers. Thus, the problem can be further decom-

posed into QU subproblems

$$\max_{\{a_{ik}^{(u)}\}} \quad \tilde{\Phi}_k^{(u)} \quad \text{s.t.} \quad \sum_{i=1}^{N_R} |a_{ik}^{(u)}|^2 \gamma_{ik}^{(u)} = P_{ku}, \forall k, u, \tag{4.29}$$

where we have exploited the fact that at optimality, the constraint in (4.27) is met with equality. Problem (4.29) shows that, for a given relay power allocation $\{P_{ku}\}$, each user has to maximize its own SINR/virtual SINR at each subcarrier in order to maximize the weighted ABR of the whole network.

Using (4.21), the problem in (4.29) can be written compactly as

$$\max_{\{\mathbf{a}_k^{(u)}\}} \quad \frac{\mathbf{a}_k^{(u)} \mathbf{Z}_k^{(u)} \mathbf{a}_k^{(u)T}}{\mathbf{a}_k^{(u)} \mathbf{D}_k^{(u)} \mathbf{a}_k^{(u)T} + \sigma_d^2}, \quad \text{s.t.} \quad \mathbf{a}_k^{(u)} \mathbf{J}_k^{(u)} \mathbf{a}_k^{(u)T} = P_{ku}, \forall k, u, \tag{4.30}$$

where we have defined $\mathbf{a}_k^{(u)} = [|a_{1k}^{(u)}|, ..., |a_{N_R k}^{(u)}|]$, $\mathbf{D}_k^{(u)} = \text{diag}\{\beta_{ik}^{(u)}, \ \forall i\}$, $\mathbf{Z}_k^{(u)} = \mathbf{z}_k^{(u)} \mathbf{z}_k^{(u)T}$, $\mathbf{z}_k^{(u)} = [|\hat{h}_{1k}^{(u)} \hat{g}_{1k}^{(u)}|, \ |\hat{h}_{2k}^{(u)} \hat{g}_{2k}^{(u)}|, ..., \ |\hat{h}_{N_R k}^{(u)} \hat{g}_{N_R k}^{(u)}|]^T$, $\mathbf{J}_k^{(u)} = \text{diag}\{\gamma_{ik}^{(u)}, \ \forall i\}$, and $\text{diag}\{\cdot\}$ is a diagonal matrix whose diagonal entries are given in the bracket. Next, we introduce a unit norm vector, $\tilde{\mathbf{a}}_k^{(u)}$, to rewrite $\mathbf{a}_k^{(u)}$ as

$$\mathbf{a}_k^{(u)} = \sqrt{P_{ku}} \tilde{\mathbf{a}}_k^{(u)} \mathbf{J}_k^{(u)-1/2}. \tag{4.31}$$

With (4.31) we can recast problem (4.30) as

$$\max_{\{\tilde{\mathbf{a}}_k^{(u)}\}} \quad \frac{P_{ku} \tilde{\mathbf{a}}_k^{(u)} \tilde{\mathbf{Z}}_k^{(u)} \tilde{\mathbf{a}}_k^{(u)T}}{P_{ku} \tilde{\mathbf{a}}_k^{(u)} \tilde{\mathbf{D}}_k^{(u)} \tilde{\mathbf{a}}_k^{(u)T} + \sigma_d^2} \quad \text{s.t.} \quad \tilde{\mathbf{a}}_k^{(u)} \tilde{\mathbf{a}}_k^{(u)T} = 1, \forall k, u, \tag{4.32}$$

where $\tilde{\mathbf{Z}}_k^{(u)} = \mathbf{J}_k^{(u)-1/2} \mathbf{Z}_k^{(u)} \mathbf{J}_k^{(u)-T/2}$, $\tilde{\mathbf{D}}_k^{(u)} = \mathbf{J}_k^{(u)-1/2} \mathbf{D}_k^{(u)} \mathbf{J}_k^{(u)-T/2}$. Note that (4.32) is a generalized Rayleigh quotient maximization problem [100]. The optimal solution of

$\tilde{\mathbf{a}}_k^{(u)}$ is thus given by

$$\tilde{\mathbf{a}}_k^{(u)} = c_k^{(u)} \sqrt{P_{ku}} \tilde{\mathbf{z}}_k^{(u)T} (P_{ku} \tilde{\mathbf{D}}_k^{(u)} + \sigma_d^2 \mathbf{I}_{N_R})^{-1}, \tag{4.33}$$

where $\tilde{\mathbf{z}}_k^{(u)} = \tilde{\mathbf{J}}_k^{(u)-1/2} \mathbf{z}_k^{(u)}$, and $c_k^{(u)}$ is a normalization factor that guarantees the unit norm of $\tilde{\mathbf{a}}_k^{(u)}$. The optimal form of $\mathbf{a}_k^{(u)}$ can be obtained by inserting (4.33) into (4.31). Explicitly, the expressions for $|a_{ik}^{(u)}|$ and $c_k^{(u)}$ are given by

$$|a_{ik}^{(u)}| = \frac{c_k^{(u)} \sqrt{P_{ku}} |\hat{h}_{ik}^{(u)} \hat{g}_{ik}^{(u)}|}{\beta_{ik}^{(u)} P_{ku} + \sigma_d^2 \gamma_{ik}^{(u)}} \quad \text{and} \quad c_k^{(u)} = \left(\sum_{i=1}^{N_R} \frac{|\hat{h}_{ik}^{(u)} \hat{g}_{ik}^{(u)}|^2 \gamma_{ik}^{(u)}}{\left[\beta_{ik}^{(u)} P_{ku} + \sigma_d^2 \gamma_{ik}^{(u)} \right]^2} \right)^{-1/2}. \tag{4.34}$$

Remark 4.1: The expression of $|a_{ik}^{(u)}|$ in (4.34) specifies the portion of the available power P_{ku} that the rBF filter should distribute across different relays. Since the noise and estimation error variances appear in the denominator of (4.34), less power is allocated to the relays with larger noise and CSI estimation error variances. In this way, the rBF filter structure in (4.34) effectively reduces the negative effects of noise and CSI estimation error amplification.

Using the optimal structure of the rBF filter in (4.34), the maximum value of $\tilde{\Phi}_k^{(u)}$, which is denoted by $\bar{\Phi}_k^{(u)}$, can be derived as

$$\bar{\Phi}_k^{(u)} = P_{ku} \mathrm{tr} \left[\left(P_{ku} \tilde{\mathbf{D}}_k^{(u)} + \sigma_d^2 \mathbf{I}_{N_R} \right)^{-1} \tilde{\mathbf{Z}}_k^{(u)} \right] = \sum_{i=1}^{N_R} \frac{P_{ku} C_{ik}^{(u)}}{P_{ku} A_{ik}^{(u)} + B_{ik}^{(u)}}, \tag{4.35}$$

where $A_{ik}^{(u)} = \beta_{ik}^{(u)}, B_{ik}^{(u)} = \sigma_d^2 \gamma_{ik}^{(u)}$, and $C_{ik}^{(u)} = |\hat{h}_{ik}^{(u)} \hat{g}_{ik}^{(u)}|^2$.

Remark 4.2: It is interesting to investigate how the CSI error variances influence the value of $\bar{\Phi}_k^{(u)}$. As P_{ku} remains an unknown factor, we assume for the moment equal power allocation at the relays, i.e., $P_{ku} = \frac{P_R}{N_c} = 1, \forall k, u$, where we have assumed

$P_R = N_c$ (Recall from Section II that we have also assumed $P_S = N_c$). Then, the denominator of (4.35) is simplified to $(|\hat{g}_{ik}^{(u)}|^2+\sigma_r^2)\sigma_{e,h_i}^2+(|\hat{h}_{ik}^{(u)}|^2+\sigma_d^2)\sigma_{e,g_i}^2+\sigma_{e,h_i}^2\sigma_{e,g_i}^2+\sigma_d^2|\hat{g}_{ik}^{(u)}|^2+\sigma_r^2|\hat{h}_{ik}^{(u)}|^2+\sigma_r^2\sigma_d^2$, from which we observe that when $(|\hat{g}_{ik}^{(u)}|^2+\sigma_r^2) > (|\hat{h}_{ik}^{(u)}|^2+\sigma_d^2)$, σ_{e,h_i}^2 will result in a larger SINR degradation (and thus a larger ABR loss) than σ_{e,g_i}^2. Otherwise, σ_{e,g_i}^2 will have a more negative effect on the ABR performance.

Now, we have solved the subproblem in (4.27) by finding the optimal $|a_{ik}^{(u)}|$. The remaining task is to solve the optimal power allocation problem given in (4.28).

4.4.2 Optimal Power Allocation for the Agg-PC

With $\bar{\Phi}_k^{(u)}$ given in (4.35), (4.28) can be restated as

$$\max_{\{P_{ku}\}} \quad f_1^{[\mathrm{Y}]}(\mathbf{p}), \quad \text{s.t.} \quad \sum_{u,k} P_{ku} \le P_R, P_{ku} \ge 0, \forall k, u, \tag{4.36}$$

with objective function

$$f_1^{[\mathrm{Y}]}(\mathbf{p}) = \begin{cases} \sum_{u=1}^{U} \tau_u \sum_{k=1}^{Q} \log_2\left(\bar{\Phi}_k^{(u)}+1\right), & \text{Y=OFDMA} \\ -\sum_{u=1}^{U} \tau_u Q \log_2\left(\frac{1}{Q}\sum_{k=1}^{Q}\left(\bar{\Phi}_k^{(u)}+1\right)^{-1}\right), & \text{Y=SC-FDMA} \end{cases} \tag{4.37}$$

where $\mathbf{p} = [\mathbf{p}_1^T, \ldots, \mathbf{p}_U^T]^T$ with $\mathbf{p}_u = [P_{1u}, \ldots, P_{Qu}]^T$.

Proposition 4.1. *Optimization problem (4.36) is a convex problem w.r.t. the optimization variables $P_{ku}, \forall k, u$.*

Since (4.36) is a convex problem, and the constraints are affine and feasible, Slater's condition is satisfied, which implies that strong duality holds for the primal and the dual problems [101]. This allows us to solve the original problem by solving

the dual problem. Let us write the Lagrangian of (4.36) as

$$\mathcal{L}^{[\mathrm{Y}]} = f_1^{[\mathrm{Y}]}(\mathbf{p}) + \lambda \left[P_R - \sum_{u=1}^{U} \sum_{k=1}^{Q} P_{ku} \right] + \sum_{u=1}^{U} \sum_{k=1}^{Q} \zeta_{ku} P_{ku}, \tag{4.38}$$

where λ is the Lagrange multiplier associated with the average aggregate relay power constraint, and $\{\zeta_{ku}, \forall k, u\}$ are the Lagrange multipliers associated with the non-negative power constraints on each subcarrier. Using (4.38), the dual problem of (4.36) can be written as

$$\min_{\lambda} \max_{\{P_{ku}, \zeta_{ku}\}} \mathcal{L}^{[\mathrm{Y}]}. \tag{4.39}$$

Applying the KKT conditions to the inner maximization problem in (4.39), we obtain the following system of equations for the optimal $\{P_{ku}\}$ for a given λ,

$$\Xi^{[\mathrm{Y}]} \frac{\partial \bar{\Phi}_k^{(u)}}{\partial P_{ku}} - \lambda + \zeta_{ku} = 0, \tag{4.40}$$

$$\zeta_{ku} P_{ku} = 0, \tag{4.41}$$

$$\zeta_{ku} \geq 0, \quad P_{ku} \geq 0 \quad \forall k, u, \tag{4.42}$$

where

$$\Xi^{[\mathrm{Y}]} = \begin{cases} -\tau_u (\bar{\Phi}_k^{(u)} + 1)^{-1} / \ln 2, & \mathrm{Y{=}OFDMA}, \\ \frac{-\tau_u (\bar{\Phi}_k^{(u)} + 1)^{-2}}{(\ln 2) \sum_{j=1}^{Q} (\bar{\Phi}_j^{(u)} + 1)^{-1}}, & \mathrm{Y{=}SC\text{-}FDMA}, \end{cases} \quad \text{and}$$

$$\frac{\partial \bar{\Phi}_k^{(u)}}{\partial P_{ku}} = \sum_{i=1}^{N_R} \frac{B_{ik}^{(u)} C_{ik}^{(u)}}{(P_{ku} A_{ik}^{(u)} + B_{ik}^{(u)})^2}.$$

Eq. (4.40) reveals that the derivative of the Lagrangian w.r.t. P_{ku} should be equal to zero at optimality. Eq. (4.41) is the complementary slackness condition for the

constraint $P_{ku} \geq 0$. Eq. (4.42) indicates the non-negativity of the primal and dual variables. From (4.41), we note that if $\zeta_{ku}^{\star} > 0$, then we must have $P_{ku}^{\star} = 0$. Otherwise, $P_{ku}^{\star}$ admits a strictly positive solution which can be obtained by solving (4.40) with $\zeta_{ku}^{\star} = 0$. Note that in both cases, there is no need to explicitly determine $\{\zeta_{ku}^{\star}\}$. In practice, the bisection method can be used to find the root of (4.40), i.e., we increase the value of P_{ku} if $\Xi^{[\mathrm{Y}]}\frac{\partial \bar{\Phi}_k^{(u)}}{\partial P_{ku}} - \lambda > 0$ and decrease the value of P_{ku} if $\Xi^{[\mathrm{Y}]}\frac{\partial \bar{\Phi}_k^{(u)}}{\partial P_{ku}} - \lambda < 0$, the procedure continues until $\left|\Xi^{[\mathrm{Y}]}\frac{\partial \bar{\Phi}_k^{(u)}}{\partial P_{ku}} - \lambda\right| \leq \epsilon_0$, where ϵ_0 is a tolerance factor that controls the accuracy of the bisection method. Based on the above discussion, we can iteratively find the optimal $\{P_{ku}, \forall k, u\}$ for a given λ. For the special case of $N_R = 1$, we can obtain the closed-from solution for OFDMA as

$$P_{ku} = \left[\frac{-B_{1k}^{(u)}(2A_{1k}^{(u)} + C_{1k}^{(u)}) + \sqrt{\Gamma_k^{(u)}}}{2A_{1k}^{(u)}(A_{1k}^{(u)} + C_{1k}^{(u)})}\right]^{+}, \tag{4.43}$$

where $\Gamma_k^{(u)} = [B_{1k}^{(u)}]^2(2A_{1k}^{(u)} + C_{1k}^{(u)}) - 4A_{1k}^{(u)}(A_{1k}^{(u)} + C_{1k}^{(u)})\left([B_{1k}^{(u)}]^2 - \frac{\tau_u B_{1k}^{(u)} C_{1k}^{(u)}}{\lambda \ln 2}\right)$. For SC-FDMA, the solution for $N_R = 1$ can be derived as

$$P_{ku} = \frac{1}{A_{1k}^{(u)} + C_{1k}^{(u)}}\left[\sqrt{\frac{\tau_u B_{1k}^{(u)} C_{1k}^{(u)}}{([\lambda \ln 2]\sum_{j=1}^{Q}(\bar{\Phi}_j^{(u)} + 1)^{-1}}} - B_{1k}^{(u)}\right]^{+}. \tag{4.44}$$

Remark 4.3: From (4.43) and (4.44), one can observe that P_{ku} has a water-filling structure which is an increasing function of τ_u. Therefore, more power will be allocated to the subcarriers of users with larger weighting factor. This is intuitive because with larger τ_u, the contribution of user u to the value of the objective function increases, and thus the user is given more power for all of its subcarriers. In addition, compared to the bisection method, the closed-form solutions in (4.43) and (4.44) provide a direct and fast way for computing the power allocation variables for a

Table 4.1: Algorithm 4.1 for finding the optimal power allocation across users and subcarriers of the rBF filters under the Agg-PC. ϵ is a small constant to control the accuracy of the subgradient method, e.g., $\epsilon = 10^{-4}$. L is the maximum iteration number, e.g., $L = 50$.

1	Initialize $\lambda^{[1]}$ and $P_{ku}^{[1]}, \forall k, u$. Set $P_{ku}^{rec} = P_{ku}^{[1]}, \forall k, u$, $\ell = 1$.

2	Repeat
	for $u = 1 : U$, $k = 1 : Q$
	Find $P_{ku}^{[\ell+1]}$ from (4.40) using the bisection method or from the closed-form solutions (4.43), (4.44) using P_{ku}^{rec} and $\lambda^{[\ell]}$.
	end for
	Set $P_{ku}^{rec} = P_{ku}^{[\ell+1]}, \forall k, u$. Update $\lambda^{[\ell+1]}$ using (4.45).
	Stop when $\lvert\lambda^{[\ell+1]} - \lambda^{[\ell]}\rvert < \epsilon$, or $\ell > L$.
	$\ell = \ell + 1$.

3	P_{ku}^{rec}, $\forall k, u$, is the optimal solution.

given λ, which translates into a faster convergence of the power allocation algorithm. To obtain the optimal Lagrange multiplier λ, we can solve the outer minimization problem in (4.39) by using the following subgradient method [122]

$$\lambda^{[\ell+1]} = \left[\lambda^{[\ell]} - \varepsilon \left(P_R - \sum_{u=1}^{U}\sum_{k=1}^{Q} P_{ku}^{[\ell+1]}\right)\right]^{+}, \tag{4.45}$$

where ε is the step size adopted in the subgradient method, $\lambda^{[\ell]}$ is the value of λ in the ℓth iteration of the subgradient method, and $P_{ku}^{[\ell+1]}$ is the solution of P_{ku} for a given $\lambda^{[\ell]}$. The algorithm to obtain the optimal power allocation for the Agg-PC is summarized in Table 4.1.

4.5 Optimal rBF filters for Ind-PCs

To solve the problem in (4.25) under Ind-PCs, we inspect the common term in the objective functions for OFDMA and SC-FDMA, $(\tilde{\Phi}_k^{(u)} + 1)^{-1}$, which can be rewritten

as

$$(\tilde{\Phi}_k^{(u)}+1)^{-1} = 1 - \frac{\left(\sum_{i=1}^{N_R}\left|\hat{g}_{ik}^{(u)}\hat{h}_{ik}^{(u)}\right|\left|a_{ik}^{(u)}\right|\right)^2}{X_{ku}}.$$

where $X_{ku} = \left(\sum_{i=1}^{N_R}\left|\hat{g}_{ik}^{(u)}\hat{h}_{ik}^{(u)}\right|\left|a_{ik}^{(u)}\right|\right)^2 + \sum_{i=1}^{N_R}|a_{ik}^{(u)}|^2\beta_{ik}^{(u)} + \sigma_d^2$. We can compactly rewrite $(\tilde{\Phi}_k^{(u)}+1)^{-1}$ and X_{ku} using $\mathbf{a}_k^{(u)}$ and $\mathbf{Z}_k^{(u)}$ defined in the previous section as

$$(\tilde{\Phi}_k^{(u)}+1)^{-1} = 1 - \frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}, \tag{4.46}$$

$$X_{ku} = \mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)} + \mathbf{a}_k^{(u)T}\mathbf{D}_k^{(u)}\mathbf{a}_k^{(u)} + \sigma_d^2. \tag{4.47}$$

Next, by considering X_{ku} as an additional optimization variable, we arrive at the following optimization problem, which is equivalent to problem (4.25) under Ind-PCs,

$$\max_{\{\{\mathbf{a}_k^{(u)}\},\{X_{ku}\}}} \quad f_2^{[\mathrm{Y}]}(\mathbf{a},\mathbf{X}) \tag{4.48a}$$

$$\text{s.t.} \quad \sum_{u=1}^{U}\sum_{k=1}^{Q}|a_{ik}^{(u)}|^2\gamma_{ik}^{(u)} - P_R^{(i)} \le 0, \quad \forall i, \tag{4.48b}$$

$$\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)} + \mathbf{a}_k^{(u)T}\mathbf{D}_k^{(u)}\mathbf{a}_k^{(u)} + \sigma_d^2 - X_{ku} \le 0, \quad \forall k,u, \tag{4.48c}$$

where $\mathbf{a} = [[\mathbf{a}^{(1)}]^T,\ldots,[\mathbf{a}^{(U)}]^T]^T$ with $\mathbf{a}^{(u)} = [[\mathbf{a}_1^{(u)}]^T,\ldots,[\mathbf{a}_Q^{(u)}]^T]^T$, $\mathbf{X} = [\mathbf{X}_1^T,\ldots,\mathbf{X}_U^T]^T$ with $\mathbf{X}_u = [X_{1u},\ldots,X_{Qu}]^T$, and

$$f_2^{[\mathrm{Y}]}(\mathbf{a},\mathbf{X}) = \begin{cases} -\sum_{u=1}^{U}\sum_{k=1}^{Q}\tau_u\log_2\left(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}\right), & \text{Y=OFDMA} \\ -\sum_{u=1}^{U}\tau_u Q\log_2\left[\frac{1}{Q}\sum_{k=1}^{Q}\left(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}\right)\right], & \text{Y=SC-FDMA.} \end{cases}$$

Note that in (4.48c) we have relaxed the equality constraint to an inequality constraint to make the constraint convex. This does not affect the final solution since at optimality constraint (4.48c) will be satisfied with equality.

Proposition 4.2. *Optimization problem (4.48) is a reverse-convex (convex maximization) problem with convex constraints.*

Proof. The proof is given in Appendix-E □

Various algorithms have been developed to solve reverse-convex problems with convex constraints, cf. [117], [118]. The problem at hand involves $(N_R + 1)N_c$ real optimization variables, i.e., considering practical values of N_c and N_R, the problem dimension is quite large. To strike a balance between system performance and computational complexity, we adopt a low-complexity suboptimal method called constrained convex concave procedure (CCCP), cf. [52], [120], [119], to find a local optimal solution of problem (4.48) . The key idea of CCCP is to approximate the non-convex part (which is differentiable) in the original problem with its first-order Taylor series expansion around some feasible point. This gives rise to a convex problem whose feasible region lies in a subset of the original non-convex problem. By solving the resulting convex problem, an improved feasible solution of the original problem can be obtained, which is then used as the new starting point for the Taylor series expansion. This process continues until the value of the objective function converges. The solution of the last convex problem is then chosen as the solution to the original problem. Based on the above discussions and the first-order Taylor series expansion

of $f_2^{[\mathrm{Y}]}(\mathbf{a},\mathbf{X})$,

$$f_2^{[\mathrm{Y}]}(\mathbf{a},\mathbf{X}) \approx f_2^{[\mathrm{Y}]}(\hat{\mathbf{a}},\hat{\mathbf{X}}) + \sum_{u=1}^{U}\sum_{k=1}^{Q}\left[\frac{\partial f_2^{[\mathrm{Y}]}(\hat{\mathbf{a}},\hat{\mathbf{X}}))]}{\partial(\Phi_k^{(u)}+1)^{-1}}\right]\left(\frac{-2[\hat{\mathbf{a}}_k^{(u)T}]\mathbf{Z}_k^{(u)}}{\hat{X}_{ku}}\left(\mathbf{a}_k^{(u)}-[\hat{\mathbf{a}}_k^{(u)}]\right)\right.$$
$$\left.+\frac{[\hat{\mathbf{a}}_k^{(u)\dagger}]\mathbf{Y}_k^{(u)}\hat{\mathbf{a}}_k^{(u)}}{(\hat{X}_{ku})^2}\left(X_{ku}-\hat{X}_{ku}\right)\right), \tag{4.49}$$

where $\{\hat{\mathbf{a}}, \hat{\mathbf{X}}\}$ is a feasible solution to the original problem, and

$$\begin{aligned}\frac{\partial f_2^{[\mathrm{OFDMA}]}(\mathbf{a},\mathbf{X})}{\partial(\Phi_k^{(u)}+1)^{-1}} &= -\tau_u\left(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}\right)^{-1}\\ \frac{\partial f_2^{[\mathrm{SC-FDMA}]}(\mathbf{a},\mathbf{X})}{\partial(\Phi_k^{(u)}+1)^{-1}} &= -\tau_u\left[\frac{1}{Q}\sum_{j=1}^{Q}\left(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}\right)\right]^{-1},\end{aligned}$$

we arrive at the following optimization problem

$$\begin{aligned}\max_{\{\mathbf{a}_k^{(u)},X_{ku}\}}\quad & [f_2^{[\mathrm{Y}]}(\mathbf{a},\mathbf{X})]^{(\ell)} + \sum_{u=1}^{U}\sum_{k=1}^{Q}\left[\frac{\partial f_2^{[\mathrm{Y}]}(\mathbf{a},\mathbf{X}))]}{\partial(\Phi_k^{(u)}+1)^{-1}}\right]^{(\ell)}\left(\frac{-2[\mathbf{a}_k^{(u)T}]^{(\ell)}\mathbf{Z}_k^{(u)}}{X_{ku}^{(\ell)}}\right.\\ &\left.\left(\mathbf{a}_k^{(u)}-[\mathbf{a}_k^{(u)}]^{(\ell)}\right)+\frac{[\mathbf{a}_k^{(u)T}]^{(\ell)}\mathbf{Z}_k^{(u)}[\mathbf{a}_k^{(u)}]^{(\ell)}}{(X_{ku}^{(\ell)})^2}\left(X_{ku}-X_{ku}^{(\ell)}\right)\right)\\ \text{s.t.}\quad & \text{(4.48b), (4.48c)}\end{aligned} \tag{4.50}$$

where $(\cdot)^{(\ell)}$ is the value of the function or variable obtained in the ℓth iteration of the CCCP. Since the problem in (4.50) is a convex problem with a linear objective function and convex constraints, it can be solved efficiently using standard convex optimization tools, e.g., CVX or SeDuMi [124], [125]. The algorithm for finding the optimal amplitudes of the rBF filter coefficients under Ind-PCs is summarized in Algorithm 4.2 in Table 4.2. It is worth mentioning that due to the non-convex nature of problem (4.48), the convergence of Algorithm 4.2 to the global optimum is not

guaranteed. However, since problem (4.48) admits a strictly feasible solution, e.g., all-zero rBF filter coefficient amplitudes, and the sequence of the objective functions in problem (4.50) increases monotonically with the iteration number ℓ and is upper-bounded by infinity, the convergence of Algorithm 4.2 to a stationary point of the original problem (4.48) is guaranteed. Moreover, as the objective function in problem (4.48) is twice continuously differentiable and strictly convex in the optimization variables, the stationary point obtained by Algorithm 4.2 can not be a saddle point but has to be a local optimum of problem (4.48). In Section 4.6, we will investigate the convergence behavior of Algorithm 4.2 numerically.

Remark 4.4: It is interesting to compare the solutions of the rBF amplitudes obtained for the two considered power constraints. For Ind-PCs, the power allocation across relays, users, and subcarriers is performed simultaneously through the rBF filter coefficient amplitude optimization in (4.50), which involves solving a sequence of convex optimization problems with $(N_R + 1)N_c$ real optimization variables. For the Agg-PC, the power allocation across relays is obtained from the closed-form expression given in (4.34), and the power allocation among users and subcarriers is obtained by solving a single convex optimization problem in (4.36) with N_c real optimization variables. Therefore, the complexity of Algorithm 4.2 for the Ind-PCs is higher than that of Algorithm 4.1 for the Agg-PC, especially for large values of N_R.

4.6 Simulation Results

In this section, we evaluate the performance of the proposed robust rBF and dEQ schemes for OFDMA and SC-FDMA systems using simulations. All TD channels are modeled as uncorrelated Rayleigh block fading channels whose CIR is generated based

Table 4.2: Algorithm 4.2 for finding the optimal amplitudes of the rBF filter coefficients under Ind-PCs. ϵ is a small constant to control the accuracy of the CCCP, e.g., $\epsilon = 10^{-4}$. L is the maximum iteration number, e.g., $L = 50$.

1	Initialize $[\mathbf{a}_k^{(u)}]^{[1]}$, $X_{ku}^{[1]}, \forall k, u$. Calculate $[f_2^{[\mathrm{Y}]}(\mathbf{a}, \mathbf{X})]^{[1]}$. Set $\ell = 1$.
2	Repeat Find $[\mathbf{a}_k^{(u)}]^{[\ell+1]}$ and $X_{ku}^{[\ell+1]}, \forall k, u$, by solving the convex problem in (4.50). Calculate the new objective function value $[f_2^{[\mathrm{Y}]}(\mathbf{a}, \mathbf{X})]^{[\ell+1]}$. Stop when $\lvert[f_2^{[\mathrm{Y}]}(\mathbf{a}, \mathbf{X})]^{[\ell+1]} - [f_2^{[\mathrm{Y}]}(\mathbf{a}, \mathbf{X})]^{[\ell]}\rvert < \epsilon$, or $\ell > L$. $\ell = \ell + 1$.
3	$[\mathbf{a}_k^{(u)}]^{[\ell+1]}$, $\forall k, u$, is the optimal solution.

on the power delay profile $p[n] = \frac{1}{\sigma_t} \sum_{l=0}^{L-1} e^{-n/\sigma_t} \delta[n-l]$ [126], where $\sigma_t = 2$, which corresponds to moderately frequency-selective fading. For convenience, we assume that the length of all multipath channels is equal to 16, i.e., $L_{g,u,i} = L_{h,i} = 16, \forall u, i$. The received SNR at relay and destination is defined as $\mathrm{SNR_r} = \frac{P_S}{\sigma_r^2 N_c}$ and $\mathrm{SNR_r} = \frac{P_R}{\sigma_d^2 N_c}$, respectively. For convenience, we assume that $P_S = P_R = N_c$. We further define $(\mathrm{SNR})_{\mathrm{ref}}$ as the value of SNR when $\mathrm{SNR_r} = \mathrm{SNR_d}$. The proposed rBF with optimal power allocation is denoted as rBF-OPA. Three baseline schemes are considered: a naive-AF scheme, optimal rBF with equal power allocation (rBF-EPA), and non-robust rBF-OPA. For the naive-AF scheme, the destination employs the optimal equalizer in (4.19) while the relays perform simple power amplification to meet the average power constraint. For rBF-EPA, the relays only align the phases of the estimated end-to-end channels and distribute the power uniformly across frequencies, users, and relays. For the non-robust scheme, both the relays and the destination treat the estimated CSI as the actual CSI for the design of the dEQ and rBF filters. Unless specified otherwise, we set $\sigma_{g,i}^2 = \sigma_{h,i}^2 = 0.1, \forall i$, the number of users is $U = 4$, the data block size of each user is $Q = 16$, uniform user weighting is adopted, i.e., $\tau_u = 1/U$, and equal relay power budgets are assumed for the Ind-PCs, i.e., $P_R^{(i)} = P_R/N_R, \forall i$. All simulation results are obtained by averaging over 10,000 realizations of the estimated channels and CSI errors.

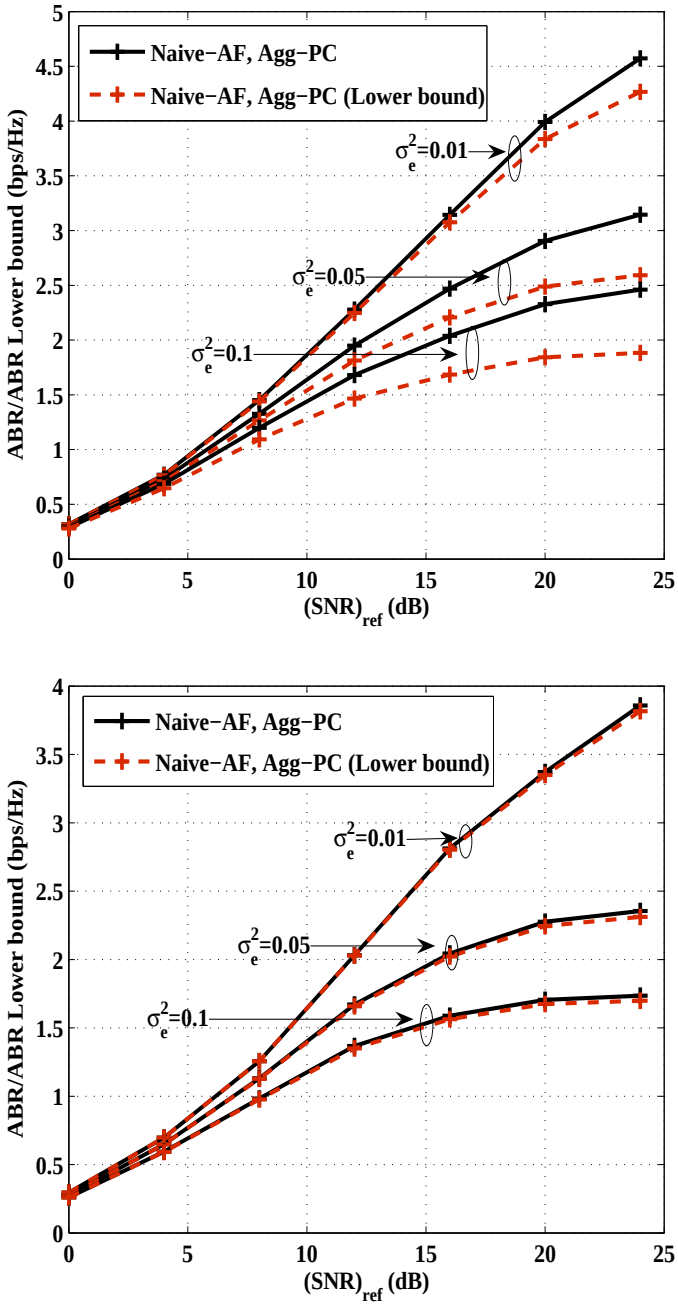

Figure 4.2: Average ABR and ABR lower bound for multiuser multi-relay systems, two relays, and $\sigma^2_{e,g_i} = \sigma^2_{e,h_i} = \sigma^2_e,\ \forall i$. **Upper**: OFDMA. **Lower**: SC-FDMA.

4.6.1 Tightness of the ABR Lower Bound

In Fig. 4.2, we show the actual average ABRs and the corresponding ABR lower bounds for two-relay OFDMA and SC-FDMA systems employing the naive-AF scheme and the Agg-PC. As can be seen, for OFDMA, the bound is close to the actual ABR when the CSI error variance is small and when the SNR is low. On the other hand, for SC-FDMA, the lower bound is tight over the entire SNR range and for all considered CSI error variances. The different tightness of the lower bound may be due to the fact that, for SC-FDMA, the inverse Fourier transform after the FDE has an averaging effect on the SINR of the TD symbols at the input of the slicer (as indicated by the factor $1/Q$ inside the logarithm in (4.14)), which reduces the fluctuation of the ABR compared to that of OFDMA. Note that although the lower bound for OFDMA is not tight for high SNR and large CSI error variances, adopting the lower bound for optimization leads to substantial performance gains compared to un-optimized and non-robust relay systems, as will be shown in the following figures. We also note that as the CSI error variances increase, the ABR degrades dramatically for both OFDMA and SC-FDMA. This illustrates the necessity of a robust design of the rBF and dEQ filters.

4.6.2 Convergence of Algorithm 4.2

As explained in Section 4.5, Algorithm 4.2, which is used to compute the optimal rBF filter coefficient amplitudes for the Ind-PCs, involves solving a sequence of convex optimization problems. Therefore, the complexity of the algorithm depends on the number of iterations required to obtain the optimal solution. In Fig. 4.3, we numerically investigate the convergence behavior of the algorithm by examining the values of the objective function (the ABR lower bound) versus the iteration number, ℓ. In

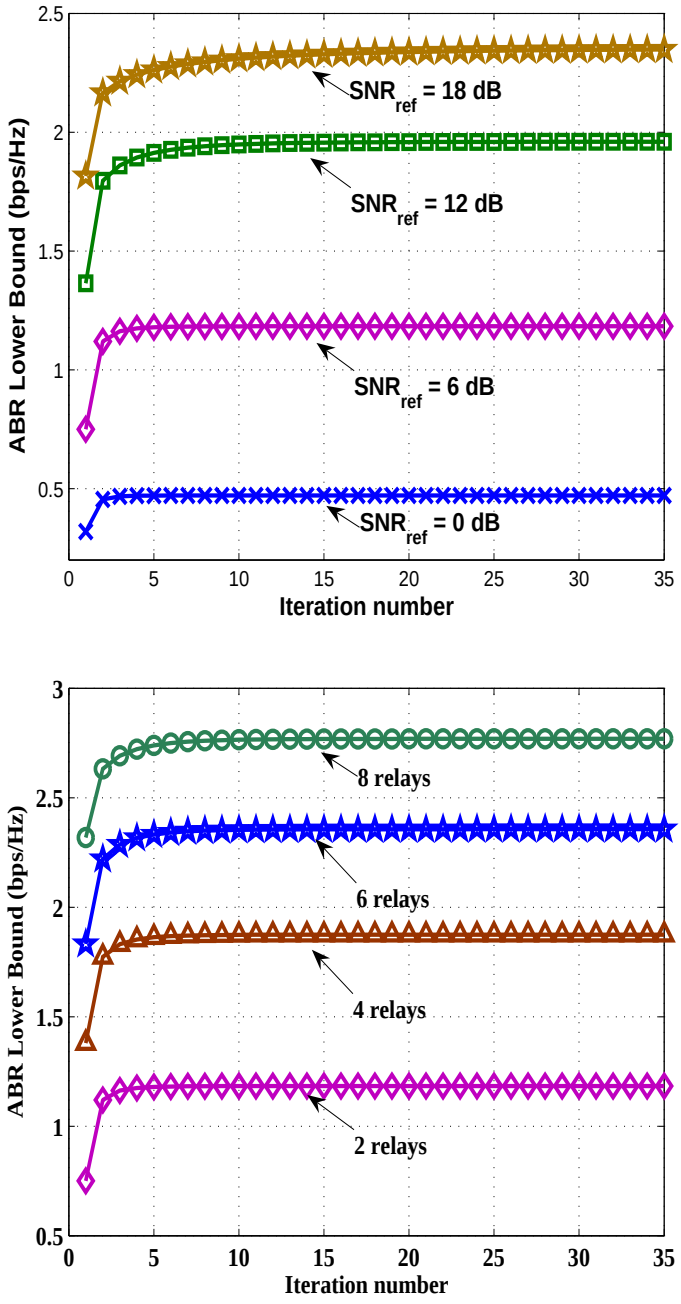

Figure 4.3: Objective function value versus the iteration number in Algorithm II for SC-FDMA, $\sigma^2_{e,g_i} = \sigma^2_{e,h_i} = 0.1$, $\forall i$. **Upper**: SNR_{ref}=$\{0, 6, 12, 18\}$ dB, and N_R=2. **Lower**: N_R=$\{2, 4, 6, 8\}$, and SNR_{ref}=6 dB.

the upper sub-figure, we study a two-relay SC-FDMA system operating at different SNR_{ref}. In the lower sub-figure, we examine a multi-relay SC-FDMA system operating at $\text{SNR}_{\text{ref}} = 6$ dB for different numbers of relays. We observe that as SNR_{ref} and N_R increase, more iterations are required for the algorithm to converge. However, for all the considered cases, the objective value is already close to the optimal value after 5 iterations. In the following, we set the maximum number of iterations for the algorithm equal to 20 to achieve near-optimal performance.

4.6.3 Comparison of the Robust and Non-robust rBF Schemes

In Fig. 4.4, we examine the ABR performance of the proposed rBF and dEQ schemes for the Agg-PC for SC-FDMA systems with one and two relays. We also include the performance of non-robust rBF/dEQ and non-robust rBF with robust dEQ for $\sigma_e^2 = 0.1$. As expected, for larger CSI error variances, the performance of both the robust and the non-robust rBF/dEQ schemes degrades, as the CSI used for designing the involved filters becomes increasingly inaccurate. However, the proposed robust scheme achieves a considerable performance gain compared to the non-robust schemes especially for large CSI error variances. Furthermore, the non-robust schemes become less sensitive to the CSI errors as the number of relays increases, because of the additional spatial diversity provided by the relays. For one- and two-relay systems with $\sigma_e^2 = 0.1$, we also observe that robust dEQ can compensate for a considerable amount of the ABR loss suffered by non-robust rBF and non-robust dEQ.

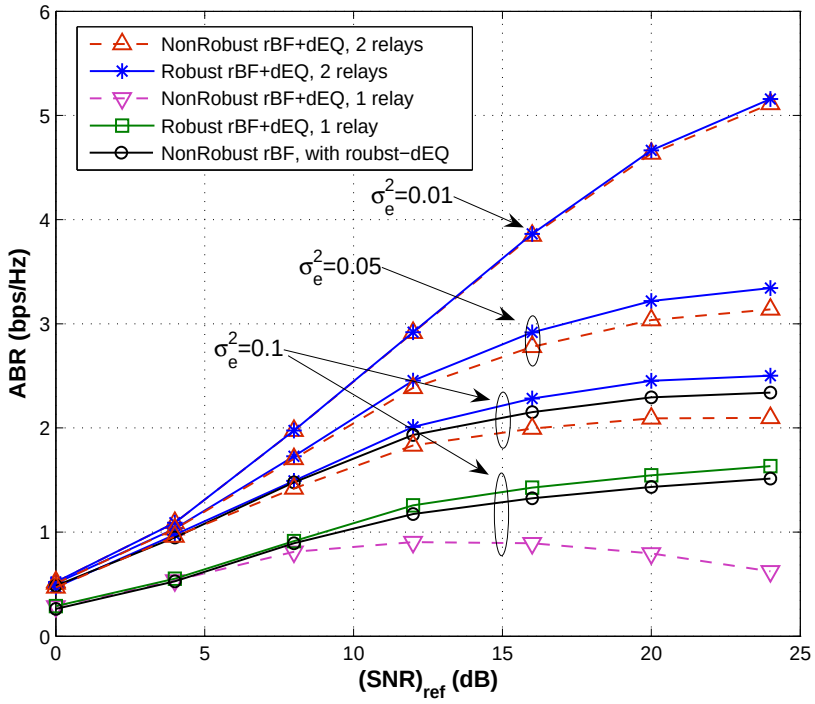

Figure 4.4: ABR of multi-relay SC-FDMA systems with robust and non-robust rBF schemes, two relays, and $\sigma^2_{e,g_i} = \sigma^2_{e,h_i} = \sigma^2_e,\ \forall i$.

4.6.4 Comparison of Optimal and Suboptimal rBF Schemes for Agg-PC and Ind-PCs

In Fig. 4.5, we study the weighted ABR of the optimal and suboptimal rBF schemes for a two-relay SC-FDMA system using uniform and non-uniform user weighting, respectively. The non-uniform weighting factors are chosen to be $\{0.9, 0.06, 0.03, 0.01\}$, which is an extreme case where one of the users has much higher priority than the other users. From both sub-figures, we can see that rBF-OPA with Ind-PCs performs only slightly worse than rBF-OPA with the Agg-PC, regardless of the choice of the user weighting factors. Also, suboptimal rBF-EPA outperforms the naive-AF scheme by a large margin, which underlines the importance of the channel coeffi-

cient phase alignment for rBF filter design. In the upper sub-figure, suboptimal rBF-EPA achieves a similar performance as rBF-OPA for both types of power constraints. This is because with uniform user weighting, all users contribute equally to the system ABR. Thus, distributing the power uniformly across users, subcarriers, and relays does not cause a significant performance degradation. However, if the weighting factors of the users are non-uniform, cf. the lower sub-figure, rBF-EPA suffers from a considerable ABR loss compared to rBF-OPA.

4.6.5 Effect of Different CSI Error Variances

In Fig. 4.6, we investigate the effect of different values of σ^2_{e,g_i} and σ^2_{e,h_i} on the ABR performance of the considered optimal and suboptimal rBF schemes for the Agg-PC. We consider two types of CSI errors: $\{\sigma^2_{e,g_i}, \sigma^2_{e,h_i}\} = \{0.5, 0\},\ \forall i$, and $\{\sigma^2_{e,g_i}, \sigma^2_{e,h_i}\} = \{0, 0.5\},\ \forall i$. This corresponds to the situation where the CSI in one hop contains large errors while the CSI in the other hop is perfectly known. We increase SNR_r and SNR_d simultaneously in the upper sub-figure, and fix $\mathrm{SNR}_r = 12$ dB in the lower sub-figure. The results show that when $\mathrm{SNR}_r = \mathrm{SNR}_d$, almost identical ABR performance is achieved for both types of CSI errors. This is in accordance with the analysis in Remark 4.2 of Section 4.41, where it was shown that when the noise level of the two hops are identical, σ^2_{e,g_i} and σ^2_{e,h_i} will have a similar effect on the ABR. However, when $\mathrm{SNR}_r \gg \mathrm{SNR}_d$, the ABR with $\{\sigma^2_{e,g_i}, \sigma^2_{e,h_i}\} = \{0.5, 0\},\ \forall i$, is considerably lower than that with $\{\sigma^2_{e,g_i}, \sigma^2_{e,h_i}\} = \{0, 0.5\},\ \forall i$, and when $\mathrm{SNR}_r \ll \mathrm{SNR}_d$, the reverse holds true. This reveals that the link with the lower SNR is more robust to CSI imperfection than the link with the higher SNR, which is again in accordance with Remark 4.2. Moreover, from both figures, we also observe that, in contrast to Fig. 4.5, the performance gain achieved by rBF-OPA over rBF-EPA is more prominent in the

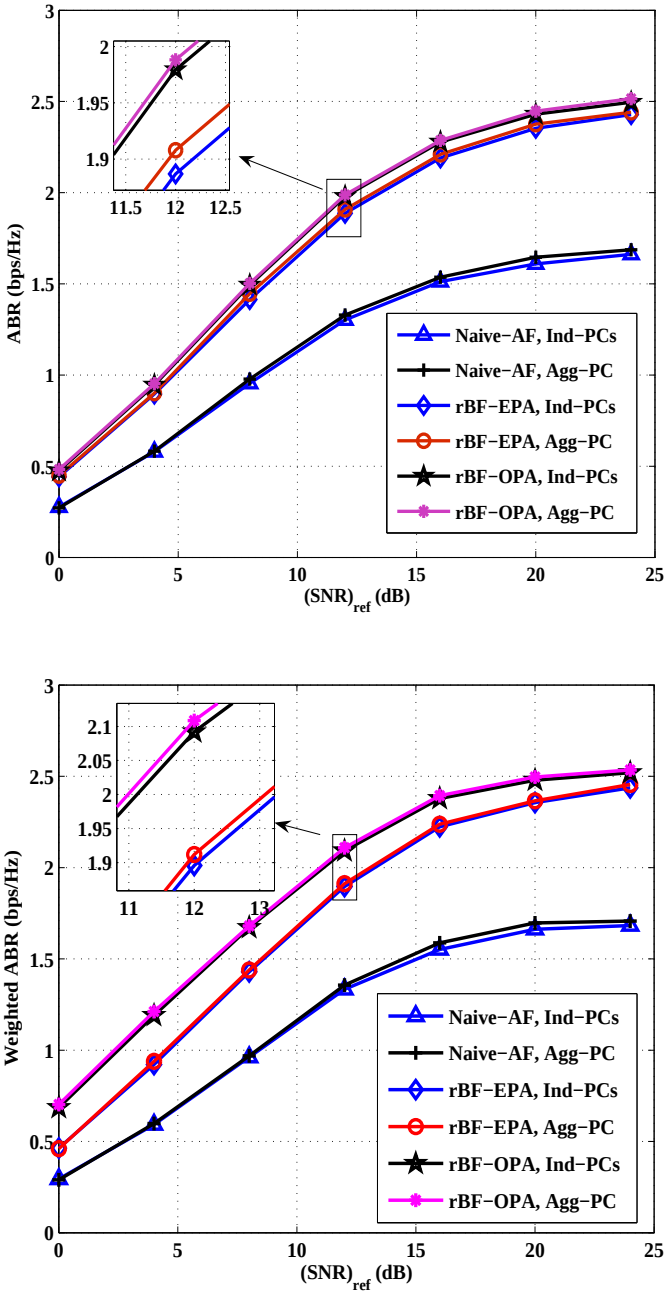

Figure 4.5: ABR and weighted ABR of multi-relay SC-FDMA systems with optimal/suboptimal rBF schemes, two relays, and $\sigma^2_{e,g_i} = \sigma^2_{e,h_i} = 0.1$, $\forall i$. **Upper**: Uniform weighting. **Lower**: Non-uniform weighting.

high SNR region. This implies that if the CSI errors in one of the two hops dominate, performing OPA is very beneficial for the ABR performance.

4.6.6 Effect of U and N_R on the Network ABR

For multiuser systems, the impact of the number of users on the ABR of the network is an important consideration. In Fig. 4.7, we evaluate the network ABR versus the number of users for two-relay OFDMA and SC-FDMA systems operating at $\text{SNR}_{\text{ref}} = 12$ dB. The total number of subcarriers, N_c, is fixed to 64. From the upper sub-figure, we observe that for OFDMA, the ABRs of all the considered schemes do not scale with the number of users in the network. In contrast, in the lower sub-figure, the ABRs for SC-FDMA increase as more users are accommodated in the network, e.g., the ABR increases from 1.9 bps/Hz to 2.1 bps/Hz for rBF-OPA when U increases from 1 to 16. This observation can be explained as follows: Since the ABR for OFDMA is a linear function of per-subcarrier ABRs, as long as N_c is fixed, the network ABR will not increase notably with the number of users. For SC-FDMA, the ABR varies with the number of subcarriers assigned to each user. Consider two extreme cases: When $U = 1$, the SC-FDMA system reduces to a single-user SC-FDE system, whose ABR is lower than that of the correponding OFDM system. When $U = 64$, the SC-FDMA system is essentially equivalent to an OFDMA system since only one subcarrier is assigned to each user. In between, as U increases, the ABR of the SC-FDMA system gradually approaches that of the OFDMA system as the number of subcarriers assigned to each user decreases.

As we consider multi-relay systems, it is also of interest to investigate how the number of relays affects the ABR of the network. In Fig. 4.8, we plot the ABR as a function of the number of relays for a multi-relay system operating at $\text{SNR}_{\text{ref}} = 10$

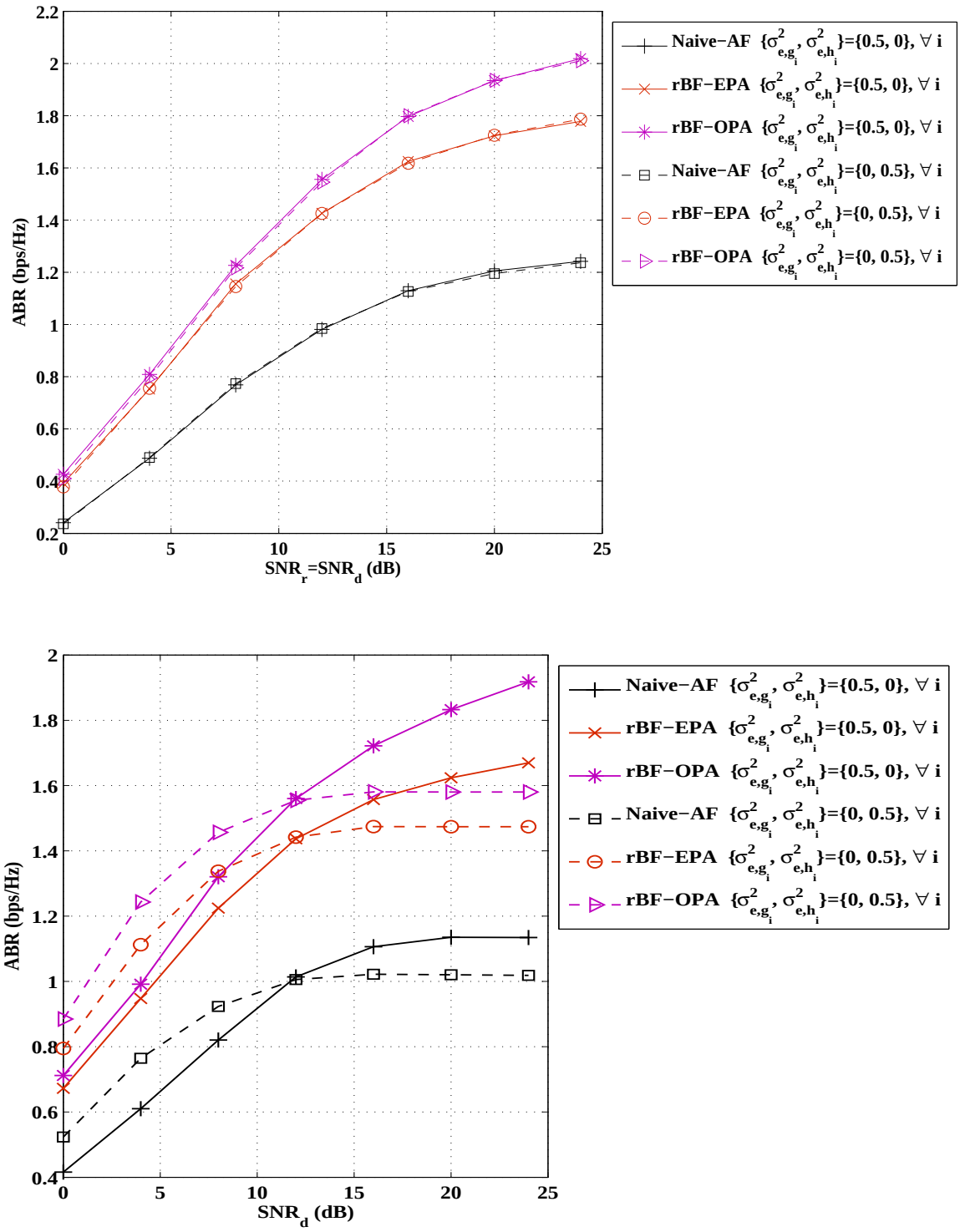

Figure 4.6: ABR of SC-FDMA systems with optimal/suboptimal rBF schemes for the Agg-PC, two relays, and $\{\sigma^2_{e,g_i}, \sigma^2_{e,h_i}\} = \{0.5, 0\}$ or$\{0, 0.5\}, \forall i$. **Upper**: $\text{SNR}_r = \text{SNR}_d$. **Lower**: $\text{SNR}_r = 12$ dB.

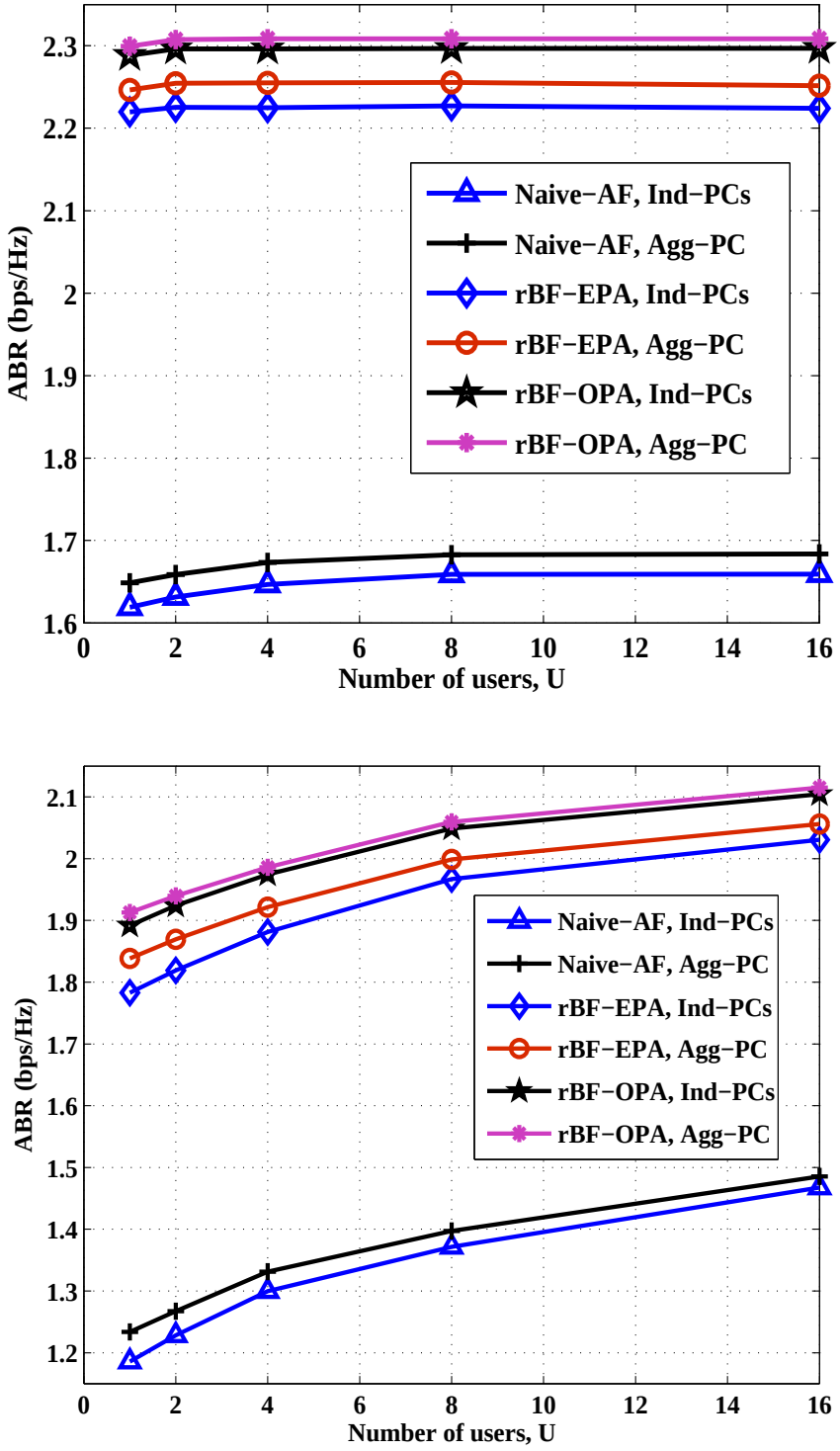

Figure 4.7: ABR of multiuser multi-relay systems with optimal/suboptimal rBF schemes, two relays, $(\text{SNR})_{\text{ref}} = 12$ dB, and $\sigma^2_{e,g_i} = \sigma^2_{e,h_i} = 0.1$, $\forall i$. **Upper**: OFDMA. **Lower**: SC-FDMA.

dB. We show the results for SC-FDMA systems with both equal and unequal relay power budgets (rPB), where in the latter case $P_R^{(i)}$ is randomly generated such that $\sum_i P_R^{(i)} = P_R$. From the figure, we observe that the ABRs for the naive-AF schemes increase only slightly with N_R, while remarkable ABR performance gains are achieved for rBF-OPA and rBF-EPA when more relays are deployed. This is due to the fact that the channel coefficient phase alignment in rBF can take full advantage of the increased spatial diversity for large numbers of relays. Moreover, the performance gap between the schemes using Ind-PCs and those using the Agg-PC is enlarged as N_R increases. This phenomenon becomes more prominent when nonequal relay power budgets are adopted. This is because with the increase of N_R, the Agg-PC can offer increasingly more degrees of freedom in allocating the relay transmit powers as compared to the Ind-PCs.

4.7 Conclusion

In this chapter, we investigated robust rBF and dEQ filter design for multi-relay OFDMA and SC-FDMA systems based on a statistical CSI error model. A lower bound on the weighted ABR of the network was maximized under an Agg-PC and Ind-PCs, respectively. First, the optimal dEQ filters and the phases of the optimal rBF filter coefficients were derived for both types of constraints. Then, for the Agg-PC, a two-step solution for the optimal amplitudes of the rBF filter coefficients was proposed by employing primal decomposition. In the first step, for each subcarrier of the users, the optimal rBF allocates power across relays by taking into account the CSI error variances. In the second step, a convex power allocation procedure further distributes the power across subcarriers and among users to enhance the weighed ABR performance. For Ind-PCs, the optimal amplitudes of the rBF filter coefficients

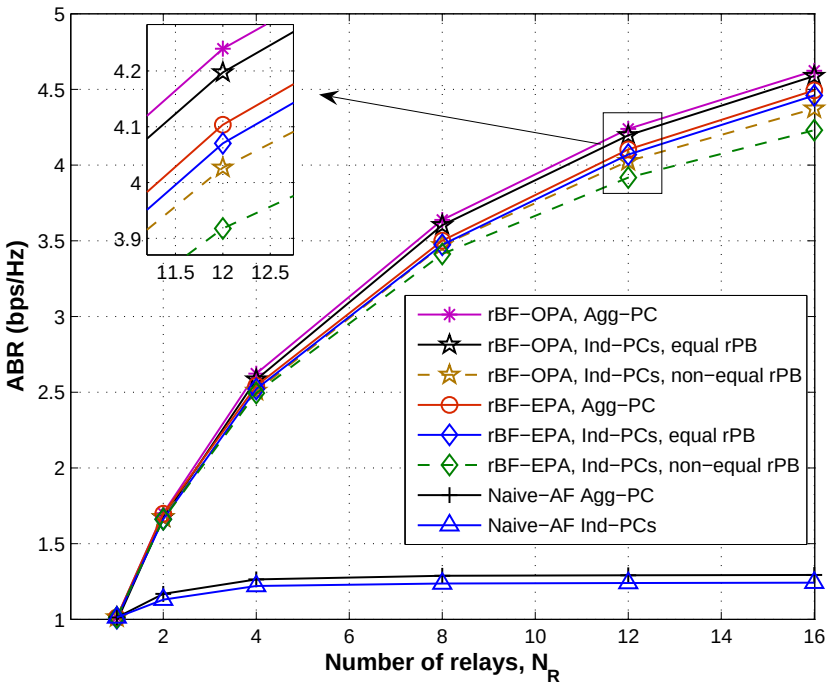

Figure 4.8: ABR of multi-relay SC-FDMA systems with optimal/suboptimal rBF schemes and equal/unequal relay power budgets (rPB), $(\text{SNR})_{\text{ref}} = 10$ dB, and $\sigma^2_{e,g_i} = \sigma^2_{e,h_i} = 0.1,\ \forall i$.

were obtained by solving a sequence of convex optimization problems. Simulation results showed that the proposed robust rBF schemes outperform conventional non-robust and naive-AF relaying schemes. Furthermore, our results revealed that the suboptimal rBF-EPA scheme achieved near-optimal performance when uniform user weighting was employed and the CSI error variances in both hops were similar. On the other hand, if non-uniform user weighting was adopted and/or the CSI error variances in the two hops were quite different, the optimal rBF-OPA scheme enjoyed a considerable performance gain over rBF-EPA.

Chapter 5

Robust Transceiver Design for SC-FDE Multi-hop Full-Duplex DF Relay Systems

5.1 Introduction

In the previous chapter, we have investigated the transceiver design for a dual-hop multiple parallel relay network. In practical wireless systems, multi-hop serial relay network is another promising candidate with various potential applications. For example, in rural areas, where the cell sizes of cellular systems are relatively large, multi-hop transmission is beneficial for ensuring the QoS of cell-edge users [14, 18]. Multi-hop relay systems have also found application in device-to-device communications, where the number of wireless devices that can potentially serve as intermediate relaying nodes is typically large [130, 129]. Another application for multi-hop relaying is millimeter-wave communications due to the severe channel attenuation at high frequencies [131]. Conventionally, multi-hop relay systems are operated in the half-duplex mode, which results in a low spectrum efficiency, especially when the number of hops is large. Hence, considering the recent advances in hardware technology for interference cancellation [74, 75, 76, 132], full-duplex relays (FDRs) that can transmit

and receive signals simultaneously are drawing a growing interest from the research community.

In this chapter, we propose a robust transceiver design for a multi-hop SC-FDE based DF FDR system assuming imperfect CSI at the transceivers. Thereby, we adopt the statistical error model because of its suitability for modeling CSI errors originating from non-ideal channel estimation [133]. Since the MSE is an important performance metric for systems employing equalization [126], we adopt the minimization of the sum MSE and the maximum MSE across the different hops as our design objectives. Separate transmit node power constraints are imposed because source and relays usually have their own power supplies and thus cannot share powers. Moreover, separate power constraints also reflect the hardware limitations and emission regulations of the transmitting nodes. To solve the formulated problems, we first derive the optimal equalization filters at the receiving nodes, which are identical for both considered optimization criteria and take the form of robust Wiener filters. With the optimal equalization filters in hand, the transmit precoding problems become non-convex power allocation problems in the frequency domain. In a first approach, we propose a sequential geometric programming (sGP) based method to solve the power allocation problems, where the condensation technique [89] is applied to transform the objective function into a posynomial. Then, by successively solving a sequence of standard GP problems, improved approximate solutions for the original problems are obtained. The sGP based power allocation schemes require a central node to collect the CSI of all hops, leading to a significant signalling overhead. In addition, only numerical solutions can be obtained for the sGP subproblems, which does not provide physical insights. Hence, motivated by the observation that by fixing the powers of all other nodes, the power optimization of a specific node can be

formulated as either a convex problem or a difference of convex problems, we also investigate an alternating optimization (AO) method [135], where closed-form solutions are obtained in each iteration step and efficient algorithms can be developed by exploiting the (partial) convexity of the subproblems. Our simulation results reveal that the AO based method and the sGP based method yield almost identical performances in terms of MSE and achievable rate. However, the AO based algorithm has the advantage that it requires significant less signalling overhead for systems with a large number of hops and facilitates a semi-distributed implementation if the relays can employ directional antennas to avoid the backward interference.

The remainder of the chapter is organized as follows. In Section 5.2, we introduce the system model for the considered SC-FDE based DF FDR system. In Section 5.3, the optimization problems for sum MSE and maximum MSE minimization are formulated and the optimal FDE filter is derived. In Section 5.4, the sGP and AO based power allocation schemes are developed. Simulation results are provided in Section 5.5, and some conclusions are drawn in Section 5.6.

5.2 System Model

We consider a multi-hop DF FDR system that consists of one source node, S, $M-1$ relay nodes, $R_i, i = 1, \ldots, M-1$, and one destination node, D, as shown in Fig. 5.1. The relay nodes are equipped with two antennas[21], one for signal transmission and one for signal reception. The source and the destination are equipped with a single antenna. For convenience, we denote the source as node 1, the ith relay as node $(i+1)$, and the destination as node $(M+1)$. Due to the large path loss and shadowing, there

[21] Although equipped with two antennas, the relay cannot employ multi-antenna techniques, e.g., beamforming, since there is only a single transceiver in each relay.

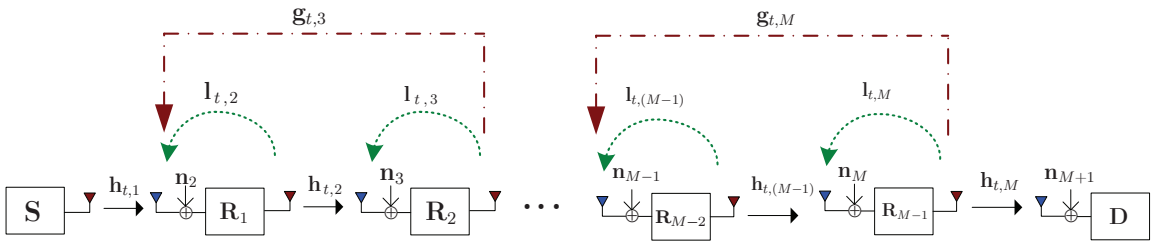

Figure 5.1: Block diagram of an M-hop full-duplex relay system. The backward interference exists only for $M \geq 3$.

are no direct links between any non-adjacent nodes. However, not only LI (dotted arrows in Fig. 5.1) will be present for $M \geq 3$, but also backward interference (dash-dotted arrows in Fig. 5.1) between adjacent relays .

5.2.1 Signal and Channel Model

The information symbol vector generated at the source is given by $\mathbf{s}^{[n]} = [s_1^{[n]}, \ldots, s_{N_c}^{[n]}]^T$, where n is the index of the transmit data block, N_c is the data block size, and the $s_m^{[n]}, \forall m, n$, are modeled as i.i.d. random variables with zero mean and unit variance. We assume error-free decoding at the relays due to the use of capacity-achieving channel codes and adoption of a suitable transmission rate which is strictly lower than the channel capacity. At transmit node i, either the original signal (for the source node) or the error-free decoded signal (for the relay nodes) is converted into the FD using an N_c-point FFT. The FD signal is then processed by a diagonal FD precoding (one-tap filter) matrix $\mathbf{P}_{f,i} = \text{diag}\{p_{k,i}, \forall k\} \in \mathbb{C}^{N_c \times N_c}$. Subsequently, the precoded signal is converted back into the TD using an N_c-point IFFT, resulting in the transmitted signal

$$\mathbf{x}_i = \mathbf{F}_{N_c}^{\dagger} \mathbf{P}_{f,i} \mathbf{F}_{N_c} \mathbf{s}^{[n-i+1]}, \quad i = 1, \ldots, M, \tag{5.1}$$

where $\mathbf{s}^{[n]}$ is assumed to be independent of $\mathbf{s}^{[n']}, \forall n' \neq n$, and $\mathbf{F}_{N_c}$ is the DFT matrix of size N_c. Note that the combined effect of $\mathbf{F}_{N_c}^{\dagger}\mathbf{P}_{f,i}\mathbf{F}_{N_c}$ is essentially a circular filtering of the signal in the TD. Before transmission, $\mathbf{x}_i$ is appended with a CP, which is identical to the last $N_{\mathrm{cp},i}$ symbols of the data block, where $N_{\mathrm{cp},i}$ is chosen to be larger than the lengths of the CIRs of the transmit channel between node i and node $(i+1)$, the LI channel of node i, and the backward channel from node i to node $(i-1)$. At the receiving node, the CP is first removed prior to any signal processing. The received signal at node $(i+1)$ after CP removal can be written as

$$\mathbf{y}_{(i+1)} = \mathbf{H}_{t,i}\mathbf{x}_i + \mathbf{L}_{t,(i+1)}\mathbf{x}_{(i+1)} + \mathbf{G}_{t,(i+2)}\mathbf{x}_{(i+2)} + \mathbf{n}_{(i+1)}, \quad i = 1, \ldots, M, \tag{5.2}$$

where $\mathbf{H}_{t,i} \in \mathbb{C}^{N_c \times N_c}$ and $\mathbf{G}_{t,i} \in \mathbb{C}^{N_c \times N_c}$ are the TD CIR matrices from node i to node $(i+1)$ and that from node $(i+1)$ to node i, respectively, $\mathbf{L}_{t,i} \in \mathbb{C}^{N_c \times N_c}$ is the TD CIR matrix of the LI channel at node i, and $\mathbf{n}_i$ is the TD AWGN vector at node i, whose entries are i.i.d. with zero mean and variance $\sigma_{n_i}^2$. Note that there is no backward interference and LI at the destination, and there is also no backward interference at the last relay, i.e., $\mathbf{G}_{t,i} = \mathbf{L}_{t,i} = \mathbf{0}_{N_c \times N_c}, \forall i \geq M+1$. Since the CP transforms the linear convolution of the CIR and the signal into a circular convolution, the TD channel matrices $\mathbf{H}_{t,i}$, $\mathbf{G}_{t,i}$, and $\mathbf{L}_{t,i}$ are circulant matrices, i.e., they can be expressed as $\mathbf{H}_{t,i} = \mathrm{circ}\{[\mathbf{h}_{t,i}^T, \mathbf{0}_{1\times(N_c-L_{h_i})}]^T\}$, $\mathbf{G}_{t,i} = \mathrm{circ}\{[\mathbf{g}_{t,i}^T, \mathbf{0}_{1\times(N_c-L_{g_i})}]^T\}$, and $\mathbf{L}_{t,i} = \mathrm{circ}\{[\mathbf{l}_{t,i}^T, \mathbf{0}_{1\times(N_c-L_{l_i})}]^T\}$, where $\mathbf{h}_{t,i} \in \mathbb{C}^{L_{h_i}\times 1}$, $\mathbf{g}_{t,i} \in \mathbb{C}^{L_{g_i}\times 1}$, and $\mathbf{l}_{t,i} \in \mathbb{C}^{L_{l_i}\times 1}$ contain the CIR coefficients of the forward, backward, and LI channels, respectively, and are given by $\mathbf{h}_{t,i} = [h_{t,i}^{[1]}, \ldots, h_{t,i}^{[L_{h_i}]}]$, $\mathbf{g}_{t,i} = [g_{t,i}^{[1]}, \ldots, g_{t,i}^{[L_{g_i}]}]$, and $\mathbf{l}_{t,i} = [l_{t,i}^{[1]}, \ldots, l_{t,i}^{[L_{l_i}]}]$, respectively. Here, L_{h_i}, L_{g_i}, and L_{l_i} are the lengths of the corresponding CIRs. Since circulant matrices can be eigen-decomposed by the DFT matrix, we have $\mathbf{Q}_{t,i} = \mathbf{F}_{N_c}^{\dagger}\mathbf{Q}_{f,i}\mathbf{F}_{N_c}$, where $\mathbf{Q}_{f,i} \in \{\mathbf{H}_{f,i}, \mathbf{G}_{f,i}, \mathbf{L}_{f,i}\}$ are diagonal matrices containing

the corresponding FD channel coefficients.

5.2.2 Channel Estimation Error Model

In this work, we assume that the channel frequency responses are estimated at the receivers via training symbols sent by the corresponding transmitters. Assuming a linear MSE channel estimator, the channel estimation errors can be modelled as Gaussian random variables with known statistics [133]. Thus, we can rewrite the actual FD channel matrices as

$$\mathbf{Q}_{f,i} = \hat{\mathbf{Q}}_{f,i} + \Delta\mathbf{Q}_{f,i}, \tag{5.3}$$

where $\hat{\mathbf{Q}}_{f,i} \in \{\hat{\mathbf{H}}_{f,i}, \hat{\mathbf{G}}_{f,i}, \hat{\mathbf{L}}_{f,i}\}$ are the estimated channel frequency response matrices, which are diagonal matrices with the main diagonal entries given by $\hat{q}_{k,i} \in \{\hat{h}_{k,i}, \hat{g}_{k,i}, \hat{l}_{k,i}\}$, $k = 1, \ldots, N_c$, respectively. $\Delta\mathbf{Q}_{f,i} \in \{\Delta\mathbf{H}_{f,i}, \Delta\mathbf{G}_{f,i}, \Delta\mathbf{L}_{f,i}\}$ are the corresponding diagonal channel estimation error matrices, whose main diagonal entries are complex Gaussian random variables with zero mean and variances $\sigma^2_{e,h,k,i}$, $\sigma^2_{e,g,k,i}$ and $\sigma^2_{e,l,k,i}$, $k = 1, \ldots, N_c$, respectively. Without loss of generality, we assume $\sigma^2_{e,q,k,i} = \sigma^2_{e,q,i}, q \in \{h, g, l\}, \forall k$, in the following for notational simplicity.

5.2.3 Processing at the Receiving Nodes

We assume that the ith relay knows the estimated FD LI and backward channels, $\hat{\mathbf{L}}_{f,(i+1)}$ and $\hat{\mathbf{G}}_{f,(i+2)}$, and the transmitted data symbol vectors, $\mathbf{x}_{(i+1)}$ and $\mathbf{x}_{(i+2)}$. Thus, it can perform partial LI and backward interference cancellation by subtracting $\hat{\mathbf{L}}_{t,(i+1)}\mathbf{x}_{(i+1)} + \hat{\mathbf{G}}_{t,(i+2)}\mathbf{x}_{(i+2)}$ from $\mathbf{y}_{(i+1)}$, where $\hat{\mathbf{L}}_{t,i} = \mathbf{F}^{\dagger}_{N_c}\hat{\mathbf{L}}_{f,i}\mathbf{F}_{N_c}$ and $\hat{\mathbf{G}}_{t,i} =$

$\mathbf{F}_{N_c}^{\dagger}\hat{\mathbf{G}}_{f,i}\mathbf{F}_{N_c}$, resulting in the following effective received signal

$$\begin{aligned}\bar{\mathbf{y}}_{(i+1)} &= \mathbf{y}_{(i+1)} - \hat{\mathbf{L}}_{t,(i+1)}\mathbf{x}_{(i+1)} - \hat{\mathbf{G}}_{t,(i+2)}\mathbf{x}_{(i+2)} \\ &= \mathbf{H}_{t,i}\mathbf{x}_i + \Delta\mathbf{L}_{t,(i+1)}\mathbf{x}_{(i+1)} + \Delta\mathbf{G}_{t,(i+2)}\mathbf{x}_{(i+2)} + \mathbf{n}_{(i+1)}, \end{aligned} \tag{5.4}$$

where $\Delta\mathbf{L}_{t,i} = \mathbf{F}_{N_c}^{\dagger}\Delta\mathbf{L}_{f,i}\mathbf{F}_{N_c}$ and $\Delta\mathbf{G}_{t,(i+1)} = \mathbf{F}_{N_c}^{\dagger}\Delta\mathbf{G}_{f,(i+1)}\mathbf{F}_{N_c}$. Signals $\bar{\mathbf{y}}_{(i+1)}$ are converted into the FD, equalized by FD equalization matrices $\mathbf{W}_{f,(i+1)} = \text{diag}\{w_{k,(i+1)}, \forall k\} \in \mathbb{C}^{N_c \times N_c}$, and then converted back to the TD. Hence, the equalized TD signal and the error vector at receiving node $(i+1)$ can be written as

$$\hat{\mathbf{y}}_{(i+1)} = \mathbf{F}_{N_c}^{\dagger}\mathbf{W}_{f,(i+1)}\mathbf{F}_{N_c}\bar{\mathbf{y}}_{(i+1)} \quad \text{and} \quad \mathbf{e}_{(i+1)} = \hat{\mathbf{y}}_{(i+1)} - \mathbf{s}^{[n-i+1]}, \quad i = 1, \ldots, M, \tag{5.5}$$

respectively. Subsequently, the signal is demodulated and decoded. At the relay, the recovered signal is further re-encoded, re-modulated, and sent over the channel to the next receiving node.

5.3 Problem Formulation and Optimal Equalization Filters

For the design of equalization schemes, the MSE is an important and commonly used performance criterion [126]. Therefore, in this section, we formulate the optimization problem for the transmit precoding and receive equalization filters as the minimization of the sum MSE and the maximum MSE across the different hops, respectively.

5.3.1 Problem Formulation

Consider the conditional MSE matrix at node i, which is defined as the auto-correlation matrix of the error vector, i.e.,

$$\mathbf{E}_{t,(i+1)}(\Delta\mathbf{H}_{t,i}, \Delta\mathbf{L}_{t,(i+1)}, \Delta\mathbf{G}_{t,(i+2)}) = \mathbb{E}_{\{\mathbf{s}^{[n-i+1]}, \mathbf{n}_{(i+1)}\}}[\mathbf{e}_{(i+1)}\mathbf{e}_{(i+1)}^{\dagger}], \tag{5.6}$$

where the expectation is carried out with respect to the signal and the noise. Using (5.1)-(5.5), the conditional MSE matrix can be written as

$$\mathbf{E}_{t,(i+1)}(\Delta\mathbf{H}_{t,i}, \Delta\mathbf{L}_{t,(i+1)}, \Delta\mathbf{G}_{t,(i+2)}) = \mathbf{F}_{N_c}^{\dagger}\mathbf{E}_{f,(i+1)}(\Delta\mathbf{H}_{f,i}, \Delta\mathbf{L}_{f,(i+1)}, \Delta\mathbf{G}_{f,(i+2)})\mathbf{F}_{N_c}, \tag{5.7}$$

where $i = 1, \ldots, M$, and $\mathbf{E}_{f,(i+1)}(\Delta\mathbf{H}_{f,i}, \Delta\mathbf{L}_{f,(i+1)}, \Delta\mathbf{G}_{f,(i+2)})$ is the conditional FD MSE matrix, which can be written as $\mathbf{E}_{f,(i+1)}(\Delta\mathbf{H}_{f,i}, \Delta\mathbf{L}_{f,(i+1)}, \Delta\mathbf{G}_{f,(i+2)}) =$

$$\begin{aligned}
&\mathbf{W}_{f,(i+1)}\mathbf{H}_{f,i}\mathbf{P}_{f,i}\mathbf{P}_{f,i}^{\dagger}\mathbf{H}_{f,i}^{\dagger}\mathbf{W}_{f,(i+1)}^{\dagger} - \mathbf{W}_{f,(i+1)}\mathbf{H}_{f,i}\mathbf{P}_{f,i} - \mathbf{P}_{f,i}^{\dagger}\mathbf{H}_{f,i}^{\dagger}\mathbf{W}_{f,(i+1)}^{\dagger} + \mathbf{I}_{N_c} \\
&+ \mathbf{W}_{f,(i+1)}\Big(\Delta\mathbf{L}_{f,(i+1)}\mathbf{P}_{f,(i+1)}\mathbf{P}_{f,(i+1)}^{\dagger}\Delta\mathbf{L}_{f,(i+1)}^{\dagger} + \Delta\mathbf{G}_{f,(i+2)}\mathbf{P}_{f,(i+2)}\mathbf{P}_{f,(i+2)}^{\dagger}\Delta\mathbf{G}_{f,(i+2)}^{\dagger} \\
&+ \sigma_{n_{(i+1)}}^2\mathbf{I}_{N_c}\Big)\mathbf{W}_{f,(i+1)}^{\dagger},
\end{aligned} \tag{5.8}$$

where $\Delta\mathbf{L}_{f,i} = \Delta\mathbf{G}_{f,i} = \mathbf{P}_{f,i} = \mathbf{0}_{N_c\times N_c}, \forall i \geq M+1$. Here, the arguments of $\mathbf{E}_{t,(i+1)}(\Delta\mathbf{H}_{t,i}, \Delta\mathbf{L}_{t,(i+1)}, \Delta\mathbf{G}_{t,(i+2)})$ and $\mathbf{E}_{f,(i+1)}(\Delta\mathbf{H}_{f,i}, \Delta\mathbf{L}_{f,(i+1)}, \Delta\mathbf{G}_{f,(i+2)})$ stress that these MSE matrices are conditioned on the TD and FD channel estimation error matrices, respectively. By taking the average of $\mathbf{E}_{t,(i+1)}(\Delta\mathbf{H}_{t,i}, \Delta\mathbf{L}_{t,(i+1)}, \Delta\mathbf{G}_{t,(i+2)})$ with respect to the CSI error matrices, we obtain the unconditional MSE matrices as

$$\hat{\mathbf{E}}_{t,(i+1)} = \mathbf{F}_{N_c}^{\dagger}\hat{\mathbf{E}}_{f,(i+1)}\mathbf{F}_{N_c}, \quad i = 1, \ldots, M, \tag{5.9}$$

where $\hat{\mathbf{E}}_{f,(i+1)}$ is the unconditional FD MSE matrix, which can be written as

$$\begin{aligned}\hat{\mathbf{E}}_{f,(i+1)} =&\mathbb{E}_{\{\Delta\mathbf{H}_{f,i},\Delta\mathbf{L}_{f,(i+1)},\Delta\mathbf{G}_{f,(i+2)}\}}\left[\mathbf{E}_{f,(i+1)}(\Delta\mathbf{H}_{f,i},\Delta\mathbf{L}_{f,(i+1)},\Delta\mathbf{G}_{f,(i+2)})\right]\\ =&\mathbf{W}_{f,(i+1)}\Big(\mathbf{P}_{f,i}[\hat{\mathbf{H}}_{f,i}\hat{\mathbf{H}}_{f,i}^{\dagger}+\sigma_{e,h_i}^2\mathbf{I}_{N_c}]\mathbf{P}_{f,i}^{\dagger}+\sigma_{e,l_{(i+1)}}^2\mathbf{P}_{f,(i+1)}\mathbf{P}_{f,(i+1)}^{\dagger}+\sigma_{n_{(i+1)}}^2\mathbf{I}_{N_c}\\ &+\sigma_{e,g_{(i+2)}}^2\mathbf{P}_{f,(i+2)}\mathbf{P}_{f,(i+2)}^{\dagger}\Big)\mathbf{W}_{f,(i+1)}^{\dagger}-\mathbf{W}_{f,(i+1)}\hat{\mathbf{H}}_{f,i}\mathbf{P}_{f,i}-\mathbf{P}_{f,i}^{\dagger}\hat{\mathbf{H}}_{f,i}^{\dagger}\mathbf{W}_{f,(i+1)}^{\dagger}+\mathbf{I}_{N_c}.\end{aligned} \tag{5.10}$$

Note that $\hat{\mathbf{E}}_{t,(i+1)}$ is a circulant matrix, cf. (5.9). Therefore, it has identical diagonal entries given by $[\hat{\mathbf{E}}_{t,(i+1)}]_{kk}=\frac{1}{N_c}\mathrm{tr}\left[\hat{\mathbf{E}}_{f,(i+1)}\right],\forall k$, which represent the MSE of the TD symbols. Assuming separate transmit node power constraints, the problems of sum MSE and maximum MSE minimization can be formulated as

$$\min_{\{\mathbf{W}_{f,(i+1)},\mathbf{P}_{f,i}\}} f_{\mathrm{X}}(\mathbf{W},\mathbf{P}) \qquad \text{s.t.}\ \ \mathrm{tr}\left(\mathbb{E}[\mathbf{x}_i\mathbf{x}_i^{\dagger}]\right)\le\bar{P}_i,\ \ i=1,\ldots,M, \tag{5.11}$$

where $\mathbf{W}=\{\mathbf{W}_{f,(i+1)},\forall i\}$, $\mathbf{P}=\{\mathbf{P}_{f,i},\forall i\}$, $\bar{P}_i$ is the transmit power limit for node i, and

$$f_{\mathrm{sumMSE}}(\mathbf{W},\mathbf{P})=\sum_{i=1}^{M}\mathrm{tr}\left[\hat{\mathbf{E}}_{f,(i+1)}\right],\ \ f_{\mathrm{maxMSE}}(\mathbf{W},\mathbf{P})=\max_{i=1,\ldots,M}\mathrm{tr}\left[\hat{\mathbf{E}}_{f,(i+1)}\right]. \tag{5.12}$$

5.3.2 Optimal Equalization Filters and Power Allocation Problems

The optimization of the equalization matrices, $\{\mathbf{W}_{f,(i+1)},\forall i\}$, in (5.11) is independent of the power constraint and decoupled among different nodes. Therefore, we can first obtain the optimal equalization matrices by separately minimizing $\mathrm{tr}\left[\hat{\mathbf{E}}_{f,(i+1)}\right]$ with respect to $\mathbf{W}_{f,(i+1)}$. By taking the derivative of $\mathrm{tr}\left[\hat{\mathbf{E}}_{f,(i+1)}\right]$ with respect to $\mathbf{W}_{f,(i+1)}$

and setting the result to zero, the optimal FDE filter at node i is obtained as

$$\mathbf{W}^{\star}_{f,(i+1)} = \mathbf{P}^{\dagger}_{f,i}\hat{\mathbf{H}}^{\dagger}_{f,i}\left(\hat{\mathbf{H}}_{f,i}\mathbf{P}_{f,i}\mathbf{P}^{\dagger}_{f,i}\hat{\mathbf{H}}^{\dagger}_{f,i} + \mathbf{K}_{f,(i+1)}\right)^{-1}, \tag{5.13}$$

where $\mathbf{K}_{f,(i+1)}$ is the covariance matrix of the end-to-end noise given by $\mathbf{K}_{f,(i+1)} = \sigma^2_{e,h_i}\mathbf{P}_{f,i}\mathbf{P}^{\dagger}_{f,i} + \sigma^2_{e,l_{(i+1)}}\mathbf{P}_{f,(i+1)}\mathbf{P}^{\dagger}_{f,(i+1)} + \sigma^2_{e,g_{(i+2)}}\mathbf{P}_{f,(i+2)}\mathbf{P}^{\dagger}_{f,(i+2)} + \sigma^2_{n_{(i+1)}}\mathbf{I}_{N_c}$. The optimal FDE filter in (5.13) has the form of a Wiener filter that takes the CSI estimation errors into account. Substituting $\mathbf{W}^{\star}_{f,(i+1)}$ into (5.10) and applying the matrix inversion lemma [100], the optimized FD MSE matrix at node i can be written as

$$\mathbf{\Psi}_{f,(i+1)} = \left(\hat{\mathbf{H}}_{f,i}\mathbf{P}_{f,i}\mathbf{K}^{-1}_{f,(i+1)}\mathbf{P}^{\dagger}_{f,i}\hat{\mathbf{H}}^{\dagger}_{f,i} + \mathbf{I}_{N_c}\right)^{-1}. \tag{5.14}$$

The diagonal entries of $\mathbf{W}^{\star}_{f,i}$, $w_{k,i} = [\mathbf{W}^{\star}_{f,i}]_{kk}$, and $\mathbf{\Psi}_{f,i}$, $\Psi_{k,i} = [\mathbf{\Psi}_{f,i}]_{kk}$, represent the equalization coefficient and the MSE at node i on frequency tone k, respectively. These entries are given by

$$\begin{aligned} w_{k,(i+1)} &= \frac{\hat{h}^{*}_{k,i}p^{*}_{k,i}}{(|\hat{h}_{k,i}|^2 + \sigma^2_{e,h_i})|p_{k,i}|^2 + \sigma^2_{e,l_{(i+1)}}|p_{k,(i+1)}|^2 + \sigma^2_{e,g_{(i+2)}}|p_{k,(i+2)}|^2 + \sigma^2_{n_{(i+1)}}}, \\ \Psi_{k,(i+1)} &= \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma^2_{e,h_i}P_{k,i} + \sigma^2_{e,l_{(i+1)}}P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}}P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}} + 1\right)^{-1}, \end{aligned} \tag{5.15}$$

where $i = 1, \ldots, M$, and $P_{k,i} = |p_{k,i}|^2$ is the transmit power allocated to the kth frequency tone at node i. With the FD MSE given in (5.15), and by rewriting the power constraint in scalar form, (5.11) reduces to the following transmit power allocation problem

$$\min_{\{P_{k,i}\}} \; f_{\mathrm{X}}(\mathbf{p}) \quad \text{s.t.} \quad \sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \quad P_{k,i} \geq 0, \quad i = 1, \ldots, M, \forall k, \tag{5.16}$$

where the constraint $P_{k,i} \geq 0$ reflects the fact that power cannot be negative, and

$$f_{\text{sumMSE}}(\mathbf{p}) = \sum_{i=1}^{M}\sum_{k=1}^{N_c} \Psi_{k,(i+1)}, \quad f_{\text{maxMSE}}(\mathbf{p}) = \max_{i=1,\ldots,M} \sum_{k=1}^{N_c} \Psi_{k,(i+1)}, \tag{5.17}$$

with $\mathbf{p} = \{P_{k,i}, \forall k, i\}$. We note that once we have obtained the optimal $P_{k,i}$, the corresponding precoding coefficient $p_{k,i}$ can be recovered by taking the square root of $P_{k,i}$ and multiplying the result with an arbitrary phase term. However, solving problem (5.16) directly is very challenging as the power variables of different hops are coupled with each other, which makes the objective function a highly non-convex function with respect to $\{P_{k,i}\}$.

5.4 Solutions to the Power Allocation Problems

In this section, we propose two different approaches to handle the difficult non-convex problem in (5.16). The first method is a centralized scheme which employs the sequential GP algorithm to iteratively find improved approximate solutions to (5.16). In the second approach, we propose an alternating optimization based method to find a stationary point solution to (5.16) by solving convex programming and difference of convex programming problems in an alternating manner.

5.4.1 A Sequential Geometric Programming (sGP) Approach

We first consider solving problem (5.16) by using a GP approach. Since the objective function is not a posynomial, the standard GP algorithm can not be applied. In a first attempt, we handle this difficulty by ignoring the constant 1 in the expression

for $\Psi_{k,(i+1)}$ in (5.15), which leads to the following approximation

$$\begin{aligned}\Psi_{k,(i+1)} &\approx \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma^2_{e,h_i} P_{k,i} + \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}}\right)^{-1} \\ &= |\hat{h}_{k,i}|^{-2} \left(\sigma^2_{e,h_i} + \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} P^{-1}_{k,i} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} P^{-1}_{k,i} + \sigma^2_{n_{(i+1)}} P^{-1}_{k,i}\right).\end{aligned} \tag{5.18}$$

By adopting (5.18) in (5.17), the objective function becomes a posynomial and we can solve the resulting GP problem using efficient convex optimization tools, such as CVX [101]. We refer to this GP approach as GP-I method. It is noted that the GP-I method leads to the minimization of a very loose upper bound for the original problem as the approximation in (5.18) is tight only when $\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma^2_{e,h_i} P_{k,i} + \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}} \gg 1$, i.e., for high SNR and negligible CSI errors. To also achieve a high performance for low-to-medium SNR and non-zero CSI errors, we invoke the sequential GP (sGP) approach [134]. To this end, we rewrite $\Psi_{k,(i+1)}$ in (5.15) as

$$\Psi_{k,(i+1)} = \frac{\sigma^2_{e,h_i} P_{k,i} + \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}}{(|\hat{h}_{k,i}|^2 + \sigma^2_{e,h_i}) P_{k,i} + \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}} = \frac{\rho_{1,k,i}(\mathbf{p})}{\rho_{2,k,i}(\mathbf{p})}, \tag{5.19}$$

from which we can observe that $\Psi_{k,(i+1)}$ is a ratio of two posynomials. In order to transform $\Psi_{k,(i+1)}$ into a posynomial, we apply the arithmetic-geometric mean inequality to lower bound the denominator, $\rho_{2,k,i}(\mathbf{p})$, with a monomial, i.e.,

$$\begin{aligned}&\underbrace{(|\hat{h}_{k,i}|^2 + \sigma^2_{e,h_i}) P_{k,i}}_{\mu_{1,k,i}(\mathbf{p})} + \underbrace{\sigma^2_{e,l_{(i+1)}} P_{k,(i+1)}}_{\mu_{2,k,i}(\mathbf{p})} + \underbrace{\sigma^2_{e,g_{(i+2)}} P_{k,(i+2)}}_{\mu_{3,k,i}(\mathbf{p})} + \underbrace{\sigma^2_{n_{(i+1)}}}_{\mu_{4,k,i}(\mathbf{p})} \\ &= \sum_{z=1}^{4} \mu_{z,k,i}(\mathbf{p}) \geq \prod_{z=1}^{4} \left(\frac{\mu_{z,k,i}(\mathbf{p})}{\alpha_{z,k,i}(\hat{\mathbf{p}})}\right)^{\alpha_{z,k,i}(\hat{\mathbf{p}})} = \hat{\rho}_{2,k,i}(\mathbf{p}),\end{aligned} \tag{5.20}$$

where $\mu_{1,k,i}(\mathbf{p}) = (|\hat{h}_{k,i}|^2+\sigma^2_{e,h_i})P_{k,i}$, $\mu_{2,k,i}(\mathbf{p}) = \sigma^2_{e,l_{(i+1)}}P_{k,(i+1)}$, $\mu_{3,k,i}(\mathbf{p}) = \sigma^2_{e,g_{(i+2)}}P_{k,(i+2)}$, and $\mu_{4,k,i}(\mathbf{p}) = \sigma^2_{n_{(i+1)}}$ are the monomial terms in the denominator, and $\alpha_{z,k,i}(\hat{\mathbf{p}}) = \frac{\mu_{z,k,i}(\hat{\mathbf{p}})}{\sum_{z=1}^{4}\mu_{z,k,i}(\hat{\mathbf{p}})}, \forall z$, are the weighting factors for these monomial terms, with $\hat{\mathbf{p}}$ being a feasible point of the original problem. Eq. (5.20) is known as the single condensation technique in sGP [134]. Due to the inequality in (5.20), replacing $\rho_{2,k,i}(\mathbf{p})$ with $\hat{\rho}_{2,k,i}(\mathbf{p})$ will lead to the minimization of an upper bound for the original problem (5.16). The resulting new problem is a standard GP and can be solved efficiently. Since this GP problem has the same feasible set as problem (5.16), its optimal solution, $\mathbf{p}^{[\mathrm{gp}]}$, is also feasible for (5.16). Therefore, by condensing $\rho_{2,k,i}(\mathbf{p})$ with the newly obtained $\mathbf{p}^{[\mathrm{gp}]}$ and solving the resulting GP problem again, the objective value can be further decreased and an improved approximate solution to the original problem can be found. This process continues until the objective value of the GP problem converges. Specifically, in the ℓth iteration of the sGP, we have to solve the following GP problem

$$\min_{\{P_{k,i}\}} \quad f_{\mathrm{X}}^{[\ell]}(\mathbf{p}) \quad \text{s.t.} \quad \sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \quad P_{k,i} \geq 0, \forall k,\ i = 1, \ldots, M, \tag{5.21}$$

where

$$f^{[\ell]}_{\mathrm{sumMSE}}(\mathbf{p}) = \sum_{i=1}^{M}\sum_{k=1}^{N_c}\frac{\rho_{1,k,i}(\mathbf{p})}{\rho^{[\ell]}_{2,k,i}(\mathbf{p})}, \quad f^{[\ell]}_{\mathrm{maxMSE}}(\mathbf{p}) = \max_{i=1,\ldots,M}\sum_{k=1}^{N_c}\frac{\rho_{1,k,i}(\mathbf{p})}{\rho^{[\ell]}_{2,k,i}(\mathbf{p})}, \tag{5.22}$$

and $\rho^{[\ell]}_{2,k,i}(\mathbf{p}) = \prod_{z=1}^{4}\left(\frac{\mu_z(\mathbf{p})}{\alpha_{z,k,i}(\mathbf{p}^{[\ell]})}\right)^{\alpha_{z,k,i}(\mathbf{p}^{[\ell]})}$. In Table 5.1, we summarize the proposed sGP algorithm.

Remark 5.1: The sGP algorithm is guaranteed to converge to a stationary point (KKT point) of the original non-convex problem since it satisfies the conditions for the convergence of the successive approximation method [101]. Since the GP problems in

Table 5.1: Algorithm 5.1 for optimization of $\{P_{k,i}, \forall k, i\}$ based on sequential GP. ϵ_1 is a small constant which controls the accuracy of the algorithm, e.g. $\epsilon_1 = 10^{-4}$.

1	**Input** $\{\|\hat{h}_{k,i}\|^2, \sigma^2_{e,h_i}, \sigma^2_{e,l_i}, \sigma^2_{e,g_i}, \forall k, i\}$. **Initialize** $\{P^{[0]}_{k,i}, \forall k, i\}$, $\text{OBJ}^{[0]}_{\text{gp}} = f^{[0]}_{\text{X}}(\mathbf{p})$, and $m = 1$.
2	**Repeat** Solve the GP problems in (5.21) to find the optimal $\{P^{[m]}_{k,i}, \forall k, i\}$ using $\{P^{[m-1]}_{k,i}, \forall k, i\}$. Update $\text{OBJ}^{[m]}_{\text{gp}} = f^{[m]}_{\text{X}}(\mathbf{p})$. If $\|\text{OBJ}^{[m]}_{\text{gp}} - \text{OBJ}^{[m-1]}_{\text{gp}}\| < \epsilon_1$, go to step 3. Otherwise, set $m = m + 1$.
3	**Output** $\{P^{[m]}_{k,i}, \forall k, i\}$

each iteration involve $(M-1)N_c$ optimization variables, the overall complexity of the sGP algorithm is $\mathcal{O}\left(I(M-1)^3 N_c^3\right)$, where I is the number of iterations required for the sGP algorithm to converge[22]. Note that the central node solving the GP problem in (5.21) has to know the CSI of all hops, which implies that the sGP based power allocation scheme is a centralized scheme requiring significant information exchange between the central node and the transmitting nodes.

5.4.2 An Alternating Optimization (AO) Approach

The sGP based approach presented in the last section is a centralized scheme which requires the availability of global CSI at a central node. Moreover, as each subproblem of the sGP algorithm can be solved only numerically, it is difficult to get any physical insights regarding the problem structure. In this section, motivated by the observation that with the powers of all other nodes being fixed, the power optimization of a specific node can be formulated as either a convex problem or difference of convex problems, we propose an AO based approach to solve the transmit power allocation problem. It will be shown that by exploiting the (partial) convexity of the subproblems, analytical power allocation solutions can be obtained, which finally leads to an efficient algorithm with reduced signaling overhead compared to the sGP

[22]Solving an n-dimensional GP problem with interior point methods requires the calculation of the inverse of the Hessian matrix, which entails a complexity of $\mathcal{O}\left(n^3\right)$ [101].

approach.

Sum MSE Minimization

We first consider the sum MSE minimization problem in (5.16), which can be explicitly stated as

$$\min_{\{P_{k,i}\}} \quad \sum_{i=1}^{M-1}\sum_{k=1}^{N_c}\left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + \sigma_{e,l_{(i+1)}}^2 P_{k,(i+1)} + \sigma_{e,g_{(i+2)}}^2 P_{k,(i+2)} + \sigma_{n_{(i+1)}}^2} + 1\right)^{-1} \tag{5.23a}$$

$$\text{s.t.} \quad \sum_{k=1}^{N_c} P_{k,i} \le \bar{P}_i, \quad P_{k,i} \ge 0, \quad \forall k, i = 1, \ldots, M. \tag{5.23b}$$

Proposition 5.1. *For given $\{P_{k,i}, \forall k\}, i \neq 1$, the problem in (5.23) with respect to the source transmit power, $\{P_{k,1}, \forall k\}$, is a convex optimization problem, and the optimal solution is given by*

$$P_{k,1} = \frac{\sigma_{e,l_2}^2 P_{k,2} + \sigma_{e,g_3}^2 P_{k,3} + \sigma_{n_2}^2}{|\hat{h}_{k,1}|^2 + \sigma_{e,h_1}^2}\left(\sqrt{\frac{|\hat{h}_{k,1}|^2}{\lambda_1(\sigma_{e,l_2}^2 P_{k,2} + \sigma_{e,g_3}^2 P_{k,3} + \sigma_{n_2}^2)}} - 1\right)^+, \tag{5.24}$$

where λ_1 is the Lagrange multiplier chosen to satisfy the transmit power constraint with equality, i.e., $\sum_k P_{k,1} = \bar{P}_1$. In practice, a simple bisection procedure can be used to iteratively identify the optimal λ_1.

Proof. Please refer to Appendix-F. □

Remark 5.2: The above solution for $\{P_{k,1}\}$ indicates that regardless of the power level at the relay, the source node always transmits with the maximum possible power. This is intuitive because the objective function is an increasing function with respect to the source transmit powers, and since the transmission of the source node does not cause any interference, it is optimal to fully utilize the available power budget.

Proposition 5.2. *For given* $\{P_{k,j}, \forall k\}, j \neq i$, *the problem in (5.23) can be transformed as a difference of convex (d.c.) programming problem with respect to* $\{P_{k,i}\}, i = 2, \ldots, M$, *as follows*

$$\begin{aligned} \min_{\{P_{k,i}\}} \quad & g_i(\mathbf{p}_i) - \underbrace{\sum_{u=1}^{2} \psi_{i,u}(\mathbf{p}_i)}_{f_i(\mathbf{p}_i)} \\ \text{s.t.} \quad & \sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \; P_{k,i} \geq 0, \quad i = 2, \ldots, M, \forall k, \end{aligned} \tag{5.25}$$

where

$$g_i(\mathbf{p}_i) = \sum_{k=1}^{N_c} \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma^2_{e,h_i} P_{k,i} + T_{k,i}} + 1 \right)^{-1}, \quad \psi_{i,u}(\mathbf{p}_i) = \sum_{k=1}^{N_c} \left(\frac{B_{k,i,u}}{A_{k,i,u} P_{k,i} + C_{k,i,u}} - 1 \right), \tag{5.26}$$

with $T_{k,i} = \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}$, $A_{k,i,1} = \sigma^2_{e,l_i}$, $A_{k,i,2} = \sigma^2_{e,g_i}$, $B_{k,i,1} = |\hat{h}_{k,(i-1)}|^2 P_{k,(i-1)}, B_{k,i,2} = |\hat{h}_{k,(i-2)}|^2 P_{k,(i-2)}$, $C_{k,i,1} = (|\hat{h}_{k,(i-1)}|^2 + \sigma^2_{e,h_{(i-1)}}) P_{k,(i-1)} + \sigma^2_{e,g_{(i+1)}} P_{k,(i+1)} + \sigma^2_{n_i}$, $C_{k,i,2} = (|\hat{h}_{k,(i-2)}|^2 + \sigma^2_{e,h_{(i-2)}}) P_{k,(i-2)} + \sigma^2_{e,l_{(i-1)}} P_{k,(i-1)} + \sigma^2_{n_{(i-1)}}$, *and* $\mathbf{p}_i = \{P_{k,i}, \forall k\}$.

Proof. Please refer to Appendix-G. □

There are many global optimum-achieving techniques for solving d.c. problems in the optimization literature [136]. However, these techniques are based on the branch-and-bound procedure, whose computational complexity becomes prohibitively high when dealing with large dimensional problems. As the dimension of the considered problem is determined by the number of frequency tones, which can be as large as 2048 in broadband systems like LTE-A [14, 16] finding the global optimum does not seem feasible in practice. Furthermore, since problem (5.25) is only a subproblem of

the proposed AO scheme, a low-complexity solution is desirable in order to keep the overall complexity of the algorithm low. Thus, in the following, we aim at deriving a low-complexity algorithm which is guaranteed to attain a stationary point (KKT point) of problem (5.25).

Proposition 5.3. *For given $\{P_{k,j}, \forall k\}, j \neq i$, a KKT point of the d.c. programming problem (5.25) with respect to $\{P_{k,i}, \forall k\}, i = 2, \ldots, M$, can be obtained by sequentially solving the following convex optimization problem,*

$$\begin{aligned} \min_{\{P_{k,i}\}} \quad & \hat{h}_i^{[\ell]}(\mathbf{p}_i) \triangleq g_i(\mathbf{p}_i) + \sum_{u=1}^{2} \sum_{k=1}^{N_c} D_{k,i,u}^{[\ell]} P_{k,i} \\ \text{s.t.} \quad & \sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \;\; P_{k,i} \geq 0, \quad \forall k, \end{aligned} \tag{5.27}$$

where $D_{k,i,u}^{[\ell]} = \frac{A_{k,i,u} B_{k,i,u}}{(A_{k,i,u} P_{k,i}^{[\ell]} + C_{k,i,u})^2}$, and $P_{k,i}^{[\ell]}$ is the solution of problem (5.27) for a given $P_{k,i}^{[\ell-1]}$. The global optimal solution of convex problem (5.27) is given by

$$P_{k,i} = \frac{T_{k,i}}{|\hat{h}_{k,i}|^2 + \sigma_{e,h_i}^2} \left(\sqrt{\frac{|\hat{h}_{k,i}|^2}{[\lambda_i + \sum_{u=1}^{2} D_{k,i,u}^{[\ell]}] T_{k,i}}} - 1 \right)^+, \tag{5.28}$$

where the optimal Lagrange multiplier λ_i can be obtained by using the subgradient method

$$\lambda_i^{[\kappa+1]} = \left[\lambda_i^{[\kappa]} + \varepsilon_0^{[\kappa]} \left(\sum_{k=1}^{N_c} P_{k,i} - \bar{P}_i \right) \right]^+, \tag{5.29}$$

where $\varepsilon_0^{[\kappa]}$ is the step size adopted in the κth iteration of the subgradient method.

Proof. Please refer to Appendix-H. □

Remark 5.3: We can show that for the optimum solution of problem (5.27), the

relay transmit power constraint is not necessarily met with equality, i.e., the relay may not use its full power budget[23]. Consider the relevant KKT conditions for the Lagrangian of problem (5.27),

$$\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}} + \lambda_i + \sum_{u=1}^{2} D_{k,i,u}^{[\ell]} - \beta_{k,i} = 0, \tag{5.30}$$

$$\lambda_i (\sum_{k=1}^{N_c} P_{k,i} - \bar{P}_i) = 0, \tag{5.31}$$

$$\lambda_i \geq 0, \beta_{k,i} \geq 0, \tag{5.32}$$

where $\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}}$ is given in (G.2). Since $\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}} < 0$ and $\beta_{k,i} \geq 0$, in order for (5.30) to hold, we must have $\lambda_i + \sum_{u=1}^{2} D_{k,i,u}^{[\ell]} > 0$. For $\sigma_{e,l_i}^2 = 0$ and $\sigma_{e,g_{(i+1)}}^2 = 0$ (zero loopback and zero backward interference), we have $\sum_{u=1}^{2} D_{k,i,u}^{[\ell]} = 0$. In this case, λ_i should take a strictly positive value. From the complementary slackness condition (5.31), $\lambda_i > 0$ implies that $\sum_{k=1}^{N_c} P_{k,i} = \bar{P}_i$, i.e., the transmit power budget of the relay is fully utilized when there is no loopback/backward interference. On the other hand, when $\sigma_{e,l_i}^2 \neq 0$ and/or $\sigma_{e,g_{(i+1)}}^2 \neq 0$, $\sum_{u=1}^{2} D_{k,i,u}^{[\ell]}$ takes a strictly positive value. Therefore, λ_i can be zero as long as $\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}} + \sum_{u=1}^{2} D_{k,i,u}^{[\ell]} - \beta_{k,i} = 0$ holds. In this case, according to (5.31), we have $\sum_{k=1}^{N_c} P_{k,i} < \bar{P}_i$. Thus, if there is non-negligible loopback/backward interference, node i may not fully use its maximum power budget in order to limit the loopback/backward interference.

Exploiting Propositions 5.1 and 5.3, we can perform AO of $\mathbf{p}_i$ to find a KKT point solution to the original non-convex problem (5.23). Specifically, for given initial $\mathbf{p}_j^{[0]}, \forall j \neq i$, we can obtain the global optimal point $\mathbf{p}_i^{[1]}, i = 1$, by solving the convex problem (F.1) (in Appendix F), and the KKT point $\mathbf{p}_i^{[1]}, i \geq 2$ by solving the d.c. problem (5.25). Substituting the updated value of $\mathbf{p}_i^{[1]}$ into the objective function,

[23] A similar observation has been made in [80] for a different optimization criterion.

Table 5.2: Algorithm 5.2 for optimization of $\{P_{k,i}, \forall k, i\}$ based on AO. ϵ_2 is a small constant which controls the accuracy of the algorithm, e.g. $\epsilon_2 = 10^{-4}$.

1	**Input** $\{\|\hat{h}_{k,i}\|^2, \sigma^2_{e,l_i}, \sigma^2_{e,h_i}, \sigma^2_{e,g_i}, \forall i,k\}$. **Initialize** $\{\mathbf{p}_i^{[0]}, \forall i\}$, $\text{OBJ}_{\text{ao}}^{[0]} = f_{\text{X}}^{[0]}(\mathbf{p})$, and $m = 1$.
2	**Repeat** Solve convex problem (F.1) with respect to $\mathbf{p}_1^{[m]}$ using $\mathbf{p}_i^{[m-1]}, \forall i \neq 1$. **For** $i = 2 : M$ Solve d.c. problem (5.25) or (5.34) with respect to $\mathbf{p}_i^{[m]}$ using $\mathbf{p}_j^{[m-1]}, \forall j \neq i$. **End** Update $\text{OBJ}_{\text{ao}}^{[m]} = f_{\text{X}}^{[m]}(\mathbf{p})$. If $\|\text{OBJ}_{\text{ao}}^{[m]} - \text{OBJ}_{\text{ao}}^{[m-1]}\| < \epsilon_2$, go to step 3. Otherwise, set $m = m + 1$.
3	Output $\{\mathbf{p}_i^{[m]}, \forall i\}$.

we can conduct another round of sequential optimization to obtain $\mathbf{p}_i^{[2]}$. This process continues until the value of the objective function converges, which is guaranteed since the objective function is lower bounded by zero, and the update of $\{\mathbf{p}_i\}$ in each iteration monotonically decreases or maintains the objective value. Furthermore, since the solutions for $\{\mathbf{p}_i\}$ in each iteration are the KKT points for convex problem (F.1) and d.c. problem (5.25), respectively, it can be shown that the final solution of the AO algorithm is also a KKT point of the original non-convex problem (5.23). The AO algorithm is summarized in Table 5.2[24].

Maximum MSE Minimization

We now proceed to solve the maximum MSE minimization problem in (5.16). By introducing an auxiliary optimization variable t, the problem can be written in an

[24]The algorithm is unified for both sum MSE and maximum MSE minimization.

epigraph form as follows

$$\min_{\{\{P_{k,i}\},t\}} \quad t \tag{5.33a}$$

$$\text{s.t.} \quad \sum_{k=1}^{N_c} \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma^2_{e,h_i} P_{k,i} + \sigma^2_{e,l_{(i+1)}} P_{k,(i+1)} + \sigma^2_{e,g_{(i+2)}} P_{k,(i+2)} + \sigma^2_{n_{(i+1)}}} + 1 \right)^{-1} \leq t, \tag{5.33b}$$

$$\sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \quad P_{k,i} \geq 0, \quad , i = 1, \ldots, M, k = 1, \ldots, N_c, \tag{5.33c}$$

which is a non-convex problem due to the non-convex constraint in (5.33b). Thus, finding the global optimal solution with polynomial time complexity is very difficult. Motivated by a similar observation as in the last subsection, we apply the AO method to solve also this non-convex problem. First, for fixed $\{P_{k,i}, \forall k\}, i = 2, \ldots, M-1$, the auxiliary variable in problem (5.33) becomes a dummy variable and thus can be eliminated. Consequently, for the first hop, the problem becomes equivalent to the sum MSE minimization, and the solution of $P_{k,1}$ is given by (5.24). Next, for given $\{P_{k,j}, \forall k\}, j \neq i, i \geq 2$, the non-convex constraint (5.33b) becomes a reverse-convex constraint with respect to $\{P_{k,i}, \forall k\}$. Hence, problem (5.33) with variable $\{P_{k,i}, \forall k\}$ for optimization can be written as

$$\min_{\{\{P_{k,i}\},t\}} \quad t \quad \text{s.t.} \quad g_i(\mathbf{p}_i) \leq t \quad -\psi_{i,1}(\mathbf{p}_i) \leq t, \quad -\psi_{i,2}(\mathbf{p}_i) \leq t, \tag{5.34a}$$

$$\sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \quad P_{k,i} \geq 0, \forall k, \tag{5.34b}$$

where $g_i(\mathbf{p}_i)$ and $\{\psi_{i,u}(\mathbf{p}_i), u = 1, 2\}$ are given in (5.26). It is well known that a convex minimization problem with additional reverse-convex constraints is essentially a special case of the d.c. programming problem [136]. Therefore, following a similar

approach as in the last subsection, we have the following result.

Proposition 5.4. *For given $\{P_{k,j}, \forall k\}, j \neq i$, a KKT point of the d.c. programming problem (5.34a) with respect to $\{P_{k,i}\}$ can be obtained by sequentially solving the following convex optimization problem*

$$\min_{\{\{P_{k,i}\},t\}} \quad t \tag{5.35a}$$

$$\text{s.t.} \quad \sum_{k=1}^{N_c} \left(E_{k,i,u}^{[\ell]} + D_{k,i,u}^{[\ell]} (P_{k,i} - P_{k,i}^{[\ell]}) \right) \leq t, \quad u = 1, 2, \tag{5.35b}$$

$$\sum_{k=1}^{N_c} \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1 \right)^{-1} \leq t, \tag{5.35c}$$

$$\sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \quad P_{k,i} \geq 0, \forall k, \tag{5.35d}$$

where $E_{k,i,u}^{(\ell)} = 1 - \frac{B_{k,i,u}}{A_{k,i,u} P_{k,i}^{[\ell]} + C_{k,i,u}}$, and $P_{k,i}^{[\ell]}$ is the solution of problem (5.35) for a given $P_{k,i}^{[\ell-1]}$. Furthermore, the solution of $\{P_{k,i}\}$ is given by

$$P_{k,i} = \frac{T_{k,i}}{|\hat{h}_{k,i}|^2 + \sigma_{e,h_i}^2} \left(\sqrt{\frac{(1 - \sum_{u=1}^{2} \gamma_u) |\hat{h}_{k,i}|^2}{[\lambda_i + \sum_{u=1}^{2} \gamma_u D_{k,i,u}^{[\ell]}] T_{k,i}}} - 1 \right)^+, \tag{5.36}$$

where the Lagrange multipliers λ_i and γ_u are obtained from (5.29) and

$$\gamma_u^{[\kappa+1]} = \left[\gamma_u^{[\kappa]} - \varepsilon_u^{[\kappa]} \left(\sum_{k=1}^{N_c} D_{k,i,u}^{[\ell]} (P_{k,i} - P_{k,i}^{(\ell)}) + E_{k,i,u}^{[\ell]} - \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1 \right)^{-1} \right) \right]^+, \tag{5.37}$$

respectively, where $\varepsilon_u^{[\kappa]}$ is the step size parameter for γ_u adopted in the κth iteration of the subgradient method.

Proof. Please refer to Appendix-I. □

Corollory 5.1. *After the convergence of the AO algorithm for solving problem (5.33), the MSEs of different nodes are not necessarily balanced. In addition, the MSE of the ith hop is always larger than or equal to that of the $(i-1)$th hop.*

Proof. Please refer to Appendix-J. □

Remark 5.4: Corollary 5.1 indicates that the MSE of the different hops may not be identical after convergence. This happens when the $(i-1)$th hop has a much better channel quality than the ith hop, which makes it impossible to reduce the MSE of the ith hop to the level of the $(i-1)$th hop even when node i fully utilizes its power budget for the power allocation. Moreover, the MSE of the ith hop is always no less than that of the $(i-1)$th hop. This is intuitive because node i can control the loopback/backward interference to the previous hops through power allocation, and we can always scale down its transmit power to increase the MSE of the ith hop while decreasing the MSEs of the $(i-1)$th hop, if this reduces the maximum MSE value.

5.4.3 Complexity and Signalling Overhead for AO

For both considered objective functions, the source power allocation problem is convex, while the relay power allocation problems are d.c. problems. The overall complexity of the AO algorithm is thus dependent on the number of inner iterations of the d.c. problem and the number of outer iterations of the AO algorithm. Thereby, we assume that the ith outer iteration is executed at node i. Given that the complexity of solving the convex problems in (5.27), (F.1), and (5.35) is $\mathcal{O}(N_c^2)$ [101], the complexity of the AO algorithm is $\mathcal{O}\left((J_1\sum_{i=1}^{M-1}K_i + J_2)N_c^2\right)$, where J_1 and J_2 are the numbers of d.c. problems and convex problems that have to be solved in the outer iterations, respectively, and K_i is the number of inner iterations needed

for solving the d.c. problem for the ith relay[25]. The signalling overhead required at the different nodes for executing the AO algorithm is summarized in Table 5.3. Note that node i only needs to solve a local optimization problem with respect to $\{P_{k,i}, \forall k\}$ using the local estimated CSI and the CSI and power feedback from the nodes that are one and two hops away. Specifically, due to the coupling of the transmissions in adjacent hops, knowledge of the power allocation vectors at node $(i-1)$ and node $(i+1)$ is required at node i. Moreover, for FDR systems with more than two hops, due to the backward interference from the subsequent relay to the previous relay, the transmit powers of node $(i+2)$ and node $(i-2)$ have also to be known by node i. Therefore, for a moderate number of hops, e.g., three hops, the feedback overhead of the AO algorithm is not significantly less than that of the sGP scheme. However, for a large number of hops, e.g., twenty hops, the reduction in the signalling and CSI overhead becomes prominent. Furthermore, if the relays can employ directional antennas which make the backward interference negligible, the feedback of the powers and CSI from node $(i+2)$ and node $(i-2)$ can be avoided. In this case, the information exchange required for the AO algorithm is limited to neighboring nodes, which further facilitates a semi-distributed implementation of the algorithm.

5.5 Simulation Results

In this section, we provide simulation results for the considered FDR system using the proposed algorithms. The input data block length is $N_c = 64$. The received signal-to-noise ratios at node i is defined as $\text{SNR}_i = \frac{\bar{P}_{(i-1)}}{N_c \sigma_{n_i}^2}, i = 2, \ldots, M+1$. For the simulation results, we adopt the reference SNR as $\text{SNR}_{\text{ref}} = \text{SNR}_i, \forall i$, unless specified

[25]From extensive experiments, we have found that $K_i = 2, \forall i$ is sufficient for the considered d.c. algorithm to converge.

Table 5.3: Required signalling overhead at different nodes for execution of the AO algorithm.

Node	Required CSI for power allocation
1	$\{\lvert\hat{h}_{k,1}\rvert^2, P_{k,2}, P_{k,3}, \forall k\}, \sigma^2_{e,h_1}, \sigma^2_{e,l_2}, \sigma^2_{e,g_3}$
$\vdots$	$\vdots$
i	$\{\lvert\hat{h}_{k,(i-2)}\rvert^2, \lvert\hat{h}_{k,(i-1)}\rvert^2, \lvert\hat{h}_{k,i}\rvert^2, \forall k\}$, $\sigma^2_{e,h_{(i-2)}}, \sigma^2_{e,h_{(i-1)}}, \sigma^2_{e,h_i}$
$\vdots$	$\vdots$
M	$\{\lvert\hat{h}_{k,(M-2)}\rvert^2, \lvert\hat{h}_{k,(M-1)}\rvert^2, \lvert\hat{h}_{k,M}\rvert^2, \forall k\}$,
$M+1$	N/A

otherwise. All the TD channel vectors are modeled as uncorrelated Rayleigh block fading with power delay profile: $p[n] = \frac{1}{\sigma_t}\sum_{l=0}^{L-1} e^{-n/\sigma_t}\delta[n-l]$ [126], where $\sigma_t = 2$, which corresponds to moderately frequency-selective fading. For convenience, we assume that the length of all multipath channels is equal to 16. The reference CSI error variance is defined as $\sigma_e^2 = \sigma^2_{e,h_i} = \sigma^2_{e,g_i} = \sigma^2_{e,l_i}, \forall i$.

We also propose two robust baseline schemes, namely, naive power allocation (PA) and equal PA (EPA), which both employ the robust equalizers in (5.13). For naive PA, different nodes optimize their own MSE performance by taking into account the CSI errors of the transmit channels, the LI channels, and the backward interference channels, but ignoring the interferences caused to other nodes. In this way, they only need to solve standard convex optimization problems which entails a lower complexity than solving d.c. problems. For EPA, the precoders distribute their transmit powers uniformly across frequency tones, and thus power optimization is not required. In addition, we consider non-robust naive PA and non-robust EPA, which both employ equalization filters and transmit precoders that are designed under the assumption that the estimated CSI is perfect.

5.5.1 Convergence of the Proposed Algorithms

In Fig. 5.2, we examine the convergence of the AO algorithm for the proposed PA scheme based on sum MSE minimization. In the upper sub-figure, a two-hop FDR system with different initial values of $\mathbf{p}_i^{[0]}$ and SNRs is considered and the CSI error is set to $\sigma_e^2 = 0.03$. As can be observed, for all investigated initial values for $\mathbf{p}_i^{[0]}$, the AO algorithm converges within two iterations to the same sum MSE value. This suggests that for two-hop FDR systems performing optimization iteratively between the source and the relay is not necessary as two iterations already lead to a near-optimal performance. In the lower sub-figure, the results for a three-hop FDR system with different CSI errors are shown and the SNR is set to $\text{SNR}_{\text{ref}} = 32$ dB. The figure reveals that for large CSI errors (e.g., $\sigma_e^2 = 0.07$) in the high SNR regime, additional iterations are beneficial to further decrease the sum MSE value. Therefore, unlike the two-hop case, for multihop systems, performing AO iteratively between different nodes may be necessary to achieve optimal performance.

In Fig. 5.3, we examine the hop-wise MSE versus the outer iteration number of the AO algorithm for a three-hop FDR system and maximum MSE minimization, where $\sigma_e^2 = 0.03$. As can be observed from the figure, after convergence, the MSE of the later hops are always no better than that of the earlier hops. In the upper sub-figure, where the SNR of the first hop is much larger than those of the later hops, we observe that the MSEs of different hops are not equal after convergence. This is because the first hop has a much better channel than the second and third hops, and it is impossible to adjust the power at the relays to balance the MSEs of the second and third hops with that of the first hop. Nevertheless, the MSEs of the second and third hops are still balanced since they have similar channel conditions due to the identical hop SNRs. In the lower sub-figure, where the SNRs of all three hops are

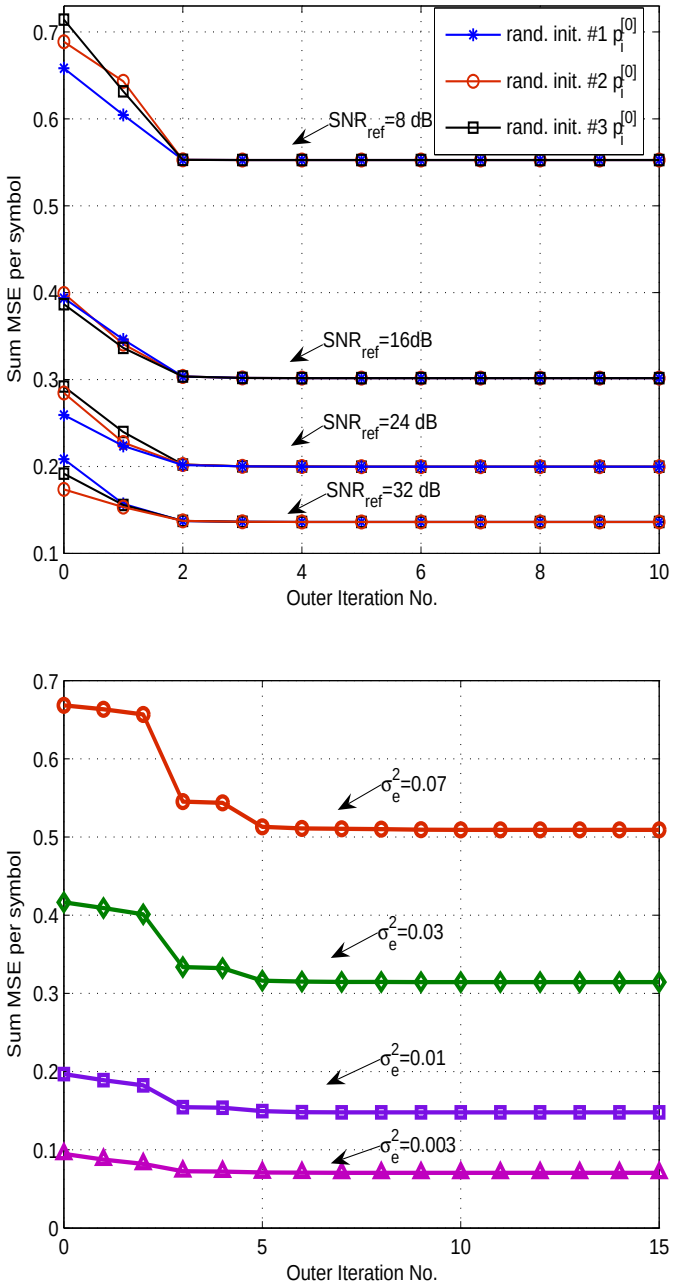

Figure 5.2: Convergence of the AO algorithm for multi-hop FDR systems employing sum MSE minimization. **Upper:** Two-hop FDR systems with different SNR values and $\sigma_e^2 = 0.03$. **Lower:** Three-hop FDR systems with different CSI errors and $\text{SNR}_{\text{ref}} = 32$ dB.

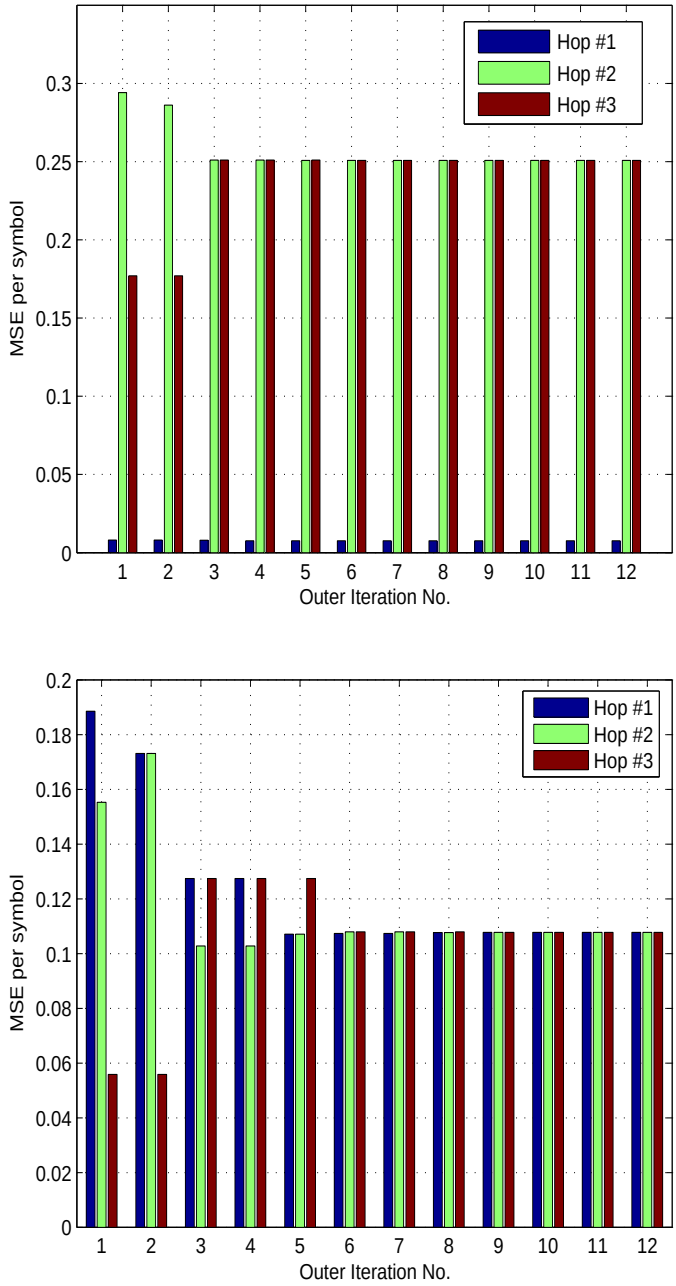

Figure 5.3: Hop-wise MSE versus outer iteration number for a three-hop FDR system employing maximum MSE minimization. **Upper:** MSE per symbol of different hops with $\mathrm{SNR}_2 = 32$ dB, $\mathrm{SNR}_3 = \mathrm{SNR}_4 = 16$ dB. **Lower:** MSE per symbol of different hops with $\mathrm{SNR}_{\mathrm{ref}} = 32$ dB.

identical, we can see the MSEs are balanced after convergence of the algorithm due to the similar channel conditions in all hops.

In Fig. 5.4, we compare the convergence of the AO and sGP algorithms. At first glance, it seems that the sGP algorithm has a faster convergence rate than the AO algorithm. However, it is worth noting that in each iteration of the sGP algorithm, we have to solve a GP problem with $(M-1)N_c$ optimization variables using numerical optimization solvers, while in each step of AO algorithm, we only need to solve a convex or a d.c. problem with N_c optimization variables using analytical subproblem solutions. Taking this into account when considering the computational complexity of the two algorithms, we can infer that the complexity of the AO algorithm is indeed lower than that of sGP algorithm. In fact, as verified by extensive simulations, the run time that the AO algorithm needs to converge is significantly lower than that of the sGP algorithm. Furthermore, from Fig. 5.4, we observe that for different SNRs and CSI estimation errors, both algorithms finally converge to the same objective value within five and eight iterations for two-hop and three-hop FDR channels, respectively. This implies that for both of these two algorithms the FDR schemes have identical performance in terms of MSE and achievable rate, which will be confirmed by the subsequent simulation results.

5.5.2 MSE Performance

In Fig. 5.5, we show the sum MSE per symbol of a two-hop FDR system for the proposed AO/sGP based power allocation and the baseline schemes. For low SNRs, the robust and non-robust schemes have a similar performance. However, as the SNR increases, the MSE of the robust schemes decrease monotonically while the MSE of the non-robust schemes starts to increase at some SNR value. This is because,

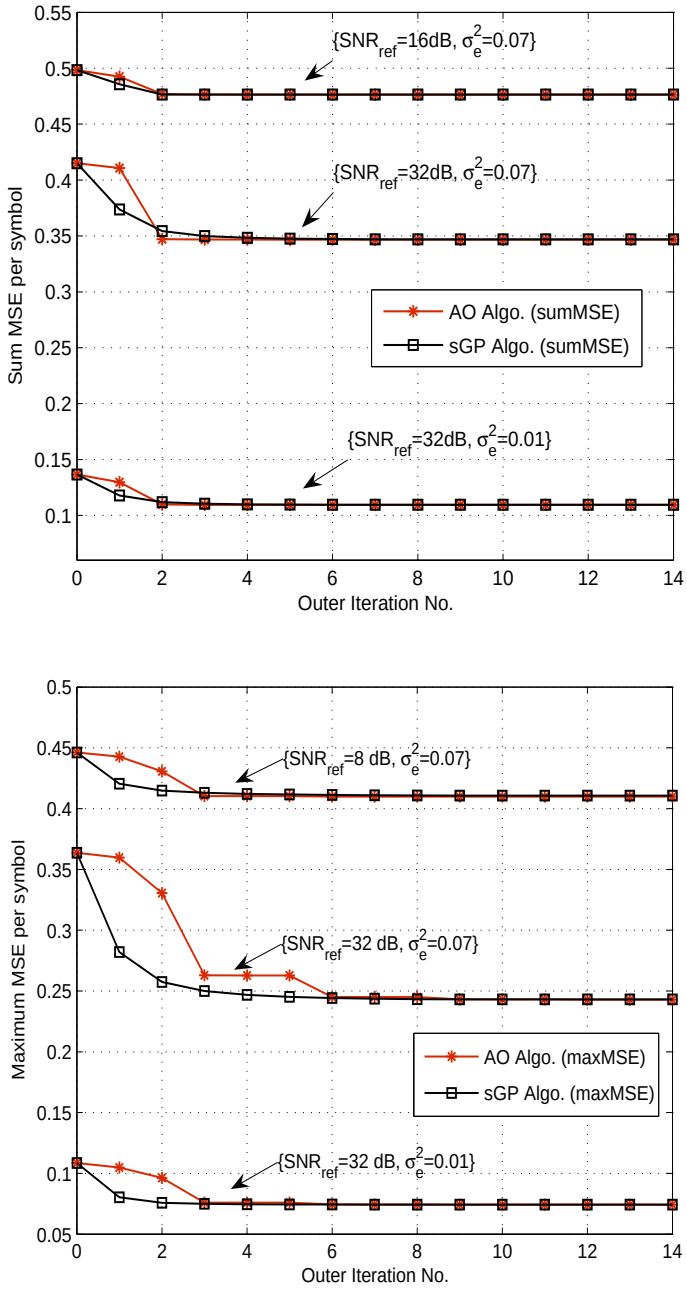

Figure 5.4: Convergence comparison of the AO and sGP algorithms. **Upper:** Two-hop FDR system employing sum MSE minimization. **Lower:** Three-hop FDR system employing maximum MSE minimization.

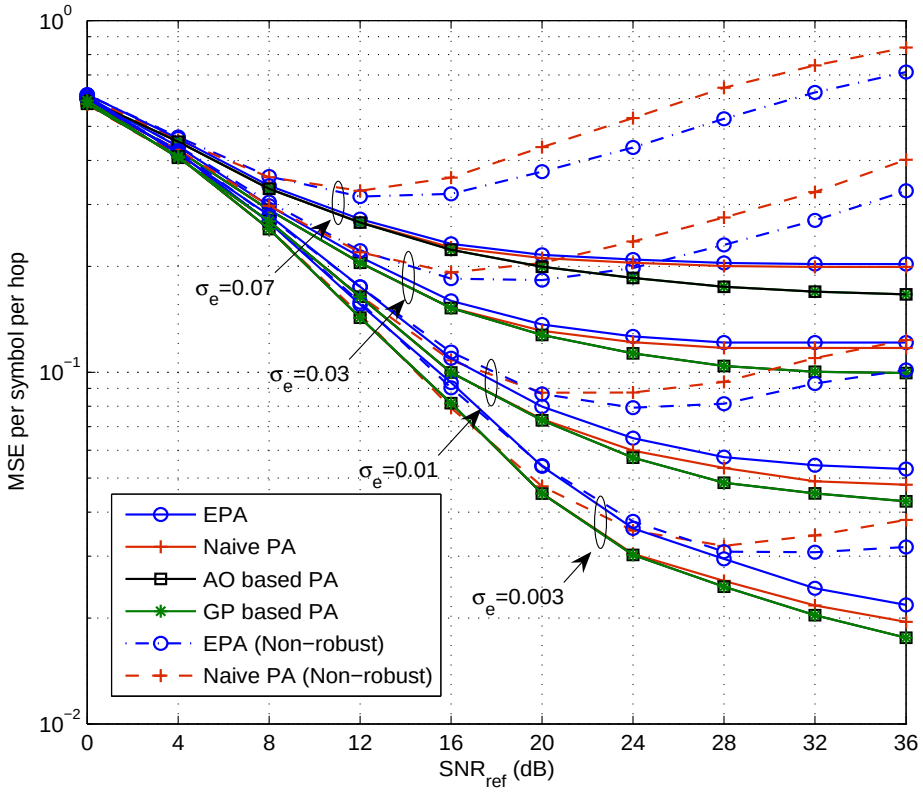

Figure 5.5: Sum hop MSE per symbol for a two-hop DF FDR system employing SC-FDE.

at high SNR, the CSI error variance dominates the MSE, and since the non-robust equalizers fail to take the resulting strong interference into account, their performance is severely degraded. Also, the proposed robust naive-PA and EPA schemes both exhibit a performance comparable to the proposed AO/sGP algorithm for low-to-medium SNRs and in the presence of large CSI errors. For small CSI errors, e.g., $\sigma_e^2 = 0.003$, the naive PA scheme achieves a near-optimal performance. Since these baseline schemes entail a lower complexity, they are attractive alternatives to the optimal algorithm for these SNRs and CSI error variances.

In Fig. 5.6, we compare the maximum MSE per symbol of the proposed AO/sGP based power allocation schemes and the GP-I scheme for a two-hop FDR system under

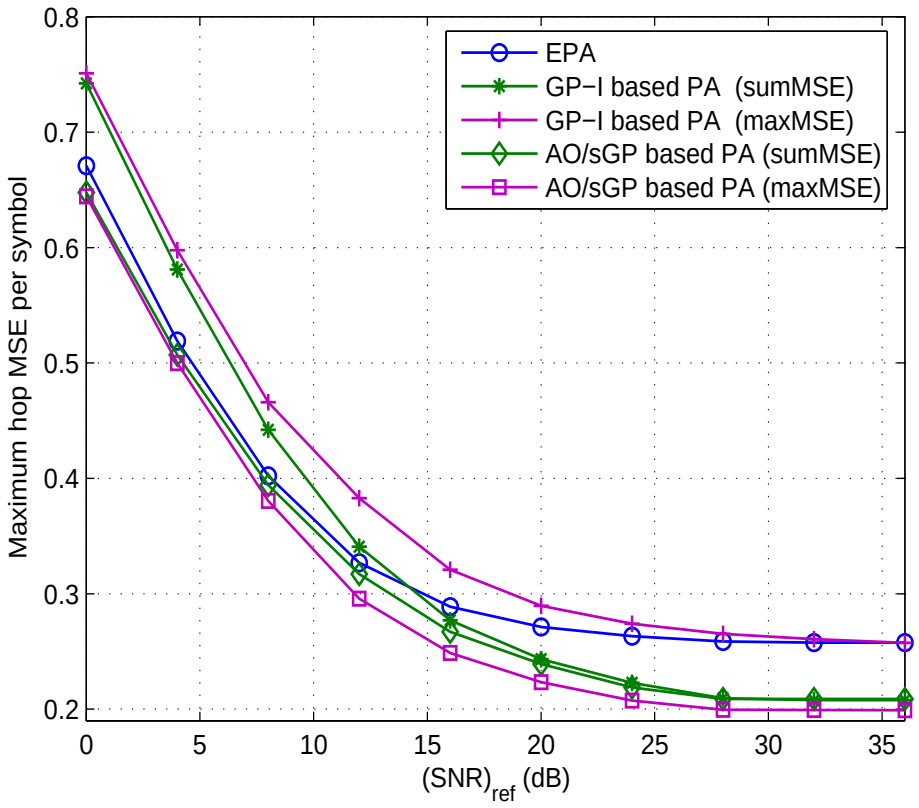

Figure 5.6: Maximum hop MSE per symbol for a two-hop DF FDR system employing SC-FDE, $\sigma_e^2 = 0.07$.

both considered optimization criteria, where $\sigma_e^2 = 0.07$. It can be seen that due to the use of the loose upper bound, the performance of the GP-I scheme based on sum MSE minimization suffers from a considerable performance loss in low-to-medium SNRs, and the performance of the GP-I scheme for maximum MSE minimization is even worse than that of EPA over the entire considered SNR range. On the other hand, the proposed AO and sGP schemes achieve for both criteria a better performance than EPA in terms of the maximum hop MSE.

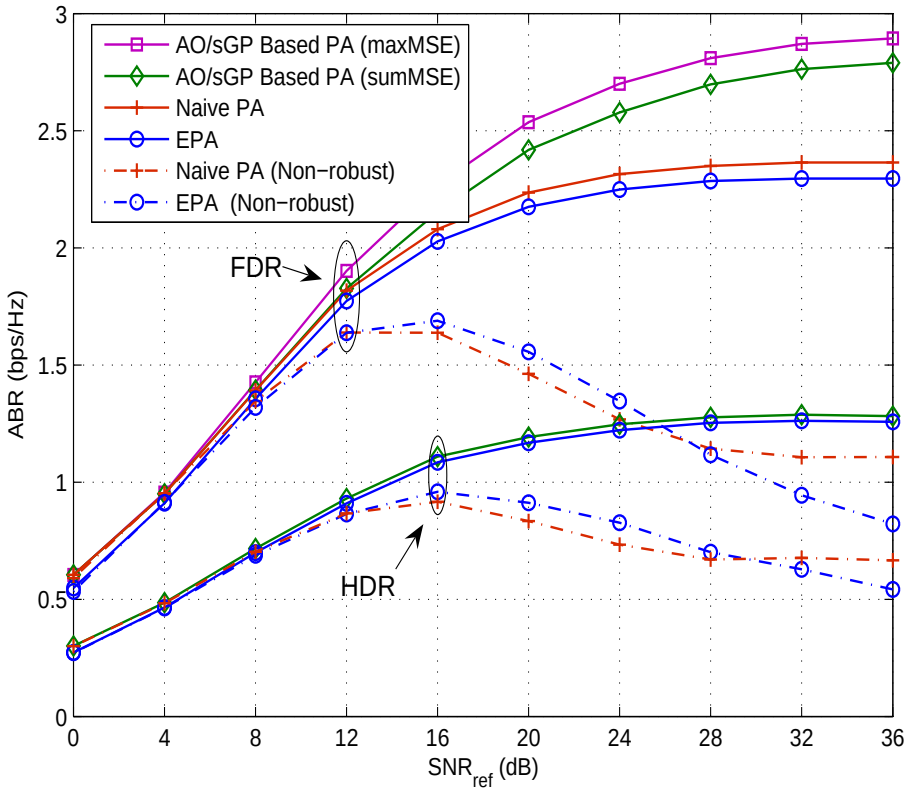

Figure 5.7: ABR of three-hop DF FDR and HDR systems employing SC-FDE, $\sigma_e^2 = 0.03$.

5.5.3 ABR Performance

For communication systems, the ABR is the ultimate performance metric. Therefore, in Fig. 5.7, we investigate the ABRs of three-hop SC-FDE based FDR relay systems, where the ABR is obtained by evaluating $\mathbb{E}_{\{\mathbf{H}_{f,i},\mathbf{L}_{f,(i+1)},\mathbf{G}_{f,(i+2)}\}}\{\min_{i=1,\ldots,M} \log_2(1+\text{SINR}_{(i+1)})\}$ via Monte-Carlo simulations. The SINR for node i is given by $\text{SINR}_i = \left[\frac{1}{N_c}\text{tr}\left(\mathbf{E}_{t,i}\right)\right]^{-1} - 1$, where we have exploited the relationship between the MSE and the SINR [96]. As expected from Fig. 5.7, the proposed robust designs achieve a larger ABR compared to the non-robust schemes, and the robust baseline schemes achieve a similar ABR as the optimized PA scheme for low-to-medium SNRs but require a

lower complexity. Similar to the MSE performance, the sGP and AO based schemes also achieve identical ABRs. For both schemes, minimizing the maximum MSE yields a higher ABR than minimizing the sum MSE since the ABR is a monotonic function of the MSE for SC-FDE systems and the ABR for DF relaying is determined by the worst-case hop ABR. We also observe that the proposed FDR system enjoys a significantly higher ABR than a HDR system if the LI and backward interference is handled appropriately. However, if the transceivers do not take the CSI errors into account, the FDR system can perform even worse than the HDR system in medium-to-high SNRs. This illustrates the importance of the proposed robust algorithms for optimizing FDR systems.

5.6 Conclusions

In this chapter, we have addressed the sum MSE and maximum MSE minimization problems for multi-hop DF FDR systems employing SC-FDE. To this end, first the optimal equalization filters at the receiving nodes and the corresponding MSE matrices were derived taking into account the imperfect CSI knowledge. Subsequently, the transmit power allocation matrices were obtained based on sGP and AO approaches, respectively. Numerical results show that both approaches yield practically identical MSE and ABR performances. In addition, the proposed robust transceiver design outperforms non-robust FDR systems and robust HDR systems by a considerable margin, especially in the case of high SNR and large channel estimation errors, i.e., the interference limited operating region.

Chapter 6

Summary of Thesis and Future Research Topics

In this final chapter, we summarize our results and highlight the contributions of this thesis in Section 6.1. In Section 6.2, we also illustrate ideas for future research.

6.1 Summary of Results

This thesis has focused on transceiver optimization for broadband cooperative communication systems with different types and topologies of the relay network.

In Chapter 2, we considered transceiver design for single-user broadband SC-FDE in a wireless network consisting of one single-antenna source, one single-antenna destination, and multiple multiantenna relays. Adopting the minimum MSE as optimality criterion, the optimal frequency-domain LE and DFE receivers were derived and corresponding objective functions for relay BF matrix optimization are specified. For a sum relay power constraint, we obtain the structure of the optimal relay BF matrices in closed form. While the structure of the optimal relay BF matrices is identical for LE and DFE as well as for an idealized matched filter receiver, the solution of the remaining power allocation problem depends on the adopted receiver. The power allocation problem is shown to be convex for all considered receivers and an efficient numerical algorithm for finding the optimal power allocation is provid-

ed. Furthermore, to reduce complexity, two suboptimal power allocation schemes assigning identical powers to all relays and/or frequencies are proposed and shown to lead to only a small loss in performance and a remarkable robustness against imperfect channel state information. Furthermore, for optimal rBF, simple FD-LE receivers approach the performance of the idealized MF receiver as the numbers of relays and/or relay antennas increase, making more complex nonlinear trellis-based receivers unnecessary.

In Chapter 3, we proposed a joint transceiver design for SC-FDE based MIMO relay systems. To this end, we first derived the optimal minimum mean-square error linear and decision-feedback frequency-domain equalization filters at the destination along with the corresponding error covariance matrices at the output of the equalizer. Subsequently, we formulated the source and relay precoding matrix design problem as the minimization of a family of Schur-convex and Schur-concave functions of the mean-squared errors at the output of the equalizer under separate power constraints for the source and the relay. By exploiting properties of the MSE matrix and results from majorization theory, we derived the optimal structures of the source and relay precoding matrices, which allows us to transform the matrix optimization problem into a scalar power optimization problem. Adopting a high SNR approximation for the objective function, we obtained the global optimal solution for the power allocation variables. Our results show that the proposed SC-FDE relaying schemes outperform the corresponding OFDM schemes in terms of both coded and uncoded BER for fixed modulation and coding rates. However, the performance gap between SC-FDE and OFDM relay systems decreases when the number of source/relay/destination antennas is larger than the number of data streams. Assuming Gaussian signalling and ideal channel coding, SC-FDE and OFDM attain similar achievable bit rates.

Furthermore, we have shown that the proposed suboptimal power allocation schemes can reduce the system complexity and feedback overhead at the expense of a moderate performance degradation, especially in case of coded transmission, making them promising candidates for practical relay systems.

In Chapter 4, we studied the robust design of the rBF and dEQ filters for broadband multiuser multi-relay networks employing SC-FDMA and OFDMA. Thereby, we considered the realistic case where only imperfect channel state information was available for rBF and dEQ filter optimization. Our goal was to maximize a lower bound for the weighted ABR of the network, subject to either Ind-PCs or an Agg-PC. We first derived the optimal dEQ filters and the phases of the optimal rBF filter coefficients, which are independent of the power constraints. For the Agg-PC, the amplitude optimization of the rBF filter coefficients was decomposed into two subproblems, which corresponded to the optimization of the power allocation across the relays and the power allocation across the users and subcarriers, respectively. We obtained a closed-form structural solution for the first subproblem by fixing the powers across users and subcarriers, and the global optimal solution for the second subproblem. For the Ind-PCs, the corresponding optimization problem was formulated as a reverse-convex problem with convex constraints. Subsequently, the constrained convex concave procedure was applied to approximate the original non-convex problem with a sequence of convex problems, which can be efficiently solved using convex optimization techniques. Simulation results validated the excellent performance of the proposed robust rBF and dEQ filter designs and showed their superiority compared to conventional non-robust and naive relaying schemes.

In Chapter 5, we tackled the problems of sum MSE and maximum MSE minimization for multi-hop DF FDR systems employing SC-FDE. To this end, first the

optimal equalization filters at the receiving nodes and the corresponding MSE matrices were derived taking into account the imperfect CSI knowledge. Subsequently, the transmit power allocation matrices were obtained based on sGP and AO approaches, respectively. In the sGP approach, we used the condensation technique to transform the objective function into a posynomial and then solved a sequence of standard GP problems. The sGP approach required global channel knowledge at a central node and the involved subproblems admited only numerical solutions. To gain further insight into the structure of the problem, we also consider an AO approach for power allocation where convex programming problems and difference of convex programming problems are solved in an alternating manner. The resulting AO algorithm admits closed-form solutions in each iteration step and requires less signaling overhead compared to the centralized sGP scheme. Numerical results show that both approaches yield practically identical MSE and ABR performances. In addition, the proposed robust transceiver design outperforms non-robust FDR systems and robust HDR systems by a considerable margin, especially in the case of high SNR and large channel estimation errors, i.e., the interference limited operating region.

6.2 Future Work

Broadband cooperative relay system design is a current research area with still many open problems for both single and multi-user systems. In the following, we provide several possible extensions of the current work based on the results obtained in this thesis.

6.2.1 Transceiver Design for Broadband Multiuser MIMO Relay Systems.

One possible extension of the work in Chapter 3 is the transceiver design for broadband multiuser MIMO relay systems. Thereby, different users can adopt orthogonal transmission schemes, such as SC-FDMA, to completely avoid multiuser interference, or employ non-orthogonal multiple access schemes, such as spatial-division multiple access (SDMA), to increase the spectral efficiency. In the former case, the transceiver design problem is similar to the problem addressed in Chapter 3. The main difference lies in that at the relay node, a joint transmit power constraint should be imposed for the signals from all users. This problem can be tackled in a layered manner by using the primal decomposition technique, where a master problem for the power allocation among users is solved in the upper-level, while subproblems in the lower-level for the source/relay precoding matrix design with given allocated user powers can be handled similarly as in Chapter 3. Specifically, we can exploit majorization theory and convex optimization to solve these subproblems. On the other hand, for the case of non-orthogonal SDMA, the corresponding transceiver design problem becomes more complicated as now we have excessive multiuser and multi-stream interferences in the system, which leads to a highly nonconvex problem structure. One possible solution is to impose some specific structure on the source and relay precoding matrices, which simplifies the problem by reducing the dimension of the system design parameters. For example, the block-diagonalization technique [137, 138] which was originally proposed for point-to-point multiuser systems, can be employed to decouple the signal transmission of different users in the spatial domain. Afterwards, scalar-based convex/non-convex optimization techniques become applicable to the resulting power allocation problem. Nevertheless, deriving the optimal scheme which does not im-

pose any structural constraint on the multiuser transceivers [139, 140] for broadband SC-FDE/OFDM systems is an interesting and challenging topic for future work.

6.2.2 Transceiver Design for Heterogenous Cooperative Cognitive Radio Networks.

Next generation wireless communication systems are envisioned to be able to accommodate heterogenous users and may adopt completely new multiple access schemes. For example, in cooperative communication scenarios, the relay stations may simultaneously receive signals from both SC-FDMA-based and OFDMA-based users. If the link CSI and the corresponding information on the signal formats of different users are available at the relays, they can perform dynamic resource allocation, e.g., adapt their transmit power and/or data rate for different users, in order to achieve the optimal network performance. Furthermore, such heterogenous cooperative networks can also be deployed for secondary-user systems in cognitive radio communication networks [141, 142]. In this case, the secondary-user system should be able to adapt to different signalling schemes for the legitimate primary-users in order to better guarantee the quality of service for the primary-user network [143]. We expect that the unified framework developed in Chapter 4 for cooperative multiuser transceiver schemes can serve as the basis for building the aforementioned heterogenous networks.

6.2.3 Robust Transceiver Design for MIMO Full-duplex AF/DF Relay Systems.

The robust transceiver design problem addressed in Chapter 5 only considers single-stream transmission with single-antenna source and destination. It is possible to extend the results to MIMO full-duplex relay systems where multiple transmit and

receive antennas are deployed at the source, the destination, and the full-duplex relays. In this case, we will have matrix optimization variables, instead of scalar power allocation variables, at all the nodes of the network [92]. In order to solve the resulting problem, matrix decomposition techniques and majorization theory are important and necessary mathematical tools to reduce the dimension of the matrix design parameters. Judging from the non-convexity of the scalar problem encountered in Chapter 5, we expect that the resulting problem for MIMO full-duplex DF relay systems will also be a highly non-convex problem whose optimal solution requires high computational complexity. Thus, low-complexity suboptimal solutions seem to be a more viable option for practical implementations. Furthermore, compared to the DF relaying scheme, the AF relaying scheme has the advantage of lower complexity and less signal processing delay. However, the robust transceiver design for full-duplex multi-hop AF relay systems is a challenging problem even for a network with single-antenna terminals. This is due to the accumulating effect of the interference-plus-noise amplification at the AF relays, which couples the design parameters of all hops [67]. The deployment of multiple antennas at the full-duplex AF relays can further complicate the problem structure [78]. These are important and challenging problems which merit consideration in future work.

Bibliography

[1] M. Rumney, ed. *LTE and the evolution to 4G wireless: Design and measurement challenges*. John Wiley And Sons, 2013.

[2] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Commun. Mag.*, vol. 52, pp. 36–43, May 2014.

[3] J. N. Laneman and G. W. Wornell, "Distributed space–time block coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[4] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062– 3080, Dec. 2004.

[5] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 29–43, Jan. 2005.

[6] D. Gesbert, et al, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 281–302, Mar. 2003.

[7] A. J. Paulraj, et al. "An overview of MIMO communications-a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, pp. 198–218, Feb. 2004.

[8] G. L. Stuber, et al, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, pp. 271–294, Jan. 2004.

[9] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power Allocation," *IEEE J. Sel. Areas Commun.,* vol. 17, pp. 1747–1758, Oct. 1999.

[10] D. Falconer, S. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.

[11] N. Benvenuto and S. Tomasin, "On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter," *IEEE Trans. Commun.*, vol. 50, pp. 947–955, Jun. 2002.

[12] F. Pancaldi, G. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 37–56, Sept. 2008.

[13] N. Benvenuto, R. Dinis, D. Falconer, and S. Tomasin, "Single carrier modulation with nonlinear frequency domain equalization: An idea whose time has come–again," *Proceedings of the IEEE*, vol. 98, pp. 69–96, Jan. 2010.

[14] D. Astély, E. Dahlman, A. Furuskär, Y. Jading, N. Lindström, and S. Parkvall, "LTE: the evolution of mobile broadband," *IEEE Commun. Mag.*, vol. 47, pp. 44–51, 2009.

[15] E. Dahlman, P. Stefan, and S. Johan, *4G: LTE/LTE-advanced for mobile broadband*, Academic Press, 2013.

[16] Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); Overall description; Stage 2, TS Standard 36.300, Sept. 2014.

[17] Q. Li, X. Lin, J. Zhang, and W. Roh, "Advancement of MIMO technology in WiMAX: From IEEE 802.16d/e/j to 802.16m," *IEEE Commun. Mag.*, vol. 47, pp. 100–107, 2009.

[18] IEEE draft standard P802.16j/D9, "Amendment to IEEE standard for local and metropolitan area networks, Part 16: Air interface for fixed and mobile broadband wireless access systems – Multi-hop relay specification," Feb. 2009.

[19] S. W. Peters and R. W. Heath, "The future of WiMAX: Multihop relaying with IEEE 802.16j," *IEEE Commun. Mag.,* vol. 47, pp. 104–111, Jan. 2009.

[20] Z. Ren, S. Chen, B. Hu, and W. Ma, "Energy-efficient resource allocation in downlink OFDM wireless systems with proportional rate constraints," *IEEE Trans. Veh. Technol.* vol. 63. pp. 2139–2150, May 2014.

[21] I. Wong and B. Evans, "Optimal downlink OFDMA resource allocation with linear complexity to maximize ergodic rates," *IEEE Trans. Wireless Commun.,* vol. 7, pp. 962–971, Mar. 2008.

[22] D. W. K. Ng and R. Schober, "Resource allocation and scheduling in multi-cell OFDMA systems with decode-and-forward relaying," *IEEE Trans. Wireless Commun.,* vol. 7, pp. 2246–2258, Jul. 2011.

[23] D. W. K. Ng, E. S. Lo, and R. Schober, "Secure resource allocation and scheduling for OFDMA decode-and-forward relay Networks," *IEEE Trans. Wireless Commun.,* vol. 10, pp. 3528–3540, Oct. 2011.

[24] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation for secure OFDMA systems," *IEEE Trans. Veh. Technol.,* vol. 61, pp. 2572–2585, Jun. 2012.

[25] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.,* vol. 11, pp. 3292–3304, Sept. 2012.

[26] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Trans. Wireless Commun.,* vol. 11, pp. 3618–3631, Oct. 2012.

[27] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. Wireless Commun.,* vol. 12, pp. 3412–3427, Jul. 2013.

[28] D. W. K. Ng, E. S. Lo, and R. Schober, "Wireless information and power transfer: Energy efficiency optimization in OFDMA systems," *IEEE Trans. Wireless Commun.,* vol. 12, pp. 6352–6370, Dec. 2013.

[29] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals,", *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, Jun. 2008.

[30] Y. Yao and G. B. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 53, pp. 173–183, Jun. 2005.

[31] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1099–1109, Aug. 2004.

[32] P. Larsson, "Large–scale cooperative relaying network with optimal combining under aggregate relay power constraint", *Future Telecommunications Conf.*, Dec. 2003.

[33] Z. Yi and I. Kim, "Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE J. Sel. Areas Commun*, vol. 25, pp. 447–458, Feb. 2007.

[34] G. Zheng, K.-K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: optimum and distributed implementation," *IEEE Signal Process. Lett.*, vol. 16, pp. 257–260, Apr. 2009.

[35] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, pp. 4306–4316, Aug. 2008.

[36] Y. Jing and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Inform. Theory*, vol. 55, pp. 2499–2517, Jun. 2009.

[37] J. M. Paredes and A. B. Gershman, "Relay network beamforming and power control using maximization of mutual information," *IEEE Trans. on Wireless Commun.*, vol. 10, pp. 4356–4365, Dec. 2011.

[38] J. H. Choi, "MMSE-based distributed beamforming in cooperative relay networks," *IEEE Trans. Commun.*, vol. 59, pp. 1346–1356, May 2011.

[39] L. Zhang, W. Liu, and J. Li, "Low-complexity distributed beamforming for relay networks with real-valued implementation," *IEEE Trans. Signal Process.*, vol. 61, pp. 5039–5048, Oct. 2013.

[40] W. Dang, N. Tao, H. Mu, and J. Huang, "Subcarrier-pair based resource allocation for cooperative multi-relay OFDM systems", *IEEE Trans. on Wireless Commun.*, vol. 9, pp. 1640–1649, May 2010.

[41] H. Chen, A. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming in relay networks with frequency selective fading," *IEEE Trans. Signal Process.*, vol. 58, pp. 1251–1262, Mar. 2010.

[42] Y. Liang, A. Ikhlef, W. H. Gerstacker, and R. Schober, "Cooperative filter-and-forward beamforming for frequency-selective channels with equalization," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 228–239, Jan. 2011.

[43] W. Cheng, M. Ghoghob, Q. Huang, D. Ma, and J. Wei, "Cooperative beamforming for OFDM-based amplify-and-forward relay networks," *Physical Communications*, vol. 4, pp. 305–312, Dec. 2011.

[44] Y. Liang and R. Schober, "Cooperative amplify-and-forward beamforming for OFDM systems with multiple relays," *IEEE International Conference on Communications*, Jun. 2009.

[45] T. T. Pham, H. H. Nguyen, and H. D. Tuan, "Power allocation in MMSE relaying over frequency-selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 58, pp. 3330–3343, Nov. 2010.

[46] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory,* vol. 46, pp. 933-946, Mar. 2000.

[47] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory,* vol. 52, pp. 2203–2214, May 2006.

[48] M. Payaro, A. Pascual-Iserte, and M. A. Lagunas, "Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization," *IEEE J. Sel. Areas Commun.*, vol. 25, pp.1390–1401, Jul. 2007.

[49] H Shen, J. Wang, and B.C. Levy, "Robust optimization for amplify-and-forward MIMO relaying from a worst-case perspective," *IEEE Trans. Signal Process.*, vol. 61, pp. 5458–5471, Nov. 2013.

[50] C. Jeong, B. Seo, S. Lee, H. Kim, and I. Kim, "Relay precoding for non-regenerative MIMO relay systems with partial CSI feedback," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1698–1711, May 2012.

[51] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-Relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, pp. 5473–5484, Nov. 2011.

[52] H. H. Kha, H. D. Tuan, H. H. Nguyen, and T. T. Pham, "Optimization of cooperative beamforming for SC-FDMA multi-user multi-relay networks by tractable D.C. programming," *IEEE Trans. Signal Process.*, vol. 61, pp. 467–479, Jan. 2013.

[53] J. K. Cavers, "Single-user and multiuser adaptive maximal ratio transmission for Rayleigh channels," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 2043–2050, Jun. 2000.

[54] H. Bolcskei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commn.*, vol. 60, pp. 225–234, Feb. 2002.

[55] L. G. Ordoez, D. P. Palomar, and J. R. Fonollosa, "High-SNR analytical performance of spatial multiplexing MIMO systems with CSI," *IEEE Trans. Signal Process.*, vol. 55, pp. 5447-5463 Nov. 2007.

[56] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory,* vol.51, pp. 29-43, Jan. 2005.

[57] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1774-1786, May 2007.

[58] O. Munoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in non-regenerative relays with channel state information," *IEEE Trans. on Signal Process.*, vol. 6, pp. 2593-2604, Jun. 2007.

[59] Z. Fang , Y. Hua, and J.C. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," *Sensor Array and Multichannel Processing, 2006. Fourth IEEE Workshop on*, pp. 239-243, 2006.

[60] I. Hammerstrom and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless. Commun.*, vol. 6, pp. 2798–2802, Aug. 2007.

[61] R. Mo and Y. Chew, "Precoder design for non-regenerative MIMO relay systems," *IEEE Trans. Wireless. Commun.*, vol. 8, pp. 5041-5049, Oct. 2009.

[62] F. S. Tseng, C. T. Lin, and W. R. Wu, "Optimum transceiver designs in two-hop amplify-and-forward MIMO relay systems with SIC receivers," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 985–997, Mar. 2015.

[63] R. Mo and Y. Chew, "MMSE-based joint source and relay precoding design for amplify-and-forward MIMO relay networks," *IEEE Trans. Wireless. Commun.*, vol. 8, pp. 4668-4676, Sept. 2009.

[64] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837-4851, Dec. 2009.

[65] D. Kim, Y. Sung, and J. Chung, "Filter-and-forward relay design for MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 62, pp. 2329–2339, Jul. 2014.

[66] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 6068-6077, Dec. 2009.

[67] Y. Rong, "Optimal linear non-regenerative multi-hop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2268–2279, July 2010.

[68] C. Xing, S. Ma, Y. C. Wu, and T. S. Ng, "Transceiver design for dual-hop non-regenerative MIMO-OFDM relay systems under channel uncertainties", *IEEE Trans. Signal Process.*, vol. 58, pp. 6325–6339, Dec. 2010.

[69] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE Trans. on Signal Process.*, vol. 57, pp. 2785-2796, May 2011.

[70] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, Sept. 2003.

[71] D. P. Palomar, "Convex primal decomposition for multicarrier linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 53, pp. 4661–4674, Jul. 2005.

[72] A. Tajer, and A. Nosratinia, "Diversity order in ISI channels with single-carrier frequency-domain equalizers," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1022–1032, May 2010.

[73] A. H. Mehana, and A. Nosratinia, "Single-carrier frequency-domain equalizer with multi-antenna transmit diversity," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 388–397, Jan. 2013.

[74] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: challenges and opportunities," *IEEE J. Select. Areas Commun.*, vol. 59, pp. 1637–1652, Sept. 2014.

[75] A. Sahai, "Wireless full-duplex: from practice to theory," Doctoral Thesis, Rice University. 2014.

[76] A. K. Khandani, "Full-duplex wireless transmission with self-interference cancellation." U.S. Patent Application 13/893, 288.

[77] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of full-duplex: design and real-time implementation," Department of Electrical and Computer Engineering, Rice University, Technical Report TREE1104, Jul. 2011.

[78] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.,* vol. 59, pp. 5983–5993, Dec. 2011.

[79] B. Chun and H. Park, "A spatial-domain joint-nulling method of self interference in full-duplex relays," *IEEE Commun. Lett.*, vol. 16, pp. 436–438, Apr. 2012.

[80] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 3074–3085, Dec. 2011.

[81] H. Kim, S. Lim, H. Wang, and D. Hong, "Optimal power allocation and outage analysis for cognitive full-duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3754– 3765, Oct. 2012.

[82] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Trans. Wireless Commun.,* vol. 12, pp. 2498–2511, May 2013.

[83] I. Krikidis, H. Suraweera, S. Yang, and K. Berberidis, "Full-duplex relaying over block fading channels: A diversity perspective", *IEEE Trans. Wireless Commun.*, vol. 11, pp. 4524–4535, Dec. 2012.

[84] T. Baranwal, D. Michalopoulos, and R. Schober, "Outage analysis of multi-hop full-duplex relaying", *IEEE Commun. Lett.*, vol. 17, pp. 63–66, Jan. 2013.

[85] T. Kwon, S. Lim, S. Choi, and D. Hong, "Optimal duplex mode for DF relay in terms of the outage probability," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 3628–3634, Sept. 2010

[86] I. Krikidis, H. Suraweera, P. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4381–4393, Dec. 2012.

[87] H. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 913–927, Feb. 2014.

[88] W. Li, J. Lilleberg, and K. Rikkinen, "On rate region analysis of half-and full-duplex OFDM communication links," *IEEE J. Select. Areas Commun.,* vol. 32, pp. 1688–1698, Sept. 2014.

[89] W. Zhang, M. Hsieh, U.Mitra, and M. Chiang, "Optimization of amplify-and-forward multicarrier two-hop transmission." *IEEE Trans. Commun.*, vol. 59, pp. 1434–1445, May 2011.

[90] P. Wu and R. Schober, "Cooperative beamforming for single-carrier frequency-domain equalization systems with multiple relays", *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2276–2286, Jun. 2012.

[91] E. Antonio-Rodriguez, R. Lopez-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Adaptive self-interference cancellation in wideband full-duplex decode-and-forward MIMO relays," *in Proc. IEEE Int. Workshop on Signal Process. Advances in Wireless Commun.,* Jun. 2013.

[92] D. Ng, E. Lo, and R. Schober, "Dynamic resource allocation in MIMO-OFDMA systems with full-duplex and hybrid relaying," *IEEE Trans. Commun.*, vol. 60, pp. 1291–1304, May 2012.

[93] P. Duhamel and M. Vetterli, "Fast Fourier transforms: a tutorial review and a state of the art", *Signal Processing*, vol. 19, pp.259–299, 1990.

[94] A. Gusmao, R. Dinis, and N. Esteves, "On frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE Commun. Lett.*, vol. 51, pp. 1029–1033, 2003.

[95] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, pp. 2921–2936, 2000.

[96] J. N. Cioffi, G. P. Dudevoir, N. V. Eyuboglu, and G. D. Forney, " On the convergence of the generalized DFE to the MMSE-DFE," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM'01)*, Nov. 2001.

[97] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback ," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1334–1347, 1994.

[98] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications.* New York: Kluwer Academic, 1999.

[99] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. 25, pp. 772–781, 1978.

[100] R. A. Horn and C. R. Johnson. *Matrix Analysis*, Cambridge University Press, 1985.

[101] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[102] T. S. Rappaport, *Wireless Communications: Principles and Practice.* Prentice Hall, 2002.

[103] U. Dang, M. Ruder, R. Schober, and W. Gerstacker, "MMSE Beamforming for SC-FDMA Transmission over MIMO ISI Channels," *EURASIP Journal on Advances in Signal Processing*, vol. 2011.

[104] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of water-filling solutions," *IEEE Trans. Signal Process.*, vol. 53, pp. 686–695, Feb. 2005.

[105] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," *Advances in Soft Computing*, Berlin, Germany: Springer, 2002, pp. 187–195.

[106] N. Al-Dhahir, and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, pp. 2921–2936, Oct. 2000.

[107] J. K. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Information Theory.*, vol. 51, pp. 154-172, Jan. 2005.

[108] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Process.*, vol. 53, pp. 3791-3803, Oct. 2005.

[109] P. Wu, R. Schober, V. K. Bhargava, "Joint Transceiver Design for MIMO Relay Systems Employing SC-FDE," *IEEE GlobeCom, Anaheim, U.S.A*, Dec. 2012.

[110] A. W. Marshall, I. Olkin, and B. Arnold, *Inequalities: theory of majorization and its applications.* New York: Academic, 1979.

[111] D. Bertsekas. *Nonlinear Programming.* Athena Scientic, 2008.

[112] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2006.

[113] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[114] T. Svensson, T. Frank, D. Falconer, M. Sternad, E. Costa, and A. Klein, "B-IFDMA: a power efficient multiple access scheme for non-frequency-adaptive transmission," *Proc. Mobile and Wireless Communications Summit*, 2007.

[115] R. Zhang, Y. C. Liang, R. Narasimhan, and J. M. Cioffi, "Approaching MIMO-OFDM capacity with per-antenna power and rate feedback," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1284–1297, Sep. 2007.

[116] J. Cioffi, G. Dudevoir, M. Eyuboglu, and G. Forney Jr., "MMSE decision-feedback equalizers and coding. I. Equalization results," *IEEE Trans. on Commun.* vol. 43, pp. 2582–2594, Sept. 1995.

[117] H. Tuy, "Convex programs with an additional reverse convex constraint," *Journal of Optimization Theory and Applications*, vol. 52, pp. 463–486, Mar. 1987.

[118] R. Horst and N. V. Thoai, "D.C. programming: overview," *Journal of Optimization Theory and Applications*, vol. 103, pp. 1–43, Oct. 1999.

[119] B. K. Sriperumbudur and G. R. G. Lanckriet, "On the convergence of the concave-convex procedure," *Neural Inf. Process. Syst.*, pp. 1–9, 2009.

[120] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Signal Process.*, vol. 60, pp. 2962–2973, Feb. 2012.

[121] D. P. Palomar, "Convex primal decomposition for multicarrier linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 53, pp. 4661–4674, Dec. 2005.

[122] S. Boyd and A. Mutapcic, "Subgradient methods," notes for EE364, Standford University, Winter 2006-07.

[123] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.,* vol. 54, pp. 1310–1321, Jul. 2006.

[124] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," Online, 2009, available: *http://www. stanford.edu/boyd/cvx.*

[125] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11–12, pp. 625–653, 1999.

[126] J. G. Proakis, *Digital Communications.* McGraw-Hill, Singapore, 4th edition, 2001.

[127] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[128] D. Erik, S. Parkvall, and J. Skold, *4G: LTE/LTE-advanced for mobile broadband.* Academic Press, 2013.

[129] H. Nishiyama, M. Ito, and N. Kato, "Relay-by-smartphone: realizing multi-hop device-to-device communications," *IEEE Commun. Mag.*, vol. 52, pp. 56–65, Apr. 2014.

[130] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Commun. Mag.,* vol. 52, pp. 36–43, May 2014.

[131] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G

cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.,* vol. 52, pp. 106–113, Feb. 2014.

[132] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range", *IEEE J. Select. Areas Commun.,* vol. 30, pp. 1541–1553, Sept. 2012.

[133] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Process.,* vol. 57, pp. 1141–1150, Mar. 2009.

[134] M. Chiang, "Geometric programming for communication systems," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1-2, pp. 1–156, Aug. 2005.

[135] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," *in Advances in Soft Computing*. Berlin, Germany: Springer, pp. 187–195, 2002.

[136] R. Horst and N. V. Thoai, "D.C. programming: Overview," *Journal of Optimization Theory and Applications*, vol. 103, pp. 1–43, Jan. 1999.

[137] D. H. Nguyen, H. Nguyen-Le, and T. Le-Ngoc, "Block-diagonalization precoding in a multiuser multicell MIMO system: Competition and coordination," *IEEE Trans. on Wireless Commun.*, vol. 13, pp. 968–981, Feb. 2014.

[138] Q. H. Spencer and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, no. 52, pp. 461–471, Jan. 2004.

[139] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, no. 53, pp. 18–28, Jan. 2004.

[140] L. Liu, R. Zhang, and K.C. Chua, "Achieving global optimality for weighted sum-rate maximization in the K-user gaussian interference channel with multiple antennas," *IEEE Trans. on Wireless Commun.*, vol. 11, pp. 1933–1945, May. 2012.

[141] S. Haykin, "Cognitive radio: brain–empowered wireless communications," *IEEE J. Sel. Areas Commun*, vol. 23, pp. 201–220, 2005.

[142] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint beamforming and power control for multiple access channels in cognitive radio networks," *IEEE J. Sel. Areas Commun.,* vol. 26, pp. 38–51, 2008.

[143] X. Kang, H. Garg, Y. C. Liang, and R. Zhang, "Optimal power allocation for OFDM-based cognitive radio with new primary transmission protection criteria," *IEEE Trans. on Wireless Commun.*, vol. 9, pp. 2066–2075, 2010.

Appendix A

Proof of Proposition 2.1

First, we rewrite $\mathbf{u}_{ik}^{\dagger}$ as $\mathbf{u}_{ik}^{\dagger} = \text{vec}(\mathbf{h}_{ik}^{\dagger}\mathbf{g}_{ik}^{\dagger}) = \mathbf{g}_{ik}^{*} \otimes \mathbf{h}_{ik}^{\dagger}$. Next, to simplify the inverse matrix $\mathbf{\Delta}_{ik}^{-1}$, we recall a result for the Kronecker product [99]. In particular, for all matrices $\mathbf{A} \in \mathbb{C}^{N\times N}$ and $\mathbf{B} \in \mathbb{C}^{N\times N}$, the following equality holds:

$$(\mathbf{A} \oplus \mathbf{B})^{-1} \quad = \quad \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{(\mathbf{a}_i \otimes \mathbf{b}_j)(\bar{\mathbf{a}}_i \otimes \bar{\mathbf{b}}_j)^{\dagger}}{\lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B})}, \tag{A.1}$$

where $\mathbf{A} \oplus \mathbf{B} = \mathbf{I}_N \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{I}_N$, $\lambda_i(\mathbf{X})$ is the ith eigenvalue of matrix $\mathbf{X}$, and $\mathbf{x}_i$ and $\bar{\mathbf{x}}_i$ are the eigenvectors that correspond to the ith eigenvalues of matrices $\mathbf{X}$ and $\mathbf{X}^{\dagger}$, respectively. Hence, we can rewrite $\mathbf{\Delta}_{ik}^{-1}$ as

$$\begin{aligned}\mathbf{\Delta}_{ik}^{-1} &= \left[\sigma_{n_2}^2(\sigma_s^2\mathbf{g}_{ik}^{*}\mathbf{g}_{ik}^{T} + \sigma_{n_1}^2\mathbf{I}_N) \oplus \sigma_{n_1}^2 P_k \mathbf{h}_{ik}^{\dagger}\mathbf{h}_{ik}\right]^{-1} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{(\mathbf{v}_i \otimes \mathbf{u}_j)(\bar{\mathbf{v}}_i \otimes \bar{\mathbf{u}}_j)^{\dagger}}{\lambda_i(\sigma_{n_1}^2 P_k \mathbf{h}_{ik}^{\dagger}\mathbf{h}_{ik}) + \lambda_j(\sigma_{n_2}^2(\sigma_s^2\mathbf{g}_{ik}^{*}\mathbf{g}_{ik} + \sigma_{n_1}^2\mathbf{I}_N))}.\end{aligned} \tag{A.2}$$

It is easy to verify that the largest eigenvalue of $\sigma_{n_1}^2 P_k \mathbf{h}_{ik}^{\dagger}\mathbf{h}_{ik}$ and the corresponding eigenvector are given by $\lambda_1 = \sigma_{n_1}^2 P_k ||\mathbf{h}_{ik}||^2$ and $\mathbf{v}_1 = \frac{\mathbf{h}_{ik}^{\dagger}}{||\mathbf{h}_{ik}||}$, respectively. Likewise, the largest eigenvalue of $\sigma_{n_2}^2(\sigma_s^2\mathbf{g}_{ik}^{*}\mathbf{g}_{ik}^{T} + \sigma_{n_1}^2\mathbf{I}_N)$ and the corresponding eigenvector are $\lambda_1' = \sigma_{n_2}^2(\sigma_s^2||\mathbf{g}_{ik}||^2 + \sigma_{n_1}^2)$ and $\mathbf{u}_1 = \frac{\mathbf{g}_{ik}^{*}}{||\mathbf{g}_{ik}||}$, respectively. By noting that except for $\mathbf{u}_1 \otimes \mathbf{v}_1 = \frac{1}{||\mathbf{h}_{ik}||||\mathbf{g}_{ik}||}\mathbf{g}_{ik}^{*} \otimes \mathbf{h}_{ik}^{\dagger}$, all other terms in the summation of (A.2) are orthogonal

to $\mathbf{u}_{ik}^{\dagger}$, we obtain

$$\boldsymbol{\Delta}_{ik}^{-1}\mathbf{u}_{ik}^{\dagger} = \frac{(\mathbf{g}_{ik}^{*}\mathbf{g}_{ik}^{T}\mathbf{g}_{ik}^{*}) \otimes (\mathbf{h}_{ik}^{\dagger}\mathbf{h}_{ik}\mathbf{h}_{ik}^{\dagger})}{||\mathbf{h}_{ik}||^{2}\,||\mathbf{g}_{ik}||^{2}(\lambda_1 + \lambda_1')} = \frac{\mathbf{g}_{ik}^{*} \otimes \mathbf{h}_{ik}^{\dagger}}{(\lambda_1 + \lambda_1')}, \tag{A.3}$$

which is the result in (2.32).

Appendix B

Proof of Theorem 3.1

We first provide some relevant definitions and lemmas that will be used in the proof.

Definition 1 [110, 1.A.1]: Given two $N\times 1$ real vectors $\mathbf{x},\mathbf{y}\in\mathbb{R}^N$, let $x_{[1]},\cdots,x_{[N]}$ and $y_{[1]},\cdots,y_{[N]}$ denote the components of $\mathbf{x}$ and $\mathbf{y}$ sorted in decreasing order. Then, $\mathbf{x}$ is majorized by $\mathbf{y}$, or $\mathbf{x}\prec\mathbf{y}$, if $\sum_{i=1}^{k}x_{[i]}\leq\sum_{i=1}^{k}y_{[i]}$ for $k<N$ and $\sum_{i=1}^{N}x_{[i]}=\sum_{i=1}^{N}y_{[i]}$. Vector $\mathbf{x}$ is weakly majorized by $\mathbf{y}$, or $\mathbf{x}\prec_w\mathbf{y}$, if $\sum_{i=1}^{k}x_{[i]}\leq\sum_{i=1}^{k}y_{[i]},\forall k$.

Definition 2 [110, 3.A.1]: A real function f is Schur-convex if for $\mathbf{x}\prec\mathbf{y}$, we have $f(\mathbf{x})\leq f(\mathbf{y})$. Similarly, f is Schur-concave if for $\mathbf{x}\prec\mathbf{y}$, we have $f(\mathbf{x})\geq f(\mathbf{y})$.

Lemma 1 [110, 9.B.1]: For a Hermitian matrix $\mathbf{A}$ with $\mathrm{diag}[\mathbf{A}]$ and $\lambda(\mathbf{A})$ denoting vectors containing the main diagonal elements and the eigenvalues of $\mathbf{A}$ arranged in decreasing order, respectively, we have $\mathrm{diag}[\mathbf{A}]\prec\lambda(\mathbf{A})$.

Lemma 2 [110, 9.H.2]: For M matrices $\mathbf{A}_i\in\mathbb{C}^{N\times N},i=1,\cdots,M$, let $\mathbf{B}=\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_M$. Then, $\sigma(\mathbf{B})\prec_w\sigma(\mathbf{A}_1)\odot\sigma(\mathbf{A}_2)\odot\cdots\sigma(\mathbf{A}_M)$, where $\sigma(\mathbf{X})$ denotes the vector containing the singular values of matrix $\mathbf{X}$ arranged in decreasing order and $\odot$ denotes the element-wise product of two vectors.

Lemma 3 [110, 3.A.8]:: A real function f satisfies $\mathbf{x}\prec_w\mathbf{y}\Rightarrow f(\mathbf{x})\leq f(\mathbf{y})$ if and only if f is Schur-convex and increasing.

Lemma 4 [110, 9.H.1]: For two Hermitian positive semidefinite matrices $\{\mathbf{A},\mathbf{B}\}\in\mathbb{C}^{N\times N}$ with eigenvalues $\lambda_{\mathbf{A},i},\lambda_{\mathbf{B},i}$, arranged in the same order, we have $\mathrm{tr}(\mathbf{AB})\geq\sum_{i=1}^{N}\lambda_{\mathbf{A},i}\lambda_{\mathbf{B},N-i+1}$.

Lemma 5 [110, p.7]: For a vector $\mathbf{x}\in\mathbb{R}^{N\times 1}$, we have $\sum_{i=1}^{N}(x_{[i]}/N)\mathbf{1}\prec\mathbf{x}$, where

$\mathbf{1}$ is the all-ones vector.

Lemma 6: For Hermitian matrices $\mathbf{A}_k$ and $\hat{\mathbf{A}}_k$, $k = 1, 2, \cdots, N_c$, if $\text{diag}[\mathbf{A}_k] \prec_w \text{diag}[\hat{\mathbf{A}}_k]$, we have $\text{diag}[\sum_{k=1}^{N_c} \mathbf{A}_k] \prec_w \text{diag}[\sum_{k=1}^{N_c} \hat{\mathbf{A}}_k]$.

Proof. According to Definition 1, if $\text{diag}[\mathbf{A}_k] \prec_w \text{diag}[\hat{\mathbf{A}}_k]$, we have $\sum_{i=1}^{j} a_{[ki]} \leq \sum_{i=1}^{j} \hat{a}_{[ki]}$ for $j \leq N$, $\forall k$, where $a_{[ki]}$ and $\hat{a}_{[ki]}$ are the ith largest diagonal entries of $\mathbf{A}_k$ and $\hat{\mathbf{A}}_k$, respectively. By taking the summation over k, we have $\sum_{k=1}^{N_c} \sum_{i=1}^{j} a_{[ki]} \leq \sum_{k=1}^{N_c} \sum_{i=1}^{j} \hat{a}_{[ki]}$ for $j \leq N$. Since the summations over k and j are exchangeable, we can write $\sum_{i=1}^{j} b_{[i]} \leq \sum_{i=1}^{j} \hat{b}_{[i]}$ for $j \leq N$, where $b_{[i]} = \sum_{k=1}^{N_c} a_{[ki]}$ and $\hat{b}_{[i]} = \sum_{k=1}^{N_c} \hat{a}_{[ki]}$ are the ith largest diagonal entries of $[\sum_{k=1}^{N_c} \mathbf{A}_k]$ and $[\sum_{k=1}^{N_c} \hat{\mathbf{A}}_k]$, respectively. By Definition 1, we obtain $\text{diag}[\sum_{k=1}^{N_c} \mathbf{A}_k] \prec_w \text{diag}[\sum_{k=1}^{N_c} \hat{\mathbf{A}}_k]$. □

Lemma 7 [110, 9.B.2]: For a diagonal matrix $\mathbf{D} \in \mathbb{C}^{M \times M}$, there is a unitary matrix $\mathbf{U}$ such that $\mathbf{A} = \mathbf{U}^\dagger \mathbf{D} \mathbf{U}$ has identical diagonal entries equal to $\text{tr}(\mathbf{D})/M$.

Lemma 8 [66]: If $f(\mathbf{x})$ is Schur-concave with respect to $\mathbf{x}$, and $\mathbf{y} = \mathbf{1} - \mathbf{x}$, where $\mathbf{1}$ is a vector of all ones, then $f(\mathbf{1} - \mathbf{y})$ is also Schur-concave with respect to $\mathbf{y}$.

We now set out to prove the optimal structure of the source and relay precoding matrices when $f(\text{diag}[\hat{\mathbf{E}}])$ is a Schur-concave increasing function w.r.t. $\text{diag}[\hat{\mathbf{E}}]$. Let us begin with the core term in the expression for $\hat{\mathbf{E}}$ in (3.13), which is given by

$$\mathbf{\Psi}_k^{-1} = \mathbf{I}_M - \mathbf{\Upsilon}_k, \tag{B.1}$$

where $\mathbf{\Upsilon}_k = \sigma_s^2 \mathbf{Q}_k^\dagger (\sigma_s^2 \mathbf{Q}_k \mathbf{Q}_k^\dagger + \sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d})^{-1} \mathbf{Q}_k$, and we have applied the matrix inversion lemma. Using (B.1), the MSE matrix in (3.16) can be expressed as

$$\hat{\mathbf{E}} = \frac{\sigma_s^2}{N_c} \sum_{k=0}^{N_c-1} \mathbf{\Psi}_k^{-1} = \frac{\sigma_s^2}{N_c} (N_c \mathbf{I}_M - \sum_{k=0}^{N_c-1} \mathbf{\Upsilon}_k). \tag{B.2}$$

By defining the following terms

$$\mathbf{X}_k = \sigma_s \mathbf{G}_k \mathbf{P}_k, \quad \mathbf{Y}_k = \mathbf{H}_k \mathbf{A}_k, \quad \mathbf{J}_k = \mathbf{Y}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{1/2}, \tag{B.3}$$

and using (3.4), we can rewrite $\boldsymbol{\Upsilon}_k$ as

$$\boldsymbol{\Upsilon}_k = \mathbf{X}_k^\dagger (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2} \mathbf{J}_k^\dagger (\mathbf{J}_k \mathbf{J}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d})^{-1} \mathbf{J}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2} \mathbf{X}_k. \tag{B.4}$$

Next, using the following SVDs

$$\begin{aligned} &\mathbf{X}_k = \mathbf{U}_X^{(k)} \boldsymbol{\Lambda}_X^{(k)} \mathbf{V}_X^{(k)}, \quad \mathbf{J}_k = \mathbf{U}_J^{(k)} \boldsymbol{\Lambda}_J^{(k)} \mathbf{V}_J^{(k)}, \\ &(\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2} = \mathbf{U}_X^{(k)} (\boldsymbol{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \boldsymbol{\Omega}^{(k)}, \end{aligned} \tag{B.5}$$

where $\mathbf{U}_X^{(k)} \in \mathbb{C}^{N_r \times M}$, $\mathbf{U}_J^{(k)} \in \mathbb{C}^{N_d \times M}$, $\{\mathbf{V}_J^{(k)}, \boldsymbol{\Omega}^{(k)}\} \in \mathbb{C}^{M \times N_r}$, $\{\boldsymbol{\Lambda}_X^{(k)}, \mathbf{V}_X^{(k)}, \boldsymbol{\Lambda}_J^{(k)}\} \in \mathbb{C}^{M \times M}$, $\boldsymbol{\Omega}^{(k)}$ is an arbitrary unitary matrix, and the diagonal entries of $\boldsymbol{\Lambda}_X^{(k)}$ and $\boldsymbol{\Lambda}_J^{(k)}$ are both sorted in decreasing order, we can further rewrite (B.4) as

$$\begin{aligned} \boldsymbol{\Upsilon}_k = \mathbf{V}_X^{(k)\dagger} \boldsymbol{\Lambda}_X^{(k)} \mathbf{Q}_2^{(k)\dagger} (\boldsymbol{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \mathbf{Q}_1^{(k)\dagger} (\mathbf{I}_M + \\ \sigma_u^2 \boldsymbol{\Lambda}_J^{(k)-2})^{-1} \mathbf{Q}_1^{(k)} (\boldsymbol{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \mathbf{Q}_2^{(k)} \boldsymbol{\Lambda}_X^{(k)} \mathbf{V}_X^{(k)}, \end{aligned} \tag{B.6}$$

where $\mathbf{Q}_1^{(k)} = \mathbf{V}_J^{(k)} \mathbf{U}_X^{(k)}$ and $\mathbf{Q}_2^{(k)} = \boldsymbol{\Omega}^{(k)} \mathbf{U}_X^{(k)}$. By applying Lemmas 1 and 2 to (B.6), we obtain

$$\mathrm{diag}[\boldsymbol{\Upsilon}_k] \prec \lambda(\boldsymbol{\Upsilon}_k) \prec_w \mathrm{diag}[\underbrace{(\mathbf{I}_M + \sigma_v^2 \boldsymbol{\Lambda}_X^{(k)-2})^{-1}}_{\mathbf{D}_1^{(k)}} \underbrace{(\mathbf{I}_M + \sigma_u^2 \boldsymbol{\Lambda}_J^{(k)-2})^{-1}}_{\mathbf{D}_2^{(k)}}]. \tag{B.7}$$

Therefore, $\mathrm{diag}[\boldsymbol{\Upsilon}_k]$ is majorized by $\mathrm{diag}[\mathbf{D}_1^{(k)} \mathbf{D}_2^{(k)}]$ when $\mathbf{V}_X^{(k)} = \boldsymbol{\Xi}_1$, $\mathbf{Q}_2^{(k)} = \boldsymbol{\Xi}_2$,

$\mathbf{Q}_1^{(k)} = \mathbf{\Xi}_3$, where $\mathbf{\Xi}_i \in \mathbb{C}^{M\times M}, \forall i$, are arbitrary diagonal matrices with unit norm diagonal elements. Without loss of generality, we can choose $\mathbf{\Xi}_i = \mathbf{I}_M, \forall i$. Hence, we have

$$\mathbf{V}_X^{(k)} = \mathbf{I}_M, \quad \mathbf{\Omega}^{(k)} = \mathbf{U}_X^{(k)\dagger}, \quad \mathbf{V}_J^{(k)} = \mathbf{U}_X^{(k)\dagger}. \tag{B.8}$$

From (B.7) and Lemma 6, we have

$$\text{diag}[\sum_{k=0}^{N_c-1} \mathbf{\Upsilon}_k] \prec_w \text{diag}[\sum_{k=0}^{N_c-1} \mathbf{D}_1^{(k)}\mathbf{D}_2^{(k)}]. \tag{B.9}$$

Recall that the objective function $f(\text{diag}[\hat{\mathbf{E}}])$ is a Schur-concave increasing function w.r.t. $\text{diag}[\hat{\mathbf{E}}]$. Based on Lemma 8, $f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{\Upsilon}_k])$ is then a Schur-concave decreasing function w.r.t. $\text{diag}[\sum_{k=0}^{N_c-1}\mathbf{\Upsilon}_k]$. By adding the minus sign, $-f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{\Upsilon}_k])$ becomes a Schur-convex increasing function w.r.t. $\text{diag}[\sum_{k=0}^{N_c-1}\mathbf{\Upsilon}_k]$. Therefore, by using (B.9) and Lemma 3, we deduce that $-f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{\Upsilon}_k]) \leq -f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{D}_1^{(k)}\mathbf{D}_2^{(k)}])$, which is equivalent to

$$f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{\Upsilon}_k]) \geq f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{D}_1^{(k)}\mathbf{D}_2^{(k)}]). \tag{B.10}$$

Consequently, with the help of the matrices in (B.8), the value of the objective function can be reduced to $f(\text{diag}[\sigma_s^2\mathbf{I}_M - \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{D}_1^{(k)}\mathbf{D}_2^{(k)}])$.

In the following, we derive the structure of the optimal source and relay precoding matrices minimizing the transmit power consumption of the source and the relay. For simplicity of notation, we only consider the case when $N_s = N_r = N_d$. The proof can be easily extended to the case where N_s, N_r, and N_d have different values. From

(B.3) and (B.5), we have $\sigma_s \mathbf{U}_G^{(k)} \mathbf{\Lambda}_G^{(k)} \mathbf{V}_G^{(k)\dagger} \mathbf{P}_k = \mathbf{U}_X^{(k)} \mathbf{\Lambda}_X^{(k)} \mathbf{V}_X^{(k)}$, which can be used to express the source transmit power consumption at frequency tone k as

$$\sigma_s^2 \mathrm{tr}(\mathbf{P}_k \mathbf{P}_k^\dagger) = \mathrm{tr}(\mathbf{\Lambda}_G^{(k)-1} \underbrace{\mathbf{U}_G^{(k)\dagger} \mathbf{U}_X^{(k)}}_{\mathbf{Q}_3^{(k)\dagger}} \mathbf{\Lambda}_X^{(k)2} \underbrace{\mathbf{U}_X^{(k)\dagger} \mathbf{U}_G^{(k)}}_{\mathbf{Q}_3^{(k)}} \mathbf{\Lambda}_G^{(k)-1}) \geq \mathrm{tr}(\bar{\mathbf{\Lambda}}_G^{(k)-2} \mathbf{\Lambda}_X^{(k)2}), \quad \text{(B.11)}$$

where the inequality follows from Lemma 4, and the diagonal matrix $\bar{\mathbf{\Lambda}}_G^{(k)} \in \mathbb{C}^{M \times M}$ contains the M largest singular values of $\mathbf{G}_k$. Therefore, in order to minimize the source transmit power, we need to choose $\mathbf{Q}_3^{(k)} = [\mathbf{I}_M \;\; \mathbf{0}_{M \times (N_s - M)}]$, i.e., $\mathbf{U}_X^{(k)} = \bar{\mathbf{U}}_G^{(k)}$, where $\bar{\mathbf{U}}_G^{(k)}$ contains the M left-most columns of $\mathbf{U}_G^{(k)}$. Recalling from (B.8) that $\mathbf{V}_X^{(k)} = \mathbf{I}_M$, the source matrix can be expressed as

$$\mathbf{P}_k = \frac{1}{\sigma_s} \bar{\mathbf{V}}_G^{(k)} \bar{\mathbf{\Lambda}}_G^{(k)-1} \mathbf{\Lambda}_X^{(k)} = \bar{\mathbf{V}}_G^{(k)} \mathbf{\Lambda}_P^{(k)}, \quad \text{(B.12)}$$

where $\bar{\mathbf{V}}_G^{(k)}$ contains the M left-most columns of $\mathbf{V}_G^{(k)}$ and $\mathbf{\Lambda}_P^{(k)} \triangleq \frac{1}{\sigma_s} \bar{\mathbf{\Lambda}}_G^{(k)-1} \mathbf{\Lambda}_X^{(k)}$. Next, from (B.3) and (B.5) we obtain $\mathbf{U}_H^{(k)} \mathbf{\Lambda}_H^{(k)} \mathbf{V}_H^{(k)\dagger} \mathbf{A}_k = \mathbf{J}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r})^{-1/2}$. Using this result, we can express the relay transmit power consumption at frequency tone k as

$$\begin{aligned} \mathrm{tr}(\mathbf{A}_k (\mathbf{X}_k \mathbf{X}_k^\dagger + \sigma_v^2 \mathbf{I}_{N_r}) \mathbf{A}_k^\dagger) &= \mathrm{tr}(\mathbf{\Lambda}_H^{(k)-1} \mathbf{U}_H^{(k)\dagger} \mathbf{J}_k \mathbf{J}_k^\dagger \mathbf{U}_H^{(k)} \mathbf{\Lambda}_H^{(k)-1}) \\ &= \mathrm{tr}(\mathbf{\Lambda}_H^{(k)-1} \underbrace{\mathbf{U}_H^{(k)\dagger} \mathbf{U}_J^{(k)}}_{\mathbf{Q}_4^{(k)\dagger}} \mathbf{\Lambda}_J^{(k)2} \underbrace{\mathbf{U}_J^{(k)\dagger} \mathbf{U}_H^{(k)}}_{\mathbf{Q}_4^{(k)}} \mathbf{\Lambda}_H^{(k)-1}) \geq \mathrm{tr}(\bar{\mathbf{\Lambda}}_H^{(k)-2} \mathbf{\Lambda}_J^{(k)2}), \end{aligned} \quad \text{(B.13)}$$

where the inequality follows from Lemma 4, and the diagonal matrix $\bar{\mathbf{\Lambda}}_H^{(k)} \in \mathbb{C}^{M \times M}$ contains the M largest singular values of $\mathbf{H}_k$. In (B.13), equality holds for $\mathbf{Q}_4^{(k)} = [\mathbf{I}_M \; \mathbf{0}_{M \times (N_r - M)}]$, i.e., $\mathbf{U}_J^{(k)} = \bar{\mathbf{U}}_H^{(k)}$, where $\bar{\mathbf{U}}_H^{(k)}$ contains the M left-most columns of $\mathbf{U}_H^{(k)}$.

Recalling from (B.8) that $\mathbf{U}_X^{(k)\dagger} = \mathbf{\Omega}^{(k)} = \mathbf{V}_J^{(k)}$, and from (B.11) that $\mathbf{U}_X^{(k)} = \bar{\mathbf{U}}_G^{(k)}$, we obtain for $\mathbf{A}_k$ the expression

$$\mathbf{A}_k = \bar{\mathbf{V}}_H^{(k)} \bar{\mathbf{\Lambda}}_H^{(k)-1} \mathbf{\Lambda}_J^{(k)} (\mathbf{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2} \bar{\mathbf{U}}_G^{(k)\dagger} = \bar{\mathbf{V}}_H^{(k)} \mathbf{\Lambda}_A^{(k)} \bar{\mathbf{U}}_G^{(k)\dagger}, \tag{B.14}$$

where $\bar{\mathbf{V}}_H^{(k)}$ and $\bar{\mathbf{U}}_G^{(k)}$ contain the M left-most columns of $\mathbf{V}_H^{(k)}$ and $\mathbf{U}_G^{(k)}$, respectively, and $\mathbf{\Lambda}_A^{(k)} \triangleq \bar{\mathbf{\Lambda}}_H^{(k)-1} \mathbf{\Lambda}_J^{(k)} (\mathbf{\Lambda}_X^{(k)2} + \sigma_v^2 \mathbf{I}_M)^{-1/2}$. Hence, we have proved that the expressions for the source and relay precoding matrices given in (B.12) and (B.14) minimize the objective function $f(\text{diag}[\hat{\mathbf{E}}])$, cf. (B.10), as well as the transmit power consumption at the source and the relay, cf. (B.11) and (B.13).

Now, we turn our attention to the case when $f(\text{diag}[\hat{\mathbf{E}}])$ is a Schur-convex increasing function w.r.t. $\text{diag}[\hat{\mathbf{E}}]$. From Lemma 5 we know that $\frac{1}{M}\text{tr}(\hat{\mathbf{E}})\mathbf{1} \prec \text{diag}[\hat{\mathbf{E}}]$. Combining this fact and Definition 2, we obtain the inequality

$$f(\text{diag}[\hat{\mathbf{E}}]) \geq f(\frac{1}{M}\text{tr}(\hat{\mathbf{E}})\mathbf{1}), \tag{B.15}$$

where equality holds when the diagonal entries of $\hat{\mathbf{E}}$ are all equal to $\frac{1}{M}\text{tr}(\hat{\mathbf{E}})$. In the following, we show that by applying a unitary rotation to the source precoding matrix, we can achieve this equality. Applying in (3.13) the eigenvalue decomposition,

$$\sigma_s^2 \mathbf{Q}_k^\dagger \left(\sigma_v^2 \mathbf{H}_k \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k^\dagger + \sigma_u^2 \mathbf{I}_{N_d} \right)^{-1} \mathbf{Q}_k = \mathbf{U}_E^{(k)} \mathbf{\Lambda}_E^{(k)} \mathbf{U}_E^{(k)\dagger}, \tag{B.16}$$

where $\mathbf{U}_E^{(k)} \in \mathbb{C}^{M \times M}$, $\mathbf{\Lambda}_E^{(k)} \in \mathbb{C}^{M \times M}$, we obtain $\mathbf{\Psi}_k^{-1} = \mathbf{U}_E^{(k)} (\mathbf{I}_M + \mathbf{\Lambda}_E^{(k)})^{-1} \mathbf{U}_E^{(k)\dagger}$. Let us consider the feasible source precoding matrix $\bar{\mathbf{P}}_k = \mathbf{P}_k \mathbf{U}_E^{(k)} \mathbf{V}_0$, where $\mathbf{U}_E^{(k)} \mathbf{V}_0$ is a unitary matrix and thus does not affect the power constraints. Replacing $\mathbf{P}_k$ with

$\bar{\mathbf{P}}_k$ in (3.13) and using (B.16), we can obtain

$$\mathbf{\Psi}_k^{-1} = \mathbf{V}_0^{\dagger}(\mathbf{I}_M + \mathbf{\Lambda}_E^{(k)})^{-1}\mathbf{V}_0, \tag{B.17}$$

which allows us to express the MSE matrix as

$$\hat{\mathbf{E}} = \frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\mathbf{\Psi}_k^{-1} = \frac{\sigma_s^2}{N_c}\mathbf{V}_0^{\dagger}\sum_{k=0}^{N_c-1}(\mathbf{I}_M + \mathbf{\Lambda}_E^{(k)})^{-1}\mathbf{V}_0. \tag{B.18}$$

Since $\sum_{k=0}^{N_c-1}(\mathbf{I}_M + \mathbf{\Lambda}_E^{(k)})^{-1}$ is the sum of N_c diagonal matrices, it is also a diagonal matrix. Based on Lemma 7, we conclude that there exists a unitary matrix $\mathbf{V}_0$ such that $\hat{\mathbf{E}}$ has identical diagonal elements given by $\frac{1}{M}\mathrm{tr}(\hat{\mathbf{E}})$. Since the objective function is an increasing function w.r.t. its arguments, minimizing the original Schur-convex objective function is now equivalent to minimizing $\frac{1}{M}\mathrm{tr}(\hat{\mathbf{E}})$, which is a Schur-concave function. Therefore, the optimal structures of $\mathbf{P}_k$ and $\mathbf{A}_k$ are given by (B.12) and (B.14), respectively. Furthermore, as the resulting $\mathbf{U}_E^{(k)}$ can be shown to be an identity matrix, the source precoding matrix for Schur-convex functions is given by $\bar{\mathbf{P}}_k = \mathbf{P}_k\mathbf{V}_0 = \bar{\mathbf{V}}_G^{(k)}\mathbf{\Lambda}_P^{(k)}\mathbf{V}_0$.

Appendix C

Proof of Proposition 3.1

To show that the objective function in (3.40) is jointly convex w.r.t. $P_{r,km}$ and $P_{s,km}$ when Φ_{km} is approximated by $\tilde{\Phi}_{km}$, let us first examine the elements of the Hessian matrix, $\nabla\mathbf{F} = \begin{bmatrix} \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{s,km}^2} & \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{s,km}\partial P_{r,km}} \\ \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{r,km}\partial P_{s,km}} & \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{r,km}^2} \end{bmatrix}$, of $\tilde{\Phi}_{km}$ w.r.t. $P_{r,km}$ and $P_{s,km}$:

$$\begin{aligned} \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{s,km}^2} &= \frac{-2\sigma_v^2 P_{r,km}^2 h_{km}^4 g_{km}^4}{\left[\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 P_{s,km} g_{km}^2\right]^3} \\ \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{r,km}\partial P_{s,km}} &= \frac{2\sigma_u^2\sigma_v^2 P_{r,km} P_{s,km} h_{km}^4 g_{km}^4}{\left[\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 P_{s,km} g_{km}^2\right]^3}, \\ \frac{\partial^2\tilde{\Phi}_{km}}{\partial P_{r,km}^2} &= \frac{-2\sigma_u^2 P_{s,km}^2 g_{km}^4 h_{km}^4}{\left[\sigma_v^2 P_{r,km} h_{km}^2 + \sigma_u^2 P_{s,km} g_{km}^2\right]^3}. \end{aligned} \tag{C.1}$$

One can verify that the trace and the determinant of $\nabla\mathbf{F}$ are given by

$$\mathrm{tr}(\nabla\mathbf{F}) = \omega_1 + \omega_2 \leq 0, \quad \text{and} \quad \det(\nabla\mathbf{F}) = \omega_1\omega_2 = 0, \tag{C.2}$$

where ω_1 and ω_2 are the eigenvalues of $\nabla\mathbf{F}$. It can be inferred from (C.2) that one of ω_1 and ω_2 is zero and the other one is non-positive. Since the Hessian matrix only has non-positive eigenvalues, we conclude $\tilde{\Phi}_{km}$ in (3.55) is jointly concave w.r.t. $P_{r,km}$ and $P_{s,km}$. Now, we are ready to prove the convexity of the objective function. For

the AMSE criterion, we have

$$\begin{aligned}\frac{\partial f_{\text{AMSE}}(\tilde{\mathbf{\Phi}})}{\partial \tilde{\Phi}_{km}} &= -\frac{\sigma_s^2}{N_c}\tilde{\Phi}_{km}^{-2} < 0,\\ \frac{\partial f^2_{\text{AMSE}}(\tilde{\mathbf{\Phi}})}{\partial^2 \tilde{\Phi}_{km}} &= \frac{2\sigma_s^2}{N_c}\tilde{\Phi}_{km}^{-3} > 0,\end{aligned} \tag{C.3}$$

where $\tilde{\mathbf{\Phi}} = \{\tilde{\Phi}_{km}, \forall, k, m\}$. Therefore, $f_{\text{AMSE}}(\tilde{\mathbf{\Phi}})$ is convex decreasing w.r.t. $\tilde{\Phi}_{km}$, and by the composition rule [101], we can deduce that $f_{\text{AMSE}}(\tilde{\mathbf{\Phi}})$ is jointly convex w.r.t. $P_{r,km}$ and $P_{s,km}$. For the GMSE criterion, we rewrite the objective function as

$$f_{\text{GMSE}}(\tilde{\mathbf{\Phi}}) = \sum_{m=1}^{M} \log_2\left(\frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\tilde{\Phi}_{km}^{-1}\right) = \sum_{m=1}^{M} \log_2\left(\frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\exp\left(-\log\tilde{\Phi}_{km}\right)\right). \tag{C.4}$$

Since $-\log\tilde{\Phi}_{km}$ is the composition of a convex decreasing function, $-\log x$, and a concave function, $\tilde{\Phi}_{km}$, from the composition rule, it is a joint convex function w.r.t. $P_{r,km}$ and $P_{s,km}$. On the other hand, note that $\log_2\left(\frac{\sigma_s^2}{N_c}\sum_{k=0}^{N_c-1}\exp y_k\right)$ is a convex increasing function w.r.t. y_k [101], therefore, by using the composition rule, $f_{\text{GMSE}}(\tilde{\mathbf{\Phi}})$ is a jointly convex function w.r.t. $P_{r,km}$ and $P_{s,km}$. Based on the above results, along with the fact that the power constraints are all affine w.r.t. $P_{r,km}$ and $P_{s,km}$, we conclude that the problem in (3.40) with Φ_{km} approximated by $\tilde{\Phi}_{km}$ is a convex optimization problem.

Appendix D

Proof of Proposition 4.1

The Hessian matrix of $\bar{\Phi}_k^{(u)}$ w.r.t. $P_{ku}, \forall k, u$ is a diagonal matrix whose diagonal entries are given by $\frac{\partial^2 \bar{\Phi}_k^{(u)}}{\partial P_{ku}^2} = \sum_{i=1}^{N_R} \frac{-A_{ik}^{(u)} B_{ik}^{(u)} C_{ik}^{(u)}}{(P_{ku} A_{ik}^{(u)} + B_{ik}^{(u)})^3} \leq 0$. Therefore, $\bar{\Phi}_k^{(u)}$ is a concave function in $P_{ku}, \forall k, u$. For OFDMA, since $\log_2(x)$ is a concave increasing function, according to the composition rule [101], $\log_2(\bar{\Phi}_k^{(u)} + 1)$ is a concave function, implying that $f_1^{[\mathrm{OFDMA}]}(\mathbf{p})$ is concave w.r.t. $P_{ku}, \forall k, u$. For SC-FDMA, we rewrite $\log_2 \left[\frac{1}{Q} \sum_{k=1}^{Q} (\bar{\Phi}_k^{(u)} + 1)^{-1} \right]$ as $\log_2 \left[\frac{1}{Q} \sum_{k=1}^{Q} \exp(-\log(\bar{\Phi}_k^{(u)} + 1)) \right]$. Since $-\log[(\bar{\Phi}_k^{(u)} + 1)]$ is convex w.r.t. $(\bar{\Phi}_k^{(u)} + 1)$, and $\log_2 \left(\frac{1}{Q} \sum_{k=1}^{Q} \exp y_k \right)$ is convex and increasing w.r.t. y_k [101], the composition of the two is also convex. Thus, $f_1^{[\mathrm{SC-FDMA}]}(\mathbf{p})$ is a concave function in $P_{ku}, \forall k, u$. On the other hand, the constraints are affine functions in $P_{ku}, \forall k, u$. Hence, problem (5.33) is a concave maximization problem with affine constraints, i.e., (5.33) is a convex optimization problem.

Appendix E

Proof of Proposition 4.2

We first show that the objective function $f_2^{[\mathrm{Y}]}(\mathbf{a}, \mathbf{X})$ is a convex function in its optimization variables $\{\mathbf{a}, \mathbf{X}\}$. Given that $\mathbf{Z}_k^{(u)}$ is a Hermitian positive semidefinite matrix and $X_{ku} \geq 0, \forall k, u$, it can be shown that $\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}$ is convex in $\mathbf{a}_k^{(u)}$ and X_{ku} [101], which further implies that $(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}})$ is a concave function in $\mathbf{a}_k^{(u)}$ and X_{ku}. For OFDMA, the composite function of $-\log(\cdot)$ and $(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}})$ is convex due to the composition rule and the fact that $-\log(\cdot)$ is a convex decreasing function. The convexity of $f_2^{[\mathrm{OFDMA}]}(\mathbf{a}, \mathbf{X})$ follows since the weighted sum of convex functions is also convex. On the other hand, for SC-FDMA, $\frac{1}{Q}\sum_{k=1}^{Q}\left(1-\frac{\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}}{X_{ku}}\right)$ is concave in $\mathbf{a}_k^{(u)}$ and X_{ku} because it is an average of Q concave functions. Based on this fact and by applying the same argument as for OFDMA, the convexity of $f_2^{[\mathrm{SC-FDMA}]}(\mathbf{a}, \mathbf{X})$ follows.

Next, we consider the convexity of the constraints in problem (4.48). The first constraint (4.48b) is convex in $\{|a_{ik}^{(u)}|\}$ because the second-order derivative of the function on the left hand side of (4.48b) is strictly greater than zero w.r.t. $|a_{ik}^{(u)}|, \forall k, u$. For the second constraint, given that $\mathbf{Z}_k^{(u)}$ and $\mathbf{D}_k^{(u)}$ are Hermitian positive semidefinite matrices, $\mathbf{a}_k^{(u)T}\mathbf{Z}_k^{(u)}\mathbf{a}_k^{(u)}$ and $\mathbf{a}_k^{(u)T}\mathbf{D}_k^{(u)}\mathbf{a}_k^{(u)}$ are both convex functions in $\mathbf{a}_k^{(u)}$. This, along with the fact that $-X_{ku}$ is an affine mapping of X_{ku} shows that the second constraint is convex in $\{\mathbf{a}_k^{(u)}, X_{ku}\}$.

Appendix F

Proof of Proposition 5.1

For fixed $\{P_{k,i}, \forall k, i \geq 2\}$, we observe that the term $\sum_{i=2}^{M} \sum_{k=1}^{N_c} \Psi_{k,(i+1)}$ in the objective function of (5.23) becomes a constant. Thus, the problem with respect to $\{P_{k,1}\}$ becomes

$$\min_{\{P_{k,1}\}} \; g_0(\mathbf{p}_1) \quad \text{s.t.} \quad \sum_{k=1}^{N_c} P_{k,1} \leq \bar{P}_1, \; P_{k,1} \geq 0, \forall k, \tag{F.1}$$

where $g_0(\mathbf{p}_1) = \sum_{k=1}^{N_c} \Psi_{k,2}$, and $\mathbf{p}_1 = \{P_{k,1}, \forall k\}$. From the first and second order derivatives of $g_0(\mathbf{p}_1)$ with respect to $\{P_{k,1}\}$,

$$\frac{\partial g_0(\mathbf{p}_1)}{\partial P_{k,1}} = \frac{-|\hat{h}_{k,1}|^2}{\left[(|\hat{h}_{k,1}|^2 + \sigma^2_{e,h_1})P_{k,1} + \sigma^2_{e,l_2} P_{k,2} + \sigma^2_{e,g_3} P_{k,3} + \sigma^2_{n_2}\right]^2} < 0,$$

$$\frac{\partial^2 g_0(\mathbf{p}_1)}{\partial P^2_{k,1}} = \frac{2|\hat{h}_{k,1}|^2(|\hat{h}_{k,1}|^2 + \sigma^2_{e,h_1})}{\left[(|\hat{h}_{k,1}|^2 + \sigma^2_{e,h_1})P_{k,1} + \sigma^2_{e,l_2} P_{k,2} + \sigma^2_{e,g_3} P_{k,3} + \sigma^2_{n_2}\right]^3} > 0,$$

we find that $g_0(\mathbf{p}_1)$ is a monotonically decreasing convex function. As the power constraints are affine functions, it follows that problem (F.1) is a convex problem. Thus, from the Karush-Kuhn-Tucker (KKT) conditions [101], we can obtain the solution of $(F.1)$ given in (5.24).

The Lagrangian of the problem is given by

$$\mathcal{L}_1(\mathbf{p}_1, \lambda_1, \boldsymbol{\beta}_1) = \sum_{k=1}^{N_c} \Psi_{k,2} + \lambda_1 (\sum_{k=1}^{N_c} P_{k,1} - \bar{P}_1) - \sum_{k=1}^{N_c} \beta_{k,1} P_{k,1}, \tag{F.2}$$

where λ_1 and $\{\beta_{k,1}\}$ are the Lagrange multipliers associated with the corresponding constraints in (F.1). The Karush-Kuhn-Tucker (KKT) conditions [101] can then be derived as

$$\frac{\partial \mathcal{L}_1(\mathbf{p}_1, \lambda_1, \boldsymbol{\beta}_1)}{\partial P_{k,1}} = \frac{\partial g_0(\mathbf{p}_1)}{\partial P_{k,1}} + \lambda_1 - \beta_{k,1} = 0, \tag{F.3a}$$

$$\lambda_1 (\sum_{k=1}^{N_c} P_{k,1} - \bar{P}_1) = 0, \tag{F.3b}$$

$$\lambda_1 \geq 0, \quad (\sum_{k=1}^{N_c} P_{k,1} - \bar{P}_1) \leq 0, \tag{F.3c}$$

$$\beta_{k,1} P_{k,1} = 0, \ \beta_{k,1} \geq 0, \ P_{k,1} \geq 0. \tag{F.3d}$$

Since $\frac{\partial g_0(\mathbf{p}_1)}{\partial P_{k,1}} < 0, \beta_{k,1} \geq 0$, in order for (F.3a) to hold, we must have $\lambda_1 > 0$. This implies that the transmit power constraint is met with equality, i.e., $\sum_{k=1}^{N_c} P_{k,1} = \bar{P}_1$, due to the complementary slackness condition in (F.3b). Furthermore, from (F.3d), we observe that if $\beta_{k,1} = 0$, then $P_{k,1} > 0$, otherwise $P_{k,1} = 0$. From (F.3a), we obtain $\beta_{k,1} = \lambda_1 + \frac{\partial g_0(\mathbf{p}_1)}{\partial P_{k,1}}$. Therefore, the positive solution of $P_{k,1}$ can be found by solving the equation $\lambda_1 + \frac{\partial g_0(\mathbf{p}_1)}{\partial P_{k,1}} = 0$ for a given λ_1. If the equation has no positive solution, we set $P_{k,1} = 0$. Based on the above discussions, we obtain the solution for $P_{k,1}$ given in (5.24).

Appendix G

Proof of Proposition 5.2

For given $\{P_{k,j}, \forall k\}, j \neq i$, the objective function with respect to $\{P_{k,i}\}, i = 2, \ldots, M$ is a difference of two convex functions. To show this, let use rewrite the objective function of problem (5.16) as $t_i(\mathbf{p}_i) + g_i(\mathbf{p}_i)$, where

$$t_i(\mathbf{p}_i) = \sum_{u=1}^{2} \sum_{k=1}^{N_c} \left(1 - \frac{B_{k,i,u}}{A_{k,i,u} P_{k,i} + C_{k,i,u}}\right), \quad g_i(\mathbf{p}_i) = \sum_{k=1}^{N_c} \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1\right)^{-1}, \tag{G.1}$$

with $\{T_{k,i}, A_{k,i,u}, B_{k,i,u}, C_{k,i,u}, \forall k, i, u\}$ and $\mathbf{p}_i$ defined in (5.26). Next, we examine the first-order and second-order derivatives of $t_i(\mathbf{p}_i)$ and $g_i(\mathbf{p}_i)$, respectively, as follows:

$$\begin{aligned}
&\frac{\partial t_i(\mathbf{p}_i)}{\partial P_{k,i}} = \sum_{u=1}^{2} \frac{A_{k,i,u} B_{k,i,u}}{\left(A_{k,i,u} P_{k,i} + C_{k,i,u}\right)^2} > 0, \quad \frac{\partial^2 t_i(\mathbf{p}_i)}{\partial P_{k,i}^2} = \sum_{u=1}^{2} \frac{-2 A_{k,i,u}^2 B_{k,i,u}}{\left(A_{k,i,u} P_{k,i} + C_{k,i,u}\right)^3} < 0, \\
&\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}} = \frac{-|\hat{h}_{k,i}|^2 T_{k,i}}{\left[(|\hat{h}_{k,i}|^2 + \sigma_{e,h_i}^2) P_{k,i} + T_{k,i}\right]^2} < 0, \\
&\frac{\partial^2 g_i(\mathbf{p}_i)}{\partial P_{k,i}^2} = \frac{2|\hat{h}_{k,i}|^2 (|\hat{h}_{k,i}|^2 + \sigma_{e,h_i}^2) T_{k,i}}{\left[(|\hat{h}_{k,i}|^2 + \sigma_{e,h_i}^2) P_{k,i} + T_{k,i}\right]^3} > 0.
\end{aligned} \tag{G.2}$$

Therefore, $t_i(\mathbf{p}_i)$ and $g_i(\mathbf{p}_i)$ are concave increasing and convex decreasing functions with respect to $\{P_{k,i}\}$, respectively. By letting $f_i(\mathbf{p}_i) = -t_i(\mathbf{p}_i)$, we obtain the difference of convex (d.c.) programming problem formulation in (5.25).

Appendix H

Proof of Proposition 5.3

We first note the following relationship between the convex function $f_i(\mathbf{p}_i)$ and its first order Taylor series expansion at a feasible point $P_{k,i}^{[\ell]}$:

$$f_i(\mathbf{p}_i) \geq \sum_{u=1}^{2}\sum_{k=1}^{N_c}\left[\frac{B_{k,i,u}}{A_{k,i,u}P_{k,i}^{[\ell]}+C_{k,i,u}} - 1 - \sum_{k=1}^{N_c}\frac{A_{k,i,u}B_{k,i,u}(P_{k,i}-P_{k,i}^{[\ell]})}{(A_{k,i,u}P_{k,i}^{[\ell]}+C_{k,i,u})^2}\right] = f_i^{[\ell]}(\mathbf{p}_i).$$

If we replace $f_i(\mathbf{p}_i)$ with $f_i^{[\ell]}(\mathbf{p}_i)$ in problem (5.25), we obtain a convex majorant of the objective function as

$$\begin{aligned} h_i^{[\ell]}(\mathbf{p}_i) \triangleq & g_i(\mathbf{p}_i) - f_i^{[\ell]}(\mathbf{p}_i) \\ = & g_i(\mathbf{p}_i) - \sum_{u=1}^{2}\sum_{k=1}^{N_c}\left[\frac{B_{k,i,u}}{A_{k,i,u}P_{k,i}^{[\ell]}+C_{k,i,u}} - 1 + \frac{A_{k,i,u}B_{k,i,u}(P_{k,i}-P_{k,i}^{[\ell]})}{(A_{k,i,u}P_{k,i}^{[\ell]}+C_{k,i,u})^2}\right]. \end{aligned} \quad \text{(H.1)}$$

Since $h_i^{[\ell]}(\mathbf{p}_i) \geq g_i(\mathbf{p}_i) - f_i(\mathbf{p}_i), \forall\ \mathbf{p}_i$, the following convex problem minimizes the upper bound of d.c. problem (5.25)

$$\min_{\{P_{k,i}\}} \ h_i^{[\ell]}(\mathbf{p}_i), \quad \text{s.t.} \ \sum_{k=1}^{N_c} P_{k,i} \leq \bar{P}_i, \ P_{k,i} \geq 0, \ \ \forall k. \quad \text{(H.2)}$$

By further dropping the constant terms that are irrelevant to the optimization, we can simplify the above convex problem as in (5.27). Let $P_{k,i}^{[\ell+1]}$ be the optimal solution of (5.27). By replacing $P_{k,i}^{[\ell]}$ with $P_{k,i}^{[\ell+1]}$ (i.e., $D_{k,i,u}^{[\ell]}$ becomes $D_{k,i,u}^{[\ell+1]}$) and solving the resulting convex problem again, an improved feasible solution $P_{k,i}^{[\ell+2]}$ for the original

d.c. problem (5.25) can be found. This procedure is then repeated until the value of the objective function in (5.25) converges.

We now show that by using the above iterative procedure, a stationary point (KKT point) of the d.c. problem can be obtained. First, we note that the d.c. problem and the convex problem share the same feasible set and are both strictly feasible. Thus, any point $\mathbf{p}_i^{[\ell]}$ which is feasible for problem (5.27) is also feasible for problem (5.25). Therefore, the sequence $\{\hat{h}_i^{[\ell]}(\mathbf{p}_i)\}$ generated by the convex problem monotonically decreases as the number of iterations ℓ increases. Since $\hat{h}_i^{[\ell]}(\mathbf{p}_i)$ is lower bounded by zero, the convergence of the sequence $\{\hat{h}^{[\ell]}(\mathbf{p}_i)\}$ and thus the convergence of the iterative algorithm for the d.c. problem is guaranteed. Due to the convexity of the problem, there is a unique correspondence between the sequences $\{\hat{h}_i^{[\ell]}(\mathbf{p}_i)\}$ and $\{\mathbf{p}_i^{[\ell]}\}$. Thus, the sequence $\{\mathbf{p}_i^{[\ell]}\}$ also converges to a limit point denoted by $\mathbf{p}_i^\star$. Furthermore, it can be verified that the limit point $\mathbf{p}_i^\star$ with the optimal Lagrange multipliers $\{\lambda_i^\star, \beta_{k,i}^\star\}$ also satisfies the KKT conditions of the original d.c. problem and is thus a stationary point (KKT point) of problem (5.25). From convex analysis, the limit point $\mathbf{p}_i^\star$ with the optimal Lagrange multipliers $\{\lambda_i^\star, \beta_{k,i}^\star\}$ satisfies the KKT conditions

$$\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}^\star} + \lambda_i^\star + \sum_{u=1}^{2} \frac{A_{k,i,u} B_{k,i,u} P_{k,i}^\star}{(A_{k,i,u} P_{k,i}^\star + C_{k,i,u})^2} - \beta_{k,i}^\star = 0 \tag{H.3a}$$

$$\lambda_i^\star (\sum_k P_{k,i}^\star - \bar{P}_i) = 0, \quad \beta_{k,i}^\star P_{k,i}^\star = 0. \tag{H.3b}$$

For the original d.c. problem, we know that a stationary point, $P_{k,i}^{\text{stat}}$, satisfies the

following KKT conditions

$$\frac{\partial g_i(\mathbf{p}_i)}{\partial P_{k,i}^{\text{stat}}} + \zeta_i + \sum_{u=1}^{2} \frac{A_{k,i,u} B_{k,i,u} P_{k,i}^{\text{stat}}}{(A_{k,i,u} P_{k,i}^{\text{stat}} + C_{k,i,u})^2} - \eta_{k,i} = 0 \tag{H.4a}$$

$$\zeta_i (\sum_k P_{k,i}^{\text{stat}} - \bar{P}_i) = 0, \quad \eta_{k,i} P_{k,i}^{\text{stat}} = 0, \tag{H.4b}$$

where ζ_i and $\eta_{k,i}$ are the Lagrangian multipliers for the constraints of the d.c. problem. By choosing $\zeta_i = \lambda_i^\star$, and $\eta_{k,i} = \beta_{k,i}^\star$, the above two sets of KKT conditions become identical. Therefore, the limit point $\mathbf{p}_i^\star$ with the associated Lagrange multipliers $\{\lambda_i^\star, \beta_{k,i}^\star\}$ satisfies the KKT conditions of the d.c. problem and is thus a stationary point (KKT point). It is known that a stationary point can be a saddle point, a local minima and a local maxima. Due to the fact that the objective function of the d.c. problem is non-convex, we can not guaranteed that the obtained solution is a local minima. Instead, we claim we can at least find a stationary point solution to the d.c. problem. In the following, we provide a closed-form solution for problem (5.27). To this end, we write its Lagrangian as

$$\begin{aligned}\mathcal{L}(\mathbf{p}_i, \lambda_i, \boldsymbol{\beta}_i) = &\sum_{k=1}^{N_c} \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1 \right)^{-1} + \sum_{k,u} D_{k,i,u}^{[\ell]} P_{k,i} + \lambda_i (\sum_{k=1}^{N_c} P_{k,i} - \bar{P}_i) \\ &- \sum_{k=1}^{N_c} \beta_{k,i} P_{k,i},\end{aligned}$$

where λ_i and $\{\beta_{k,i}, \forall k, i\}$ are the Lagrange multipliers associated with the corresponding constraints in (5.27). The dual problem of (5.27) can be written as

$$\max_{\lambda_i} \quad \min_{\{P_{k,i}\}, \{\beta_{k,i}\}} \quad \mathcal{L}(\mathbf{p}_i, \lambda_i, \boldsymbol{\beta}_i) = \max_{\lambda_i} \quad g(\lambda_i), \tag{H.5}$$

where $g(\lambda_i)$ is the dual function. The dual problem can be further decomposed into

N_c subproblems as follows

$$\min_{\{P_{k,i}\},\{\beta_{k,i}\}} \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1 \right)^{-1} + \left(\sum_{u=1}^{2} D_{k,i,u}^{[\ell]} + \lambda_i - \beta_{k,i} \right) P_{k,i}, \quad \forall k. \tag{H.6}$$

Applying the KKT conditions to the subproblem in (H.6), the solution for $P_{k,i}$ given in (5.28) is obtained. The optimal Lagrange multiplier λ_i can be found by solving the maximization problem with respect to λ_i. It can be verified that the subgradient of λ_i is given by $\sum_{k=1}^{N_c} P_{k,i} - \bar{P}_i$. Therefore, we can adopt the subgradient method in (5.29) for updating λ_i.

Appendix I

Proof of Proposition 5.4

From convex analysis, the following inequality holds between the convex function $\psi_{i,u}(\mathbf{p}_i)$ and its first-order Taylor series expansion at a feasible point $\{P_{k,i}^{[\ell]}\}$,

$$\psi_{i,u}(\mathbf{p}_i) \geq \sum_{k=1}^{N_c} \left(\frac{B_{k,i,u}}{A_{k,i,u}\hat{P}_{k,i} + C_{k,i,u}} - 1 - D_{k,i,u}^{[\ell]}(P_{k,i} - P_{k,i}^{[\ell]}) \right) = \psi_{i,u}^{[\ell]}(\mathbf{p}_i), \tag{I.1}$$

Since $\psi_{i,u}^{[\ell]}(\mathbf{p}_i)$ is a convex function with respect to $\mathbf{p}_i$, replacing $\psi_{i,u}(\mathbf{p}_i)$ with $\psi_{i,u}^{[\ell]}(\mathbf{p}_i)$ in constraint (5.34a) leads to a stricter constraint which requires that the upper bound of the MSE is no larger than t. The resulting problem is given in (5.35), which is a convex optimization problem whose feasible set lies in the subset of the original d.c. problem. Therefore, the optimal solution of (5.35), denoted as $\{\mathbf{p}_i^{[\ell+1]}\}$, is also feasible for the d.c. problem. Then, by applying the first-order Taylor series expansion at $\{\mathbf{p}_i^{[\ell+1]}\}$ and repeating the above process, the objective value can be further decreased and a better approximate solution for the original d.c. problem (5.34) can be found. Similar to the case of sum MSE minimization, we can show that by iteratively solving (5.35), a stationary point (KKT point) of the d.c. problem (5.34) can be obtained. The details are omitted here for brevity. It can be verified that Slater's condition is satisfied for problem (5.35). Thus, we can solve (5.35) by solving its dual problem. Upon rearranging the terms, the Lagrangian of (5.35) can

be written as

$$\begin{aligned}
&\mathcal{L}\left(\mathbf{p}_i, \lambda_i, \gamma_u, \boldsymbol{\beta}_i, t\right) \\
&= \sum_{k=1}^{N_c} \left[\gamma_3 \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1 \right)^{-1} + (\sum_{u=1}^{2} \gamma_u D_{k,i,u}^{[\ell]} - \beta_{k,i} + \lambda_i) P_{k,i} + \sum_{u=1}^{2} \gamma_u (E_{k,i,u}^{[\ell]} \right. \\
&\quad \left. + D_{k,i,u}^{[\ell]} P_{k,i}^{[\ell])}) \right] + \left(1 - \sum_{u=1}^{3} \gamma_u \right) t - \lambda_i \hat{P}_2, \qquad \text{(I.2)}
\end{aligned}$$

where $\{\gamma_u, u = 1, 2, 3\}$ are the Lagrangian multiplier associated with the constraints (5.35b) and (5.35c), respectively. The dual problem of (5.35) is thus given by

$$\max_{\{\lambda_i, \gamma_u\}} \quad \min_{t, \{P_{k,i}\}, \{\beta_{k,i}\}} \quad \mathcal{L}\left(\mathbf{p}_i, \lambda_i, \gamma_u, \boldsymbol{\beta}_i, t\right) = \max_{\lambda_i, \gamma_u} \quad g(\lambda_i, \gamma_u). \qquad \text{(I.3)}$$

Note that the inner minimization problem is separable with respect to t and $P_{k,i}$, respectively. For t, we have the following unconstrained minimization problem

$$\min_{t} \quad \left(1 - \sum_{u=1}^{3} \gamma_u \right) t \qquad \text{(I.4)}$$

If $1 - \sum_{u=1}^{3} \gamma_u \neq 0$, the value of the dual function $g(\lambda_i, \gamma_u)$ is unbounded from below as t tends to negative infinity. Therefore, to make the minimization problem bounded from below, we should have $1 - \sum_{u=1}^{3} \gamma_u = 0$. On the other hand, the inner minimization with respect to $\{P_{k,i}\}$ can be decomposed into N_c subproblems as

$$\min_{\{P_{k,i}\}, \{\beta_{k,i}\}} \gamma_3 \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1 \right)^{-1} + \left(\sum_{u=1}^{2} \gamma_u D_{k,i,u}^{[\ell]} + \lambda_i - \beta_{k,i} \right) P_{k,i}. \qquad \text{(I.5)}$$

Applying the KKT conditions to (I.5) and using the fact that $\gamma_3 = 1 - \sum_{u=1}^{2} \gamma_u$, which is the optimal condition for t, we obtain the solution of $P_{k,i}$ given in (5.36).

Next, we determine the optimal Lagrange multipliers by solving the dual problem in (I.3) using the subgradient method.

Lemma 1: The subgradients of the Lagrange multipliers (dual variables) λ_i and γ_1 are given by $(\sum_{k=1}^{N_c} P_{k,i} - \bar{P}_i)$ and $\sum_{k=1}^{N_c} D_{k,i,u}^{[\ell]}(P_{k,i} - P_{k,i}^{(\ell)}) + E_{k,i}^{[\ell]} - \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i} + T_{k,i}} + 1\right)^{-1}$, respectively.

Proof. According to the definition in [101], s is a subgradient of a function $f(x)$ (not necessarily convex) at x if

$$f(y) \geq f(x) + s(y - x), \forall y. \tag{I.6}$$

Let $\{\tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2\}$ and $\{\hat{\lambda}_i, \hat{\gamma}_1, \hat{\gamma}_2\}$ be two feasible solutions of the dual problem, $\max_{\lambda_i, \gamma_u} g(\lambda_i, \gamma_u)$. Assume that $\{\tilde{\mathbf{p}}_i, \tilde{\boldsymbol{\beta}}_2, \tilde{t}\}$ and $\{\hat{\mathbf{p}}_i, \hat{\boldsymbol{\beta}}_i, \hat{t}\}$ are the optimal solutions of the primal variables corresponding to these two sets of dual variables. Then, we have

$$\begin{aligned}
g(\tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2) = & \min_{t, \{P_{k,i}\}, \{\tilde{\beta}_{k,i}\}} \quad \mathcal{L}\left(\mathbf{p}_i, t, \boldsymbol{\beta}_i, \tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2\}\right) = \mathcal{L}\left(\tilde{\mathbf{p}}_i, \tilde{t}, \tilde{\boldsymbol{\beta}}_2, \tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2\right) \\
& \overset{(a)}{\leq} \mathcal{L}\left(\hat{\mathbf{p}}_i, \hat{t}, \hat{\boldsymbol{\beta}}_i, \tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2\right) \\
& = \mathcal{L}\left(\hat{\mathbf{p}}_i, \hat{t}, \hat{\boldsymbol{\beta}}_i, \hat{\lambda}_i, \hat{\gamma}_1, \hat{\gamma}_2, \right) + \mathcal{L}\left(\hat{\mathbf{p}}_i, \hat{t}, \hat{\boldsymbol{\beta}}_i, \tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2\right) - \mathcal{L}\left(\hat{\mathbf{p}}_i, \hat{t}, \hat{\boldsymbol{\beta}}_i, \hat{\lambda}_i, \hat{\gamma}_1, \hat{\gamma}_2\right) \\
& = \mathcal{L}\left(\hat{\mathbf{p}}_i, \hat{t}, \hat{\boldsymbol{\beta}}_i, \hat{\lambda}_i, \hat{\gamma}_1, \hat{\gamma}_2, \right) + (\tilde{\lambda}_i - \hat{\lambda}_i)\left(\sum_{k=1}^{N_c} \hat{P}_{k,i} - \bar{P}_i\right) + (\tilde{\beta}_i - \hat{\beta}_i)\hat{P}_{k,i} \\
& \quad + (\tilde{\gamma}_1 - \hat{\gamma}_1)\left[\sum_{k=1}^{N_c}\left(\frac{|\hat{h}_{k,i}|^2 \hat{P}_{k,i}}{\sigma_{e,h_i}^2 \hat{P}_{k,i} + T_{k,i}} + 1\right)^{-1} - t\right] \\
& \quad + (\tilde{\gamma}_2 - \hat{\gamma}_2)\left[\sum_{k=1}^{N_c}\left(E_{k,i}^{[\ell]} + D_{k,i,u}^{[\ell]}(\hat{P}_{k,i} - P_{k,i}^{[\ell]})\right) - t\right],
\end{aligned} \tag{I.7}$$

where inequality (a) is due to the fact that $\{\tilde{\mathbf{p}}_i, \tilde{\boldsymbol{\beta}}_2, \tilde{t}\}$ is the global minimizer of the Lagrangian for the given set of dual variables $\{\tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2\}$. Next, by substituting the

optimality condition for t, i.e., $\gamma_2 = 1 - \gamma_1$, into the (I.7), we get

$$g(\tilde{\lambda}_i, \tilde{\gamma}_1, \tilde{\gamma}_2) \leq \mathcal{L}\left(\hat{\mathbf{p}}_i, \hat{t}, \hat{\boldsymbol{\beta}}_i, \hat{\lambda}_i, \hat{\gamma}_1, \hat{\gamma}_2, \right) + (\tilde{\lambda}_i - \hat{\lambda}_i)\left(\sum_{k=1}^{N_c} \hat{P}_{k,i} - \bar{P}_i\right) + (\tilde{\beta}_i - \hat{\beta}_i)\hat{P}_{k,i}$$
$$+ (\tilde{\gamma}_1 - \hat{\gamma}_1)\left[\sum_{k=1}^{N_c}\left(E_{k,i}^{[\ell]} + D_{k,i,u}^{[\ell]}(\hat{P}_{k,i} - P_{k,i}^{[\ell]})\right) - \sum_{k=1}^{N_c}\left(\frac{|\hat{h}_{k,i}|^2\hat{P}_{k,i}}{\sigma_{e,h_i}^2\hat{P}_{k,i} + T_{k,i}} + 1\right)^{-1}\right]. \quad \text{(I.8)}$$

According to the definition of subgradient [101], it can be verified that $\sum_{k=1}^{N_c}\Big[D_{k,i,u}^{[\ell]}(P_{k,i} - P_{k,i}^{(\ell)}) - E_{k,i,u}^{[\ell]} - \left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{A_{k,i,u}P_{k,i}+T_{k,i}} + 1\right)^{-1}\Big]$ is a subgradient of $\gamma_u, u = 1, 2$. Consequently, we can use the subgradient method in (5.37) to obtain the optimal dual variables. This completes the proof. □

Appendix J

Proof of Corollary 5.1

Consider the KKT conditions for the original Lagrangian in (I.2) as follows

$$\frac{-\gamma_3|\hat{h}_{k,i}|^2T_{k,i}}{\left[(|\hat{h}_{k,i}|^2+\sigma^2_{e,h_i})P_{k,i}+T_{k,i}\right]^2}+\lambda_i+\sum_{u=1}^{2}\gamma_u D^{[\ell]}_{k,i,u}-\beta_{k,i}=0, \tag{J.1a}$$

$$1-\sum_{u=1}^{3}\gamma_u=0, \tag{J.1b}$$

$$\lambda_i(\sum_{k=1}^{N_c}P_{k,i}-\bar{P}_i)=0, \tag{J.1c}$$

$$\beta_{k,i}P_{k,i}=0, \tag{J.1d}$$

$$\gamma_3\left[\sum_{k=1}^{N_c}\left(\frac{|\hat{h}_{k,i}|^2P_{k,i}}{\sigma^2_{e,h_i}P_{k,i}+T_{k,i}}+1\right)^{-1}-t\right]=0, \tag{J.1e}$$

$$\gamma_u\left[\sum_{k=1}^{N_c}\left(E^{[\ell]}_{k,i,u}+D^{[\ell]}_{k,i,u}(P_{k,i}-P^{[\ell]}_{k,i})\right)-t\right]=0,\ u=1,2. \tag{J.1f}$$

Now, we investigate whether the constraints (5.35b) and (5.35c) are active simultaneously at the optimal point. In other words, we are interested in whether the optimized MSEs of three consecutive hops are equal. Note that due to (J.1b), γ_u should take values between zero and one. First, we observe that $\gamma_3=1$, or equivalently, $\gamma_1=\gamma_2=0$ holds provided that $\lambda_i=\beta_{k,i}-Y_{k,i}$, where $Y_{k,i}=\frac{|\hat{h}_{k,i}|^2T_{k,i}}{[(|\hat{h}_{k,i}|^2+\sigma^2_{e,h_i})P_{k,i}+T_{k,i}]^2}$. Then, by the

complementary slackness conditions (J.1e) and (J.1f), we have

$$\sum_{k=1}^{N_c}\left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i}+T_{k,i}}+1\right)^{-1}=t,\quad \sum_{k=1}^{N_c}\left(E_{k,i,u}^{[\ell]}+D_{k,i,u}^{[\ell]}(P_{k,i}-P_{k,i}^{[\ell]})\right)<t, u=1,2 \tag{J.2}$$

This means that the MSE of the three hops can be different at the optimum of problem (5.35). Next, we observe that $\gamma_3 > 0$ is required, which can be proved by contradiction. Assume that at optimality, we have $\gamma_3 = 0$, then from (J.1a) we can deduce that $\beta_{k,i} = \lambda_i + \sum_{u=1}^{2}\gamma_u D_{k,i,u}^{[\ell]}, \forall k$, which is strictly positive due to the fact that $\lambda_i \geq 0$, $D_{k,i,1}^{[\ell]} > 0$ if $\sigma_{e,l_i}^2 > 0$, $D_{k,i,2}^{[\ell]} > 0$ if $\sigma_{e,g_i}^2 > 0$, and that $\gamma_1 + \gamma_2 = 1 > 0$. From the complementary slackness condition (J.1d), we obtain $P_{k,i} = 0, \forall k$, i.e., there is no transmission on the ith hop. Since $\gamma_3 = 0$, by (J.1e) it follows that

$$\sum_{k=1}^{N_c}\left(\frac{|\hat{h}_{k,i}|^2 P_{k,i}}{\sigma_{e,h_i}^2 P_{k,i}+T_{k,i}}+1\right)^{-1}=N_c<t, \tag{J.3}$$

where we have substituted $P_{k,i} = 0, \forall k$. On the other hand, since $\gamma_1 + \gamma_2 \geq 0$, at least one of γ_1 and γ_2 should be strictly positive. Assume that γ_1 is positive, we have from (J.1f) that

$$\sum_{k=1}^{N_c}\left(E_{k,i,1}^{[\ell]}+D_{k,i,1}^{[\ell]}(P_{k,i}-P_{k,i}^{[\ell]})\right)=N_c-\sum_{k=1}^{N_c}\left(\frac{B_{k,i,1}}{\sigma_{e,l_i}^2 P_{k,i}^{[\ell]}+C_{k,i,1}}+D_{k,i,1}^{(\ell)}P_{k,i}^{(\ell)}\right)=t<N_c. \tag{J.4}$$

It is clear from (J.3) and (J.4) that there is a contradiction for the value of t. Therefore, we conclude that $\gamma_3^\star > 0$. From the complementary slackness condition (J.1e), it can be inferred that the MSE in the ith hop always equals t, which is always larger than or equal to the MSEs of the $(i-1)$th and $(i-2)$th hops due to the constraint in (5.35b). Note that although the MSE of the $(i-1)$th hop is not necessarily larger

than that of the $(i-2)$th hop, the recursive nature of the AO algorithm will guarantee that the MSEs in the later hops are always larger than those in the earlier hops after the convergence of the algorithm, i.e., the ith hop will always have a larger MSE than the $(i-1)$th hop, $i = 2, \ldots, M$.