Channel Estimation and Non-Linear Transceiver Designs for MIMO OFDM Relay Systems

Author:
Andrew Paul Millar

Supervisor:
Dr Stephan Weiss

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Electronic and Electrical Engineering
University of Strathclyde

March 31, 2014

Dedicated to

my parents, my brother,
and in loving memory of my grandparents.

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: _____

Date: _____

Acknowledgements

There are a number of people that without whom the writing of this thesis would certainly not have been possible, and I am happy to offer them my sincerest thanks.

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr Stephan Weiss, for his ever enthusiastic encouragement, tremendous support, and the confidence he has always shown in me throughout the full duration of my studies. Despite his busy and hectic schedule, Stephans door has always been open and he has always been available for an impromptu chat. I have really enjoyed our fruitful and stimulating discussions on various research topics, and his keen insight and profound knowledge has been invaluable to my research. I am also grateful to Stephan for sending me on some fantastic international conferences, and I particularly enjoyed the conferences in Aalborg and Barcelona that we attended together. I would finally like to thank Stephan for his constant support throughout my studies. Whenever I have doubted my own ability or progress he has always reassured me and given me confidence. I will be eternally grateful for his kindness and support, and everything else he has done for me throughout my PhD. I could not imagine having a better supervisor.

I would also like to thank Professor Malcolm Macleod of Qinetiq Malvern for agreeing to be my second supervisor and for providing a reference for me before I started my PhD.

I would like to express my great thanks to Dr Mikel Mendicute, of the University of Mondragon, and Professor John Soraghan, from the University of Strathclyde, for kindly agreeing to examine this thesis. I am thankful to them for making the viva such an (unexpectedly) enjoyable experience. Their insightful comments and suggestions have greatly helped improve the quality of this thesis. I would like to further thank John for guiding me through the corrections to the thesis.

The writing of this thesis would not have been possible without the generous support from a University of Strathclyde scholarship, as well as a Motorola University Research Grant. I would particularly like to thank Professor Robert Stewart,

from the University of Strathclyde, for making it possible for me to continue my studies through a Motorola University Research Grant. Without this source of extra funding it would not have been possible for me to complete my studies.

There are a many number of close friends from the University of Strathclyde that I owe great thanks. I would especially like to thank Dr Joseph Jackson and Tapiwa Mutasa for their wonderful friendship and all the great times we have had together. I have thoroughly enjoyed our weekly table tennis games and the in depth analysis that followed in the pub. I have also enjoyed the many hours we spent discussing complex mathematical problems. I would also like to thank Rui Gongzhang for participating in our table tennis matches. Great thanks also go to my friends Dr Paul Murray, Gregour Bolton, and Dr Rahul Summan who have all been at the University of Strathclyde with me since our undergraduate studies and helped me immensely throughout this period. I would also like to thank all past and present members of the Centre of Excellence in Signal and Image Processing department for making the time I have spent studying such an enjoyable and exhilarating experience. It has truly been a pleasure working within this group and I have made some wonderful friends throughout my time here.

My biggest thanks must go to my family and loved ones. I would particularly like to thank my parents. Without their constant love and support I would never have reached this stage in my life. They have always encouraged me to do well and kept me on the right path. The last five years have been both very exciting and very stressful. In times of stress my parents have encouraged me, in times of despair they have consoled me, and in times of joy they have celebrated with me. I will never be able to truly express how grateful I am for everything that they do for me, and how thankful I am that they are my parents. I would also like to thank Bob and Alison for their support and encouragement, and for all the fun times we have shared together. Nights spent with them at the farm and at their home have been a source of much needed fun in times of great stress. I would also like to thank the rest of my family - all my aunties, uncles, and cousins who have always encouraged me. I would finally like to thank Hannah for her love, support and constant reassurance. Without her I would certainly never have made it through the final years of the PhD and I am extremely thankful to have her in my life. Meeting her has been the highlight of this journey.

Abstract

Multiple-input multiple output (MIMO) systems deploy multiple antennas at either end of a communication link and can provide significant benefits compared to traditional single antenna systems, such as increased data rates through spatial multiplexing gain, and/or improved link reliability through diversity techniques. Recently, the natural extension of utilising multiple antennas in relay networks, known as MIMO relaying, has attracted significant research attention due to the fact that the benefits of MIMO can be coupled with extended network coverage through the use of relaying devices. This thesis concentrates on the design and analysis of different aspects of MIMO relay systems communicating over frequency selective channels with the use of orthogonal frequency division multiplexing (OFDM).

The first focus of this thesis is on the development of training based channel estimation algorithms for two-hop MIMO OFDM relaying. In the first phase of channel estimation the relay-destination channel is estimated using conventional point-to-point MIMO estimation techniques. In the second phase, the source sends known training symbols to the relay, which precodes the received symbols and forwards them to the destination. In order to estimate the source-relay channel at the destination, an iterative algorithm is derived, which involves sequentially solving a number of convex optimisation problems and has guaranteed convergence. Since the proposed iterative algorithm may be too computationally complex for practical systems, a simplified approach is also derived where the channel estimation processors can be calculated in closed form.

Under the assumption of perfect channel state information (CSI), we then develop non-linear transceiver designs for MIMO OFDM relay systems, focusing specifically on decision feedback equalisation (DFE) and Tomlinson Harashima precoding (THP). The optimal source and relay precoding matrices are derived that minimise the arithmetic mean square error (MSE) subject to source and relay transmission power constraints, when either a zero forcing (ZF) or minimum

mean square error (MMSE) equaliser is used at the destination. Simulation results demonstrate that the proposed non-linear solutions outperform linear transceivers in terms of bit error rate (BER) and MSE.

For the case of imperfect CSI at all nodes, robust DFE and THP transceivers are then considered that aim to minimise the expected artithmetic MSE subject to the source and relay transmission power constraints. The channel estimation errors are modelled as being drawn from matrix variate Gaussian distributions with known mean and covariance. The source and relay precoder structures are derived for the case that the optimal MMSE equaliser is used at the destination. The derived precoder structures are shown to be optimal for the special case that the channel estimation errors are uncorrelated. Simulation results demonstrate the robustness of the proposed algorithms to channel estimation errors.

Contents

List of Author's Publications

Journal Publications

1. Andrew P. Millar, Stephan Weiss, and Robert W. Stewart "THP Transceiver Design for MIMO Relaying with Direct Link and Partial CSI," *IEEE Communications Letters*

2. Andrew P. Millar, Stephan Weiss, and Robert W. Stewart "Precoder Design for MIMO Relay Networks with Direct Link and Decision Feedback Equalisation," *IEEE Communications Letters,*

3. Andrew P. Millar, Stephan Weiss, and Robert W. Stewart "Channel Estimation Algorithms for Dual-Hop MIMO OFDM Relaying Networks," *in preparation to be submitted to IEEE Trans. Signal Processing.*

Conference Publications

1. Andrew P. Millar, Stephan Weiss, and Robert W. Stewart "Robust Transceiver Design for MIMO Relay Systems with Tomlinson Harashima Precoding," *20th European Signal Processing Conference, EUSIPCO 2012,* Bucharest, Romania, August 2012.

2. Andrew P. Millar, Stephan Weiss, and Robert W. Stewart "ZF DFE Transceiver Design for MIMO Relay Systems with Direct Source-Destination Link," *19th European Signal Processing Conference, EUSIPCO 2011,* Barcelona, Spain, August 2011.

3. Andrew P. Millar, Stephan Weiss, and Robert W. Stewart "Tomlinson Harashima Precoding Design for Non-regenerative MIMO Relay Networks," *73th IEEE Vehicular Technology Conference, VTC 2011,* Budapest, Hungary, May 2011.

4. Andrew P. Millar and Stephan Weiss "Transceiver Design for Non-regenerative MIMO Relay Systems with Decision Feedback Detection," *18th European Signal Processing Conference, EUSIPCO 2010,* Aalborg, Denmark, August 2010.

5. Andrew P. Millar and Stephan Weiss "Design of Block Based Linear MMSE Precoding and Equalisation for Broadband MIMO Relay Networks," *17th European Signal Processing Conference, EUSIPCO 2009,* Glasgow, Scotland, August 2009.

6. Andrew P. Millar and Stephan Weiss "Design of ZF and MMSE Filterbanks in Relay Networks with Average and Peak Power Constraints," *15th IEEE Workshop on Statistical Signal Processing, SSP 2009,* Cardiff, Wales, August 2009.

7. Stephan Weiss, Nicola Moret, Andrew P. Millar, Andrea Tonello, and Robert W. Stewart "Initial Results on an MMSE Precoding and Equalisation Approach to MIMO PLC Channels," *IEEE International Symposium on Power Line Communications and its Applications, ISPLC 2011,* Udine, Italy, April 2011.

8. Stephan Weiss, Andrew P. Millar, and Robert W. Stewart "Inversion of Parahermitian Matrices," *18th European Signal Processing Conference, EUSIPCO 2010,* Aalborg, Denmark, August 2010.

9. Stephan Weiss, Andrew P. Millar, Robert W. Stewart, and Malcolm Macleod "Performance of Transmultiplexers Based on Oversampled Filterbanks under Variable Oversampling Ratios," *18th European Signal Processing Conference, EUSIPCO 2010,* Aalborg, Denmark, August 2010.

10. Stephan Weiss, Paul Yarr, Waleed Al-Hanafy, Andrew P. Millar, and Chi-Hieu Ta "An Oversampled Modulated Filterbank Transmultiplexer with Precoding and Equalisation," *3rd Workshop on Power Line Communications, WSPLC 2009,* Udine, Italy, October 2009.

11. Waleed Al-Hanafy, Andrew P. Millar, Chi Hieu Ta, and Stephan Weiss "Broadband SVD and Non-Linear Precoding and Equalisation Applied to Broadband MIMO Channels," *42nd Asilomar Conference on Signals, Systems and Computers, ACSSC 2008,* Pacific Grove, USA, October 2008.

Abbreviations

1G	first generation
3,9G	pre-fourth generation
4G	fourth generation
AF	amplify forward
AMSE	arithmetic mean square error
AWGN	additive white Gaussian noise
BER	bit error rate
CP	cyclic prefix
CSI	channel state information
DFE	decision feedback equalisation/equaliser
DFT	discrete Fourier transform
DSP	digital signal processing
EVD	eigenvalue decomposition
FIR	finite impulse response
GP	geometric program/programming
IBI	interblock interference
IDFT	inverse discrete Fourier transform
ISI	inter-symbol interference
KKT	Karush Kuhn Tucker
LS	least squares
MI	mutual information
MIMO	multiple-input multiple-output

MISO	multiple-input single-output
MMSE	minimum mean square error
M-QAM	M-ary quadrature amplitude modulation
MSE	mean square error
OFDM	orthogonal frequency division multiplexing
PA	power allocation
QPSK	quadrature phase shift keying
SDP	semi-definite program/programming
SIMO	single-input multiple-output
SINR	signal to interference noise ratio
SISO	single-input single-output
SLS	scaled least squares
SNR	signal to noise ratio
SOCP	second order conic program/programming
SRP	source and relay precoded
ST	space-time
STBC	space-time block coding
s.t.	subject to
SVD	singular value decomposition
THP	Tomlinson Harashima precoding/precoder
UPA	uniform power allocation
VP	vector precoding
w.l.o.g.	without loss of generality
w.r.t.	with respect to
ZF	zero forcing

Mathematical Notations

$\mathbb{R}$	set of real numbers
$\mathbb{R}^M$	set of M dimensional vectors with real entries
$\mathbb{R}_+^M$	set of M dimensional vectors with non-negative real entries
$\mathbb{R}_{++}^M$	set of M dimensional vectors with positive real entries
$\mathbb{R}^{M\times N}$	set of $M \times N$ dimensional matrices with real entries
$\mathbb{C}$	set of complex numbers
$\mathbb{C}^M$	set of M dimensional vectors with complex entries
$\mathbb{C}^{M\times N}$	set of $M \times N$ dimensional matrices with complex entries
$\mathbb{Z}$	set of integer numbers
$\mathbb{Z}^M$	set of M dimensional vectors with integer entries
$\mathbb{H}^{N\times N}$	set of $N \times N$ Hermitian matrices
$\mathbb{H}_+^{N\times N}$	set of $N \times N$ Hermitian positive semi-definite matrices
$\mathbb{H}_{++}^{N\times N}$	set of $N \times N$ Hermitian positive definite matrices
a	scalars are denoted by lower case normal font
a_i	ith element of vector $\boldsymbol{a}$
$a_{[i]}$	ith largest element of vector $\boldsymbol{a}$
$\{a_i\}_{i=1}^N$	the set of scalars $a_1, a_2, ..., a_N$
$\boldsymbol{a}$	vectors are denoted by lower case bold font
$[\boldsymbol{a}]_i$	ith element of vector $\boldsymbol{a}$
$[x_{i,j}]_{i,j}^{I,J}$	the vector $[x_{1,1}, ..., x_{1,J}, x_{2,1}, ..., x_{2,J}, ..., x_{I,1}, ..., x_{I,J}]^T$
$\boldsymbol{A}$	matrices are denoted by upper case bold font
$[\boldsymbol{A}]_{ij}$	element located in the ith row and jth column of matrix $\boldsymbol{A}$

$\{.\}^T$	transpose
$\{.\}^H$	Hermitian transpose
$\{.\}^*$	complex conjugate
$\mathrm{tr}\{\boldsymbol{A}\}$	trace of matrix $\boldsymbol{A}$
$\lvert\boldsymbol{A}\rvert$	determinant of matrix $\boldsymbol{A}$
$\lVert\boldsymbol{A}\rVert_F$	Frobenious norm of matrix $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	inverse of non-singular matrix $\boldsymbol{A}$
$\{\boldsymbol{A}\}^\dagger$	pseudo-inverse of singular matrix $\boldsymbol{A}$
$\bigotimes_{n=1}^{N} \boldsymbol{A}_n$	the matrix multiplication $\boldsymbol{A}_1\boldsymbol{A}_2...\boldsymbol{A}_N$
$\mathrm{rank}\{\boldsymbol{A}\}$	the rank of matrix $\boldsymbol{A}$
$\boldsymbol{A} \preceq \boldsymbol{B}$	$\boldsymbol{B}-\boldsymbol{A}$ is positive semi-definite
$\boldsymbol{A} \succeq \boldsymbol{B}$	$\boldsymbol{A}-\boldsymbol{B}$ is positive semi-definite
$\lambda_{\max}(\boldsymbol{A})$	maximum eigenvalue of matrix $\boldsymbol{A}$
$\mathrm{diag}[\{\boldsymbol{A}_i\}_{i=1}^N]$	block diagonal matrix with ith diagonal block given by $\boldsymbol{A}_i$
$\mathrm{vec}[.]$	vectorisation operator
$\otimes$	Kronecker product
$\boldsymbol{I}_N$	identity matrix of dimension $N \times N$
$\boldsymbol{0}_{N\times M}$	$N \times M$ dimensional zero matrix
$\boldsymbol{1}_N$	N dimensional column vector with every element equal to 1
$\mathbb{E}\{.\}$	statistical expectation
$\Re\{.\}$	returns real part of argument
$\Im\{.\}$	returns imaginary part of argument
$\max(.)$	returns the maximum value of the argument
$\min(.)$	returns the minimum value of the argument
$\forall$	for all
$\lfloor . \rfloor$	floor operator
$[x]^+$	$[x]^+ \triangleq \max(x, 0)$
$\log(.)$	natural logarithm function

e	Euler's number
$\exp(x)$	e^x
π	pi
$\triangleq$	defined as
$\sim$	distributed according to
$\mathcal{CN}(\boldsymbol{A}\ ,\ \boldsymbol{B})$	matrix variate complex Gaussian distribution with mean $\boldsymbol{A}$ and covariance $\boldsymbol{B}$
$\boldsymbol{e}_i$	ith elementary unit vector
$\boldsymbol{\mathcal{F}}$	normalised DFT matrix
$\boldsymbol{a} \preceq_w^+ \boldsymbol{b}$	$\boldsymbol{a}$ is weakly additively submajorised by $\boldsymbol{b}$
$\boldsymbol{a} \preceq_w^\times \boldsymbol{b}$	$\boldsymbol{a}$ is weakly multiplicatively submajorised by $\boldsymbol{b}$

Commonly Used Symbols and Variables

N_s	number of antennas at source
N_r	number of antennas at relay
N_d	number of antennas at destination
l	frequency selective channel delay path index
L	frequency selective channel order
n	time domain signalling interval index
$\boldsymbol{\mathcal{H}}_s[l]$	source-relay MIMO channel matrix for lth delay path
$\boldsymbol{\Theta}_s[l]$	source-relay channel transmit spatial correlation matrix for lth delay path
$\boldsymbol{\Upsilon}_s[l]$	source-relay channel receive spatial correlation matrix for lth delay path
$\boldsymbol{\mathcal{H}}_{sw}[l]$	spatially white source-relay MIMO channel matrix for lth delay path
$\sigma^2_{h_s}[l]$	variance of elements in $\boldsymbol{\mathcal{H}}_{sw}[l]$
$\boldsymbol{h}_s$	source-relay channel impulse response vector

$\hat{\boldsymbol{h}}_s$	estimate of source-relay channel impulse response vector
$\boldsymbol{e}_s$	source-relay channel estimation error vector
$\boldsymbol{R}_{e_s}$	source-relay channel estimation error covariance matrix
$\boldsymbol{R}_{h_s}$	source-relay channel covariance matrix
$\boldsymbol{\mathcal{H}}_r[l]$	relay-destination MIMO channel matrix for lth delay path
$\boldsymbol{\Theta}_r[l]$	relay-destination channel transmit side spatial correlation matrix for lth delay path
$\boldsymbol{\Upsilon}_r[l]$	relay-destination channel receive side spatial correlation matrix for lth delay path
$\boldsymbol{\mathcal{H}}_{rw}[l]$	spatially white relay-destination MIMO channel matrix for lth delay path
$\sigma^2_{h_r}[l]$	variance of elements in $\boldsymbol{\mathcal{H}}_{rw}[l]$
$\boldsymbol{h}_r$	relay-destination channel impulse response vector
$\hat{\boldsymbol{h}}_r$	estimate of relay-destination channel impulse response vector
$\hat{\boldsymbol{\mathcal{H}}}_r[l]$	estimate of relay-destination lth MIMO delay path
$\boldsymbol{\mathcal{E}}_r[l]$	estimate of relay-destination lth MIMO delay path
$\boldsymbol{e}_r$	relay-destination channel estimation error vector
$\boldsymbol{R}_{e_r}$	relay-destination channel estimation error covariance matrix
$\boldsymbol{R}_{h_r}$	relay-destination channel covariance matrix
$\boldsymbol{X}$	training matrix for relay-destination channel estimation
$\boldsymbol{W}$	destination processor for relay-destination channel estimation
$\boldsymbol{S}$	training matrix for source-relay channel estimation
$\boldsymbol{G}$	relay precoder for source-relay channel estimation
$\bar{\boldsymbol{W}}$	destination processor for source-relay channel estimation
k	OFDM subcarrier index
K	total number of OFDM subcarriers
N_k	number of transmit data symbols for kth OFDM subcarrier
$\bar{N}$	total number of transmit symbols over all subcarriers

$\boldsymbol{H}_{s,k}$	source-relay OFDM channel matrix for kth subcarrier
$\boldsymbol{H}_{s}$	block diagonal source-relay OFDM channel matrix
$\boldsymbol{H}_{r,k}$	relay-destination OFDM channel matrix for kth subcarrier
$\hat{\boldsymbol{H}}_{r,k}$	estimate of relay-destination OFDM channel matrix for kth subcarrier
$\boldsymbol{E}_{r,k}$	estimation error of kth relay-destination OFDM subcarrier
$\boldsymbol{H}_{r}$	block diagonal relay-destination OFDM channel matrix
$\hat{\boldsymbol{H}}_{r}$	estimate of block diagonal relay-destination OFDM channel matrix
$\boldsymbol{E}_{r}$	estimation error of block diagonal relay-destination OFDM channel matrix
$\boldsymbol{F}_{k}$	source precoding matrix for kth subcarrier
$\boldsymbol{G}_{k}$	relay precoding matrix for kth subcarrier
$\boldsymbol{W}_{k}$	destination equaliser matrix for kth subcarrier
$\boldsymbol{B}_{k}$	DFE/THP strictly upper right triangular feedback matrix for kth subcarrier
$\boldsymbol{U}_{k}$	DFE/THP unit diagonal upper right triangular feedback matrix for kth subcarrier
$\boldsymbol{R}_{e,k}$	DFE/THP error covariance matrix for kth subcarrier
$\boldsymbol{e}_{k}$	DFE/THP error signal for kth subcarrier
$\boldsymbol{v}_{s,k}$	source-relay AWGN symbols for kth subcarrier
$\boldsymbol{v}_{s}$	ssource-relay AWGN vectors over all subcarriers
$\boldsymbol{v}_{r,k}$	relay-destination AWGN symbols for kth subcarrier
$\boldsymbol{v}_{r}$	relay-destination AWGN vectors over all subcarriers
$\sigma^2_{v_s}$	variance of source-relay AWGN symbols
$\sigma^2_{v_r}$	variance of relay-destination AWGN symbols
$\mathcal{D}(.)$	decision device operator
$\mathcal{M}(.)$	modulo device operator

$\mathcal{A}$	M-QAM symbol constellation set
$\mathcal{U}$	Voronoi region of M-QAM symbol constellation
D	side length of Voronoi region
P_s	available power budget to the source
P_r	available power budget to the relay

List of Figures

Chapter 1

Introduction

1.1 Motivation

The past couple of decades have witnessed a significant and rapid growth in wireless communications owing to improved technology, the development of advanced digital signal processing (DSP) algorithms, as well as the ever increasing demand for more sophisticated, high quality services. A typical example of such an evolution is in the mobile communications sector where systems have quickly evolved from the first generation (1G) narrowband analogue systems, through to the current pre-fourth generation (3.9G) wideband digital systems seen today. This trend of constant and rapid growth is anticipated to continue in the coming years with the next generation of wireless systems (4G and beyond) expected to have substantially greater capabilities than those of their predecessors.

One of the main requirements of current and future wireless communications is the ability to reliably support higher data rates in order to provide a wide variety of high quality applications. However, there is a fundamental limit to the maximum data rate that a channel can support reliably or error free, which is termed the channel capacity. In [1] Shannon derived the capacity for an additive white Gaussian noise (AWGN) channel in terms of the available bandwidth and the transmit signal power. It was shown that in order to increase the capacity of a single antenna communications link, commonly termed single-input single-output (SISO), one must either increase the utilised bandwidth and/or increase the transmission power. Both bandwidth and power are limited and precious resources. This has fuelled a number of innovative research efforts, with one of the primary goals being to develop spectrally efficient techniques capable of providing reliable high data rates with limited bandwidth and power consumption. The use of antenna arrays, popularly referred to as multi-antenna systems, space-time

(ST) communications, and more generally as MIMO communication [2–4], have emerged as spectrally efficient methods of dealing with the limited bandwidth availability of the wireless channel and have received significant interest from both the academic and industrial communities.

1.2 Point-to-Point MIMO Communication

MIMO systems offer a number of benefits when compared to SISO systems and are capable of providing array, diversity, and spatial multiplexing gains [2–5]. The degree to which these gains can be exploited depends on the antenna configuration and the availability of CSI. Traditionally, array and diversity gains have been utilised to combat the adverse effects of the channel and to provide robustness to fading. When CSI is available at the destination, array gain and/or diversity gain can be achieved through a coherent combination of the received signals at the destination [3]. When CSI is available at the source, these gains can also be achieved through transmit processing techniques. In the absence of CSI at the source, diversity can still be realised through space-time block coding (STBC) [5–9] techniques, although in this case no array gain can be achieved. Whilst array and/or diversity gains can be achieved in single-input multiple-output (SIMO) and multiple-input single-output (MISO) configurations, the spatial multiplexing gain can only be achieved when multiple antennas are employed at both ends of the link [9–12] i.e. in"true" MIMO systems. Spatial multiplexing techniques open up a number of parallel data pipes over which multiple independent data streams can be transmitted. This gives rise to a dramatic increase in channel capacity and, importantly, does so without bandwidth expansion. MIMO antenna configurations can simultaneously provide array, diversity, and multiplexing gains. However, due to their conflicting demands, it is not possible to fully leverage these gains at the same time and there is a fundamental trade-off of between them [13–15].

To achieve the aforementioned potential benefits of MIMO systems, appropriate transceiver designs must be utilised. The introduction of a pre-equaliser (a.k.a. precoder) at the source and/or equaliser (a.k.a. decoder) at the destination provides a convenient framework that offers a flexible trade-off between realising array, diversity, and multiplexing gains. Such a framework also encompasses beamforming in MISO and SIMO antenna configurations as special cases. The joint design of source precoding and destination equalisation is commonly referred to as joint transceiver design and has attracted significant research attention. Linear transceiver designs have been considered to minimise the arithmetic mean square error (AMSE) [16–20], maximise the channel capacity [16, 17, 20, 21],

minimise the geometric signal-interference noise ratio (GSINR) [16, 17], and minimise the bit error rate (BER) [16, 17, 22, 23]. Non-linear techniques such as DFE [24–31] and THP [32–35] have also received great interest, and have been shown to provide improved performance compared to linear transceivers.

1.3 MIMO Relay Communication

More recently the topic of MIMO relaying has garnered a lot of attention. Differing from point-to-point MIMO, where transmission is carried out directly between a source and destination, relaying networks make use of intermittent nodes to forward the data from the source to destination. In addition to the benefits provided by the use of multiple antennas, the introduction of relaying devices has been shown to increase network coverage and improve link reliability through the spatial diversity offered by the relays [36–38].

Similar to point-to-point MIMO communications, appropriate transceiver designs are needed to exploit the potential gains in MIMO relay networks. Transceiver designs can be developed by introducing precoding at the source and relay terminals and equalisation at the destination device. The transceiver design problem for MIMO relaying becomes far more involved than that for the point-to-point MIMO scenario (especially in the case of multi-hop networks) due to the increased number of variables (precoders and equalisers) as well as the additional relay power constraints that must be satisfied. The problem is further complicated by the fact that the relay transmission power depends on the transmit power of the preceeding links, which results in optimisation problems where the constraints have coupled variables [39]. Furthermore, due to relay precoding and noise propagation, the noise seen at the destination antennas is no longer white, which makes transceiver design more difficult. Thus the extension of point-to-point MIMO transceivers to the case of MIMO relaying is not straightforward. Despite the problems associated with MIMO relay designs, optimal linear transceiver solutions have been presented to minimise the AMSE [40–45], maximise the channel capacity [45–49], and optimise various other design criteria [45, 50]. Non-linear DFE and THP transceiver designs have also been considered in [51–53] and [54, 55], respectively. Despite the vast amount of research that has been conducted on MIMO relaying systems, there are still a plentiful number of open problems. The following section briefly describes the problems that we aim to solve in this thesis.

1.4 Thesis Aims and Objectives

The main aims and objectives of this thesis are:

- **MIMO OFDM Relay Channel Estimation**: In order to design appropriate transceiver designs for MIMO relaying systems, CSI is required to be available to different nodes in the network. CSI can be achieved through channel estimation algorithms. Such algorithms have been developed for narrowband MIMO relaying systems in [56–60] but have not yet been considered for MIMO OFDM relaying over frequency selective channels. The first objective of this thesis is therefore to develop channel estimation algorithms for the estimation of the channels in MIMO OFDM relay networks.

- **DFE/THP Transceiver Designs with Perfect CSI**: The main aim of this thesis will be to derive the processors for non-linear transceiver designs in two-hop MIMO relaying systems. Whilst such transceivers have been studied in [51–55] for the case of narrowband MIMO relaying, we aim to design DFE and THP transceivers for the case of MIMO relaying over frequency selective channels utilising OFDM.

- **DFE/THP Transceiver Designs with Imperfect CSI**: The previously mentioned DFE and THP designs in [51–55] all assume the availability of perfect CSI. Robust THP transceiver designs have also been considered in [61–63] for transmission over narrowband channels. The last objective of this thesis is to extend these transceiver designs to the more complicated scenario of MIMO OFDM relaying.

1.5 Thesis Contributions

The following contributions are made in Chapter 3:

- In Section 3.4 of Chapter 3 we derive an iterative algorithm for estimating the source-relay channel in a two-hop MIMO OFDM relaying system. The derivation of the iterative algorithm in this section is considered novel to the best of our knowledge.

- In Section 3.5 of Chapter 3 we develop a suboptimal but simplified source-relay channel estimation algorithm for two-hop MIMO OFDM relay systems. The derivation of the simplified algorithm in Section 3.5, as well as the proof in Section 3.8.4, is considered novel to the best of our knowledge.

The following contributions made in Chapter 4 are considered novel to the best of our knowledge:

- The design of ZF DFE/THP transceivers for two-hop MIMO OFDM relay networks with perfect CSI of all channels.
- The design of MMSE DFE/THP transceivers for two-hop MIMO OFDM relay networks with perfect CSI of all channels.

The following contributions made in Chapter 5 are considered novel to the best of our knowledge:

- The design of robust MMSE DFE/THP transceivers for two-hop MIMO OFDM relay networks with imperfect CSI of all channels.

1.6 Thesis Layout

The remainder of this thesis is organised as follows:

- **Chapter 2**: In Chapter 2 we introduce the signal model for a two-hop MIMO relay network that consists of single source, relay, and destination terminals that are each equipped with multiple antennas. In such a system, assuming the use of a half-duplex relay, the communication process between the source and destination is carried out over two orthogonal transmission stages. In the first phase the source transmits data to the relay, whilst in the second phase the relay forwards the received symbols to the destination. In our model we assume the source-relay and relay-destination channels to be frequency selective and we discuss the use of OFDM that enables the communication of data over orthogonal narrowband subcarriers. We then introduce DFE and THP transceivers, which are non-linear techniques, that aim to combat the interference for each OFDM subcarrier. For both the DFE and THP transceivers we formulate an optimisation problem to minimise the arithmetic MSE subject to transmission power constraints at the source and relay terminals. For the considered MIMO OFDM relaying system this optimisation problem shall be the basis for the derivation of optimal DFE and THP transceiver designs in Chapters 4 and 5.
- **Chapter 3**: In Chapter 3 we consider the task of MIMO OFDM relay channel estimation in order to estimate the source-relay and relay-destination channels. Channel estimation for the MIMO relay system is divided into

two phases. In the first phase the relay transmits known pilot/training symbols to the destination allowing the receiver to estimate the relay-destination channel. This is simply a point-to-point MIMO channel estimation problem for which least squares (LS) and MMSE solutions are well known and are discussed in this chapter. The estimation of the source-relay channel is carried out in the second phase and depends on the processing capabilities of the relaying device. If the relay can perform channel estimation then the source-relay channel can be estimated at the relay in a similar manner to the estimation of the relay-destination channel carried out in the first phase. On the other hand, if the relay has limited processing capabilities then the source-relay channel must be estimated by using measurements received at the destination. In such a scenario, to estimate the source-relay channel in the second phase, the source transmits known pilots to the relay, which then forwards them to the destination. The destination can then perform channel estimation by utilising the relay-destination channel estimate acquired from the first phase. Using the measurements at the destination we firstly derive an iterative source-relay channel estimation algorithm which involves sequentially solving a number of convex optimisation problems and as such is guaranteed to converge to a locally optimal solution. Since the iterative algorithm may be too computationally expensive for practical systems we also discuss a suboptimal algorithm that has reduced complexity. Simulations are provided showing the effectiveness of the proposed algorithms.

- **Chapter 4**: In this chapter we consider DFE and THP transceiver designs for MIMO OFDM relaying under the assumption that perfect CSI is available to all nodes in the network. We consider the optimisation of the processors for minimising the arithmetic MSE subject to transmission power constraints at the source and relay. The optimal source and relay precoder structures are derived when either ZF or MMSE equalisation is utilised at the destination. It is shown that, for both cases of ZF and MMSE equalisation, the optimal source and relay precoding matrices decompose the original matrix valued optimisation problem into a simpler power allocation problem that only involves scalar variables. Power allocation algorithms are discussed to solve the scalar valued optimisation problems. Simulation results are presented comparing the performance of the proposed DFE and THP transceivers in terms of both BER and MSE with various linear benchmarks proposed in the literature. It is shown that our DFE and THP designs offer improvement in both BER and MSE performance compared to linear transceivers.

- **Chapter 5**: In Chapter 5 we consider robust DFE/THP transceiver designs for dual-hop MIMO OFDM relaying systems with imperfect CSI. It is shown that for various channel estimation algorithms, the OFDM subcarrier channels can be modelled as being drawn from matrix variate Gaussian distributions with known mean and covariance. The distribution mean represents the subcarrier channel estimate, whilst the covariance accounts for the estimation error. Due to the fact that the subcarrier channels are not completely known, the direct minimisation of the arithmetic MSE cannot be conducted. We therefore formulate the robust optimisation problem for DFE/THP transceivers to minimise the expected arithmetic MSE subject to a source power constraint and an expected relay power constraint thus making the problem analytically tractable. It is shown that for general channel estimation error covariance matrices the optimal solution is difficult to obtain. However, for the special case of uncorrelated channel estimation errors the optimal source and relay precoders are identified. Simulations highlight the robustness of the proposed algorithms to channel estimation errors.

- **Chapter 6**: Finally, in Chapter 6 we summarise the main results of this thesis and draw conclusions. Based on the results obtained in this thesis, a number of possible future research topics are suggested and briefly discussed.

Chapter 2

MIMO OFDM Relay Signal Models

2.1 MIMO Relay Signal Model

Throughout this thesis we shall consider a 3 node system consisting of a source device with N_s antennas, a relay device with N_r antennas, and a destination device with N_d antennas. It is worthwhile developing the signal model for such a system here since it shall be extensively used throughout the remaining chapters of this thesis. In our MIMO relaying model we consider all channels to be frequency selective with the channel impulse response between each transmit and receive antenna being modelled as a causal finite impulse response (FIR) filter of order L (the length of the channel impulse responses are $L+1$)[1]. Assuming the use of a half-duplex relay, the transmission of data in such a system is seperated into two time slots/phases. This is due to the fact that with a half-duplex relay the relay device cannot transmit and receive in the same time slots. The use of a half-duplex relay simplifies the design of the relay system since the source and relay transmissions are orthogonal i.e. there is no interference between the signals transmitted by the source and relay devices.

2.1.1 First Phase Transmission

In the first phase of transmission the source transmits data to the relay whilst the relay remains silent as shown in Figure 2.1. The received signal at the relay device

[1]In general the impulse responses for the various channels shall be of different lengths. However we make the assumption that they have the same length for ease of notation. Furthermore, the channel responses can be made to have the same number of taps by either appending the shorter responses with zeros or truncating the longer channel impulse responses. The former method is preferred as the latter technique may result in some ISI being unaccounted for.

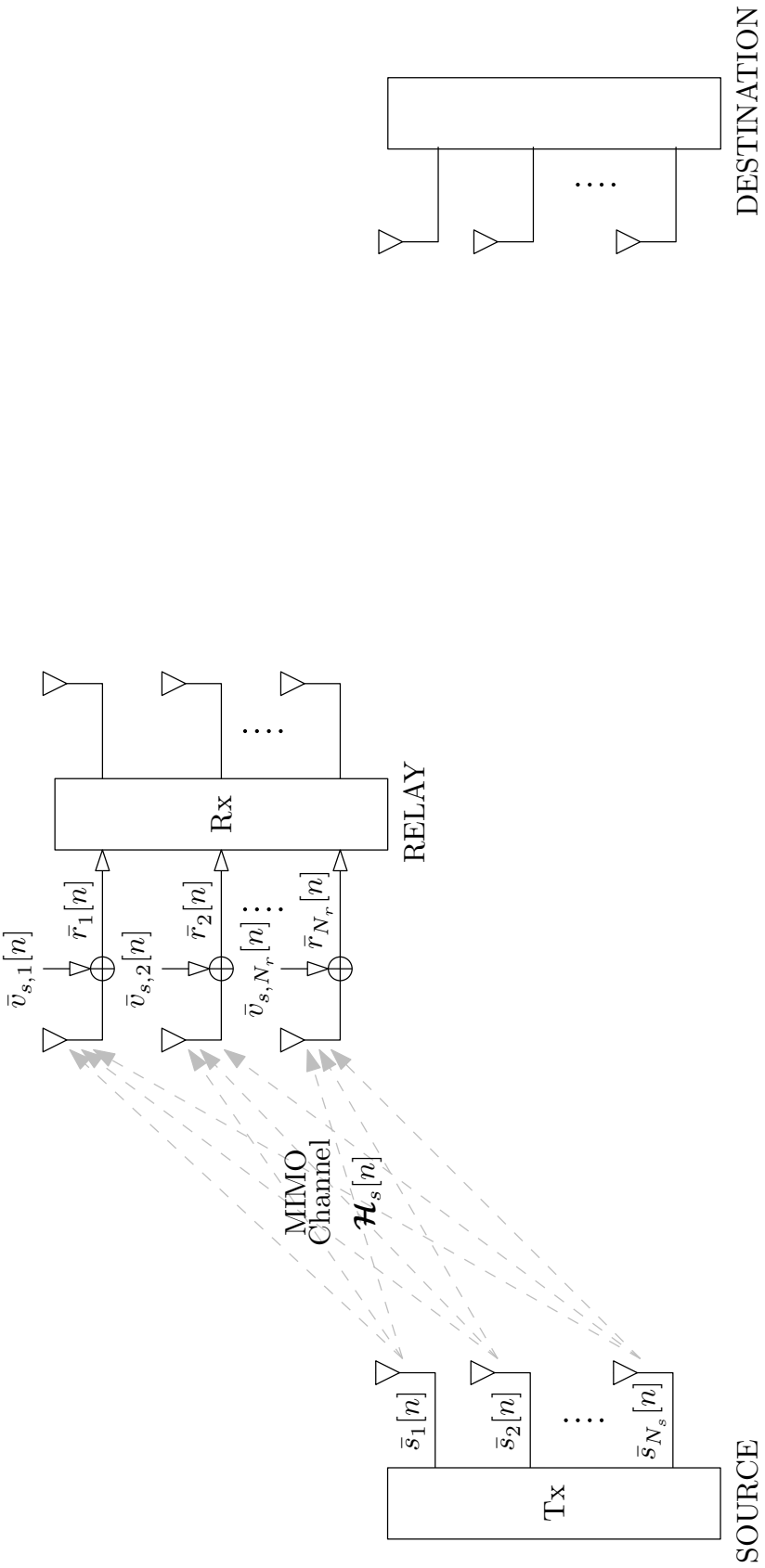

Figure 2.1: First phase transmission for a two-hop MIMO relay system.

can be written as

$$\bar{\boldsymbol{r}}[n] = \sum_{l=0}^{L} \boldsymbol{\mathcal{H}}_s[l]\bar{\boldsymbol{s}}[n-l] + \bar{\boldsymbol{v}}_s[n], \tag{2.1}$$

where $\bar{\boldsymbol{s}}[n] \triangleq [\bar{s}_1[n], ..., \bar{s}_{N_s}[n]]^T \in \mathbb{C}^{N_s}$ is the vector of transmit symbols at the nth signalling interval, with $\bar{s}_i[n] \in \mathbb{C}$ denoting the transmitted symbol from the ith source antenna. In (2.1) $\bar{\boldsymbol{r}}[n] \triangleq [\bar{r}_1[n], ..., \bar{r}_{N_r}[n]]^T \in \mathbb{C}^{N_r}$ is the vector of received symbols at the nth signalling interval, with $\bar{r}_j[n] \in \mathbb{C}$ denoting the received symbol at the jth relay antenna, and $\bar{\boldsymbol{v}}_s[n] \triangleq [\bar{v}_{s,1}[n], ..., \bar{v}_{s,N_r}[n]]^T \in \mathbb{C}^{N_r}$ contains the complex noise samples added at the relay receiver antennas. The matrices $\boldsymbol{\mathcal{H}}_s[l] \in \mathbb{C}^{N_r \times N_s}$ in (2.1) characterise the frequency selective channel between the source and relay device and are given by

$$\boldsymbol{\mathcal{H}}_s[l] = \begin{bmatrix} h_{s,11}[l] & \dots & h_{s,1N_s}[l] \\ \vdots & \ddots & \vdots \\ h_{s,N_r1}[l] & \dots & h_{s,N_rN_s}[l] \end{bmatrix}, \tag{2.2}$$

where $h_{s,ji}[l] \in \mathbb{C}$ is the lth channel tap representing the complex fading gain between the ith source antenna and the jth relay receiver antenna.

2.1.2 Second Phase Transmission

As depicted in Figure 2.2, in the second phase of transmission the source remains silent whilst the relay forwards the data received from the source in the first stage transmission to the destination device. The symbols received at the destination in this transmission stage can be written as

$$\bar{\boldsymbol{y}}[n] = \sum_{l=0}^{L} \boldsymbol{\mathcal{H}}_r[l]\bar{\boldsymbol{x}}[n-l] + \bar{\boldsymbol{v}}_r[n], \tag{2.3}$$

where $\bar{\boldsymbol{y}}[n] \triangleq [\bar{y}_1[n], ..., \bar{y}_{N_d}[n]]^T \in \mathbb{C}^{N_d}$ and $\bar{\boldsymbol{x}}[n] \triangleq [\bar{x}_1[n], ..., \bar{x}_{N_r}[n]]^T \in \mathbb{C}^{N_r}$ are the vectors of received symbols at the destination and relay transmit symbols, respectively, and $\bar{\boldsymbol{v}}_r[n] \triangleq [\bar{v}_{r,1}[n], ..., \bar{v}_{r,N_d}[n]]^T \in \mathbb{C}^{N_d}$ contains complex noise samples. The relay-destination channel matrices $\boldsymbol{\mathcal{H}}_r[l] \in \mathbb{C}^{N_d \times N_r}$ contain the complex fading gains between each transmit and receive antenna and are given by

$$\boldsymbol{\mathcal{H}}_r[l] = \begin{bmatrix} h_{r,11}[l] & \dots & h_{r,1N_s}[l] \\ \vdots & \ddots & \vdots \\ h_{r,N_r1}[l] & \dots & h_{r,N_rN_s}[l] \end{bmatrix}, \tag{2.4}$$

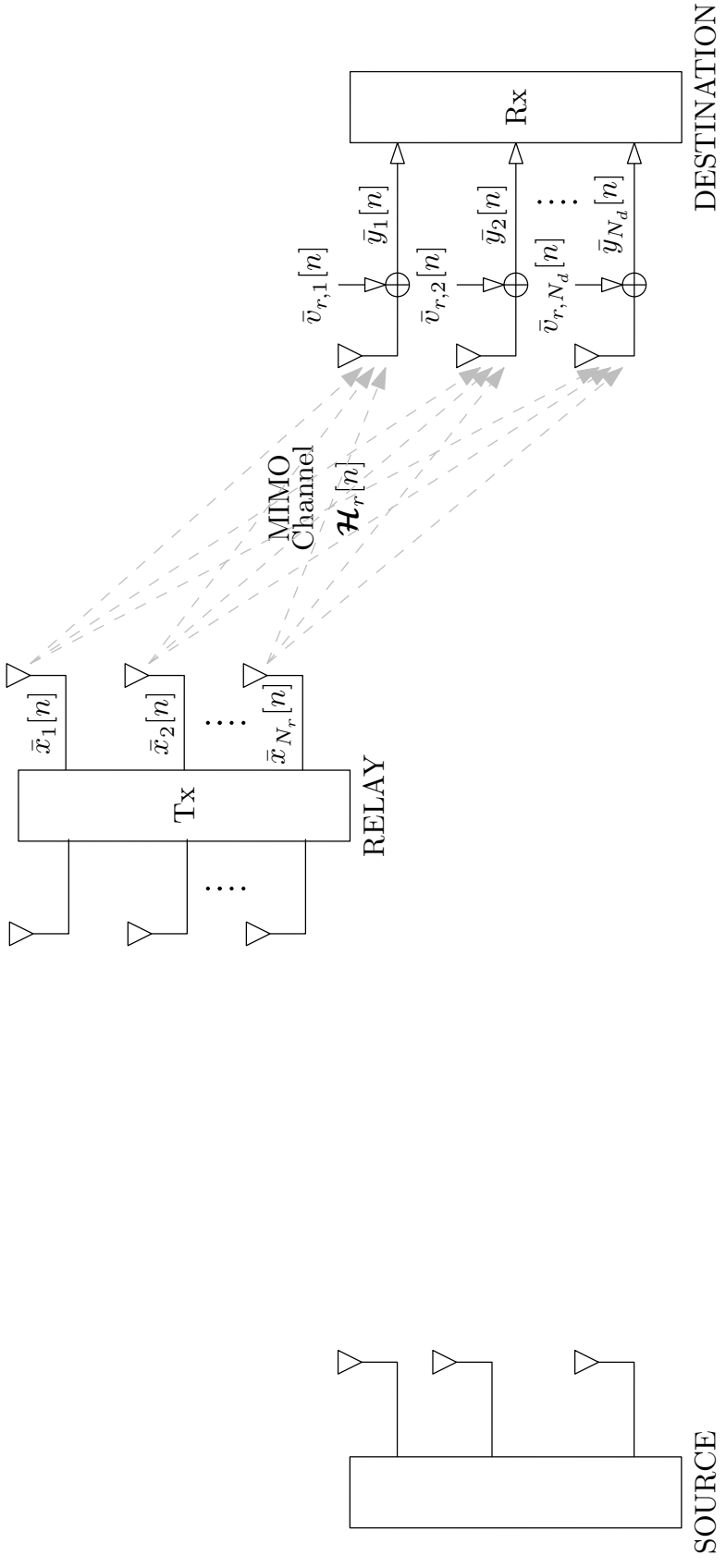

Figure 2.2: Second phase transmission for a two-hop MIMO relay system.

where $h_{r,ji}[l] \in \mathbb{C}$ is the lth channel tap representing the complex gain between the ith relay transmit antenna and jth destination antenna.

It is worth highlighting that in our signal model we have assumed that there is no direct link between the source and destination devices. In a practical system the direct link will be negligible when the source and destination devices are seperated by a substantially large enough distance or when the source-destination channel experiences strong shadowing. In fact, one of the main potential benefits of relaying is that it can provide coverage to a destination device in cases where a point-to-point transmission between the source and destination is infeasible. Thus, relays are likely to feature far more prominently in scenarios where the effect of a direct link can be assumed to be negligible. In any case, since the relay transmission is seperated into two orthogonal time slots, the effect of the direct link can be made negligible by having the destination device discard any information received from the source device in the first time slot.

2.1.3 Spatially Correlated MIMO Channels

We assume that the source-relay and relay-destination channel taps given in (2.2) and (2.4), respectively, are spatially correlated on both the transmit and receive sides and can be modelled according to the Kronecker product model (see e.g. [12, 64–66]). To define such a model let us firstly introduce the definition of a matrix variate Gaussian distribution:

Definition 1: [67] A random matrix $\boldsymbol{A} \in \mathbb{C}^{N\times M}$ is said to have a matrix variate Gaussian distribution with mean $\bar{\boldsymbol{A}} \in \mathbb{C}^{N\times M}$ and covariance matrix $\boldsymbol{B}\otimes\boldsymbol{C}$, where $\boldsymbol{B} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{C} \in \mathbb{C}^{N\times N}$, if it satisfies

$$\mathrm{vec}\left[\boldsymbol{A}\right] \sim \mathcal{CN}\left(\mathrm{vec}[\bar{\boldsymbol{A}}], \boldsymbol{B}\otimes\boldsymbol{C}\right), \tag{2.5}$$

where

$$\mathbb{E}\{\boldsymbol{A}\} = \bar{\boldsymbol{A}} \tag{2.6}$$

$$\mathbb{E}\Big\{\left(\mathrm{vec}\left[\boldsymbol{A}\right] - \mathrm{vec}\left[\bar{\boldsymbol{A}}\right]\right)\left(\mathrm{vec}\left[\boldsymbol{A}\right] - \mathrm{vec}\left[\bar{\boldsymbol{A}}\right]\right)^{H}\Big\} = \boldsymbol{B}\otimes\boldsymbol{C}. \tag{2.7}$$

The matrix variate Gaussian distribution of the random matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A} \sim \mathcal{CN}(\bar{\boldsymbol{A}}, \boldsymbol{B}\otimes\boldsymbol{C})$.

According to the Kronecker spatial correlation model the source-relay and relay-destination MIMO channel taps characterised by (2.2) and (2.4), respectively, can

be decomposed as

$$\boldsymbol{\mathcal{H}}_s[l] = \boldsymbol{\Upsilon}_s^{1/2}[l]\boldsymbol{\mathcal{H}}_{sw}[l]\boldsymbol{\Theta}_s^{\mathrm{T}/2}[l] \tag{2.8}$$

$$\boldsymbol{\mathcal{H}}_r[l] = \boldsymbol{\Upsilon}_r^{1/2}[l]\boldsymbol{\mathcal{H}}_{rw}[l]\boldsymbol{\Theta}_r^{\mathrm{T}/2}[l], \tag{2.9}$$

where $\boldsymbol{\Upsilon}_s[l] \in \mathbb{H}_{++}^{N_r \times N_r}$ and $\boldsymbol{\Upsilon}_r[l] \in \mathbb{H}_{++}^{N_d \times N_d}$ are the positive definite receive side spatial correlation matrices for the lth MIMO channel taps, with the corresponding positive definite transmit side spatial correlation matrices being represented by $\boldsymbol{\Theta}_s[l] \in \mathbb{H}_{++}^{N_s \times N_s}$ and $\boldsymbol{\Theta}_r[l] \in \mathbb{H}_{++}^{N_r \times N_r}$. The spatial correlation matrices represent the second order statistics of the MIMO channels and depend on parameters such as the average angle of arrival/departure, angle spread, antenna spacing, and wavelength [12, 64–66, 68]. Although each delay path is spatially correlated on both the transmit and receive side, we assume that the different delay paths are uncorrelated. Thus the matrices $\boldsymbol{\mathcal{H}}_{sw}[l] \in \mathbb{C}^{N_r \times N_s}$ and $\boldsymbol{\mathcal{H}}_{rw}[l] \in \mathbb{C}^{N_d \times N_r}$ in (2.8) and (2.9) are modelled as having independently identically distributed (i.i.d.) complex Gaussian entries with zero mean and variances $\sigma_{h_s}^2[l] \in \mathbb{R}_+$ and $\sigma_{h_r}^2[l] \in \mathbb{R}_+$, respectively. In other words these matrices have matrix variate complex Gaussian distibutions given by

$$\boldsymbol{\mathcal{H}}_{sw}[l] \sim \mathcal{CN}\left(\mathbf{0}_{N_r \times N_s}, \sigma_{h_s}^2[l]\boldsymbol{I}_{N_s} \otimes \boldsymbol{I}_{N_r}\right) \tag{2.10}$$

$$\boldsymbol{\mathcal{H}}_{rw}[l] \sim \mathcal{CN}\left(\mathbf{0}_{N_d \times N_r}, \sigma_{h_r}^2[l]\boldsymbol{I}_{N_r} \otimes \boldsymbol{I}_{N_d}\right). \tag{2.11}$$

Using Definition 1, the Kronecker product structures in (2.8)-(2.9), and the matrix variate Gaussian distributions in (2.10)-(2.11), it is straightforward to show that

$$\boldsymbol{\mathcal{H}}_s[l] \sim \mathcal{CN}\left(\mathbf{0}_{N_r \times N_s}, \sigma_{h_s}^2[l]\boldsymbol{\Theta}_s \otimes \boldsymbol{\Upsilon}_s\right) \tag{2.12}$$

$$\boldsymbol{\mathcal{H}}_r[l] \sim \mathcal{CN}\left(\mathbf{0}_{N_d \times N_r}, \sigma_{h_r}^2[l]\boldsymbol{\Theta}_r \otimes \boldsymbol{\Upsilon}_r\right). \tag{2.13}$$

To obtain (2.12)-(2.13) we have used the rules $\mathrm{vec}[\boldsymbol{AXB}] = (\boldsymbol{B}^T \otimes \boldsymbol{A})\mathrm{vec}[\boldsymbol{X}]$, $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC} \otimes \boldsymbol{BD})$, $(\boldsymbol{A} \otimes \boldsymbol{B})^H = (\boldsymbol{A}^H \otimes \boldsymbol{B}^H)$, as well as the fact that the spatial correlation matrices are positive definite Hermitian matrices.

2.2 MIMO OFDM Relay Signal Model

As is evident from the input-output relationships for the source-relay and relay-destination transmissions described in equations (2.1) and (2.3), respectively, the received symbols are corrupted by both spatial and temporal interferences caused by the use of multiple antennas as well as the frequency selective nature of the

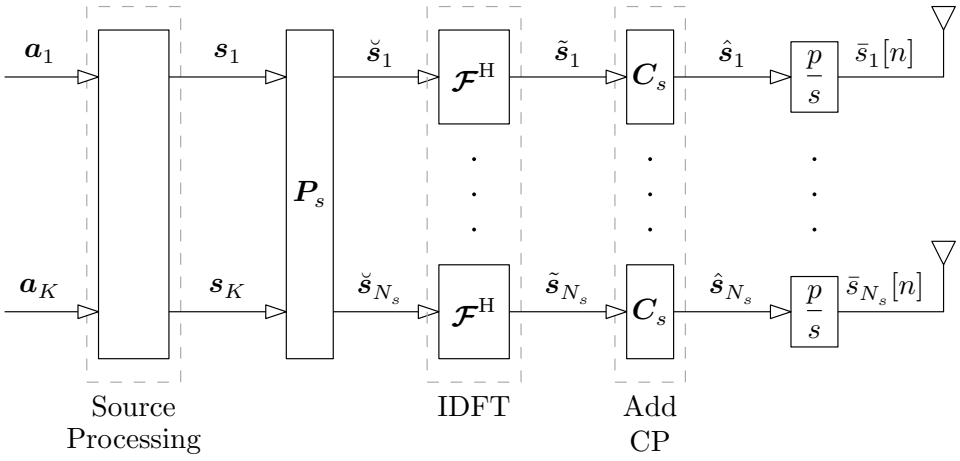

Figure 2.3: MIMO OFDM Source Transmitter structure.

wireless channels. In order to realise the potential benefits of enhanced data rates provided by spatial multiplexing of several data streams over the same communication medium, appropriate signal processing must be performed at the source, relay, and destination devices to combat these interferences. A first step is to deal with the temporal interference or ISI caused by frequency selectivity. OFDM has emerged as one of the most popular techniques in dealing with such interference and is widely regarded along with MIMO transmission to be one of the key components for realising the required high data rates of current and future generation wireless systems. OFDM must be employed for both the source-relay and relay-destination transmission stages in order to decouple the frequency selective channels into parallel non-frequency selective subcarriers. Since the OFDM transmit and receive stages for the source-relay and relay-destination transmission stages are similar, in the following we only give detailed analysis of the processing performed at the source transmission and relay receive stages as an example.

2.2.1 Source OFDM Transmission

The main processing stages for the utilisation of OFDM at the source device are depicted in Figure 2.3. Without loss of generality (w.l.o.g.) we focus on the transmission of a single OFDM block $\boldsymbol{a} \in \mathbb{C}^{\bar{N}}$ given by

$$\boldsymbol{a} \triangleq \left[\boldsymbol{a}_1^T, ..., \boldsymbol{a}_K^T\right]^T, \tag{2.14}$$

where $\boldsymbol{a}_k \in \mathbb{C}^{N_k}$ is the source transmit vector for the kth subcarrier, $1 \leq k \leq K$ denotes the subcarrier index, K is the total number of OFDM subcarriers, N_k is the number of symbols to be transmitted on the kth subcarrier, and we define $\bar{N} \triangleq \sum_{k=1}^{K} N_k$ as the total number of symbols over all subcarriers. The data vector

$\boldsymbol{a}$ is then precoded in some manner by the source device to produce the transmit vector $\boldsymbol{s} \in \mathbb{C}^{KN_s}$ given by

$$\boldsymbol{s} \triangleq \left[\boldsymbol{s}_1^T, ..., \boldsymbol{s}_K^T\right]^T, \tag{2.15}$$

where $\boldsymbol{s}_k \in \mathbb{C}^{N_s}$ is the transmit vector for the kth subcarrier. The precoded symbols are then re-arranged into N_s vectors $\breve{\boldsymbol{s}}_i \in \mathbb{C}^K$ which are to be transmitted from the ith transmit antenna. Each vector $\breve{\boldsymbol{s}}_i$ is obtained by stacking the ith element from each subcarrier vector $\boldsymbol{s}_k$ into a column vector and is given by

$$\breve{\boldsymbol{s}}_i = \left[[\boldsymbol{s}_1]_i, ..., [\boldsymbol{s}_K]_i\right]^T. \tag{2.16}$$

Here we use the notation $[\boldsymbol{s}_k]_i$ to denote the ith element of $\boldsymbol{s}_k$. By further stacking each $\breve{\boldsymbol{s}}_i$ into a single column vector $\breve{\boldsymbol{s}} \in \mathbb{C}^{KN_s}$ given by

$$\breve{\boldsymbol{s}} \triangleq \left[\breve{\boldsymbol{s}}_1^T, ..., \breve{\boldsymbol{s}}_{N_s}^T\right]^T, \tag{2.17}$$

we can see that the elements of $\breve{\boldsymbol{s}}$ are simply a permutation of those in $\boldsymbol{s}$. As shown in Figure 2.3, $\breve{\boldsymbol{s}}$ in (2.17) can equivalently be obtained from

$$\breve{\boldsymbol{s}} = \boldsymbol{P}_s \boldsymbol{s}, \tag{2.18}$$

where $\boldsymbol{P}_s \in \mathbb{R}_+^{KN_s \times KN_s}$ is a permutation matrix[2] such that (2.16) holds. Given (2.16) and (2.17) it can be shown that the permutation matrix $\boldsymbol{P}_s$ is

$$\boldsymbol{P}_s = \left[\boldsymbol{e}_1^{(s)}, \boldsymbol{e}_{K+1}^{(s)}, ..., \boldsymbol{e}_{(N_s-1)K+1}^{(s)}, \boldsymbol{e}_2^{(s)}, \boldsymbol{e}_{K+2}^{(s)}, ..., \boldsymbol{e}_{(N_s-1)K+2}^{(s)}, \boldsymbol{e}_K^{(s)}, \boldsymbol{e}_{2K}^{(s)}, ..., \boldsymbol{e}_{KN_s}^{(s)}\right], \tag{2.19}$$

where $\boldsymbol{e}_i^{(s)} \in \mathbb{R}_+^{KN_s}$ is the elementary unit vector containing all zero elements with only the ith element being equal to 1. After the permutation of the symbols, on each antenna branch, an inverse discrete Fourier transform (IDFT) is performed to transform the frequency domain symbols of $\breve{\boldsymbol{s}}_i$ into the time domain symbols $\tilde{\boldsymbol{s}}_i \in \mathbb{C}^{K \times K}$ given by

$$\tilde{\boldsymbol{s}}_i = \boldsymbol{\mathcal{F}}^H \breve{\boldsymbol{s}}_i, \tag{2.20}$$

[2] A permutation matrix is a square matrix where every row and every column contains only one element equal to 1 with all other elements being 0.

where $\boldsymbol{\mathcal{F}} \in \mathbb{C}^{K\times K}$ is the normalised K point discrete Fourier transform (DFT) matrix given by

$$\boldsymbol{\mathcal{F}} = \frac{1}{\sqrt{K}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{-j2\pi/K} & e^{-j4\pi/K} & \dots & e^{-j2\pi(K-1)/K} \\ 1 & e^{-j4\pi/K} & e^{-j8\pi/K} & \dots & e^{-j4\pi(K-1)/K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j2\pi(K-1)/K} & e^{-j4\pi(K-1)/K} & \dots & e^{-j2\pi(K-1)^2/K} \end{bmatrix}. \tag{2.21}$$

The scaling factor in (2.21) ensures that $\boldsymbol{\mathcal{F}}$ is unitary i.e. $\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{F}}^H = \boldsymbol{\mathcal{F}}^H\boldsymbol{\mathcal{F}} = \boldsymbol{I}_K$. We note that since $\boldsymbol{\mathcal{F}}$ is the normalised DFT matrix then $\boldsymbol{\mathcal{F}}^H$ is the normalised IDFT matrix. For later convenience let us now define a vector $\tilde{\boldsymbol{s}} \in \mathbb{C}^{KN_s}$ as

$$\tilde{\boldsymbol{s}} \triangleq \left[\tilde{\boldsymbol{s}}_1^T, ..., \tilde{\boldsymbol{s}}_{N_s}^T\right]^T. \tag{2.22}$$

With the definition of $\tilde{\boldsymbol{s}}$ in (2.22) and that of $\breve{\boldsymbol{s}}$ in (2.17) we can write (2.20) over all antenna branches compactly as

$$\tilde{\boldsymbol{s}} = \left(\text{diag}\{\boldsymbol{\mathcal{F}}^H\}_{i=1}^{N_s}\right)\breve{\boldsymbol{s}} \tag{2.23}$$

$$= \left(\boldsymbol{I}_{N_s} \otimes \boldsymbol{\mathcal{F}}^H\right)\breve{\boldsymbol{s}}, \tag{2.24}$$

where we have used the fact that, for a matrix $\boldsymbol{A} \in \mathbb{C}^{N\times M}$, $\boldsymbol{I}_P \otimes \boldsymbol{A}$ produces a $PN \times PM$ block diagonal matrix with $\boldsymbol{A}$ on each diagonal block.

After performing the IDFT a cyclic prefix (CP) is added to each $\tilde{\boldsymbol{s}}_i$, which acts as a guard interval to eliminate interblock interference (IBI). With the channel impulse responses being of order L, the CP is required to be of length L in order to completely eliminate IBI. The CP is added by taking the last L symbols from $\tilde{\boldsymbol{s}}_i$ and appending them to the start of the block to produce the vector $\hat{\boldsymbol{s}}_i \in \mathbb{C}^{K+L}$. The addition of the CP can be written using matrix-vector notation as

$$\hat{\boldsymbol{s}}_i = \underbrace{\begin{bmatrix} \boldsymbol{0}_{L\times K+L} & \boldsymbol{I}_L \\ \boldsymbol{I}_{K-L} & \boldsymbol{0}_{K-L\times L} \\ \boldsymbol{0}_{L\times K-L} & \boldsymbol{I}_L \end{bmatrix}}_{\boldsymbol{C}_s} \tilde{\boldsymbol{s}}_i. \tag{2.25}$$

Finally, the symbols contained in $\hat{\boldsymbol{s}}_i$ are converted from parallel to serial for transmission, resulting in the sequence $s_i[n]$, which is then transmitted from the ith source antenna to the relaying device.

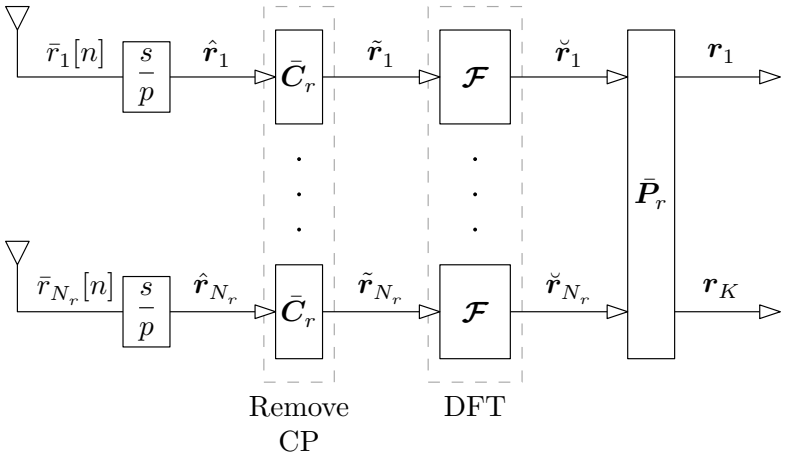

Figure 2.4: MIMO OFDM Relay Receiver structure.

2.2.2 Relay OFDM Reception

The receive side processing performed at the relay device for the first stage transmission is depicted in Figure 2.4. The relay acquires the vectors $\boldsymbol{r}[n]$, characterised by (2.1), over $n = K+L$ signalling intervals. On each relay receiver antenna branch the vectors $\hat{\boldsymbol{r}}_j \in \mathbb{C}^{K+L}$ are produced by stacking the symbols $r_j[n]$ over $K + L$ intervals into column vectors. The CP that was added by the source is then removed by discarding the first L symbols in $\hat{\boldsymbol{r}}_j$ to produce the vectors $\tilde{\boldsymbol{r}}_j \in \mathbb{C}^{K+L}$. The removal of the CP ensures the complete elimination of IBI since, assuming the CP was of length L, the IBI is restricted to the first L samples of $\hat{\boldsymbol{r}}_j$. The CP removal can be written as

$$\tilde{\boldsymbol{r}}_j = \underbrace{\begin{bmatrix} \boldsymbol{0}_{K\times L} & \boldsymbol{I}_K \end{bmatrix}}_{\bar{\boldsymbol{C}}_r} \hat{\boldsymbol{r}}_j. \tag{2.26}$$

The DFT is then performed on (2.26) to produce the frequency domain symbols $\breve{\boldsymbol{r}}_j \in \mathbb{C}^K$ given by

$$\breve{\boldsymbol{r}}_j = \boldsymbol{\mathcal{F}}\tilde{\boldsymbol{r}}_j, \tag{2.27}$$

which over all antenna branches can be written compactly as

$$\breve{\boldsymbol{r}} = \left(\mathrm{diag}\{\boldsymbol{\mathcal{F}}\}_{j=1}^{N_r}\right)\tilde{\boldsymbol{r}} \tag{2.28}$$

$$= \left(\boldsymbol{I}_{N_r} \otimes \boldsymbol{\mathcal{F}}\right)\tilde{\boldsymbol{r}}. \tag{2.29}$$

Here we define the vectors $\breve{\boldsymbol{r}} \in \mathbb{C}^{KN_r}$ and $\tilde{\boldsymbol{r}} \in \mathbb{C}^{KN_r}$ as

$$\breve{\boldsymbol{r}} \triangleq \left[\breve{\boldsymbol{r}}_1^T, ..., \breve{\boldsymbol{r}}_{N_r}^T\right]^T \tag{2.30}$$

$$\tilde{\boldsymbol{r}} \triangleq \left[\tilde{\boldsymbol{r}}_1^T, ..., \tilde{\boldsymbol{r}}_K^T\right]^T. \tag{2.31}$$

We note that to obtain (2.29) we have again used the fact that for a matrix $\boldsymbol{A} \in \mathbb{C}^{N\times M}$ we have $\boldsymbol{I}_P \otimes \boldsymbol{A} = \text{diag}\{\boldsymbol{A}\}_{i=1}^P$. The relay now undoes the permutation that was applied to the frequency domain symbols $\boldsymbol{s}$ by the source device (see (2.16)-(2.19)). Similar to the process carried out by (2.16)-(2.19), the received vector over all subcarriers $\boldsymbol{r} \in \mathbb{C}^{KN_r}$ is produced by

$$\boldsymbol{r} = \bar{\boldsymbol{P}}_r \breve{\boldsymbol{r}}, \tag{2.32}$$

where the received vector $\boldsymbol{r}$ and the permutation matrix $\bar{\boldsymbol{P}}_r \in \mathbb{R}_+^{KN_r \times KN_r}$ are defined as

$$\boldsymbol{r} \triangleq \left[\boldsymbol{r}_1^T, ..., \boldsymbol{r}_K^T\right]^T \tag{2.33}$$

$$\bar{\boldsymbol{P}}_r \triangleq \left[\boldsymbol{e}_1^{(r)}, \boldsymbol{e}_{N_r+1}^{(r)}, ..., \boldsymbol{e}_{(K-1)N_r+1}^{(r)}, \boldsymbol{e}_2^{(r)}, \boldsymbol{e}_{N_r+2}^{(r)}, ..., \boldsymbol{e}_{(K-1)N_r+2}^{(r)}, \boldsymbol{e}_{N_r}^{(r)}, \boldsymbol{e}_{2N_r}^{(r)}, ..., \boldsymbol{e}_{KN_r}^{(r)}\right]. \tag{2.34}$$

In (2.33) $\boldsymbol{r}_k$ signifies the received vector at the relay for the kth subcarrier. This concludes the OFDM processing for the source-relay transmission stage.

2.2.3 Source-Relay OFDM Subcarriers

We now show that the previously described source transmit and relay receive processing results in the frequency domain source transmit symbols effectively being transmitted over orthogonal frequency domain channels. In other words we show that the use of OFDM decomposes the frequency selective MIMO source-relay channel into a number of parallel non frequency selective subcarriers. To this end we note that the additon of the CP at the source and its removal at the destination converts the channel convolution into a block circular convolution and the input-output relationship between $\tilde{\boldsymbol{r}}$ and $\tilde{\boldsymbol{s}}$ can be expressed as

$$\bar{\boldsymbol{P}}_r \tilde{\boldsymbol{r}} = \boldsymbol{\mathcal{H}}_s \boldsymbol{P}_s^H \tilde{\boldsymbol{s}} + \tilde{\boldsymbol{v}}_s, \tag{2.35}$$

where $\tilde{\boldsymbol{v}}_s \in \mathbb{C}^{KN_r}$ is the vector of additive time domain noise samples. In (2.35) we define the block circulant matrix $\boldsymbol{\mathcal{H}}_s \in \mathbb{C}^{KN_r \times KN_s}$ as

$$\boldsymbol{\mathcal{H}}_s \triangleq \begin{bmatrix} \boldsymbol{\mathcal{H}}_s[0] & \mathbf{0}_{N_r \times N_s} & \cdots & \mathbf{0}_{N_r \times N_s} & \boldsymbol{\mathcal{H}}_s[L] & \cdots & \boldsymbol{\mathcal{H}}_s[1] \\ \boldsymbol{\mathcal{H}}_s[1] & \boldsymbol{\mathcal{H}}_s[0] & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \boldsymbol{\mathcal{H}}_s[1] & \ddots & \ddots & & \ddots & \boldsymbol{\mathcal{H}}_s[L] \\ \boldsymbol{\mathcal{H}}_s[L] & \vdots & \ddots & \ddots & \ddots & & \mathbf{0}_{N_r \times N_s} \\ \mathbf{0}_{N_r \times N_s} & \boldsymbol{\mathcal{H}}_s[L] & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \mathbf{0}_{N_r \times N_s} \\ \mathbf{0}_{N_r \times N_s} & \cdots & \mathbf{0}_{N_r \times N_s} & \boldsymbol{\mathcal{H}}_s[L] & \cdots & \boldsymbol{\mathcal{H}}_s[1] & \boldsymbol{\mathcal{H}}_s[0] \end{bmatrix}. \tag{2.36}$$

Block circulant matrices are fully defined by their first block column and have the property that they can be block diagonalised through pre multiplication with $\boldsymbol{\mathcal{F}} \otimes \boldsymbol{I}$ and post multiplication by $\boldsymbol{\mathcal{F}}^H \otimes \boldsymbol{I}$. We now note that by substituting (2.18) into (2.24) it can be shown that

$$\tilde{\boldsymbol{s}} = \left(\boldsymbol{I}_{N_s} \otimes \boldsymbol{\mathcal{F}}^H\right) \boldsymbol{P}_s \boldsymbol{s}, \tag{2.37}$$

$$= \boldsymbol{P}_s \left(\boldsymbol{\mathcal{F}}^H \otimes \boldsymbol{I}_{N_s}\right) \boldsymbol{s}, \tag{2.38}$$

where to obtain (2.38) we have utilised the structure of $\boldsymbol{P}_s$ given in (2.19). Similarly, by substituting (2.29) into (2.32) we can show the relationship

$$\tilde{\boldsymbol{r}} = \bar{\boldsymbol{P}}_r^H \left(\boldsymbol{\mathcal{F}}^H \otimes \boldsymbol{I}_{N_r}\right) \boldsymbol{r}. \tag{2.39}$$

Upon substituting (2.38) and (2.39) into (2.35) we obtain

$$\boldsymbol{r} = \underbrace{\left(\boldsymbol{\mathcal{F}} \otimes \boldsymbol{I}_{N_r}\right) \boldsymbol{\mathcal{H}}_s \left(\boldsymbol{\mathcal{F}}^H \otimes \boldsymbol{I}_{N_s}\right)}_{\boldsymbol{H}_s} \boldsymbol{s} + \underbrace{\left(\boldsymbol{\mathcal{F}} \otimes \boldsymbol{I}_{N_r}\right) \tilde{\boldsymbol{v}}_s}_{\boldsymbol{v}_s}, \tag{2.40}$$

where $\boldsymbol{v}_s \in \mathbb{C}^{KN_r}$ is the frequency domain noise vector and can be partitioned as

$$\boldsymbol{v}_s \triangleq \left[\boldsymbol{v}_{s,1}^T, ..., \boldsymbol{v}_{s,K}^T\right]^T, \tag{2.41}$$

with $\boldsymbol{v}_{s,k} \in \mathbb{C}^{N_r}$ representing the noise added on the kth subcarrier. It is worth noting here that, due to the unitary nature of $\boldsymbol{\mathcal{F}}$, the noise vector in the frequency domain has the exact same statistical properties of the noise in the time domain. Resulting from the block circulant structure of $\boldsymbol{\mathcal{H}}_s$ in (2.36), the matrix $\boldsymbol{H}_s$ in

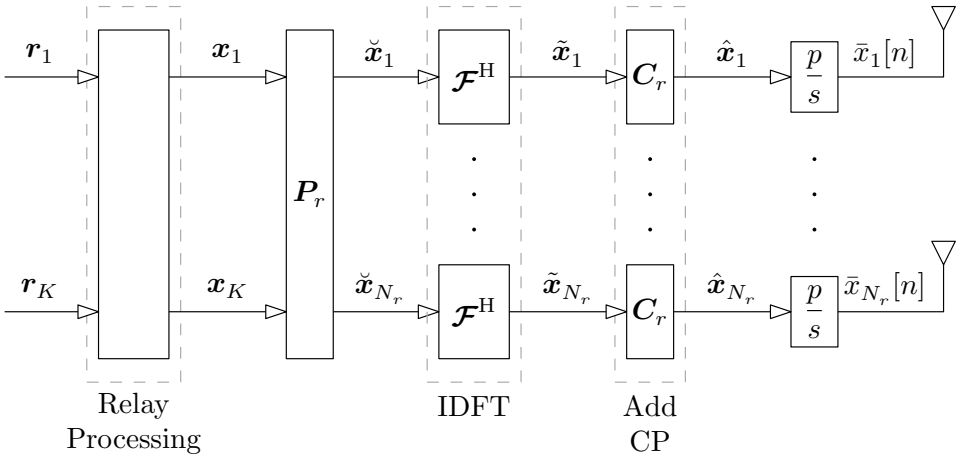

Figure 2.5: MIMO OFDM Relay Transmitter structure.

(2.40) is block diagonal and can be expressed as

$$\boldsymbol{H}_s = \mathrm{diag}\{\boldsymbol{H}_{s,k}\}_{k=1}^{K}, \tag{2.42}$$

where the matrix $\boldsymbol{H}_{s,k} \in \mathbb{C}^{N_r \times N_s}$ for the kth subcarrier is given by

$$\boldsymbol{H}_{s,k} = \sum_{l=0}^{L} \boldsymbol{\mathcal{H}}_s[l] e^{-j2\pi(k-1)l/K}. \tag{2.43}$$

It is clear from (2.40), that with the block diagonal structure of $\boldsymbol{H}_s$ in (2.42) and with the definitions of $\boldsymbol{s}$ and $\boldsymbol{r}$ in (2.15) and (2.33), that we can write

$$\boldsymbol{r}_k = \boldsymbol{H}_{s,k}\boldsymbol{s}_k + \boldsymbol{v}_{s,k}. \tag{2.44}$$

It is evident from (2.44) that the use of OFDM processing has elegantly converted the frequency selective source-relay channel transmission into the transmission of data over orthogonal frequency domain subcarriers characterised by (2.43).

2.2.4 Relay-Destination OFDM

As previously remarked, the processing performed by the relay and destination in the second stage MIMO OFDM relay transmission is similar to that for the first stage transmission just described. For completeness, the MIMO relay transmitter structure is shown in Figure 2.5 and the destination structure is depicted in Figure 2.6. The previous analysis can easily be carried out for the second stage transmission simply by replacing the signals and processors in Figures 2.3 and 2.4 with those in Figures 2.5 and 2.6, respectively. We also note that the substitutions $N_s \rightarrow N_r$ and $N_r \rightarrow N_d$ should be used. Similar to our previous analysis

the use of OFDM results in the relay transmit symbols being communicated over parallel subcarrier channels. Specifically, for the kth subcarrier, we have

$$\boldsymbol{y}_k = \boldsymbol{H}_{r,k}\boldsymbol{x}_k + \boldsymbol{v}_{r,k}, \tag{2.45}$$

where $\boldsymbol{y}_k \in \mathbb{C}^{N_d}$ is the vector of received symbols on the kth subcarrier, $\boldsymbol{x}_k \in \mathbb{C}^{N_r}$ is the symbols transmitted by the relay on the kth subcarrier, and $\boldsymbol{v}_{r,k} \in \mathbb{C}^{N_d}$ is the noise vector for the kth subcarrier. The kth relay-destination subcarrier channel matrix $\boldsymbol{H}_{r,k} \in \mathbb{C}^{N_d \times N_r}$ is given by

$$\boldsymbol{H}_{r,k} = \sum_{l=0}^{L} \boldsymbol{\mathcal{H}}_r[l] e^{-j2\pi(k-1)l/K}. \tag{2.46}$$

By making the definitions

$$\boldsymbol{y} \triangleq \left[\boldsymbol{y}_1^T, ..., \boldsymbol{y}_K^T\right]^T \tag{2.47}$$

$$\boldsymbol{x} \triangleq \left[\boldsymbol{x}_1^T, ..., \boldsymbol{x}_K^T\right]^T \tag{2.48}$$

$$\boldsymbol{v}_r \triangleq \left[\boldsymbol{v}_{r,1}^T, ..., \boldsymbol{v}_{r,K}^T\right]^T \tag{2.49}$$

$$\boldsymbol{H}_r \triangleq \text{diag}\{\boldsymbol{H}_{r,k}\}_{k=1}^K, \tag{2.50}$$

we can express the received signal in (2.45) over all subcarriers compactly as

$$\boldsymbol{y} = \boldsymbol{H}_r\boldsymbol{x} + \boldsymbol{v}_r. \tag{2.51}$$

From (2.51) (or equivalently from (2.45)) it can again be seen that the frequency selective relay-destination channel has been decoupled into the transmission of data over orthogonal subcarriers given in (2.46).

2.2.5 Equivalent MIMO OFDM Relay Model and Statistical Assumptions

From equations (2.40), (2.44), (2.45), and (2.51), we see that the MIMO OFDM relay system depicted in Figures 2.3-2.6 can be written solely in terms of frequency domain components. This leads to the simplified interpretation of the MIMO OFDM relay system shown in Figure 2.7. Throughout the remainder of this thesis we shall find it convenient to focus on this simplified model as opposed to the more complicated underlying structure given in Figures 2.3-2.6.

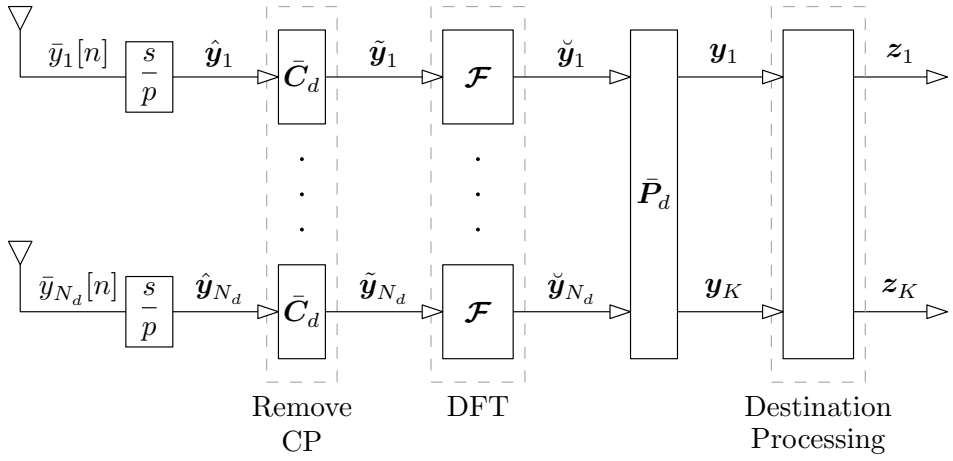

Figure 2.6: MIMO OFDM Destination Receiver structure.

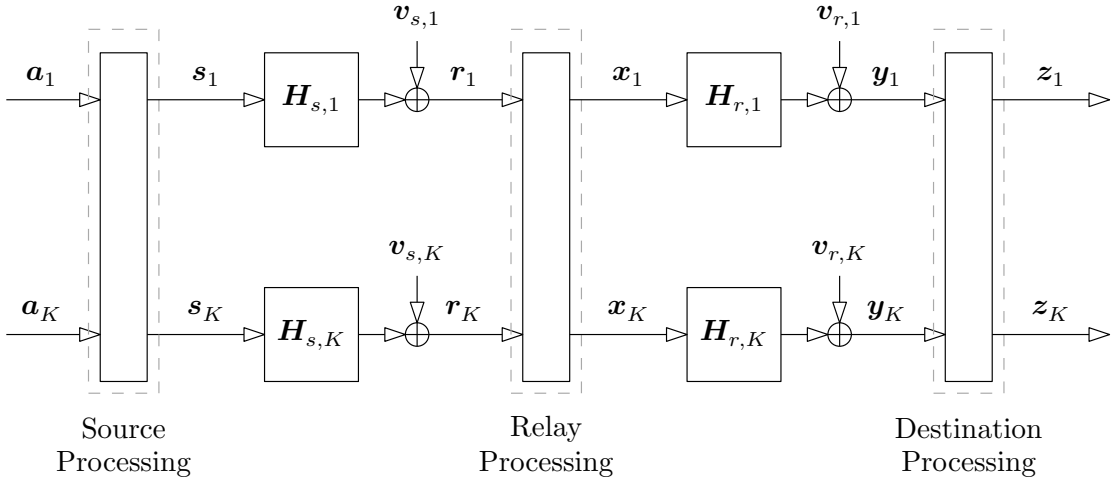

Figure 2.7: Equivalent MIMO OFDM relay model.

Hitherto we have made no particular assumptions on the statistical properties of the transmit or noise symbols. However, in the remainder of this thesis we shall find it convenient to make the following standard assumptions:

(A1.) The transmit symbols contained in $\boldsymbol{a}$, defined in (2.14), are assumed to be drawn from zero mean M-QAM signal constellations and are white with unit energy resulting in $\boldsymbol{R}_a \triangleq \mathbb{E}\{\boldsymbol{a}\boldsymbol{a}^H\} = \boldsymbol{I}_{\bar{N}}$. From the definition of $\boldsymbol{a}$ in (2.14) this means that the covariance matrix of $\boldsymbol{a}_k$ is $\boldsymbol{R}_{a,k} \triangleq \mathbb{E}\{\boldsymbol{a}_k\boldsymbol{a}_k^H\} = \boldsymbol{I}_{N_k}$. Furthermore, the cross covariance matrix between $\boldsymbol{a}_k$ and $\boldsymbol{a}_j$ is given by $\boldsymbol{R}_{a,kj} \triangleq \mathbb{E}\{\boldsymbol{a}_k\boldsymbol{a}_j^H\} = \boldsymbol{0}_{N_k \times N_j}$, $\forall k \neq j$.

(A2.) The elements of the additive noise vector $\boldsymbol{v}_s$, defined in (2.41), are drawn from i.i.d. Gaussian distributions with zero mean and variance $\sigma_{v_s}^2 \in \mathbb{R}_+$. The noise vector therefore has covariance matrix $\boldsymbol{R}_{v_s} \triangleq \mathbb{E}\{\boldsymbol{v}_s\boldsymbol{v}_s^H\} = \sigma_{v_s}^2\boldsymbol{I}_{KN_r}$ which from (2.41) implies that $\boldsymbol{R}_{v_s,k} \triangleq \mathbb{E}\{\boldsymbol{v}_{s,k}\boldsymbol{v}_{s,k}^H\} = \sigma_{v_s}^2\boldsymbol{I}_{N_r}$, as well as $\boldsymbol{R}_{v_s,kj} \triangleq \mathbb{E}\{\boldsymbol{v}_{s,k}\boldsymbol{v}_{s,j}^H\} = \boldsymbol{0}_{N_r \times N_r}$, $\forall j \neq k$.

(A3.) The elements of the additive noise vector $\boldsymbol{v}_r$, defined in (2.49), are drawn from i.i.d. Gaussian distributions with zero mean and variance $\sigma^2_{v_r} \in \mathbb{R}_+$. We thus have $\boldsymbol{R}_{v_r} \triangleq \mathbb{E}\{\boldsymbol{v}_r \boldsymbol{v}_r^H\} = \sigma^2_{v_r} \boldsymbol{I}_{KN_d}$ which from (2.49) implies that we have the covariance matrices $\boldsymbol{R}_{v_r,k} \triangleq \mathbb{E}\{\boldsymbol{v}_{r,k} \boldsymbol{v}_{r,k}^H\} = \sigma^2_{v_r} \boldsymbol{I}_{N_d}$ for each subcarrier, and the cross covariance matrices $\boldsymbol{R}_{v_r,kj} \triangleq \mathbb{E}\{\boldsymbol{v}_{r,k} \boldsymbol{v}_{r,j}^H\} = \boldsymbol{0}_{N_d \times N_d}$, $\forall j \neq k$.

(A4.) The vectors $\boldsymbol{v}_s$ and $\boldsymbol{v}_r$ are uncorrelated with each other and we have $\mathbb{E}\{\boldsymbol{v}_{s,k} \boldsymbol{v}_{r,j}^H\} = \boldsymbol{0}_{N_r \times N_d}$, $\forall k, j$, as well as $\mathbb{E}\{\boldsymbol{v}_{r,k} \boldsymbol{v}_{s,j}^H\} = \boldsymbol{0}_{N_d \times N_r}$, $\forall k, j$. Furthermore, the noise vectors are uncorrelated with the transmit vector $\boldsymbol{a}$ and we have $\mathbb{E}\{\boldsymbol{a}_k \boldsymbol{v}_{s,j}^H\} = \boldsymbol{0}_{N_k \times N_r}$, $\forall k, j$, and $\mathbb{E}\{\boldsymbol{v}_{s,k} \boldsymbol{a}_j^H\} = \boldsymbol{0}_{N_r \times N_k}$, $\forall k, j$, as well as $\mathbb{E}\{\boldsymbol{a}_k \boldsymbol{v}_{r,j}^H\} = \boldsymbol{0}_{N_k \times N_d}$, $\forall k, j$, and $\mathbb{E}\{\boldsymbol{v}_{r,k} \boldsymbol{a}_j^H\} = \boldsymbol{0}_{N_d \times N_k}$, $\forall k, j$.

Hitherto, we have established that the use of OFDM results is the transmission of data over orthogonal subcarriers as depicted by the equivalent MIMO OFDM relay model in Figure 2.7. Although OFDM has been utilised to combat temporal interference in our MIMO relay system, it is evident there is still interference present for each OFDM subcarrier. In order for the reliable detection of symbols at the destination, appropriate transceiver designs must be employed to combat the remaining interference. Transceivers for MIMO systems can be either linear or non-linear depending on the type of transmit and receive processing that is performed. It is well known for point-to-point MIMO systems that non-linear transceivers such as DFE and THP can offer significant improvement in performance when compared to linear transceiver designs [24–35]. This motivates the study of non-linear techniques for the considered MIMO OFDM relaying system.

2.3 Non-linear Transceiver Models

In this section we discuss the signal models for DFE and THP transceivers to combat the interference that still remains in the MIMO OFDM relaying system of Figure 2.7. Such transceivers can either be subcarrier non-cooperative or subcarrier cooperative depending on the structure of the global transmit and receive processors [16, 45]. For subcarrier non-cooperative transceivers the processors are restricted to operate independently for each subcarrier, whereas subcarrier cooperative transceivers allow for data to be mixed between subcarriers. As discussed in [16, 45] the subcarrier non-cooperative approach results in the most general problem formulation and subcarrier cooperative designs can be straightforwardly obtained from the subcarrier non-cooperative optimisation problem. We therefore restrict our attention to subcarrier non-cooperative DFE and THP transceivers.

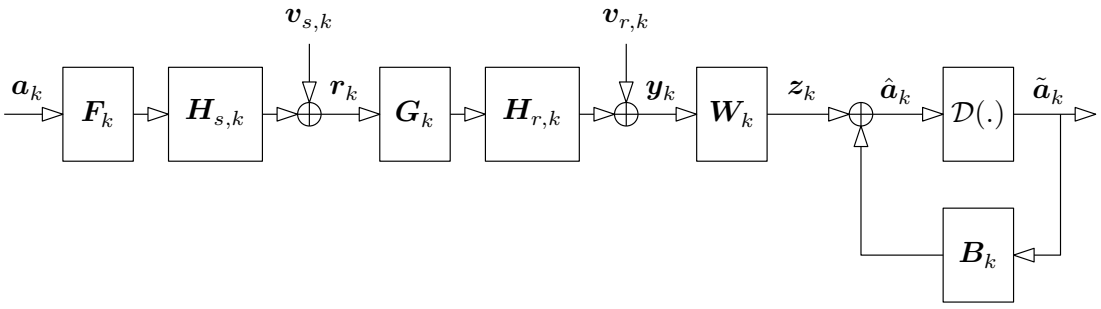

Figure 2.8: Signal Model for MIMO OFDM relay system with DFE.

In the following, we firstly discuss the DFE and THP signal models for the considered MIMO OFDM relay model and derive the error covariance matrix for these transceivers. The error covariance matrices will play a vital role in formulating the optimisation problem for deriving the optimal DFE and THP processors since most commonly used design criteria are inherently related to the elements of the error covariance matrix [16, 17, 45, 50].

2.3.1 DFE Signal Model and Error Covariance Matrix

In the DFE transceiver for the MIMO OFDM relay model in Figure 2.7 we introduce linear precoders at the source and relay, as well as a DFE at the receiver, for each subcarrier. This configuration is illustrated in Figure 2.8 for the kth subcarrier. We stress that this transceiver configuration is utilised for every subcarrier of the MIMO OFDM relay model in Figure 2.7. In the following we discuss the processing performed for a single subcarrier. At the source the symbols $\boldsymbol{a}_k$ are linearly precoded by $\boldsymbol{F}_k \in \mathbb{C}^{N_s \times N_k}$ and are transmitted over the source-relay subcarrier channel $\boldsymbol{H}_{s,k}$, which is given in (2.43). The received symbols $\boldsymbol{r}_k$ at the relay device are then given by

$$\boldsymbol{r}_k = \boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{a}_k + \boldsymbol{v}_{s,k}, \tag{2.52}$$

where $\boldsymbol{v}_{s,k}$ is the AWGN vector. At the relay device the received symbols in (2.52) are linearly precoded by $\boldsymbol{G}_k \in \mathbb{C}^{N_r \times N_r}$ and the resulting symbols are transmitted over the relay-destination subcarrier matrix $\boldsymbol{H}_{r,k}$, which is characterised in (2.46). The symbols $\boldsymbol{y}_k$ received at the destination are then given by

$$\boldsymbol{y}_k = \boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{r}_k + \boldsymbol{v}_{r,k}, \tag{2.53}$$

with $\boldsymbol{v}_{r,k}$ being the AWGN vector added by the relay-destination transmission. The symbols $\boldsymbol{y}_k$ are processed by the DFE feedforward matrix $\boldsymbol{W}_k \in \mathbb{C}^{N_k \times N_d}$ to

produce $\boldsymbol{z}_k = \boldsymbol{W}_k \boldsymbol{y}_k \in \mathbb{C}^{N_k}$, which using (2.52) and (2.53) results in

$$\boldsymbol{z}_k = \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{a}_k + \boldsymbol{v}_{s,k} \right) + \boldsymbol{v}_{r,k} \right). \tag{2.54}$$

Decision feedback detection is then performed where the DFE receiver consists of a decision detection unit $\mathcal{D}(.)$ and a strictly upper right triangular feedback matrix $\boldsymbol{B}_k \in \mathbb{C}^{N_k \times N_k}$. The action of the DFE receiver is to successively detect symbols and then subtract their interference from the remaining undetected symbols. This process is known as successive interference cancellation (SIC) and can be stated mathematically as [24–31]

$$[\tilde{\boldsymbol{a}}_k]_m = \mathcal{D} \left([\boldsymbol{z}_k]_m - \sum_{n=m+1}^{N_k} [\boldsymbol{B}_k]_{mn} \, \tilde{\boldsymbol{a}}_n \right), \qquad m = N_k, N_k - 1, ..., 1, \tag{2.55}$$

where $\tilde{\boldsymbol{a}}_k \in \mathbb{C}^{N_k}$ is the vector of detected symbols. The SIC performed in (2.55) is equivalent to successively making decision on the symbols $\hat{\boldsymbol{a}}_k \in \mathbb{C}^{N_k}$ given by

$$\begin{aligned} \hat{\boldsymbol{a}}_k &= \boldsymbol{z}_k - \boldsymbol{B}_k \tilde{\boldsymbol{a}}_k & (2.56) \\ &= \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{a}_k + \boldsymbol{v}_{s,k} \right) + \boldsymbol{v}_{r,k} \right) - \boldsymbol{B}_k \tilde{\boldsymbol{a}}_k, & (2.57) \end{aligned}$$

where we have used (2.54) to obtain (2.57). The error between the input to the decision device, given in (2.57), and the transmitted data vectors $\boldsymbol{a}_k$ provides a useful measure of performance quality for the DFE transceiver and is constructed as $\boldsymbol{e}_k \triangleq \hat{\boldsymbol{a}}_k - \boldsymbol{a}_k \in \mathbb{C}^{N_k}$ which using (2.57) can be expanded as

$$\begin{aligned} \boldsymbol{e}_k &= \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{a}_k + \boldsymbol{v}_{s,k} \right) + \boldsymbol{v}_{r,k} \right) - \boldsymbol{B}_k \tilde{\boldsymbol{a}}_k - \boldsymbol{a}_k & (2.58) \\ &= \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right) \boldsymbol{a}_k + \boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{v}_{s,k} + \boldsymbol{W}_k \boldsymbol{v}_{r,k}. & (2.59) \end{aligned}$$

To obtain (2.59), we have made the standard assumption that previous symbols have been detected correctly, i.e. $\tilde{\boldsymbol{a}}_k = \boldsymbol{a}_k$ (see e.g. [24–31]), and we have defined $\boldsymbol{U}_k \triangleq \boldsymbol{B}_k + \boldsymbol{I}_{N_k}$, which by construction is a unit diagonal upper right triangular matrix (recall that $\boldsymbol{B}_k$ is a strictly upper right triangular matrix). We now define the error covariance matrix $\boldsymbol{R}_{e,k} \triangleq \mathbb{E}\{\boldsymbol{e}_k \boldsymbol{e}_k^H\} \in \mathbb{C}^{N_k \times N_k}$, which using the error signal in (2.59) can be written as

$$\begin{aligned} \boldsymbol{R}_{e,k} &= \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right) \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right)^H \\ &\quad + \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \right) \boldsymbol{W}_k^H. \end{aligned} \tag{2.60}$$

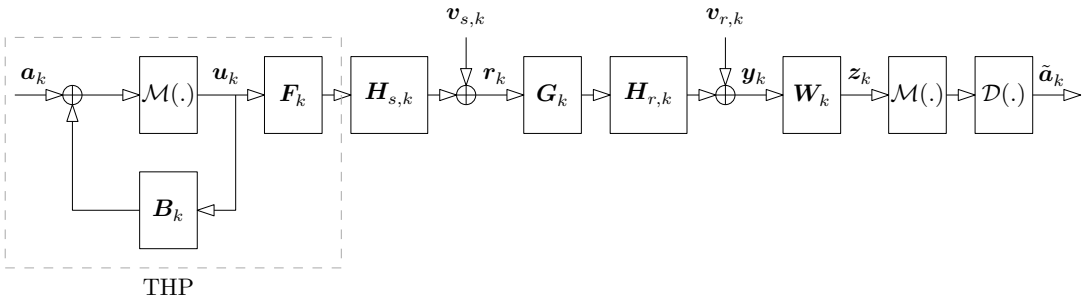

Figure 2.9: Signal Model for MIMO OFDM relay system with THP.

To obtain (2.60) we have used the assumptions that $\boldsymbol{a}_k$ are white with unit variance and are uncorrelated with the noise signals, as well as the assumption that the noise vectors $\boldsymbol{v}_{s,k}$ and $\boldsymbol{v}_{r,k}$ have elements drawn from independent complex Gaussian distributions with zero mean and variances $\sigma^2_{v_s}$ and $\sigma^2_{v_r}$ respectively (see the statistical assumptions in **(A1.)-(A4.)** of Section 2.2.5).

2.3.2 THP Signal Model and Error Covariance Matrix

A main drawback of DFE transceivers is that, due to the action of successively making decisions on symbols and using the detected symbol for interference cancellation, they can suffer from error propagation when symbols are detected incorrectly. The level of error propagation that occurs in DFE transceivers depends on the size of the data block that is to be detected, with larger data blocks having the potential for severe error propagation. THP is a non-linear precoding technique that can circumvent this problem, and does so by essentially moving the DFE feedback device to the transmitter where all signals are completely known and as such error propagation does not occur. THP was originally proposed in [69, 70] for the mitigation of ISI in frequency selective SISO systems and has since been extensively studied for point-to-point MIMO systems in e.g. [32–35].

We consider the use of THP for the MIMO OFDM relay model in Figure 2.7. For each subcarrier a TH precoder is introduced at the source along with a linear precoder at the relay and linear equalisation at the destination. This configuration is shown in Figure 2.9 for the kth subcarrier. We again emphasise that this configuration is utilised for every subcarrier. The TH precoder consists of a feedback loop comprising a strictly upper right triangular matrix[3] $\boldsymbol{B}_k$ as well as a modulo device $\mathcal{M}(.)$. The use of THP is tightly related to the signal constellation of the data symbols $\boldsymbol{a}_k$, which are assumed to be drawn from a zero mean, unit

[3]Traditionally the feedback matrix for THP is assumed to be strictly lower left triangular [33–35]. However for the purpose of formulating a unified optimisation problem for both the DFE and THP transceivers we consider the feedback matrix to be strictly upper right triangular.

variance, M-QAM signal constellation (see assumption **(A1.)** of Section 2.2.5). The Voronoi region $\mathcal{U}$ of this constellation is a square of side length D [32]. To highlight the need of the non-linear modulo operator let us firstly analyse the THP system in its absence. In this case the elements of the vector $\boldsymbol{u}_k \in \mathbb{C}^{N_k}$ are recursively computed according to [33, 34]

$$[\boldsymbol{u}_k]_m = [\boldsymbol{a}_k]_m - \sum_{n=m+1}^{N_k} [\boldsymbol{B}_k]_{mn}[\boldsymbol{u}_k]_n, \qquad m = N_k, N_k-1, ..., 1, \tag{2.61}$$

which can be equivalently written using vector/matrix notation as

$$\boldsymbol{u}_k = \boldsymbol{U}_k^{-1}\boldsymbol{a}_k. \tag{2.62}$$

In (2.62) we again define $\boldsymbol{U}_k \triangleq \boldsymbol{B}_k + \boldsymbol{I}_{N_k}$ which is a unit diagonal upper right triangular matrix since $\boldsymbol{B}_k$ is a strictly upper right triangular matrix. We now note that a direct transmission of the data vector $\boldsymbol{a}_k$ results in the average power consumption of the source being given by

$$\mathrm{tr}\Big\{\mathbb{E}\Big\{\boldsymbol{a}_k\boldsymbol{a}_k^H\Big\}\Big\} = \mathrm{tr}\big\{\boldsymbol{I}_{N_k}\big\} = N_k, \tag{2.63}$$

where we have used the statistical assumption of the data vector $\boldsymbol{a}_k$ given in **(A1.)** of Section 2.2.5. On the other hand, the average transmit power consumed by the transmission of the vector $\boldsymbol{u}_k$ in (2.62) is given by

$$\begin{aligned}
\mathrm{tr}\Big\{\mathbb{E}\Big\{\boldsymbol{u}_k\boldsymbol{u}_k^H\Big\}\Big\} &= \mathrm{tr}\Big\{\boldsymbol{U}_k^{-1}\mathbb{E}\Big\{\boldsymbol{a}_k\boldsymbol{a}_k^H\Big\}\boldsymbol{U}_k^{-\mathrm{H}}\Big\} && (2.64)\\
&= \mathrm{tr}\Big\{\boldsymbol{U}_k^{-1}\boldsymbol{U}_k^{-\mathrm{H}}\Big\} && (2.65)\\
&= N_k + \sum_{\forall m>n} \big|\big[\boldsymbol{U}_k^{-1}\big]_{mn}\big|^2 && (2.66)
\end{aligned}$$

To obtain (2.66) we have used the statistical assumption for $\boldsymbol{a}_k$ given in **(A1.)** of Section 2.2.5, as well as the fact that $\boldsymbol{U}_k$ is unit diagonal triangular. By comparing (2.66) and (2.63) it is clear that the power consumed by the transmission of $\boldsymbol{u}_k$ is higher than that for a direct transmission of $\boldsymbol{a}_k$. In order to limit this increase of transmission power the modulo operator $\mathcal{M}(.)$ is introduced into the THP feedback loop. With the inclusion of the modulo operator the symbols in $\boldsymbol{u}_k$ are now recursively computed according to

$$[\boldsymbol{u}_k]_m = \mathcal{M}\left([\boldsymbol{a}_k]_m - \sum_{n=m+1}^{N_k} [\boldsymbol{B}_k]_{mn}[\boldsymbol{u}_k]_n\right), \qquad m = N_k, N_k-1, ..., 1. \tag{2.67}$$

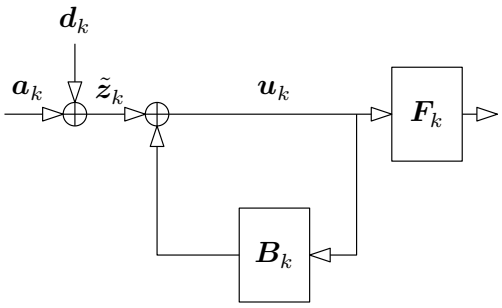

Figure 2.10: Equivalent THP transmitter structure.

The effect of the modulo operator in (2.67) is equivalent to the addition of integer multiples of D to both the real and imaginary components of $\boldsymbol{a}_k$ and is such that, after processing by the THP feedback loop, the symbols in $\boldsymbol{u}_k$ are bounded by the region $\mathcal{U}$ thus limiting the transmission power [32–35, 54]. The elements of $\boldsymbol{u}_k$ in (2.67) can therefore be equivalently written as

$$[\boldsymbol{u}_k]_m = [\boldsymbol{a}_k]_m + [\boldsymbol{d}_k]_m - \sum_{n=m+1}^{N_k} [\boldsymbol{B}_k]_{mn}[\boldsymbol{u}_k]_n, \qquad m = N_k, N_k - 1, ..., 1, \tag{2.68}$$

where the real and imaginary components of the elements in $\boldsymbol{d}_k \in \mathbb{C}^{N_k}$ are appropriate integer multiples of D such that the elements of $\boldsymbol{u}_k$ are bounded by $\mathcal{U}$. This observation leads us to the equivalent THP transmitter structure depicted in Figure 2.10. The selection of the elements in $\boldsymbol{d}_k$ is done implicitly through the action of the modulo unit. The addition of the vector $\boldsymbol{d}_k$ results in the modified data symbols $\tilde{\boldsymbol{z}}_k = \boldsymbol{a}_k + \boldsymbol{d}_k \in \mathbb{C}^{N_k}$, which are drawn from an extended version of the signal constellation $\mathcal{A}$, being fed to the feedback unit. The vector $\boldsymbol{u}_k$ in (2.68) can be written using vector/matrix notation as

$$\boldsymbol{u}_k = \boldsymbol{U}_k^{-1}\tilde{\boldsymbol{z}}_k. \tag{2.69}$$

It is important to note that, although the use of the modulo operator in THP limits the increase in transmission power, their is still a slight increase in transmit power due to the linear prefiltering by $\boldsymbol{U}^{-1}$ in (2.69). However for moderate to high M-QAM signal constellations this slight increase becomes negligible [32–35]) and we can make the statistical assumption

$$\mathbb{E}\Big\{\boldsymbol{u}_k\boldsymbol{u}_k^H\Big\} = \mathbb{E}\Big\{\boldsymbol{a}_k\boldsymbol{a}_k^H\Big\} = \boldsymbol{I}_{N_k}. \tag{2.70}$$

Before proceeding, we illustrate the effect of THP processing in Figure 2.11,

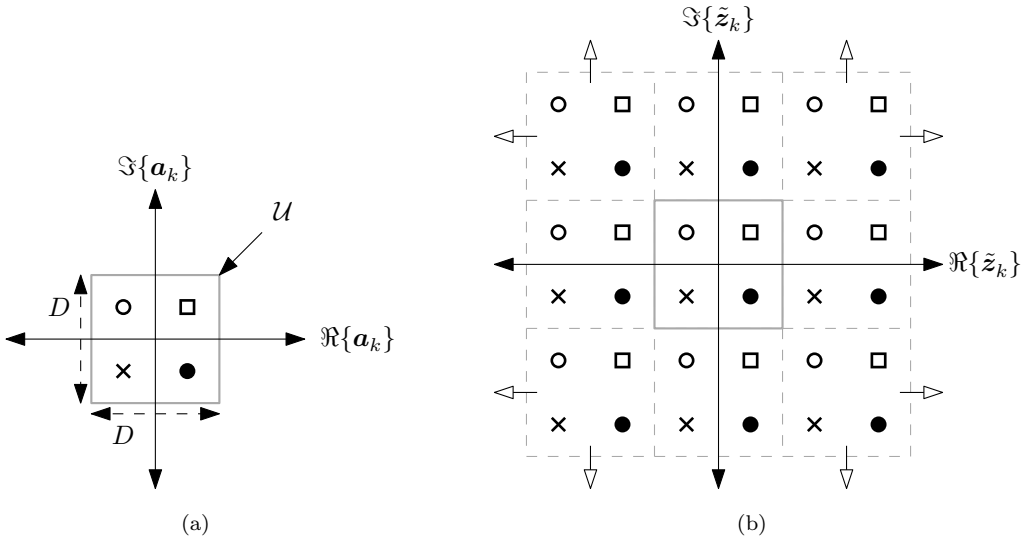

Figure 2.11: (a) Original signal constellation $\mathcal{A}$ for 4-QAM symbols. (b) Extended signal constellation

where we consider the use of a 4-QAM signal constellation. Figure 2.11 (a) depicts the original constellation $\mathcal{A}$, from which the elements of the data vector $\boldsymbol{a}_k$ are selected from, along with the Voronoi region $\mathcal{U}$ which bounds the elements of $\boldsymbol{u}_k$ that result from THP processing. Figure 2.11 (b) depicts the extended signal constellation from which the elements of the modified data symbols in $\tilde{\boldsymbol{z}}_k$ are drawn. All points that are seperated in the real and/or imaginary planes by an integer multiple of D represent the same symbol from the original signal constellation in Figure 2.11 (a).

The symbols $\boldsymbol{u}_k$ produced from TH precoding are then linearly precoded by the source precoding matrix $\boldsymbol{F}_k$ before being transmitted over the source-relay subcarrier channel $\boldsymbol{H}_{s,k}$. The resulting signal at the relay is then given by

$$\boldsymbol{r}_k = \boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{u}_k + \boldsymbol{v}_{s,k}. \tag{2.71}$$

The received symbols $\boldsymbol{r}_k$ at the relay are precoded by the relay matrix $\boldsymbol{G}_k$ and transmitted over the relay-destination subcarrier channel matrix $\boldsymbol{H}_{r,k}$ resulting in the received vector at the destination

$$\boldsymbol{y}_k = \boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{r}_k + \boldsymbol{v}_{r,k}. \tag{2.72}$$

At the destination linear equalisation is performed to produce $\boldsymbol{z}_k = \boldsymbol{W}_k \boldsymbol{y}_k$, which using (2.71) and (2.72) can be expanded as

$$\boldsymbol{z}_k = \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{u}_k + \boldsymbol{v}_{s,k} \right) + \boldsymbol{v}_{r,k} \right). \tag{2.73}$$

Ideally the symbols in (2.73) should be identical to those in $\tilde{\boldsymbol{z}}_k$. The symbols in (2.73) can then be passed to a modulo device $\mathcal{M}(.)$ which serves to compensate for the effect of the periodic extension to the original signal constellation that resulted due to the modulo operator at the transmitter. In other words the elements of $\boldsymbol{z}_k$ are modulo reduced such that they lie within the Voronoi region $\mathcal{U}$ (see e.g. [32–35, 54]). Finally the resulting symbols are quantised to the nearest symbol in the original signal constellation $\mathcal{A}$ to obtain $\tilde{\boldsymbol{a}}_k$, which is the estimate of the original data symbols $\boldsymbol{a}_k$.

We now derive the error covariance matrix for the THP transceiver. Differing from the DFE system, where an error signal was defined in terms of the source symbols $\boldsymbol{a}_k$, the error signal for the THP transceiver is defined in terms of the modified data symbols $\boldsymbol{z}_k$. The error signal for the THP transceiver is thus defined as $\boldsymbol{e}_k \triangleq \boldsymbol{z}_k - \tilde{\boldsymbol{z}}_k \in \mathbb{C}^{N_k}$, which using (2.73) and (2.69) can be expanded as

$$\boldsymbol{e}_k = \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right) \boldsymbol{u}_k + \boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{v}_{s,k} + \boldsymbol{W}_k \boldsymbol{v}_{r,k}. \tag{2.74}$$

We can now define the error covariance matrix as $\boldsymbol{R}_{e,k} \triangleq \mathbb{E}\{\boldsymbol{e}_k \boldsymbol{e}_k^H\} \in \mathbb{C}^{N_k \times N_k}$, which using the error signal in (2.74) can be written as

$$\begin{aligned} \boldsymbol{R}_{e,k} = & \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right) \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right)^H \\ & + \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \right) \boldsymbol{W}_k^H, \end{aligned} \tag{2.75}$$

where we have used the statistical assumptions in **(A1.)**-**(A4.)** of Section 2.2.5 as well as the statistical assumption in (2.70) to obtain (2.75). Interestingly, despite the differences in the DFE and THP models, both transceivers result in error covariance matrices with the same structure (c.f. (2.60) and (2.75)). This is a very appealing property since it shall allow us to derive the optimal processors for both systems using a single optimisation problem.

2.4 Minimum MSE Problem Formulation

A popular and one of the most commonly used criterion for transceiver designs that shall be considered throughout this thesis is the minimisation of the arithmetic MSE (also referred to as the minimisation of the sum MSE or more loosely simply

as the minimisation of the MSE). In this section we formulate this optimisation problem for both the DFE and THP MIMO OFDM relay transceivers, which can be done in a unified manner since the error covariance matrices for both transceivers have the same structure. The arithmetic MSE for the kth subcarrier is given by the arithmetic mean of the diagonal elements of $\boldsymbol{R}_{e,k}$ and can be written mathematically as $\sum_{i=1}^{N_k}[\boldsymbol{R}_{e,k}]_{ii}/N_k$ or equivalently as $\mathrm{tr}\{\boldsymbol{R}_{e,k}\}/N_k$. As well as minimising the arithmetic MSE we also wish to limit the average transmit power consumed by the source and relay terminals. For the DFE model in Figure 2.8 and the THP model in Figure 2.9 the transmit power consumed by the source on the kth subcarrier is

$$\mathrm{tr}\Big\{\mathbb{E}\Big\{\boldsymbol{F}_k\boldsymbol{a}_k\boldsymbol{a}_k^H\boldsymbol{F}_k^H\Big\}\Big\} = \mathrm{tr}\Big\{\mathbb{E}\Big\{\boldsymbol{F}_k\boldsymbol{u}_k\boldsymbol{u}_k^H\boldsymbol{F}_k^H\Big\}\Big\} = \mathrm{tr}\Big\{\boldsymbol{F}_k\boldsymbol{F}_k^H\Big\}, \tag{2.76}$$

where we have used the assumptions in **(A1.)** and (2.70). It can also be straightforwardly shown that the transmit power consumed by the relay on the kth subcarrier for both the DFE and THP models is

$$\mathrm{tr}\Big\{\mathbb{E}\Big\{\boldsymbol{G}_k\boldsymbol{r}_k\boldsymbol{r}_k^H\boldsymbol{G}_k^H\Big\}\Big\} = \mathrm{tr}\Big\{\boldsymbol{G}_k\left(\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)\boldsymbol{G}_k^H\Big\}, \tag{2.77}$$

where $\boldsymbol{r}_k$ is given in (2.52) for the DFE transceiver and in (2.71) for the THP transceiver. To obtain (2.77) we have also used the assumptions **(A1.)**, **(A2.)**, **(A4.)**, as well as (2.70)[4]. With the power consumed on the kth subcarrier by the source and relay terminals given in (2.76) and (2.77), respectively, we can now formulate the constrained optimisation problem to minimise the arithmetic MSE for both DFE and THP MIMO OFDM relay transceivers as

$$\min_{\boldsymbol{F}_k,\boldsymbol{G_k},\boldsymbol{W}_k,\boldsymbol{U}_k} \quad \frac{1}{K}\sum_{k=1}^{K}\frac{\mathrm{tr}\{\boldsymbol{R}_{e,k}\}}{N_k} \tag{2.78}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\mathrm{tr}\Big\{\boldsymbol{F}_k\boldsymbol{F}_k^H\Big\} \leq P_s \tag{2.79}$$

$$\sum_{k=1}^{K}\mathrm{tr}\Big\{\boldsymbol{G}_k\left(\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)\boldsymbol{G}_k^H\Big\} \leq P_r. \tag{2.80}$$

The objective function in (2.78) is the arithmetic MSE averaged over all subcarriers, whilst (2.79) and (2.80) impose the source and relay power constraints over

[4]We note that for the THP model, (2.70) is only reasonable for moderate to large M-QAM constellations and thus the power consumed in (2.76) and (2.77) are only reasonable for moderate to large M-QAM constellations. In fact for small sized constellation, THP will consume more transmit power than (2.76) and (2.77)

all subcarriers with P_s and P_r being the corresponding available power budgets.

2.5 Chapter Summary and Conclusions

This chapter introduced the signal model for data transmission in a two-hop MIMO relaying system where the source-relay and relay-destination channels are frequency selective which result in the received symbols being corrupted by ISI. The use of OFDM was then discussed and shown to convert the frequency selective MIMO relay channels into a number of narrowband subcarriers. We then discussed non-linear transceivers to deal with the interference for the MIMO OFDM subcarriers. Specifically, we introduced DFE and THP transceivers where the processors were restricted to operate independently on each subcarrier. The optimisation problem for minimising the arithmetic MSE subject to source and relay transmit power constraints was then formulated which can be solved to derive the processors for both DFE and THP transceivers. In Chapters 4 and 5 we shall revisit this problem and derive solutions to the optimisation problem. These solutions depend on the level and quality of CSI available to the source, relay, and destination channels. In order to make CSI available to these nodes channel estimation algorithms must be utilised, which shall be discussed in the next chapter.

Chapter 3

MIMO OFDM Relay Channel Estimation Algorithms

3.1 Introduction

To realise the potential higher data rates in MIMO systems through spatial multiplexing gain CSI is required, which can be obtained through channel estimation. Channel estimation is therefore a vital component in MIMO systems and has attracted significant research interest, with training based solutions being particularly well studied [66, 71–74]. In training based channel estimation, which is also referred to as supervised training, known training sequences or pilot symbols are transmitted from the source to the destination. The destination can then use the received symbols from which, under favourable conditions, an accurate channel estimate can be obtained through appropriate signal processing.

Various estimators for spatially correlated narrowband point-to-point MIMO channels are studied in [71, 72], with the various algorithms differing in terms of their computational complexity, required a priori knowledge of the channel, and estimation MSE performance. The authors firstly derive the optimal training sequences for LS channel estimation. It is shown that the LS solution does not require any a priori statistical knowledge of the MIMO channel, and that the optimal training matrix is a scaled unitary matrix. The authors suggest that the optimal LS training matrix can be selected as an appropriately chosen scaled DFT matrix. The optimal LS training algorithm does not provide the best MSE performance due to the fact that it does not utilise any a priori channel information. The authors therefore suggest another LS solution that they refer to as a scaled LS (SLS) algorithm. This algorithm requires the knowledge of the trace of the channel covariance matrix and provides an improved channel estimate compared to

the LS design, which comes at the expense of increased computational complexity. Neither the LS or SLS utilise the full channel covariance matrix. If knowledge of this matrix is available then it is shown in [71, 72] that MMSE channel estimators provide a further improvement in terms of channel estimation performance.

As well as the study of channel estimation for narrowband MIMO systems, channel estimation of frequency selective channels with the use of OFDM have also been considered in e.g. [66, 73, 74]. The optimal training design for LS channel estimation is discussed in [73] where, similar to the case of narrowband systems, it is shown that the optimal training sequences are orthogonal. For the case of MMSE channel estimation in MIMO OFDM systems, where the frequency selective delay paths are spatially correlated on both the transmit and receive sides, conditions for optimality of training sequences are derived in [74]. Whilst the conditions of optimality are established in this work for general spatial correlation, the optimal solutions are only identified for special cases of spatial correlation. Specifically, the optimal training matrix design for the cases that the transmit correlation matrices for each delay path are the same and the case that there is one dominant MIMO delay path. In both these cases the optimal transmit matrix structure is derived in [74] and only depends on a single transmit spatial correlation matrix. It is shown that the optimal training matrix structure results in a complicated power allocation problem for which there is no closed form solution in general. The authors therefore derive power allocation solutions for the high and low signal-noise ratio (SNR) regimes. For high SNR it is shown that the optimal power allocation is an equal power allocation (EPA), whilst at low SNR only a single eigenmode is used. For the case of general spatial correlation in MIMO OFDM systems, the optimal MMSE training solution has recently been derived in [66]. Using the optimality conditions established in [74], the authors of [66] derive the optimal training matrix structure for general spatial correlation. Under an appropriate change of variable the optimisation problem for the training matrix is reformulated as a semi-definite program (SDP), which is a standard convex optimisation problem and can be efficiently solved using interior point methods [39]. In [66] suboptimal solutions are also suggested based on minimising tight upper and lower bounds of the channel estimation MSE objective function. Interestingly, it is shown that the suboptimal solutions provide comparable performance to the optimal approach but at a reduced computationaly complexity.

More recently, channel estimation has also been considered for two-hop MIMO relaying systems over narrowband channels in [57, 58, 75, 76]. Since a two-hop relaying system consists of a source-relay and relay-destination channel, the channel estimation problem becomes far more involved compared to the case point-to-point

systems. For a fixed relay precoding matrix, the compound channel between the source and destination can be estimated using standard point-to-point channel estimation algorithms. However, transceiver designs based only on the knowledge of the compound channel do not utilise all degrees of freedom that the relay channel offers and can result in poor performance. Elegant channel estimation algorithms that can seperately estimate the source-relay and relay-destination channels is therefore of paramount importance in such systems. In [75] a LS channel estimation approach is suggested to estimate the source-relay and relay-destination channels from the observation of the composite MIMO channel between the source and destination. The authors in [76] also focus on deriving both channels from the observation of the compound MIMO channel at the destination. The main drawback with the algorithms in[75] and [76] is that their is a scalar ambiguity between the estimates of the source-relay and relay-destination channels. A different approach is adopted in [58], where the channel estimation process is divided into two seperate phases. In the first phase the relay sends known training symbols to the destination, allowing the receiver to estimate the relay-destination channel. In the second phase the source sends pilot symbols to the relay, which precodes the received symbols and forwards them to the destination. Under the assumption that the relay-destination channel was estimated accurately in the first phase, the authors then derive the MMSE source training matrix and relay precoder for the source-relay channel estimation problem. The optimal solution is derived for the case that the narrowband MIMO source-relay channel is uncorrelated on the receiver side. It is shown that the assumption of the relay-destination channel is estimated perfectly is only reasonable when the SNR during the first phase of channel estimation is sufficiently large. If this is not the case then the resulting channel estimation error will adversely affect the source-relay channel estimate. Due to the adverse impact that the relay-destination estimation error has on the obtained source-relay estimate, the authors of [57] propose a robust source-relay channel estimation algorithm that takes into account the channel estimation error. Similar to [58] the solution derived in [57] is only optimal for a specific case of source-relay channel covariance.

Similar to the works of [58] and [57], in this chapter we consider the task of channel estimation for MIMO relaying where the estimation process is divided into two seperate phases. However, differing from these works we consider the task of channel estimation of more general frequency selective source-relay and relay-destination channels with the use of OFDM. Furthermore, our solutions are derived for general spatial correlation of the frequency selective MIMO delay paths. For the relay-destination channel estimation problem we discuss known LS and MMSE

solutions in Section 3.2. In Section 3.3 we introduce the problem of estimating the source-relay channel at the destination device. Similar to [58] we simplify this problem by making the assumption that the relay-destination channel is estimated perfectly in the first phase of channel estimation. An iterative source-relay channel estimation solution in then derived in Section 3.4. The proposed iterative approach involves solving a number of convex optimisation problems in a sequential fashion and is shown to have guaranteed convergence. Due to its iterative nature this proposed solution may have a high computational complexity. For this reason we then consider a suboptimal strategy in Section 3.5 where all channel estimation processors can be derived in closed form. The simplified approach is derived using a high SNR approximation and is optimal for the case of high SNR in the source-relay link. In Section 3.6 simulation results are presented that demonstrate the effectiveness of the proposed solutions.

3.2 Relay-Destination Channel Estimation

For the two-hop MIMO OFDM relaying system introduced in Chapter 2, the task of channel estimation is seperated into two main phases. In the first phase the relay device sends known training symbols to the destination which allows the destination to obtain an estimate of the relay-destination channel. This is a point-to-point channel estimation problem for which solutions are well known. In this section we firstly introduce the signal model for the relay-destination channel estimation before discussing known LS and MMSE algorithms.

3.2.1 Signal Model and Problem Formulation

The signal model for estimating the relay-destination channel for the considered two-hop MIMO OFDM relaying system is depicted in Figure 3.1. For the kth subcarrier the relay transmits the training vector $\boldsymbol{x}_k \in \mathbb{C}^{N_r}$ over the relay-destination subcarrier channel $\boldsymbol{H}_{r,k} \in \mathbb{C}^{N_d \times N_r}$, resulting in the received signal $\boldsymbol{y}_k \in \mathbb{C}^{N_d}$ at the destination being given by

$$\boldsymbol{y}_k = \boldsymbol{H}_{r,k}\boldsymbol{x}_k + \boldsymbol{v}_{r,k}, \tag{3.1}$$

where $\boldsymbol{v}_{r,k} \in \mathbb{C}^{N_d}$ is the AWGN noise vector for the kth subcarrier, which we assume satisfies the statistical properties stated in **(A2.)**-**(A4.)** in Section 2.2.5 of the previous chapter. For convenience we also recall that $\boldsymbol{H}_{r,k}$ is related to the

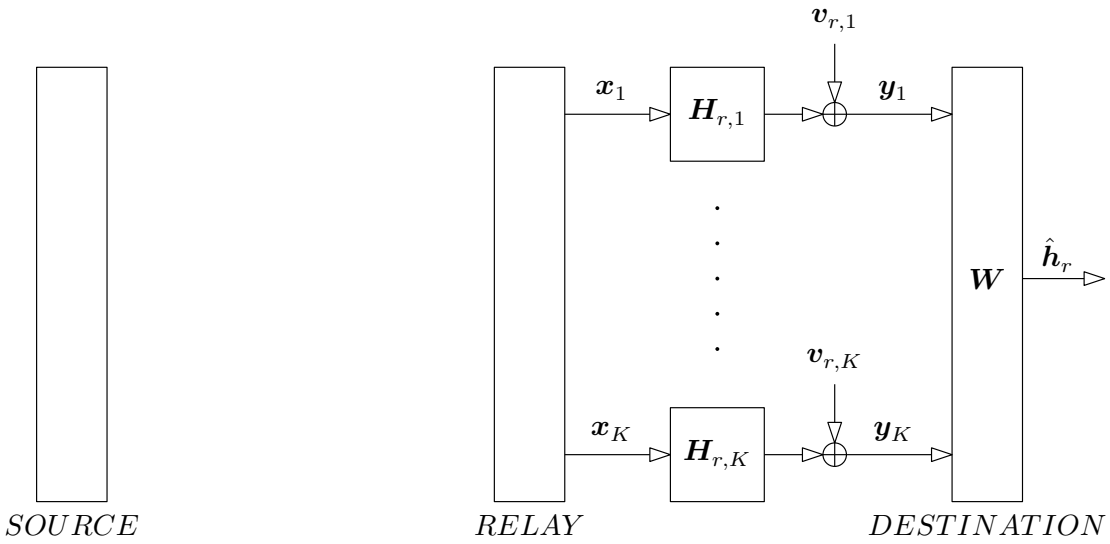

Figure 3.1: Signal model for relay-destination channel estimation.

underlying time domain channel matrices $\boldsymbol{\mathcal{H}}_r[l] \in \mathbb{C}^{N_d \times N_r}$ through the relationship

$$\boldsymbol{H}_{r,k} = \sum_{l=0}^{L} \boldsymbol{\mathcal{H}}_r[l] e^{-j2\pi(k-1)l/K}, \tag{3.2}$$

which can be equivalently written in the more compact matrix form

$$\boldsymbol{H}_{r,k} = [\boldsymbol{\mathcal{H}}_r[0], ..., \boldsymbol{\mathcal{H}}_r[L]] \begin{bmatrix} e^{-j2\pi(k-1)0/K}\boldsymbol{I}_{N_r} \\ \vdots \\ e^{-j2\pi(k-1)L/K}\boldsymbol{I}_{N_r} \end{bmatrix}. \tag{3.3}$$

We reiterate that the time domain channel matrices $\boldsymbol{\mathcal{H}}_r[l]$ are modelled according to the Kronecker spatial correlation model described in Section 2.2.5 of the previous chapter. This is an important point to stress since the inclusion of spatial correlation can significantly complicates the derivation of the optimal channel estimate. In fact, the optimal point-to-point MIMO MMSE channel estimate for the case of spatially correlated channels appears to have only recently been solved in [66]. Substituting (3.3) into (3.1) we can write $\boldsymbol{y}_k$ as

$$\boldsymbol{y}_k = [\boldsymbol{\mathcal{H}}_r[0], ..., \boldsymbol{\mathcal{H}}_r[L]] \begin{bmatrix} \boldsymbol{x}_k e^{-j2\pi(k-1)0/K} \\ \vdots \\ \boldsymbol{x}_k e^{-j2\pi(k-1)L/K} \end{bmatrix} + \boldsymbol{v}_{r,k}. \tag{3.4}$$

$$= \left(\left[\boldsymbol{x}_k^T e^{-j2\pi(k-1)0/K}, ..., \boldsymbol{x}_k^T e^{-j2\pi(k-1)L/K} \right] \otimes \boldsymbol{I}_{N_d} \right) \boldsymbol{h}_r + \boldsymbol{v}_{r,k}, \tag{3.5}$$

where we define the relay-destination channel vector $\boldsymbol{h}_r \in \mathbb{C}^{N_r N_d (L+1)}$ as

$$\boldsymbol{h}_r \triangleq \mathrm{vec}\left[\boldsymbol{\mathcal{H}}_r[0], ..., \boldsymbol{\mathcal{H}}_r[L]\right]. \tag{3.6}$$

To obtain (3.5) from (3.4) we have used the rule $\mathrm{vec}[\boldsymbol{AXB}] = (\boldsymbol{B}^T \otimes \boldsymbol{A})\mathrm{vec}[\boldsymbol{X}]$. Collecting all received subcarrier vectors $\boldsymbol{y}_k$ in (3.5) into a single vector $\boldsymbol{y} \in \mathbb{C}^{KN_d}$ defined as $\boldsymbol{y} \triangleq [\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_K^T]^T$, it can be straightforwardly shown that

$$\boldsymbol{y} = \left(\begin{bmatrix} \boldsymbol{x}_1^T e^{-j2\pi 00/K} & \ldots & \boldsymbol{x}_1^T e^{-j2\pi 0L/K} \\ \vdots & \ddots & \vdots \\ \boldsymbol{x}_K^T e^{-j2\pi (K-1)0/K} & \ldots & \boldsymbol{x}_K^T e^{-j2\pi (K-1)L/K} \end{bmatrix} \otimes \boldsymbol{I}_{N_d} \right) \boldsymbol{h}_r + \boldsymbol{v}_r, \tag{3.7}$$

where we define the total relay-destination AWGN noise vector $\boldsymbol{v}_r \in \mathbb{C}^{KN_d}$ as $\boldsymbol{v}_r \triangleq [\boldsymbol{v}_{r,1}^T, \ldots, \boldsymbol{v}_{r,K}^T]^T$. By also defining a diagonal matrix $\boldsymbol{F}_l \in \mathbb{C}^{K \times K}$ and the relay training matrix $\boldsymbol{X} \in \mathbb{C}^{K \times N_r}$ as

$$\boldsymbol{F}_l \triangleq \begin{bmatrix} e^{-j2\pi 0l/K} & & 0 \\ & \ddots & \\ 0 & & e^{-j2\pi (K-1)l/K} \end{bmatrix} \tag{3.8}$$

$$\boldsymbol{X} \triangleq [\boldsymbol{x}_1, ..., \boldsymbol{x}_K]^T, \tag{3.9}$$

we can further write (3.7) equivalently as

$$\boldsymbol{y} = \left(\underbrace{[\boldsymbol{F}_0 \boldsymbol{X}, ..., \boldsymbol{F}_L \boldsymbol{X}]}_{\boldsymbol{M}_x} \otimes \boldsymbol{I}_{N_d} \right) \boldsymbol{h}_r + \boldsymbol{v}_r. \tag{3.10}$$

From the received signal in (3.10) the task now is to estimate the time domain channel matrices $\boldsymbol{\mathcal{H}}_r[l]$ which, from the definition in (3.6), is clearly equivalent to estimating the vector $\boldsymbol{h}_r$. As depicted in Figure 3.1, in order to facilitate the computation of a channel estimate $\hat{\boldsymbol{h}}_r \in \mathbb{C}^{N_r N_d (L+1)}$ of the relay-destination channel vector $\boldsymbol{h}_r$, a linear processor $\boldsymbol{W} \in \mathbb{C}^{KN_d \times N_r N_d (L+1)}$ is employed at the destination. Using (3.10) the channel estimate $\hat{\boldsymbol{h}}_r$ is therefore obtained from

$$\hat{\boldsymbol{h}}_r = \boldsymbol{W}\boldsymbol{y} \tag{3.11}$$

$$= \boldsymbol{W}\left(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)\boldsymbol{h}_r + \boldsymbol{W}\boldsymbol{v}_r, \tag{3.12}$$

where the matrix $\boldsymbol{M}_x \in \mathbb{C}^{K \times N_r (L+1)}$ was defined in (3.10). Since the problem of computing $\hat{\boldsymbol{h}}_r$ is a standard point-to-point MIMO channel estimation problem,

LS and MMSE algorithms are well known and shall be discussed in the following sections. We shall note here that, as well as being of use to compute the relay-destination channel estimate, the algorithms covered in the following sections will also be useful in deriving our proposed suboptimal solutions for the source-relay channel estimation problem, which shall be discussed later in Section 3.5. Therefore, whilst the following LS and MMSE algorithms are well known, it is worthwhile discussing them in detail since it shall later facilitate the discussion of our proposed suboptimal source-relay channel estimation solutions.

3.2.2 LS Channel Estimation

LS algorithms for point-to-point MIMO channel estimation have been studied in e.g. [71, 72, 77]. In this section we discuss an LS solution for computing the relay-destination channel estimate $\hat{\boldsymbol{h}}_r$ in (3.12). We note that our problem formulation differs from those in [71, 72, 77] and the optimal solution is therefore reached in a slightly different manner. Nevertheless, our approach to identifying the optimal solution follows similar arguments made in [71, 72, 77] and we obtain analogous results.

The optimal processor $\boldsymbol{W}$ for the LS estimate of the relay-destination channel vector is given by

$$\begin{aligned}\boldsymbol{W} &= \left(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^{\dagger} && (3.13)\\ &= \left(\left(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^H \left(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)\right)^{-1} \left(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^H && (3.14)\\ &= \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^{-1} \left(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\right) && (3.15)\end{aligned}$$

where (3.14) is obtained from the fact that the Moore Penrose pseudo-inverse of a matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$, with $M \geq N$, is given by $\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$. To obtain (3.15) from (3.14) we have also made use of the fact that $(\boldsymbol{A}\otimes\boldsymbol{B})^H = (\boldsymbol{A}^H \otimes \boldsymbol{B}^H)$ and for matrices of commensurate dimension $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC} \otimes \boldsymbol{BD})$. Substituting (3.15) into (3.12) we obtain the channel estimate

$$\hat{\boldsymbol{h}}_r = \boldsymbol{h}_r + \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^{-1} \left(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\right) \boldsymbol{v}_r. \tag{3.16}$$

Defining an error $\boldsymbol{e}_r \in \mathbb{C}^{N_r N_d (L+1)}$ between the channel estimate $\hat{\boldsymbol{h}}_r$ and the true channel vector $\boldsymbol{h}_r$ as $\boldsymbol{e}_r = \hat{\boldsymbol{h}}_r - \boldsymbol{h}_r$, we directly have from (3.16) that

$$\boldsymbol{e}_r = \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^{-1} \left(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\right) \boldsymbol{v}_r. \tag{3.17}$$

Using the error signal in (3.17) we can now derive the relay-destination channel estimation error covariance matrix $\boldsymbol{R}_{e_r} \triangleq \mathbb{E}\{\boldsymbol{e}_r \boldsymbol{e}_r^H\} \in \mathbb{H}_+^{N_r N_d(L+1) \times N_r N_d(L+1)}$ as

$$\boldsymbol{R}_{e_r} = \sigma_{v_r}^2 \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d} \right)^{-1} \tag{3.18}$$

With the error covariance matrix given in (3.18) the objective function for minimising the MSE is given by $\mathrm{tr}\{\boldsymbol{R}_{e_r}\}$, which can be written as

$$\mathrm{tr}\{\boldsymbol{R}_{e_r}\} = \mathrm{tr}\left\{ \sigma_{v_r}^2 \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\} \tag{3.19}$$

$$= N_d \mathrm{tr}\left\{ \sigma_{v_r}^2 \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \right)^{-1} \right\}, \tag{3.20}$$

where we have used $(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = (\boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1})$ and $\mathrm{tr}\{(\boldsymbol{A} \otimes \boldsymbol{B})\} = \mathrm{tr}\{\boldsymbol{A}\}\mathrm{tr}\{\boldsymbol{B}\}$ to obtain (3.20) from (3.19). As well as minimising the MSE given by (3.20) we also wish to constrain the relay transmit power, which is given by $\mathrm{tr}\{\boldsymbol{X}\boldsymbol{X}^H\}$. We can thus formulate the constrained optimisation problem

$$\min_{\boldsymbol{X}} \quad \mathrm{tr}\left\{ \sigma_{v_r}^2 \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \right)^{-1} \right\} \tag{3.21}$$

$$\text{s.t.} \quad \mathrm{tr}\left\{ \boldsymbol{X}\boldsymbol{X}^H \right\} \leq P_r, \tag{3.22}$$

where $P_r \in \mathbb{R}_{++}$ is the power budget available to the relay during the first phase of channel estimation, and we note that $\boldsymbol{M}_x$ is a function of $\boldsymbol{X}$ through the definition in (3.10). From (3.21)-(3.22) we see that, in order to optimise the transmit training matrix $\boldsymbol{X}$, we require no prior knowledge of the relay-destination channel. As will be seen later on, this differs from the case of MMSE channel estimation algorithms, which require a priori statistical knowledge of the channel vector $\boldsymbol{h}_r$. Since the LS algorithm does not require such a priori knowledge, it is very appealing for practical systems.

Theorem 1: The optimal training matrix $\boldsymbol{X}$ that minimises the objective function in (3.21) whilst satisfying the relay power constraint in (3.22) is given by

$$\boldsymbol{X} = \sqrt{\frac{P_r}{N_r}} \boldsymbol{Q}, \tag{3.23}$$

where $\boldsymbol{Q} \in \mathbb{C}^{K \times N_r}$ is a semi-unitary matrix that satisfies the properties

$$\boldsymbol{Q}^H \boldsymbol{F}_m^H \boldsymbol{F}_n \boldsymbol{Q} = \boldsymbol{0}_{N_r \times N_r} \qquad \forall m \neq n \tag{3.24}$$

$$\boldsymbol{Q}^H \boldsymbol{F}_l^H \boldsymbol{F}_l \boldsymbol{Q} = \boldsymbol{I}_{N_r} \qquad \forall l \tag{3.25}$$

$$\boldsymbol{Q}^H \boldsymbol{Q} = \boldsymbol{I}_{N_r}. \tag{3.26}$$

Proof: See Section 3.8.1 on page 77. □

From Theorem 1 we see that in order to compute the optimal training matrix $\boldsymbol{X}$ we require to construct a semi-unitary matrix $\boldsymbol{Q}$ that satisfies (3.24)-(3.26). It is shown in [66] and [74] that the matrix $\boldsymbol{Q}$ which satisfies these properties can be constructed as follows: Let us firstly partition $\boldsymbol{Q}$ as

$$\boldsymbol{Q} = \left[\boldsymbol{q}_1, ..., \boldsymbol{q}_{N_r}\right], \tag{3.27}$$

where $\boldsymbol{q}_i \in \mathbb{C}^K$ denotes the ith column of the matrix $\boldsymbol{Q}$. Now let the first column of $\boldsymbol{Q}$ be given by

$$\boldsymbol{q}_1 = \sqrt{\frac{1}{K}} \boldsymbol{1}_K. \tag{3.28}$$

The kth element of the remaining columns of $\boldsymbol{Q}$ can then be constructed as

$$[\boldsymbol{q}_i]_k = \sqrt{\frac{1}{K}} e^{-j2\pi \lfloor K/N_r \rfloor (i-1)(k-1)/K}, \tag{3.29}$$

where we note that $K \geq N_r \lfloor K/N_r \rfloor$. With the columns of $\boldsymbol{Q}$ constructed according to (3.28) and (3.29) it can be verified that the conditions in (3.24)-(3.26) are satisfied. This concludes the LS relay-destination channel estimation design.

3.2.3 Optimal MMSE Channel Estimation

It is well known that MMSE channel estimation algorithms have the capability of providing improved channel estimates compared to LS algorithms. For the case of point-to-point MIMO OFDM systems, the optimal MMSE channel estimation of spatially correlated channels has recently been derived in [66]. In this section we discuss the optimal MMSE channel estimation algorithm proposed in [66].

The error between the channel estimate $\hat{\boldsymbol{h}}_r$ and the actual channel vector $\boldsymbol{h}_r$ is given by $\boldsymbol{e}_r = \hat{\boldsymbol{h}}_r - \boldsymbol{h}_r \in \mathbb{C}^{N_r N_d (L+1)}$, which using (3.12) results in

$$\boldsymbol{e}_r = \boldsymbol{W} \left(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right) \boldsymbol{h}_r + \boldsymbol{W} \boldsymbol{v}_r - \boldsymbol{h}_r. \tag{3.30}$$

By defining the relay-destination channel estimation error covariance matrix as $\boldsymbol{R}_{e_r} \triangleq \mathbb{E}\{\boldsymbol{e}_r \boldsymbol{e}_r^H\} \in \mathbb{H}_+^{N_r N_d(L+1) \times N_r N_d(L+1)}$, it is straightforward to show that

$$\begin{aligned}\boldsymbol{R}_{e_r} = \boldsymbol{W}\Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\mathbb{E}\Big\{\boldsymbol{h}_r \boldsymbol{h}_r^H\Big\}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{W}^H + \mathbb{E}\Big\{\boldsymbol{h}_r \boldsymbol{h}_r^H\Big\} \\ - \boldsymbol{W}\Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\mathbb{E}\Big\{\boldsymbol{h}_r \boldsymbol{h}_r^H\Big\} - \mathbb{E}\Big\{\boldsymbol{h}_r \boldsymbol{h}_r^H\Big\}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{W}^H + \boldsymbol{W}\boldsymbol{W}^H \sigma_{v_r}^2,\end{aligned} \tag{3.31}$$

In (3.31) we require to compute the expectation $\mathbb{E}\{\boldsymbol{h}_r \boldsymbol{h}_r^H\}$ which is the relay-destination channel covariance matrix. Using $\boldsymbol{h}_r = \text{vec}[\boldsymbol{\mathcal{H}}_r[0], ..., \boldsymbol{\mathcal{H}}_r[L]]$ from (3.6), as well as the matrix variate Gaussian distributions of $\boldsymbol{\mathcal{H}}_r[l]$ described in Section 2.2.5 in Chapter 2, we can show that $\boldsymbol{R}_{h_r} \triangleq \mathbb{E}\{\boldsymbol{h}_r \boldsymbol{h}_r^H\} \in \mathbb{H}_+^{N_r N_d(L+1) \times N_r N_d(L+1)}$ is given by

$$\boldsymbol{R}_{h_r} \triangleq \mathbb{E}\{\boldsymbol{h}_r \boldsymbol{h}_r^H\} = \begin{bmatrix} \sigma_{h_r}^2[0]\boldsymbol{\Theta}_r[0] \otimes \boldsymbol{\Upsilon}_r[0] & & \boldsymbol{0}_{N_r N_d \times N_r N_d} \\ & \ddots & \\ \boldsymbol{0}_{N_r N_d \times N_r N_d} & & \sigma_{h_r}^2[L]\boldsymbol{\Theta}_r[L] \otimes \boldsymbol{\Upsilon}_r[L] \end{bmatrix}, \tag{3.32}$$

where we recall that $\sigma_{h_r}^2[l]$ is the variance of the lth relay-destination MIMO delay path, with $\boldsymbol{\Theta}_r[l] \in \mathbb{C}^{N_r \times N_r}$ and $\boldsymbol{\Upsilon}_r[l] \in \mathbb{C}^{N_d \times N_d}$ being the corresponding transmit and receive side spatial correlation matrices, respectively. We assume that $\sigma_{h_r}^2[l]$, $\boldsymbol{\Theta}_r[l]$, and $\boldsymbol{\Upsilon}_r[l]$ are known to all nodes in the network. We see that, unlike the LS algorithm discussed in the previous section, the MMSE channel estimation algorithm requires prior knowledge of the relay-destination channel in terms of the covariance matrix in (3.32). Substituting (3.32) into (3.31) let us firstly write the error covariance matrix as

$$\begin{aligned}\boldsymbol{R}_{e_r} = \boldsymbol{W}\Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{R}_{h_r}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{W}^H + \boldsymbol{R}_{h_r} \\ - \boldsymbol{W}\Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{R}_{h_r} - \boldsymbol{R}_{h_r}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{W}^H + \boldsymbol{W}\boldsymbol{W}^H \sigma_{v_r}^2,\end{aligned} \tag{3.33}$$

and formulate the constrained optimisation problem for minimising the channel estimation MSE subject to the transmission power constraint at the relay as

$$\min_{\boldsymbol{X}, \boldsymbol{W}} \quad \text{tr}\Big\{\boldsymbol{R}_{e_r}\Big\} \tag{3.34}$$

$$\text{s.t.} \quad \text{tr}\Big\{\boldsymbol{X}\boldsymbol{X}^H\Big\} \leq P_r. \tag{3.35}$$

The objective function (3.34) is a function of the destination processor $\boldsymbol{W}$ and the relay transmit matrix $\boldsymbol{X}$ (note that $\boldsymbol{M}_x$ in (3.33) is a function of $\boldsymbol{X}$ through

the definition in (3.10)) which we require to optimise. In order to solve the optimisation problem in (3.34)-(3.35) we can begin by deriving the optimal MMSE processor $\boldsymbol{W}$. This is due to the fact that

$$\min_{\boldsymbol{x},\boldsymbol{y}} f(\boldsymbol{x},\boldsymbol{y}) = \min_{\boldsymbol{x}} \min_{\boldsymbol{y}} f(\boldsymbol{x},\boldsymbol{y}), \tag{3.36}$$

i.e. in solving an optimisation problem with several variables, we can find the optimal solution to the problem by firstly minimising over some variables and then optimising over the remaining ones [22, 39]. Thus to solve (3.34)-(3.35) we firstly minimise over $\boldsymbol{W}$. Since the constraint in (3.35) is independent of $\boldsymbol{W}$, the optimal solution is given by the unconstrained problem of minimising the objective function in (3.34). Furthermore, since the objective function is convex quadratic [39] in $\boldsymbol{W}$, there exists only a single minima. The optimal processor can therefore be found by setting the derivative of $\mathrm{tr}\{\boldsymbol{R}_{e_r}\}$ with respect to (w.r.t.) $\boldsymbol{W}^*$ to zero and solving the resultant equation for $\boldsymbol{W}$. This gives the optimal solution

$$\boldsymbol{W} = \boldsymbol{R}_{h_r}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big)\left(\Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{R}_{h_r}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big) + \sigma_{v_r}^2 \boldsymbol{I}_{KN_d}\right)^{-1}. \tag{3.37}$$

Substituting the optimal processor $\boldsymbol{W}$ into the error covariance matrix given in (3.33) we can write $\boldsymbol{R}_{e_r}$ more compactly as

$$\begin{aligned}\boldsymbol{R}_{e_r} = \boldsymbol{R}_{h_r} - \boldsymbol{R}_{h_r}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big)\left(\Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{R}_{h_r}\Big(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\Big) + \sigma_{v_r}^2 \boldsymbol{I}_{KN_d}\right)^{-1} \\ \times \Big(\boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\boldsymbol{R}_{h_r} \qquad (3.38)\end{aligned}$$

$$= \left(\boldsymbol{R}_{h_r}^{-1} + \frac{1}{\sigma_{v_r}^2}\Big(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\right)^{-1}, \tag{3.39}$$

where to obtain (3.39) from (3.38) we have applied the matrix inversion lemma as well as some basic Kronecker product rules. Substituting (3.39) into (3.34) we can restate the problem as

$$\min_{\boldsymbol{X}} \quad \mathrm{tr}\left\{\left(\boldsymbol{R}_{h_r}^{-1} + \frac{1}{\sigma_{v_r}^2}\Big(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\Big)\right)^{-1}\right\} \tag{3.40}$$

$$\text{s.t.} \quad \mathrm{tr}\Big\{\boldsymbol{X}\boldsymbol{X}^H\Big\} \leq P_r. \tag{3.41}$$

The optimal structure of the relay transmit matrix $\boldsymbol{X}$ is now established in the following theorem (c.f. [66]):

Theorem 2: The optimal $\boldsymbol{X}$ as the solution to (3.40)-(3.41) is given by

$$\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}, \tag{3.42}$$

where $\bar{\boldsymbol{X}} \in \mathbb{C}^{N_r \times N_r}$ is a matrix yet to be determined and $\boldsymbol{Q} \in \mathbb{C}^{K \times N_r}$ is a semi-unitary matrix that satisfies

$$\boldsymbol{Q}^H \boldsymbol{F}_m^H \boldsymbol{F}_n \boldsymbol{Q} = \boldsymbol{0}_{N_r \times N_r} \qquad \forall m \neq n \tag{3.43}$$

$$\boldsymbol{Q}^H \boldsymbol{F}_l^H \boldsymbol{F}_l \boldsymbol{Q} = \boldsymbol{I}_{N_r} \qquad \forall l \tag{3.44}$$

$$\boldsymbol{Q}^H \boldsymbol{Q} = \boldsymbol{I}_{N_r}, \tag{3.45}$$

which are precisely the same conditions as (3.24)-(3.26) and as such the matrix $\boldsymbol{Q}$ can be constucted in the same manner as for the previously discussed LS algorithm.

Proof: See Section 3.8.2 on page 78. □

We see that the optimal structure of the MMSE training matrix given in (3.42) has a more general structure than the optimal LS training matrix in (3.23). Specifically, (3.23) can be obtained from (3.42) with $\bar{\boldsymbol{X}} = \sqrt{P_r/N_r}\boldsymbol{I}_{N_r}$. This is in fact unsurprising since it is well known that the performance of LS channel estimation algorithms converge to that of MMSE algorithms with increasing SNR.

With the definition of $\boldsymbol{M}_x$ in (3.10) along with the structure of $\boldsymbol{X}$ given in (3.42) of Theorem 2, we now note that

$$\boldsymbol{M}_x^H \boldsymbol{M}_x = \begin{bmatrix} \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_0^H \boldsymbol{F}_0 \boldsymbol{Q} \bar{\boldsymbol{X}} & \dots & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_0^H \boldsymbol{F}_L \boldsymbol{Q} \bar{\boldsymbol{X}} \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_L^H \boldsymbol{F}_0 \boldsymbol{Q} \bar{\boldsymbol{X}} & \dots & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_L^H \boldsymbol{F}_L \boldsymbol{Q} \bar{\boldsymbol{X}} \end{bmatrix} \tag{3.46}$$

$$= \begin{bmatrix} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} & & \boldsymbol{0}_{N_r \times N_r} \\ & \ddots & \\ \boldsymbol{0}_{N_r \times N_r} & & \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \end{bmatrix}, \tag{3.47}$$

where to obtain (3.47) we have utilised the fact that $\boldsymbol{Q}$ satisfies the properties in (3.43)-(3.45). Substituting (3.47) and (3.42) into (3.40)-(3.41), and using the structure of the relay-destination channel covariance matrix $\boldsymbol{R}_{h_r}$ in (3.32), we can now write the problem as

$$\min_{\bar{\boldsymbol{X}}} \quad \sum_{l=0}^{L} \mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\} \tag{3.48}$$

$$\text{s.t.} \quad \mathrm{tr}\left\{ \bar{\boldsymbol{X}} \bar{\boldsymbol{X}}^H \right\} \leq P_r. \tag{3.49}$$

Furthermore, with the variable change

$$\hat{\boldsymbol{X}} = \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}}, \tag{3.50}$$

where $\hat{\boldsymbol{X}} \in \mathbb{H}_+^{N_r \times N_r}$ is a Hermitian positive semi-definite matrix to be determined, the problem in (3.48)-(3.49) is equivalent to the following problem in $\hat{\boldsymbol{X}}$

$$\min_{\hat{\boldsymbol{X}}} \quad \sum_{l=0}^{L} \mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \hat{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\} \tag{3.51}$$

$$\text{s.t.} \quad \mathrm{tr}\left\{ \hat{\boldsymbol{X}} \right\} \leq P_r \tag{3.52}$$

$$\hat{\boldsymbol{X}} \succeq \boldsymbol{0}_{N_r \times N_r}, \tag{3.53}$$

where the additional constraint in (3.53) results from the fact that $\bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}}$ is a Hermitian positive semi-definite matrix. It can be shown that the problem (3.51)-(3.52) is a convex optimisation problem and thus the optimal solution to $\hat{\boldsymbol{X}}$ is readily obtainable through convex programming [39]. In order to find the optimal solution let us firstly introduce the auxilliary matrices $\boldsymbol{Z}_l \in \mathbb{H}_+^{N_r \times N_r}$ and equivalently write the problem as

$$\min_{\hat{\boldsymbol{X}}, \{\boldsymbol{Z}_l\}_{l=0}^{L}} \quad \sum_{l=0}^{L} \mathrm{tr}\{\boldsymbol{Z}_l\} \tag{3.54}$$

$$\text{s.t.} \quad \mathrm{tr}\left\{ \hat{\boldsymbol{X}} \right\} \leq P_r \tag{3.55}$$

$$\hat{\boldsymbol{X}} \succeq \boldsymbol{0}_{N_r \times N_r} \tag{3.56}$$

$$\left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \hat{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \preceq \boldsymbol{Z}_l. \tag{3.57}$$

We can now transform (3.54)-(3.57) into a semi-definite programming (SDP) problem based on the following lemma:

Lemma 1: [39] Given a matrix $\boldsymbol{D}$ partitioned as

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^H & \boldsymbol{C} \end{bmatrix}, \tag{3.58}$$

then $\boldsymbol{D} \succeq \boldsymbol{0}$ if both $\boldsymbol{C} \succeq \boldsymbol{0}$ and $\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^H \succeq \boldsymbol{0}$. This condition is known as the Schur complement lemma.

Applying the Schur complement lemma to the constraint in (3.57), we can write the problem (3.54)-(3.57) in standard SDP form as

$$\min_{\hat{\boldsymbol{X}},\{\boldsymbol{Z}_l\}_{l=0}^{L}} \quad \sum_{l=0}^{L} \mathrm{tr}\{\boldsymbol{Z}_l\} \tag{3.59}$$

$$\text{s.t.} \quad \mathrm{tr}\Big\{\hat{\boldsymbol{X}}\Big\} \leq P_r \tag{3.60}$$

$$\hat{\boldsymbol{X}} \succeq \boldsymbol{0}_{N_r \times N_r} \tag{3.61}$$

$$\begin{bmatrix} \boldsymbol{Z}_l & \boldsymbol{I}_{N_r N_d} \\ \boldsymbol{I}_{N_r N_d} & \dfrac{1}{\sigma_{h_r}^2[l]}\boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \dfrac{1}{\sigma_{v_r}^2}\hat{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \end{bmatrix} \succeq \boldsymbol{0}_{2N_r N_d \times 2N_r N_d}. \tag{3.62}$$

The optimal solution $\hat{\boldsymbol{X}}$ to the SDP in (3.59)-(3.62) can be found using interior point algorithms [39]. Once the optimal $\hat{\boldsymbol{X}}$ is identified, we can recover $\bar{\boldsymbol{X}}$ from (3.50) as $\bar{\boldsymbol{X}} = \hat{\boldsymbol{X}}^{1/2}$, and finally the optimal relay transmit matrix $\boldsymbol{X}$ is then given by (3.42) of Theorem 2.

3.2.4 Suboptimal MMSE Channel Estimation

Whilst the algorithm in the previous section provides the optimal solution for the relay-destination channel estimation problem, it requires solving the SDP problem in (3.59)-(3.62) using interior point methods and may be too computationally complex for practical implementation. The authors of [66] therefore also suggested a suboptimal solution to the optimisation problem in (3.40)-(3.41) for which the the training matrix $\boldsymbol{X}$ can be calculated in closed form, and is thus more computationally efficient. The suboptimal solution in [66] is discussed in this section.

Assuming that the training matrix is constructed as $\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}$ as given in Theorem 2, then as discussed previously, the optimisation problem reduces to that in (3.48)-(3.49), which we repeat here for convenience

$$\min_{\bar{\boldsymbol{X}}} \quad \sum_{l=0}^{L} \mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]}\boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2}\bar{\boldsymbol{X}}^H\bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\} \tag{3.63}$$

$$\text{s.t.} \quad \mathrm{tr}\Big\{\bar{\boldsymbol{X}}\bar{\boldsymbol{X}}^H\Big\} \leq P_r. \tag{3.64}$$

To obtain a suboptimal solution to this problem that can be derived in closed form we shall consider minimising an upper bound of the objective function. It is shown in Section 3.8.3 on page 80 that an upper bound of the objective function

in (3.63) is given by

$$\sum_{l=0}^{L} \mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\}$$
$$\leq \mathrm{tr}\left\{ \left(\left(\sum_{l=0}^{L} \sigma_{h_r}^2[l] \boldsymbol{\Theta}_r[l] \right)^{-1} \otimes \bar{\boldsymbol{\Upsilon}}_r^{-1} + \frac{1}{(L+1)\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\}, \quad (3.65)$$

where we define the diagonal matrix $\bar{\boldsymbol{\Upsilon}}_r \in \mathbb{R}_{++}^{N_d \times N_d}$ as

$$\bar{\boldsymbol{\Upsilon}}_r \triangleq \begin{bmatrix} \max\left(\{\upsilon_{r,1}[l]\}_{l=0}^{L} \right) & & 0 \\ & \ddots & \\ 0 & & \max\left(\{\upsilon_{r,N_d}[l]\}_{l=0}^{L} \right) \end{bmatrix}, \quad (3.66)$$

with $\upsilon_{r,i}[l] \in \mathbb{R}_{++}$ being the ith largest eigenvalue of $\boldsymbol{\Upsilon}_r[l]$. Replacing the objective function in (3.63) with the upper bound in (3.65) we now pose the following optimisation problem

$$\min_{\bar{\boldsymbol{X}}} \quad \mathrm{tr}\left\{ \left(\left(\sum_{l=0}^{L} \sigma_{h_r}^2[l] \boldsymbol{\Theta}_r[l] \right)^{-1} \otimes \bar{\boldsymbol{\Upsilon}}_r^{-1} + \frac{1}{(L+1)\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\} \quad (3.67)$$

$$\text{s.t.} \quad \mathrm{tr}\left\{ \bar{\boldsymbol{X}} \bar{\boldsymbol{X}}^H \right\} \leq P_r. \quad (3.68)$$

The optimal solution to (3.67)-(3.68) will obviously, in general, provide a suboptimal solution to the original problem in (3.63)-(3.64) due to the fact that we are minimising an upper bound of the original objective function. However, as we will see in the following, the solution to (3.67)-(3.68) can be computed in closed form and thus this suboptimal approach is computationally efficient. Before establishing the optimal solution to (3.67)-(3.68) let us introduce the eigendecomposition

$$\left(\sum_{l=0}^{L} \sigma_{h_r}^2[l] \boldsymbol{\Theta}_r[l] \right)^{-1} = \boldsymbol{U}_t \boldsymbol{\Lambda}_t \boldsymbol{U}_t^H, \quad (3.69)$$

where $\boldsymbol{U}_t \in \mathbb{C}^{N_r \times N_r}$ is unitary and the diagonal matrix $\boldsymbol{\Lambda}_t \in \mathbb{R}_{++}^{N_r \times N_r}$ contains the positive eigenvalues $\{\lambda_{t,i}\}_{i=1}^{N_r} \in \mathbb{R}_{++}$ of $(\sum_{l=0}^{L} \sigma_{h_r}^2[l] \boldsymbol{\Theta}_r[l])^{-1}$. Substituting the

eigendecomposition (3.69) into (3.67) it is straightforward to show that

$$\mathrm{tr}\left\{\left(\left(\sum_{l=0}^{L}\sigma_{h_r}^2[l]\boldsymbol{\Theta}_r[l]\right)^{-1}\otimes\bar{\boldsymbol{\Upsilon}}_r^{-1}+\frac{1}{(L+1)\sigma_{v_r}^2}\bar{\boldsymbol{X}}^H\bar{\boldsymbol{X}}\otimes\boldsymbol{I}_{N_d}\right)^{-1}\right\} \tag{3.70}$$

$$=\mathrm{tr}\left\{\underbrace{\left(\boldsymbol{\Lambda}_t\otimes\bar{\boldsymbol{\Upsilon}}_r^{-1}+\frac{1}{(L+1)\sigma_{v_r}^2}\boldsymbol{U}_t^H\bar{\boldsymbol{X}}^H\bar{\boldsymbol{X}}\boldsymbol{U}_t\otimes\boldsymbol{I}_{N_d}\right)^{-1}}_{\bar{\boldsymbol{E}}}\right\}, \tag{3.71}$$

where we have used the fact that $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{AC}\otimes\boldsymbol{BD})$. To establish the optimal structure of $\bar{\boldsymbol{X}}$ as the solution to (3.67)-(3.68) let us recall the following lemma:

Lemma 2: [78] For a Hermitian positive definite matrix $\boldsymbol{A}\in\mathbb{H}_{++}^{N\times N}$ we have

$$\mathrm{tr}\{\boldsymbol{A}^{-1}\}\geq\sum_{i=1}^{N}\left[\boldsymbol{A}^{-1}\right]_{ii}. \tag{3.72}$$

where equality holds if $\boldsymbol{A}$ is a diagonal matrix.

From Lemma 2 it is straightforward to conclude that the objective function in (3.71) is minimised when $\bar{\boldsymbol{E}}$ in (3.71) is diagonal. This occurs when the matrix $\bar{\boldsymbol{X}}$ is given by

$$\bar{\boldsymbol{X}}=\boldsymbol{\Gamma}\boldsymbol{U}_t^H, \tag{3.73}$$

where $\boldsymbol{\Gamma}\in\mathbb{R}_+^{N_r\times N_r}$ is a diagonal matrix with non-negative diagonal elements $\{\gamma_i\}_{i=1}^{N_r}\in\mathbb{R}_+$. Substituting the structure of $\bar{\boldsymbol{X}}$ given in (3.73) into (3.67)-(3.68), and using the eigendecomposition in (3.69), the problem reduces to the scalar valued optimisation problem

$$\min_{\boldsymbol{\gamma}}\quad\sum_{i=1}^{N_r}\sum_{j=1}^{N_d}\frac{1}{\lambda_{t,i}\upsilon_{r,j}^{-1}+\gamma_i^2\left((L+1)\sigma_{v_r}^2\right)^{-1}} \tag{3.74}$$

$$\text{s.t.}\quad\sum_{i=1}^{N_r}\gamma_i^2\leq P_r \tag{3.75}$$

$$\gamma_i^2\geq 0,\qquad 1\leq i\leq N_r, \tag{3.76}$$

where $\boldsymbol{\gamma}\triangleq[\gamma_1,...,\gamma_{N_r}]^T$. Unfortunately the optimal solution to (3.74)-(3.76) is still difficult to obtain in closed form. To proceed further we again consider minimising a simpler upper bound of the objective function in (3.74). To obtain an upper

bound we firstly note that the function $f(x) = (x^{-1} + a)^{-1}$ is concave in $x > 0$ (this is a particular case of Lemma 6 given in Section 3.8.3 on page 81). From Jensens inequality [39] we can therefore obtain the upper bound

$$\begin{aligned} &\sum_{i=1}^{N_r}\sum_{j=1}^{N_d} \frac{1}{\lambda_{t,i}\upsilon_{r,j}^{-1} + \gamma_i^2\left((L+1)\sigma_{v_r}^2\right)^{-1}} \\ &\leq \sum_{i=1}^{N_r} \frac{N_d}{N_d\lambda_{t,i}a_i + \gamma_i^2\left((L+1)\sigma_{v_r}^2\right)^{-1}}, \end{aligned} \tag{3.77}$$

where we define the scalar $a_i \in \mathbb{R}_{++}$ as

$$a_i = \left(\sum_{j=1}^{N_r} \upsilon_{r,j}\right)^{-1}. \tag{3.78}$$

Replacing the objective function in (3.74) with the upper bound of (3.77) we have the simplified optimisation problem

$$\min_{\boldsymbol{\gamma}} \quad \sum_{i=1}^{N_r} \frac{N_d}{N_d\lambda_{t,i}a_i + \gamma_i^2\left((L+1)\sigma_{v_r}^2\right)^{-1}} \tag{3.79}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_r} \gamma_i^2 \leq P_r \tag{3.80}$$

$$\gamma_i^2 \geq 0, \qquad 1 \leq i \leq N_r. \tag{3.81}$$

It can be shown that the objective function in (3.79) is convex and furthermore the constraints in (3.80) and (3.81) are also convex. Thus (3.79)-(3.81) is a standard convex optimisation problem and the optimal solution can be found from the KKT conditions and is given by

$$\gamma_i^2 = (L+1)\sigma_{v_r}^2\left[\sqrt{\frac{N_d}{\mu_r(L+1)\sigma_{v_r}^2}} - N_d\lambda_{t,i}a_i\right]^+, \tag{3.82}$$

where μ_r is the Lagrangian multiplier required to be calculated to ensure that the relay power constraint in (3.80) is satisfied, and can be efficiently computed using the waterfilling algorithm in [79].

3.2.5 Relay-Destination Channel Estimation Error

Before proceeding to developing the proposed source-relay channel estimation algorithms in the next sections, we briefly note that the relay-destination channel

estimates obtained from the preceeding algorithms are in general imperfect i.e. channel estimation errors exist. The relay-destination channel estimate $\hat{\boldsymbol{h}}_r$ given by (3.12) yields the time domain estimates $\hat{\boldsymbol{\mathcal{H}}}_r[l] \in \mathbb{C}^{N_r N_d (L+1)}$ given by (c.f. (3.6))

$$\hat{\boldsymbol{h}}_r \triangleq \text{vec}\left[\hat{\boldsymbol{\mathcal{H}}}_r[0], ..., \hat{\boldsymbol{\mathcal{H}}}_r[L]\right]. \tag{3.83}$$

Similarly, the time domain channel estimation error matrices $\boldsymbol{\mathcal{E}}_r[l] \in \mathbb{C}^{N_r N_d (L+1)}$ can be obtained from

$$\boldsymbol{e}_r \triangleq \text{vec}\left[\boldsymbol{\mathcal{E}}_r[0], ..., \boldsymbol{\mathcal{E}}_r[L]\right]. \tag{3.84}$$

Analagous to (3.2) the time domain channel estimate in (3.83) and the corresponding channel estimation error in (3.84) lead to the kth OFDM subcarrier estimate $\hat{\boldsymbol{H}}_{r,k} \in \mathbb{C}^{N_d \times N_r}$ and the kth OFDM subcarrier estimation error $\boldsymbol{E}_{r,k} \in \mathbb{C}^{N_d \times N_r}$ being given by

$$\hat{\boldsymbol{H}}_{r,k} = \sum_{l=0}^{L} \hat{\boldsymbol{\mathcal{H}}}_r[l] e^{-j2\pi(k-1)l/K} \tag{3.85}$$

$$\boldsymbol{E}_{r,k} = \sum_{l=0}^{L} \boldsymbol{\mathcal{E}}_r[l] e^{-j2\pi(k-1)l/K}. \tag{3.86}$$

Based on the fact that $\boldsymbol{e}_r = \hat{\boldsymbol{h}}_r - \boldsymbol{h}_r$ we clearly have the relationship

$$\boldsymbol{H}_{r,k} = \hat{\boldsymbol{H}}_{r,k} - \boldsymbol{E}_{r,k}, \tag{3.87}$$

where $\boldsymbol{H}_{r,k}$ is the true kth subcarrier channel matrix given by (3.2). Over all subcarriers we can express this relationship compactly as

$$\boldsymbol{H}_r = \hat{\boldsymbol{H}}_r - \boldsymbol{E}_r, \tag{3.88}$$

where $\hat{\boldsymbol{H}}_r \in \mathbb{C}^{KN_d \times KN_r}$ and $\boldsymbol{E}_r \in \mathbb{C}^{KN_d \times KN_r}$ are the channel estimate and the corresponding channel estimation error matrices over all subcarriers, respectively, and are defined as

$$\hat{\boldsymbol{H}}_r \triangleq \begin{bmatrix} \hat{\boldsymbol{H}}_{r,1} & & \boldsymbol{0}_{N_d \times N_r} \\ & \ddots & \\ \boldsymbol{0}_{N_d \times N_r} & & \hat{\boldsymbol{H}}_{r,K} \end{bmatrix}, \qquad \boldsymbol{E}_r \triangleq \begin{bmatrix} \boldsymbol{E}_{r,1} & & \boldsymbol{0}_{N_d \times N_r} \\ & \ddots & \\ \boldsymbol{0}_{N_d \times N_r} & & \boldsymbol{E}_{r,K} \end{bmatrix}. \tag{3.89}$$

It will be shown in Section 3.6.1 that the relay-destination channel estimation error in (3.88) will adversely effect the accuracy of the source-relay channel estimate.

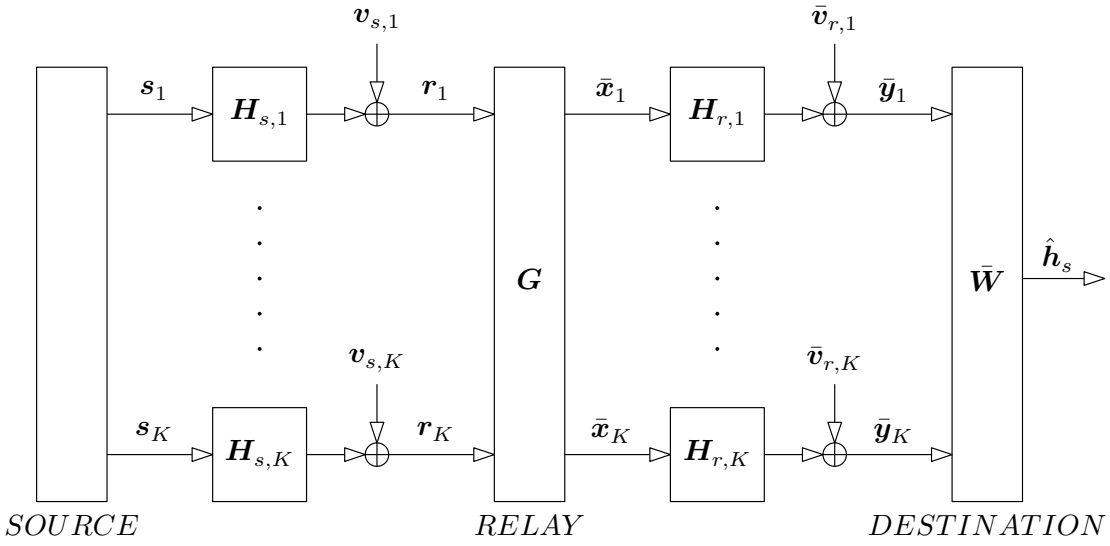

Figure 3.2: Signal model for source-relay channel estimation.

3.3 Source-Relay Channel Estimation Problem

Having discussed the problem of estimating the relay-destination channel in the first phase of channel estimation, and having characterised the resulting channel estimation error, we now focus on developing novel algorithms for estimating the source-relay channel. In this section we firstly formulate the optimisation problem for obtaining the MMSE source-relay channel estimate before deriving iterative and simplified solutions to this problem in subsequent sections.

The signal model for the source-relay channel estimation phase is depicted in Figure 3.2. Whilst the relay remains silent, on the kth subcarrier the source transmits the known training symbols $\boldsymbol{s}_k \in \mathbb{C}^{N_s}$ over the source-relay subcarrier $\boldsymbol{H}_{s,k} \in \mathbb{C}^{N_r \times N_s}$. The received vector at the relay over all subcarriers given by $\boldsymbol{r} \triangleq [\boldsymbol{r}_1^T, ..., \boldsymbol{r}_K^T]^T \in \mathbb{C}^{KN_r}$, where $\boldsymbol{r}_k \in \mathbb{C}^{N_r}$ is the vector received on the kth subcarrier, can be written as (c.f. (3.10))

$$\boldsymbol{r} = \left(\underbrace{[\boldsymbol{F}_0 \boldsymbol{S}, ..., \boldsymbol{F}_L \boldsymbol{S}]}_{\boldsymbol{M}_s} \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{h}_s + \boldsymbol{v}_s, \tag{3.90}$$

where $\boldsymbol{F}_l$ were given in (3.8) and we define the source training matrix $\boldsymbol{S} \in \mathbb{C}^{K \times N_s}$, the source-relay channel vector $\boldsymbol{h}_s \in \mathbb{C}^{N_s N_r (L+1)}$, and the source-relay AWGN noise

vector $\boldsymbol{v}_s \in \mathbb{C}^{KN_r}$ as

$$\boldsymbol{S} \triangleq [\boldsymbol{s}_1, ..., \boldsymbol{s}_K]^T \tag{3.91}$$

$$\boldsymbol{h}_s \triangleq \text{vec}\,[\boldsymbol{\mathcal{H}}_s[0], ..., \boldsymbol{\mathcal{H}}_s[L]] \tag{3.92}$$

$$\boldsymbol{v}_s \triangleq \left[\boldsymbol{v}_{s,1}^T, \ldots, \boldsymbol{v}_{s,K}^T\right]^T. \tag{3.93}$$

Here $\boldsymbol{\mathcal{H}}_s[l]$ is the lth MIMO channel tap of the source-relay channel, which was characterised in Section 2.2.5 of the previous chapter, and we also recall that the noise vector $\boldsymbol{v}_s$ is assumed to satisfy the statistical assumptions **(A3.)** and **(A4.)** made in Section 2.2.5 in Chapter 2.

It is worth remarking here that if the relay has the capability of performing channel estimation then based on the received signal given in (3.90) the relay can estimate the source-relay channel using either the LS, optimal MMSE, and suboptimal MMSE, algorithms described in Sections 3.2.2, 3.2.3, and 3.2.4, respectively. However we shall assume that, similar to [56–60], the task of channel estimation is dedicated to the destination. The relay therefore only performs a linear precoding operation on the received symbols in (3.90) and forwards the resulting precoded symbols to the destination. Using (3.90) the received symbols at the destination $\bar{\boldsymbol{y}} \in \mathbb{C}^{KN_d} \triangleq [\bar{\boldsymbol{y}}_1^T, ..., \bar{\boldsymbol{y}}_K^T]^T \in \mathbb{C}^{KN_d}$, where $\bar{\boldsymbol{y}}_k \in \mathbb{C}^{N_d}$ are the symbols received on the kth subcarrier, are

$$\bar{\boldsymbol{y}} = \boldsymbol{H}_r\boldsymbol{G}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{h}_s + \boldsymbol{H}_r\boldsymbol{G}\boldsymbol{v}_s + \bar{\boldsymbol{v}}_r, \tag{3.94}$$

where $\boldsymbol{G} \in \mathbb{C}^{KN_r \times KN_r}$ is the relay precoder, $\boldsymbol{M}_s \in \mathbb{C}^{K \times N_s(L+1)}$ was defined in (3.90), and $\bar{\boldsymbol{v}}_r \triangleq [\bar{\boldsymbol{v}}_{r,1}^T, ..., \bar{\boldsymbol{v}}_{r,K}^T]^T \in \mathbb{C}^{KN_d}$ is the collection of AWGN noise vectors over all subcarriers. The noise vector is assumed to satisfy assumptions **(A2.)**-**(A4.)** made in Section 2.2.5 of the previous chapter. At the destination the source-relay channel estimate $\hat{\boldsymbol{h}}_s \in \mathbb{C}^{N_sN_r(L+1)}$ can be obtained from the received signal in (3.94) through

$$\hat{\boldsymbol{h}}_s = \bar{\boldsymbol{W}}\hat{\boldsymbol{y}} \tag{3.95}$$

$$= \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{h}_s + \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\boldsymbol{v}_s + \bar{\boldsymbol{W}}\bar{\boldsymbol{v}}_r, \tag{3.96}$$

where $\bar{\boldsymbol{W}} \in \mathbb{C}^{KN_d \times N_sN_r(L+1)}$ is the linear processor at the destination that facilitates the computation of the source-relay channel estimate. Defining the error signal $\boldsymbol{e}_s \triangleq \hat{\boldsymbol{h}}_s - \boldsymbol{h}_s \in \mathbb{C}^{N_sN_r(L+1)}$ we have

$$\boldsymbol{e}_s = \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{h}_s + \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\boldsymbol{v}_s + \bar{\boldsymbol{W}}\bar{\boldsymbol{v}}_r - \boldsymbol{h}_s, \tag{3.97}$$

from which we can derive the source-relay channel estimation error covariance matrix $\boldsymbol{R}_{e_s} \triangleq \{\boldsymbol{e}_s\boldsymbol{e}_s^H\} \in \mathbb{H}_+^{N_sN_r(L+1)\times N_sN_r(L+1)}$ as

$$\begin{aligned}\boldsymbol{R}_{e_s} &= \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\Big(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{G}^H\boldsymbol{H}_r^H\bar{\boldsymbol{W}}^H \\ &\quad - \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\Big(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s} - \boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{G}^H\boldsymbol{H}_r^H\bar{\boldsymbol{W}}^H + \boldsymbol{R}_{h_s} \\ &\quad + \bar{\boldsymbol{W}}\boldsymbol{H}_r\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{H}_r^H\bar{\boldsymbol{W}}^H\sigma_{v_s}^2 + \bar{\boldsymbol{W}}\bar{\boldsymbol{W}}^H\sigma_{\bar{v}_r}^2. \end{aligned} \tag{3.98}$$

In (3.98) $\sigma_{v_s}^2 \in \mathbb{R}_+$ and $\sigma_{\bar{v}_r}^2 \in \mathbb{R}_+$ are the variances of the source-relay and relay-destination noise vectors, respectively. The source-relay channel covariance matrix $\boldsymbol{R}_{h_s} \triangleq \{\boldsymbol{h}_s\boldsymbol{h}_s^H\} \in \mathbb{H}_{++}^{N_sN_r(L+1)\times N_sN_r(L+1)}$ in (3.98) is given by

$$\boldsymbol{R}_{h_s} = \begin{bmatrix} \sigma_{h_s}^2[0]\boldsymbol{\Theta}_s[0] \otimes \boldsymbol{\Upsilon}_s[0] & & \boldsymbol{0}_{N_sN_r\times N_sN_r} \\ & \ddots & \\ \boldsymbol{0}_{N_sN_r\times N_sN_r} & & \sigma_{h_s}^2[L]\boldsymbol{\Theta}_s[L] \otimes \boldsymbol{\Upsilon}_s[L] \end{bmatrix}, \tag{3.99}$$

which is obtained based on the matrix variate Gaussian distributions of $\boldsymbol{\mathcal{H}}_s[l]$ described in Section 2.2.5 in Chapter 2.

Unfortunately from (3.98) we find that the channel estimation error covariance matrix depends directly on the relay-destination OFDM block diagonal channel matrix $\boldsymbol{H}_r$. According to (3.88) $\boldsymbol{H}_r$ is, in general, not completely known due to the channel estimation error $\boldsymbol{E}_r$ that results from the estimation of $\boldsymbol{H}_r$ in the first phase of channel estimation. Therefore an analytically tractable optimisation problem cannot be formulated using $\boldsymbol{R}_{e_s}$ in (3.98). A possible approach to this problem is to average the objective function over the channel estimation error matrix to obtain a robust channel estimate as done in e.g. [56] and [57] for the simpler case of narrowband MIMO relay channel estimation. However, using such an approach for MIMO OFDM relay channel estimation results in a formidable problem that is more difficult than the case of narrowband channels. Furthermore, this would require the statistics of the channel estimation error to be known to all nodes which may be unknown. For simplicity, we shall thus make the assumption that the relay-destination OFDM channel estimate $\hat{\boldsymbol{H}}_r$, obtained from the first phase of channel estimation, is sufficiently accurate that the estimation error matrix in (3.88) can be neglected i.e. we assume that $\hat{\boldsymbol{H}}_r = \boldsymbol{H}_r$. Such an assumption can be validated when the relay transmit power P_r in the first phase of channel estimation is sufficiently larger than the noise variance as shall be shown through numerical simulations in Section 3.6. With the assumption $\hat{\boldsymbol{H}}_r = \boldsymbol{H}_r$ we

define a new error covariance matrix $\hat{\boldsymbol{R}}_{e_s} \in \mathbb{H}_{++}^{N_s N_r(L+1)\times N_s N_r(L+1)}$ as

$$\begin{aligned}\hat{\boldsymbol{R}}_{e_s} \triangleq\ & \bar{\boldsymbol{W}}\hat{\boldsymbol{H}}_r\boldsymbol{G}\Big(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}^H \\ & - \bar{\boldsymbol{W}}\hat{\boldsymbol{H}}_r\boldsymbol{G}\Big(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s} - \boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}^H + \boldsymbol{R}_{h_s} \\ & + \bar{\boldsymbol{W}}\hat{\boldsymbol{H}}_r\boldsymbol{G}\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}^H\sigma_{v_s}^2 + \bar{\boldsymbol{W}}\bar{\boldsymbol{W}}^H\sigma_{\bar{v}_r}^2, \end{aligned} \tag{3.100}$$

which is obtained from (3.98) simply by using the substitution $\hat{\boldsymbol{H}}_r = \boldsymbol{H}_r$. Using (3.100) we can now formulate a tractable problem for minimising the source-relay channel MSE subject to the source and relay transmision power constraints. The transmit power consumed by the source in the second phase of channel estimation is $\mathrm{tr}\{\boldsymbol{S}\boldsymbol{S}^H\}$, whilst the transmit power consumed by the relay is $\mathrm{tr}\{\boldsymbol{G}\boldsymbol{r}\boldsymbol{r}^H\boldsymbol{G}^H\}$, where $\boldsymbol{r}$ is given in (3.90). We see that the transmit power consumed by the relay depends on the source channel vector $\boldsymbol{h}_s$ which is currently unknown. We shall thus limit the expected transmit power which can be expressed as

$$\begin{aligned} & \mathrm{tr}\Big\{\boldsymbol{G}\mathbb{E}\left\{\boldsymbol{r}\boldsymbol{r}^H\right\}\boldsymbol{G}^H\Big\} \\ = & \ \mathrm{tr}\Big\{\boldsymbol{G}\left(\Big(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\Big) + \sigma_{v_s}^2\boldsymbol{I}_{KN_r}\right)\boldsymbol{G}^H\Big\}, \end{aligned} \tag{3.101}$$

where $\boldsymbol{R}_{h_s}$ is given in (3.99). Taking into account the source and relay transmit power constraints we can formulate the optimisation problem for minimising the MSE of the source-relay channel estimate as

$$\min_{\boldsymbol{S},\boldsymbol{G},\bar{\boldsymbol{W}}} \quad \mathrm{tr}\Big\{\hat{\boldsymbol{R}}_{e_s}\Big\} \tag{3.102}$$

$$\text{s.t.} \quad \mathrm{tr}\Big\{\boldsymbol{S}\boldsymbol{S}^H\Big\} \leq P_s \tag{3.103}$$

$$\mathrm{tr}\Big\{\boldsymbol{G}\left(\Big(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\Big) + \sigma_{v_s}^2\boldsymbol{I}_{KN_r}\right)\boldsymbol{G}^H\Big\} \leq \bar{P}_r, \tag{3.104}$$

where $P_s \in \mathbb{R}_{++}$ and $\bar{P}_r \in \mathbb{R}_{++}$ are the available power budgets to the source and relay, respectively, during the second phase of channel estimation.

Even with the simplifying assumption that $\hat{\boldsymbol{H}}_r = \boldsymbol{H}_r$, the constrained optimisation problem in (3.102)-(3.104) is still a challenging problem since it is non-convex in the design variables $\boldsymbol{S}$, $\boldsymbol{G}$, and $\bar{\boldsymbol{W}}$, and obtaining the globally optimal solution is therefore difficult to achieve. In the following sections we firstly propose an iterative algorithm that is shown to be guaranteed to achieve at least a locally optimal solution to (3.102)-(3.104). Due to its iterative nature such an algorithm may be too computationally expensive for practical implementation. We therefore then proceed to deriving a suboptimal but simplified approach, which has the possibility of a reduced computational expense.

3.4 Iterative Source-Relay Channel Estimation Algorithm

The optimal solution to (3.102)-(3.104) cannot be obtained in closed form since the optimisation problem is not jointly convex in the design variables $\boldsymbol{S}$, $\boldsymbol{G}$, and $\bar{\boldsymbol{W}}$. However, when any two of the variables are fixed the resulting optimisation problem can be shown to be a convex optimisation problem for the remaining variable. This suggests that each variable can be iteratively updated until the objective function converges to (at least) a locally optimal solution. Let us denote $\boldsymbol{S}_i$, $\boldsymbol{G}_i$, and $\bar{\boldsymbol{W}}_i$ as being the variables for the ith iteration (a single iteration refers to the updating of the triplet $\{\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_i\}$ and not just to the updating of a single variable) of the proposed iterative algorithm, and define the cost function with these variables as

$$\begin{aligned}
&\Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_i\Big) \\
&\triangleq \mathrm{tr}\Big\{\bar{\boldsymbol{W}}_i \hat{\boldsymbol{H}}_r \boldsymbol{G}_i \Big(\boldsymbol{M}_{s,i} \otimes \boldsymbol{I}_{N_r}\Big) \boldsymbol{R}_{h_s} \Big(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r}\Big) \boldsymbol{G}_i^H \hat{\boldsymbol{H}}_r^H \bar{\boldsymbol{W}}_i^H\Big\} \\
&- \mathrm{tr}\Big\{\bar{\boldsymbol{W}}_i \hat{\boldsymbol{H}}_r \boldsymbol{G}_i \Big(\boldsymbol{M}_{s,i} \otimes \boldsymbol{I}_{N_r}\Big) \boldsymbol{R}_{h_s}\Big\} - \mathrm{tr}\Big\{\boldsymbol{R}_{h_s} \Big(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r}\Big) \boldsymbol{G}_i^H \hat{\boldsymbol{H}}_r^H \bar{\boldsymbol{W}}_i^H\Big\} \\
&+ \mathrm{tr}\Big\{\bar{\boldsymbol{W}}_i \hat{\boldsymbol{H}}_r \boldsymbol{G}_i \boldsymbol{G}_i^H \hat{\boldsymbol{H}}_r^H \bar{\boldsymbol{W}}_i^H \sigma_{v_s}^2\Big\} + \mathrm{tr}\Big\{\bar{\boldsymbol{W}}_i \bar{\boldsymbol{W}}_i^H \sigma_{\bar{v}_r}^2\Big\} + \mathrm{tr}\big\{\boldsymbol{R}_{h_s}\big\}, \qquad (3.105)
\end{aligned}$$

where we also define

$$\boldsymbol{M}_{s,i} \triangleq [\boldsymbol{F}_0 \boldsymbol{S}_i, ..., \boldsymbol{F}_L \boldsymbol{S}_i]. \qquad (3.106)$$

3.4.1 Updating the Source Training Matrix

We focus firstly on updating the source training matrix $\boldsymbol{S}_i$ given the variables $\boldsymbol{G}_{i-1}$ and $\boldsymbol{W}_{i-1}$. As previously mentioned, when $\boldsymbol{G}_{i-1}$ and $\boldsymbol{W}_{i-1}$ from the last iteration are fixed, the optimisation problem in (3.102)-(3.104) for $\boldsymbol{S}_i$ is convex and the optimal solution can therefore be found. The problem for finding the source training matrix on the ith iteration is to minimise $\Psi(\boldsymbol{S}_i, \boldsymbol{G}_{i-1}, \bar{\boldsymbol{W}}_{i-1})$ subject to the source and relay transmission power constraints. To proceed let us firstly write the cost function for updating the source training matrix in terms of Frobenious

norms as

$$\Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_{i-1}, \bar{\boldsymbol{W}}_{i-1}\Big) = \left\| \begin{bmatrix} \left(\bar{\boldsymbol{W}}_{i-1}\hat{\boldsymbol{H}}_r\boldsymbol{G}_{i-1}\left(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}^{1/2} - \boldsymbol{R}_{h_s}^{1/2}\right)^T \\ \left(\bar{\boldsymbol{W}}_{i-1}\left(\hat{\boldsymbol{H}}_r\boldsymbol{G}_{i-1}\boldsymbol{G}_{i-1}^H\hat{\boldsymbol{H}}_r^H\sigma_{v_s}^2 + \sigma_{\bar{v}_r}^2\boldsymbol{I}_{KN_d}\right)^{1/2}\right)^T \end{bmatrix}^T \right\|_F^2 . \tag{3.107}$$

The power constraints in (3.103) and (3.104) can also be equivalently written as

$$\|\boldsymbol{S}_i\|_F^2 \leq P_s \tag{3.108}$$

$$\left\|\boldsymbol{G}_{i-1}\big(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\big)\boldsymbol{R}_{h_s}^{1/2}\right\|_F^2 \leq \bar{P}_r - \mathrm{tr}\Big\{\boldsymbol{G}_{i-1}\boldsymbol{G}_{i-1}^H\sigma_{v_s}^2\Big\}. \tag{3.109}$$

Given the precoder $\boldsymbol{G}_{i-1}$ and the destination processor $\boldsymbol{W}_{i-1}$ the problem in (3.102)-(3.104) for the source training matrix can therefore be written as

$$\min_{\boldsymbol{S}_i} \quad \left\| \begin{bmatrix} \left(\bar{\boldsymbol{W}}_{i-1}\hat{\boldsymbol{H}}_r\boldsymbol{G}_{i-1}\left(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}^{1/2} - \boldsymbol{R}_{h_s}^{1/2}\right)^T \\ \left(\bar{\boldsymbol{W}}_{i-1}\left(\hat{\boldsymbol{H}}_r\boldsymbol{G}_{i-1}\boldsymbol{G}_{i-1}^H\hat{\boldsymbol{H}}_r^H\sigma_{v_s}^2 + \sigma_{\bar{v}_r}^2\boldsymbol{I}_{KN_d}\right)^{1/2}\right)^T \end{bmatrix}^T \right\|_F^2 \tag{3.110}$$

$$\text{s.t.} \quad \|\boldsymbol{S}_i\|_F^2 \leq P_s \tag{3.111}$$

$$\left\|\boldsymbol{G}_{i-1}\big(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\big)\boldsymbol{R}_{h_s}^{1/2}\right\|_F^2 \leq \bar{P}_r - \mathrm{tr}\Big\{\boldsymbol{G}_{i-1}\boldsymbol{G}_{i-1}^H\sigma_{v_s}^2\Big\}. \tag{3.112}$$

By introducing an auxilliary variable $t \in \mathbb{R}_+$ we can equivalantly write this optimisation problem as the following second order conic programming (SOCP) problem [39]

$$\min_{\boldsymbol{S}_i, t} \quad t \tag{3.113}$$

$$\text{s.t.} \quad \|\boldsymbol{S}_i\|_F^2 \leq P_s \tag{3.114}$$

$$\left\|\boldsymbol{G}_{i-1}\big(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\big)\boldsymbol{R}_{h_s}^{1/2}\right\|_F^2 \leq \bar{P}_r - \mathrm{tr}\Big\{\boldsymbol{G}_{i-1}\boldsymbol{G}_{i-1}^H\sigma_{v_s}^2\Big\} \tag{3.115}$$

$$\left\| \begin{bmatrix} \left(\bar{\boldsymbol{W}}_{i-1}\hat{\boldsymbol{H}}_r\boldsymbol{G}_{i-1}\left(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}^{1/2} - \boldsymbol{R}_{h_s}^{1/2}\right)^T \\ \left(\bar{\boldsymbol{W}}_{i-1}\left(\hat{\boldsymbol{H}}_r\boldsymbol{G}_{i-1}\boldsymbol{G}_{i-1}^H\hat{\boldsymbol{H}}_r^H\sigma_{v_s}^2 + \sigma_{\bar{v}_r}^2\boldsymbol{I}_{KN_d}\right)^{1/2}\right)^T \end{bmatrix}^T \right\|_F^2 \leq t. \tag{3.116}$$

Since SOCP problems are standard convex optimisation problems, we can find the optimal solution to (3.113)-(3.116) using interior point methods [39].

3.4.2 Updating the Relay Precoder

Given the training matrix $\boldsymbol{S}_i$ from solving (3.113)-(3.116), and given the variable $\boldsymbol{W}_{i-1}$ from the previous iteration, we now focus on updating the relay precoding matrix $\boldsymbol{G}_i$. The optimisation problem for updating the relay precoding matrix is

$$\min_{\boldsymbol{G}_i} \quad \Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_{i-1}\Big) \tag{3.117}$$

$$\mathrm{tr}\Big\{\boldsymbol{G}_i \underbrace{\Big(\Big(\boldsymbol{M}_{s,i} \otimes \boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r}\Big) + \sigma_{v_s}^2 \boldsymbol{I}_{KN_r}\Big)}_{\boldsymbol{T}_i} \boldsymbol{G}_i^H\Big\} \leq \bar{P}_r. \tag{3.118}$$

The objective function in (3.117) is convex w.r.t. the relay precoding matrix $\boldsymbol{G}_i$ and thus the optimisation problem is a standard convex optimisation problem. It is worthwhile mentioning that the optimisation problem for the relaying precoder in (3.117)-(3.118) can be formulated as a SOCP and solved using interior point methods. However, since there is only one constraint in the problem, a simpler solution can be derived based on the KKT conditions with reduced complexity compared to a SOCP. To this end we consider the Lagrangian associated with the problem in (3.117)-(3.118) which is given by

$$\mathcal{L}\left(\boldsymbol{G}_i, \mu_{r,i}\right) = \Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_{i-1}\Big) + \mu_{r,i}\left(\mathrm{tr}\Big\{\boldsymbol{G}_i \boldsymbol{T}_i \boldsymbol{G}_i^H\Big\} - \bar{P}_r\right), \tag{3.119}$$

where $\mu_{r,i} \in \mathbb{R}_+$ is the KKT multiplier associated with the relay power constraint in (3.118). Since (3.117)-(3.118) is a convex optimisation problem, the following KKT conditions [39] are necessary and sufficient for obtaining the optimal solution to (3.117)-(3.118):

$$\frac{\partial \mathcal{L}\left(\boldsymbol{G}_i, \mu_{r,i}\right)}{\partial \boldsymbol{G}_i^*} = \boldsymbol{0}_{KN_r \times KN_r} \tag{3.120}$$

$$\mathrm{tr}\Big\{\boldsymbol{G}_i \boldsymbol{T}_i \boldsymbol{G}_i^H\Big\} - \bar{P}_r \leq 0 \tag{3.121}$$

$$\mu_{r,i}\left(\mathrm{tr}\Big\{\boldsymbol{G}_i \boldsymbol{T}_i \boldsymbol{G}_i^H\Big\} - \bar{P}_r\right) = 0 \tag{3.122}$$

$$\mu_{r,i} \geq 0 \tag{3.123}$$

The optimal precoding matrix structure can be obtained by solving the condition in (3.120). Using (3.119), then solving (3.120) for $\boldsymbol{G}_i$ results in

$$\boldsymbol{G}_i = \Big(\hat{\boldsymbol{H}}_r^H \bar{\boldsymbol{W}}_{i-1}^H \bar{\boldsymbol{W}}_{i-1} \hat{\boldsymbol{H}}_r + \mu_{r,i} \boldsymbol{I}_{KN_r}\Big)^{-1} \hat{\boldsymbol{H}}_r^H \bar{\boldsymbol{W}}_{i-1}^H \boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r}\Big) \boldsymbol{T}_i^{-1}. \tag{3.124}$$

The remaining task now is to compute the KKT multiplier which is required to ensure that the power constraint in (3.118) is satisfied. Substituting (3.124) into (3.118) we require to compute $\mu_{r,i}$ such that

$$\begin{aligned}\mathrm{tr}\Big\{\boldsymbol{G}_i\boldsymbol{T}_i\boldsymbol{G}_i^H\Big\} &= \mathrm{tr}\Bigg\{\Big(\hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}_{i-1}^H\bar{\boldsymbol{W}}_{i-1}\hat{\boldsymbol{H}}_r + \mu_{r,i}\boldsymbol{I}_{KN_r}\Big)^{-2}\boldsymbol{C}_i\Bigg\} \\ &\leq \bar{P}_r, \end{aligned} \tag{3.125}$$

where for convenience we define

$$\boldsymbol{C}_i \triangleq \hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}_{i-1}^H\boldsymbol{R}_{h_s}\Big(\boldsymbol{M}_{s,i}^H\otimes\boldsymbol{I}_{N_r}\Big)\boldsymbol{T}_i^{-1}\Big(\boldsymbol{M}_{s,i}\otimes\boldsymbol{I}_{N_r}\Big)\boldsymbol{R}_{h_s}\bar{\boldsymbol{W}}_{i-1}\hat{\boldsymbol{H}}_r. \tag{3.126}$$

We now note that the condition in (3.122) can only be satisfied if either $\mu_{r,i} = 0$ or if $\mathrm{tr}\{\boldsymbol{G}_i\boldsymbol{T}_i\boldsymbol{G}_i^H\} - \bar{P}_r = 0$. Since (3.125) is a monotonically decreasing function of $\mu_{r,i}$ it is straightforward to see that if $\mu_{r,i} = 0$ satisfies (3.121) then it can be the only possible solution that also satisfies (3.122) and (3.123). If on the other hand $\mu_{r,i} = 0$ does not satisfy (3.121) then we require to compute some positive value of $\mu_{r,i}$ to satisfy $\mathrm{tr}\{\boldsymbol{G}_i\boldsymbol{T}_i\boldsymbol{G}_i^H\} - \bar{P}_r = 0$, such that the condition in (3.122) is met. Thus, in this case, we require to compute $\mu_{r,i}$ such that (3.125) holds with equality. An appropriate $\mu_{r,i}$ that satisfies (3.125) with equality can be computed using the method of bisection. In order to utilise the method of bisection we require an upper and lower bound for $\mu_{r,i}$. An obvious lower bound for $\mu_{r,i}$ comes directly from the KKT condition in (3.123). An upper bound for $\mu_{r,i}$ can be straightforwardly derived from (3.125) and the bounds for $\mu_{r,i}$ are given by

$$0 \leq \mu_{r,i} \leq \sqrt{\frac{\mathrm{tr}\{\boldsymbol{C}_i\}}{\bar{P}_r}}. \tag{3.127}$$

Once the value of $\mu_{r,i}$ is computed using the method of bisection, the optimal relay precoding matrix is given by (3.124).

3.4.3 Updating the Destination Processor

We lastly focus on updating the destination linear processor $\bar{\boldsymbol{W}}_i$ given the source training matrix $\boldsymbol{S}_i$ and the relay precoder $\boldsymbol{G}_i$. Since the source and relay power constraints are independent of $\bar{\boldsymbol{W}}_i$ (see (3.103) and (3.104)), the optimal destination processor can be derived by solving the unconstrained problem

$$\min_{\bar{\boldsymbol{W}}_i} \quad \Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_i\Big). \tag{3.128}$$

Furthermore, since $\Psi(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_i)$ is a convex quadratic in $\bar{\boldsymbol{W}}_i$ the optimal solution can be found by setting the derivative of $\Psi(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_i)$ w.r.t. $\bar{\boldsymbol{W}}_i^*$ to zero and solving for $\bar{\boldsymbol{W}}_i$ which results in the solution

$$\bar{\boldsymbol{W}}_i = \boldsymbol{R}_{h_s} \left(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{G}_i^H \hat{\boldsymbol{H}}_r^H \left(\hat{\boldsymbol{H}}_r \boldsymbol{G}_i \boldsymbol{T}_i \boldsymbol{G}_i^H \hat{\boldsymbol{H}}_r^H + \sigma_{\bar{v}_r}^2 \boldsymbol{I}_{KN_d} \right)^{-1}. \tag{3.129}$$

3.4.4 Summary of Iterative Algorithm and Convergence

Having discussed how to update the individual components, we now summarise the proposed iterative algorithm and prove that it is guaranteed to converge to at least a locally optimal solution of the problem. The main steps of the proposed iterative algorithm for computing the source, relay, and destination processors for source-relay channel estimation is summarised in Algorithm 1:

Algorithm 1 : Iterative algorithm to compute $\boldsymbol{S}$, $\boldsymbol{G}$, and $\boldsymbol{W}$, for source-relay channel estimation.

Initialisation: Set $i = 0$. Initialise $\boldsymbol{S}_0$ and $\boldsymbol{G}_0$ to satisfy (3.103) and (3.104) and initialise $\boldsymbol{W}_0$.

repeat
  Set $i = i + 1$.
  Update $\boldsymbol{S}_i$ by solving the SOCP in (3.113)-(3.116).
  Compute the matrix $\boldsymbol{T}_i$ in (3.118).
  Compute the matrix $\boldsymbol{C}_i$ in (3.126).
  if $\mathrm{tr}\{(\hat{\boldsymbol{H}}_r^H \boldsymbol{W}_{i-1}^H \boldsymbol{W}_{i-1} \hat{\boldsymbol{H}}_r)^{-2} \boldsymbol{C}_i\} \leq \bar{P}_r$ **then**
    Set $\mu_{r,i} = 0$.
  else
    Solve $\mathrm{tr}\{(\hat{\boldsymbol{H}}_r^H \boldsymbol{W}_{i-1}^H \boldsymbol{W}_{i-1} \hat{\boldsymbol{H}}_r + \mu_{r,i} \boldsymbol{I}_{KN_r})^{-2} \boldsymbol{C}_i\} = \bar{P}_r$ for $\mu_{r,i}$ using the method of bisection.
  end if
  Update $\boldsymbol{G}_i$ using (3.124).
  Update $\boldsymbol{W}_i$ using (3.129).
until $|\hat{\Psi}(\boldsymbol{S}_i, \boldsymbol{G}_i, \boldsymbol{W}_i) - \hat{\Psi}(\boldsymbol{S}_{i-1}, \boldsymbol{G}_{i-1}, \boldsymbol{W}_{i-1})| \leq \epsilon$.
Set $\boldsymbol{S} = \boldsymbol{S}_i$, $\boldsymbol{G} = \boldsymbol{G}_i$, and $\boldsymbol{W} = \boldsymbol{W}_i$.

The iterative algorithm begins by initialising the source training matrix, the relay precoder, and the destination processors[1]. The initialisation of the source training matrix and of the relay precoder should be such that the power constraints in (3.103) and (3.104) are satisfied with equality. On the ith iteration the source training matrix $\boldsymbol{S}_i$ is firstly updated by solving the SOCP in (3.113)-(3.116). Since this is a convex optimisation problem, after updating the training matrix it is

[1]In practice there are numerous choices for initialising the matrices $\boldsymbol{S}_0$, $\boldsymbol{G}_0$, and $\boldsymbol{W}_0$. The impact of different initialisation points shall be considered in Section 3.6

guaranteed that

$$\Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_{i-1}, \bar{\boldsymbol{W}}_{i-1}\Big) \leq \Psi\Big(\boldsymbol{S}_{i-1}, \boldsymbol{G}_{i-1}, \bar{\boldsymbol{W}}_{i-1}\Big). \tag{3.130}$$

In other words, the updating of the source training matrix can only decrease or maintain the objective function value. The relay precoding matrix $\boldsymbol{G}_i$ is then updated according to (3.124) where the parameter $\mu_{r,i}$ is computed to satisfy the relay power constraint in (3.104). Since the derivation of the relay precoding matrix resulted from solving a convex optimisation problem, the updating of $\boldsymbol{G}_i$ can only decrease or maintain the objective function value and we have the inequality

$$\Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_{i-1}\Big) \leq \Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_{i-1}, \bar{\boldsymbol{W}}_{i-1}\Big). \tag{3.131}$$

Finally, the destination processor $\boldsymbol{W}_i$ is updated according to (3.129). Again, since this processor was derived from solving a convex quadratic function it is guaranteed that

$$\Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_i\Big) \leq \Psi\Big(\boldsymbol{S}_i, \boldsymbol{G}_i, \bar{\boldsymbol{W}}_{i-1}\Big). \tag{3.132}$$

The training matrix, relay precoder, and destination processor are repeatedly updated in this fashion until some termination criterion is met e.g. the objective function value reaches some specified threshold or a maximum number of iterations is reached. Due to the fact that the updating of the individual components result in the inequalities (3.130)-(3.132) as well as the fact that the channel estimation MSE is lower bounded by 0, the algorithm is guaranteed to converge. The convergence of the proposed iterative algorithm shall also be demonstrated through numerical simulations in Section 3.6.

3.5 Simplified Source-Relay Channel Estimation Algorithm

In the previous section we established an iterative algorithm to compute the source-relay channel estimate that was shown to be guaranteed to converge to (at least) a locally optimal solution. Due to its iterative nature such an algorithm may be too computationally expensive for practical implementation since it may require a large number of iterations to converge to a reasonable solution. In this section we therefore propose a suboptimal channel estimation algorithm that has reduced computational complexity compared to the proposed iterative algorithm. Before

proceeding let us firstly define the matrix

$$\boldsymbol{T} \triangleq \left(\left(\boldsymbol{M}_{s,i} \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r}\right) + \sigma_{v_s}^2 \boldsymbol{I}_{KN_r}\right), \tag{3.133}$$

which upon substituting into $\hat{\boldsymbol{R}}_{e_s}$ in (3.100) results in

$$\begin{aligned}\hat{\boldsymbol{R}}_{e_s} = {} & \bar{\boldsymbol{W}}\hat{\boldsymbol{H}}_r\boldsymbol{G}\boldsymbol{T}\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}^H - \bar{\boldsymbol{W}}\hat{\boldsymbol{H}}_r\boldsymbol{G}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s} \\ & - \boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H\bar{\boldsymbol{W}}^H + \boldsymbol{R}_{h_s} + \bar{\boldsymbol{W}}\bar{\boldsymbol{W}}^H\sigma_{\bar{v}_r}^2.\end{aligned} \tag{3.134}$$

With the definition of $\boldsymbol{T}$ in (3.133) we can now equivalently write the optimisation problem in (3.102)-(3.104) as

$$\min_{\boldsymbol{S},\boldsymbol{G},\bar{\boldsymbol{W}}} \quad \mathrm{tr}\left\{\hat{\boldsymbol{R}}_{e_s}\right\} \tag{3.135}$$

$$\text{s.t.} \quad \mathrm{tr}\left\{\boldsymbol{S}\boldsymbol{S}^H\right\} \leq P_s \tag{3.136}$$

$$\mathrm{tr}\left\{\boldsymbol{G}\boldsymbol{T}\boldsymbol{G}^H\right\} \leq \bar{P}_r. \tag{3.137}$$

In solving the optimisation problem (3.135)-(3.137) in the following sections, we shall firstly derive the optimal destination processing matrix $\bar{\boldsymbol{W}}$. Based on a high SNR approximation for the source-relay channel we then derive a suboptimal solution for the source training matrix $\boldsymbol{S}$. With the given source training matrix we can then derive the optimal relay precoder $\boldsymbol{G}$. The suboptimality of the proposed algorithm therefore results due to the training matrix being derived under a high SNR approximation. However, we shall note that the following algorithm is asymptotically optimal with increasing SNR of the source-relay channel.

3.5.1 Optimal Destination Processor

As has previously been established, the optimal solution for $\bar{\boldsymbol{W}}$ is the solution to the unconstrained problem of minimising $\mathrm{tr}\{\hat{\boldsymbol{R}}_{e_s}\}$ and is given by

$$\bar{\boldsymbol{W}} = \boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H\left(\hat{\boldsymbol{H}}_r\boldsymbol{G}\boldsymbol{T}\boldsymbol{G}^H\hat{\boldsymbol{H}}_r^H + \sigma_{\bar{v}_r}^2\boldsymbol{I}_{KN_d}\right)^{-1}. \tag{3.138}$$

Substituting (3.138) into (3.134), and utilising the matrix inversion lemma, we can write the source-relay channel estimation error covariance matrix in a more

concentrated manner as

$$\hat{\boldsymbol{R}}_{e_s} = \underbrace{\left(\boldsymbol{R}_{h_s}^{-1} + \frac{1}{\sigma_{v_s}^2}\left(\boldsymbol{M}_s^H \boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\right)^{-1}}_{\boldsymbol{E}_1} + \underbrace{\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\left(\frac{1}{\sigma_{\bar{v}_r}^2}\boldsymbol{T}\boldsymbol{G}^H \hat{\boldsymbol{H}}_r^H \hat{\boldsymbol{H}}_r \boldsymbol{G}\boldsymbol{T} + \boldsymbol{T}\right)^{-1}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}}_{\boldsymbol{E}_2}. \tag{3.139}$$

We see from (3.139) that the error covariance matrix is the summation of the two seperate matrices $\boldsymbol{E}_1 \in \mathbb{C}^{N_s N_r(L+1)\times N_s N_r(L+1)}$ and $\boldsymbol{E}_2 \in \mathbb{C}^{N_s N_r(L+1)\times N_s N_r(L+1)}$. The matrix $\boldsymbol{E}_1$ is the MSE matrix associated with the source-relay link, whilst $\boldsymbol{E}_2$ represents the increment in the MSE due to the forwarding of the pilot symbols over the relay-destination link. Similar observations have been made in [41] and [45], in the context of linear transceiver designs for two-hop MIMO relay systems, as well as in [80] for linear transceivers designs in a multi-hop multi-user MIMO relay network. It is shown in [41] and [80] that in a high SNR environment the decomposition of the error covariance matrix into two seperate MSE matrices can greatly simplify linear transceiver designs. For the case of low SNR the resulting linear transceivers are suboptimal but nevertheless have been shown in [41] and [80] to have comparable performance to optimal solutions. In the following we shall show that, under a high SNR approximation, the decomposition of the error covariance matrix according to (3.139) can greatly simplify the source-relay channel estimation problem in the sense that it allows the relay precoder $\boldsymbol{G}$ and the training matrix $\boldsymbol{S}$ to be calculated from two seperate optimisation problems. As such the resulting simplified algorithm is non-iterative and should have reduced complexity compared to the iterative algorithm proposed in Section 3.4. Before proceeding we note that with the error covariance now given by (3.139) we can write the optimisation problem in (3.135)-(3.137) as

$$\min_{\boldsymbol{S},\boldsymbol{G}} \quad \mathrm{tr}\{\boldsymbol{E}_1\} + \mathrm{tr}\{\boldsymbol{E}_2\} \tag{3.140}$$

$$\text{s.t.} \quad \mathrm{tr}\Big\{\boldsymbol{S}\boldsymbol{S}^H\Big\} \leq P_s, \tag{3.141}$$

$$\mathrm{tr}\Big\{\boldsymbol{G}\boldsymbol{T}\boldsymbol{G}^H\Big\} \leq \bar{P}_r. \tag{3.142}$$

3.5.2 Suboptimal Source Training Matrix Design

We focus firstly on deriving the source training matrix $\boldsymbol{S}$. Since the optimal training matrix as the solution to (3.140)-(3.142) cannot be computed in closed form, and our main interest here is to derive a computationally efficient channel estimation algorithm, we adopt a suboptimal approach for which a closed form solution is available. To this end let us assume for the moment that the relay precoding matrix $\boldsymbol{G}$ can be decomposed as

$$\boldsymbol{G} = \boldsymbol{L}\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{T}^{-1}, \tag{3.143}$$

where $\boldsymbol{L} \in \mathbb{C}^{KN_r \times N_s N_r (L+1)}$ is an arbitrary matrix yet to be determined. It shall be shown later that the optimal relay precoder as the solution to (3.140)-(3.142) does indeed have the structure given in (3.143). It is interesting to note that, with the definition of the matrix $\boldsymbol{T}$ given in (3.133), then (3.143) can be written as

$$\boldsymbol{G} = \boldsymbol{L}\underbrace{\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\left(\left(\boldsymbol{M}_{s,i} \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_{s,i}^H \otimes \boldsymbol{I}_{N_r}\right) + \sigma_{v_s}^2 \boldsymbol{I}_{KN_r}\right)^{-1}}_{\boldsymbol{W}_G}. \tag{3.144}$$

We see that the relay precoding matrix is comprised of two main components. The matrix $\boldsymbol{W}_G \in \mathbb{C}^{N_s N_r (L+1) \times KN_r}$ represents the optimal MMSE receiver for the first hop channel, whilst $\boldsymbol{L}$ represents a subsequent precoding operation. Thus, the relay firstly minimises the MSE of the received symbols $\boldsymbol{r}$ (see Figure 3.2 and equation (3.90)) using the optimal MMSE receive filter $\boldsymbol{W}_G$, prior to performing a linear precoding through the transmit precoding matrix $\boldsymbol{L}$. Substituting (3.143) into $\boldsymbol{E}_2$ defined in (3.139) and using the matrix inversion lemma we can write

$$\boldsymbol{E}_2 = \left(\frac{1}{\sigma_{v_r}^2}\boldsymbol{L}^H \hat{\boldsymbol{H}}_r^H \hat{\boldsymbol{H}}_r \boldsymbol{L} + \left(\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{T}^{-1}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}\right)^{-1}\right)^{-1}. \tag{3.145}$$

It is now straightforward to show that

$$\lim_{\sigma_{v_s}^2 \to 0}\left(\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{T}^{-1}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s}\right)^{-1} = \boldsymbol{R}_{h_s}^{-1}, \tag{3.146}$$

and consequently we have

$$\lim_{\sigma^2_{v_s}\to 0} \boldsymbol{E}_2 = \left(\frac{1}{\sigma^2_{\bar{v}_r}} \boldsymbol{L}^H \hat{\boldsymbol{H}}_r^H \hat{\boldsymbol{H}}_r \boldsymbol{L} + \boldsymbol{R}_{h_s}^{-1} \right)^{-1}. \tag{3.147}$$

In other words for the case of high SNR of the source-relay channel $\boldsymbol{E}_2$ only depends on the relay transmit variable $\boldsymbol{L}$. Similarly, using the structure of the relay precoder in (3.143) it is also straightforward to show that

$$\lim_{\sigma^2_{v_s}\to 0} \mathrm{tr}\Big\{ \boldsymbol{G}\boldsymbol{T}\boldsymbol{G}^H \Big\} = \mathrm{tr}\Big\{ \boldsymbol{L}\boldsymbol{R}_{h_s}\boldsymbol{L}^H \Big\}, \tag{3.148}$$

where we again observe that for the scenario of high source-relay channel SNR the transmit power consumed by the relay depends only on the variable $\boldsymbol{L}$. With these observations we can decompose the original problem in (3.140)-(3.142) and optimise the source training matrix $\boldsymbol{S}$ by solving

$$\min_{\boldsymbol{S}} \quad \mathrm{tr}\left\{ \left(\boldsymbol{R}_{h_s}^{-1} + \frac{1}{\sigma^2_{v_s}} \left(\boldsymbol{M}_s^H \boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r} \right) \right)^{-1} \right\} \tag{3.149}$$

$$\text{s.t.} \quad \mathrm{tr}\Big\{ \boldsymbol{S}\boldsymbol{S}^H \Big\} \leq P_s. \tag{3.150}$$

It is interesting to observe that, under the high SNR approximation, the optimisation of $\boldsymbol{S}$ according to (3.149)-(3.150) does not require any knowledge of the relay-destination channel estimate $\hat{\boldsymbol{H}}_r$. Since the problem (3.149)-(3.150) has been obtained based on the high SNR approximations of (3.146) and (3.147) it is obvious that any solution to (3.149)-(3.150) will, in general, be a suboptimal solution to the original problem (3.140)-(3.142). However, since (3.149)-(3.150) is independent of $\boldsymbol{L}$, and is therefore independent of the relay precoder $\boldsymbol{G}$, it is a significantly easier problem to handle. In fact, we note that the optimisation problem has precisely the same structure as that for obtaining the MMSE solution for the relay-destination channel estimation problem (c.f. (3.40)-(3.41)). Therefore the optimal and suboptimal MMSE algorithms discussed in Sections 3.2.3 and 3.2.4, respectively, can be used to optimise the training matrix $\boldsymbol{S}$. Furthermore, if the source has no information of the covariance matrix $\boldsymbol{R}_{h_s}$, then we can set $\boldsymbol{R}_{h_s}^{-1} = \boldsymbol{0}_{N_sN_r(L+1)\times N_sN_r(L+1)}$ and then (3.149)-(3.150) reduces to a problem that has the same structure as that in (3.21)-(3.22). In this case the matrix $\boldsymbol{S}$ can be optimised using the LS algorithm discussed in Section 3.2.2.

3.5.3 Optimal Relay Precoder

We now focus on deriving the relay precoding matrix $\boldsymbol{G}$ given any training matrix $\boldsymbol{S}$. Whilst the source training matrix design was simplified by using a high SNR approximation, no such approximation shall be used in deriving the relay precoder. In other words we shall compute the optimal relay precoder as the solution to (3.140)-(3.142). We note firstly that the matrix $\boldsymbol{E}_1$ defined in (3.139) and the source power constraint in (3.150) are both independent of $\boldsymbol{G}$. Thus the optimal solution to (3.140)-(3.142) can be found from solving

$$\min_{\boldsymbol{G}} \quad \mathrm{tr}\{\boldsymbol{E}_2\} \tag{3.151}$$

$$\mathrm{tr}\Big\{\boldsymbol{G}\boldsymbol{T}\boldsymbol{G}^H\Big\} \leq \bar{P}_r. \tag{3.152}$$

To derive the optimal relay precoder let us consider the following singular value decompositions (SVD's)

$$\boldsymbol{T}^{-1/2}\left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{R}_{h_s} = \boldsymbol{U}_t\boldsymbol{\Lambda}\boldsymbol{V}_t^H \tag{3.153}$$

$$\boldsymbol{H}_r = \boldsymbol{U}_r\boldsymbol{\Delta}\boldsymbol{V}_r^H, \tag{3.154}$$

where $\boldsymbol{U}_t \in \mathbb{C}^{KN_r \times KN_r}$ and $\boldsymbol{V}_t \in \mathbb{C}^{N_sN_r(L+1)\times N_sN_r(L+1)}$ are unitary matrices, and the diagonal matrix $\boldsymbol{\Lambda} \in \mathbb{R}_+^{KN_r \times N_sN_r(L+1)}$ contains the non-zero singular values $\{\lambda_i\}_{i=1}^{R_t} \in \mathbb{R}_{++}$, with $R_t \triangleq \mathrm{rank}\{\boldsymbol{\Lambda}\}$, on its upper left main diagonal. Similarly, $\boldsymbol{U}_r \in \mathbb{C}^{KN_d \times KN_d}$ and $\boldsymbol{V}_r \in \mathbb{C}^{KN_r \times KN_r}$, are unitary and $\boldsymbol{\Delta} \in \mathbb{R}_+^{KN_d \times KN_r}$ is diagonal containing the non-zero singular values $\{\delta_i\}_{i=1}^{R_r} \in \mathbb{R}_{++}$, with $R_r \triangleq \mathrm{rank}\{\boldsymbol{\Delta}\}$, on its upper left main diagonal. The singular values in $\boldsymbol{\Lambda}$ and $\boldsymbol{\Delta}$ are assumed w.l.o.g. to be arranged in descending order.

Theorem 3: The structure of the optimal relay precoder $\boldsymbol{G}$ as the solution to the problem in (3.151)-(3.152) is given by

$$\boldsymbol{G} = \boldsymbol{V}_r\boldsymbol{\Phi}\left(\boldsymbol{\Lambda}^T\right)^\dagger \boldsymbol{V}_t^H\boldsymbol{R}_{h_s}\left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r}\right)\boldsymbol{T}^{-1}, \tag{3.155}$$

where $\boldsymbol{\Phi} \in \mathbb{R}_+^{KN_r \times KN_r}$ is a diagonal matrix satisfying $R \triangleq \mathrm{rank}\{\boldsymbol{\Phi}\} \leq \min(R_t, R_r)$, and has non-negative diagonal elements $\{\phi_i\}_{i=1}^{R} \in \mathbb{R}_+$ yet to be determined.

Proof: See Section 3.8.4 on page 83. □

Before proceeding we note that by defining $\boldsymbol{L} \triangleq \boldsymbol{V}_r\boldsymbol{\Phi}(\boldsymbol{\Lambda}^T)^\dagger\boldsymbol{V}_t^H$, the optimal relay precoder in (3.155) matches the structure that was assumed in (3.143). This validates the derivation of the suboptimal training matrix design in the previous

section. Substituting the optimal relay precoder structure of (3.155) into (3.141)-(3.142), and using the decompositions in (3.153) and (3.154), then after some straightforward deductions it can be shown that the matrix valued optimisation problem reduces to

$$\min_{\boldsymbol{\phi}} \quad \sum_{i=1}^{R} \frac{\lambda_i^2 \sigma_{\bar{v}_r}^2}{\phi_i^2 \delta_i^2 + \sigma_{\bar{v}_r}^2} \tag{3.156}$$

$$\text{s.t.} \quad \sum_{i=1}^{R} \phi_i^2 \leq \bar{P}_r \tag{3.157}$$

$$\phi_i^2, \geq 0 \quad 1 \leq i \leq R. \tag{3.158}$$

This is a standard convex optimisation problem for which the optimal solution can be found using the KKT conditions of optimality [39]. The optimal solution can be shown to be given by

$$\phi_i^2 = \left[\sqrt{\frac{\lambda_i^2 \sigma_{\bar{v}_r}^2}{\mu_r \delta_i^2}} - \frac{\sigma_{\bar{v}_r}^2}{\delta_i^2} \right]^+, \tag{3.159}$$

where $\mu_r \in \mathbb{R}_+$ is the KKT multiplier which is required to be computed to satisfy the relay power constraint in (3.157) and can be efficiently found using the waterfilling algorithm in [79].

3.6 Simulation Results

In this section we evaluate the performance of the discussed channel estimation algorithms through numerical simulations.

3.6.1 General Simulation Parameters

In all examples we consider a two-hop MIMO OFDM relaying system equipped with $N_s = N_r = N_d = 2$ antennas at the source, relay, and destination devices. The frequency selective paths between each transmit and receive antenna pair is of length $L+1 = 10$ and we assume that the MIMO channel delay paths are spatially correlated on both the transmit and receive sides, but are temporally uncorrelated. The lth MIMO channel taps of the source-relay and relay-destination channels are

therefore modelled as (see Sections 2.1.3 of Chapter 2)

$$\boldsymbol{\mathcal{H}}_s[l] = \boldsymbol{\Upsilon}_s^{1/2}[l]\boldsymbol{\mathcal{H}}_{sw}[l]\boldsymbol{\Theta}_s^{\mathrm{T}/2}[l] \tag{3.160}$$

$$\boldsymbol{\mathcal{H}}_r[l] = \boldsymbol{\Upsilon}_r^{1/2}[l]\boldsymbol{\mathcal{H}}_{rw}[l]\boldsymbol{\Theta}_r^{\mathrm{T}/2}[l]. \tag{3.161}$$

The transmit spatial correlation matrices $\boldsymbol{\Theta}_s[l]$ and $\boldsymbol{\Theta}_r[l]$, and the receive spatial correlation matrices $\boldsymbol{\Upsilon}_s[l]$ and $\boldsymbol{\Upsilon}_r[l]$, are modelled using the exponential model (see e.g. [81–84]) and have elements given by

$$[\boldsymbol{\Theta}_s[l]]_{mn} = \rho_s[l]^{|m-n|} \tag{3.162}$$

$$[\boldsymbol{\Theta}_r[l]]_{mn} = \rho_r[l]^{|m-n|} \tag{3.163}$$

$$[\boldsymbol{\Upsilon}_s[l]]_{mn} = \varrho_s[l]^{|m-n|} \tag{3.164}$$

$$[\boldsymbol{\Upsilon}_r[l]]_{mn} = \varrho_r[l]^{|m-n|}, \tag{3.165}$$

where the correlation co-efficients $\rho_s[l]$, $\rho_r[l]$, $\varrho_s[l]$, and $\varrho_r[l]$, which are selected from the interval $[0, 1]$, define the level of spatial correlation with lower values signifying low correlation and vice versa. The elements of $\boldsymbol{\mathcal{H}}_{sw}[l]$ and $\boldsymbol{\mathcal{H}}_{rw}[l]$ in (3.160) and (3.161) are drawn from i.i.d. complex Gaussian distributions with zero mean and variances $\sigma^2_{h_s}[l]$ and $\sigma^2_{h_r}[l]$, respectively, and in all simulations we set $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$. The SNR of the source-relay channel is given by $\mathrm{SNR_s} = P_s/(K\sigma^2_{v_s})$ and the SNR of the relay-destination link during the relay-destination channel estimation phase is $\mathrm{SNR_r} = P_r/(K\sigma^2_{v_r})$. Similarly, the SNR of the relay destination link during the source-relay channel estimation phase is $\overline{\mathrm{SNR}}_{\mathrm{r}} = \bar{P}_r/(K\sigma^2_{\bar{v}_r})$. Unless stated otherwise, the number of OFDM subcarriers is set at $K = 32$ in all examples.

3.6.2 Comparison of Relay-Destination Channel Estimation Algorithms

We firstly assess the performance of the optimal LS, optimal MMSE, and suboptimal MMSE algorithms discussed in Sections 3.2.2, 3.2.3, and 3.2.4, respectively, which are used to estimate the relay-destination channel in the first phase of channel estimation. We also include the performance of another suboptimal MMSE estimation algorithm that utilises an equal power allocation (EPA). For this algorithm the destination processor $\boldsymbol{W}$ is given by (3.37) whilst the relay training matrix is given by (3.23). The performance metric used to assess the quality of channel estimate obtained by these algorithms is $MSE \triangleq \mathrm{tr}\{\boldsymbol{R}_{e_r}\}/(N_r N_d(L+1))$, where

$\boldsymbol{R}_{e_r}$ is given in (3.18) and (3.39) for the LS and various MMSE algorithms, respectively. Figure 3.3 shows the MSE performance of the various algorithms against varying SNR_r with the spatial correlation co-efficients set as $\rho_r[l] = \varrho_r[l] = 0.2$, which signifies a low degree of correlation. We see that all MMSE channel estimation algorithms significantly outperform the optimal LS algorithm at low SNR_r, with the performance of the LS algorithm approaching that of the MMSE algorithms in the high SNR_r range. It can also be seen that, interestingly, the suboptimal MMSE solution provides almost the exact same performance across all SNR_r values as the optimal MMSE solution. This comparable level of performance is achieved at a substantially reduced computional complexity, making the suboptimal MMSE solution a very attractive option. We also observe from Figure 3.3 that both the suboptimal and optimal MMSE algorithms provide only a slightly improved channel estimate compared to the suboptimal MMSE EPA solution.

Figure 3.4 shows the channel estimation MSE results for the optimal LS, MMSE EPA, suboptimal MMSE, and optimal MMSE algorithms with the spatial correlation co-efficients now set as $\rho_r[l] = \varrho_r[l] = 0.8$, which signifies high spatial correlation. The same trends to the previous results can be observed. Comparing the results of Figure 3.3 and Figure 3.4 we see that the level of spatial correlation does not affect the performance of the LS algorithm. This can be explained by the fact that the LS error covariance matrix (see (3.18)) does not depend on the relay-destination channel covariance. On the other hand, the performance of all MMSE solutions is improved with higher spatial correlation. We also see from Figure 3.4 that the gap between the MMSE EPA and the MMSE optimal/suboptimal solutions increases with increased spatial correlation.

The LS and MMSE algorithms discussed in Sections 3.2.2, 3.2.3, and 3.2.4 all utilise the same training matrix structure given by $\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}$ (c.f. (3.23) and (3.42)) and only differ in their selection of $\bar{\boldsymbol{X}}$. It was shown that the structure $\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}$ is optimal for the LS and both the optimal and suboptimal MMSE solutions (note that the suboptimality of the suboptimal MMSE solution results due to the choice of $\bar{\boldsymbol{X}}$ and not due to the structure $\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}$). In our next simulation example we compare the performance of the optimal LS and MMSE algorithms to suboptimal LS and MMSE solutions that do not utilise the training matrix structure $\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}$. These suboptimal solutions simply transmit QPSK modulated random training symbols (RTS), that are of course known to the destination, and are scaled to satisfy the relay transmit power constraint with equality. The same random QPSK symbols are used for both the LS RTS and MMSE RTS algorithms. Figure 3.5 shows the MSE results for the LS and MMSE algorithms with the channel spatial correlation co-efficients set as $\rho_r[l] = \varrho_r[l] = 0.5$. It is evident

that the structure of the training matrix $\boldsymbol{X}$ significantly impacts the quality of channel estimation, with the LS and MMSE designs that utilise the optimal training matrix structure outperforming their counterparts that use random training symbols. Interestingly, the gap between the optimal LS and LS RTS algorithms remains almost constant over all $\mathrm{SNR_r}$ values, whereas the gap between the optimal MMSE and MMSE RTS solutions gets larger with increasing $\mathrm{SNR_r}$. It is also interesting to see that the optimal LS algorithm obtains a better channel estimate than the MMSE RTS solution at mid to high $\mathrm{SNR_r}$ values.

In the last example of this section we consider the impact that the number of subcarriers K has on channel estimation performance. The suboptimal MMSE algorithm discussed in Section 3.2.4 is used to demonstrate the effect of varying the number of subcarriers. The spatial correlation co-efficients are set as in the last simulation example at $\rho_r[l] = \varrho_r[l] = 0.5$. Figure 3.6 shows the performance of the suboptimal MMSE algorithms for $K = \{32, 64, 128, 256, 512, 1024\}$. It is evident from these results that increasing the number of subcarriers leads to an improved channel estimation performance. This is due to the fact that with more OFDM subcarriers there are more training symbols being used to estimate the time domain channel matrices and consequently an improved channel estimate can be obtained.

3.6.3 Comparison of Source-Relay Channel Estimation Algorithms

We now assess the performance of the proposed iterative and simplified source-relay channel estimation algorithms. In all examples the simulation parameters are chosen as in Section 3.6.1 with $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$. For the proposed simplified algorithm the source training matrix $\boldsymbol{S}$ is designed using the suboptimal MMSE algorithm discussed in Section 3.2.4.

In our first examples we consider that $\mathrm{SNR_r}$ is sufficiently high such that $\hat{\boldsymbol{H}}_r = \boldsymbol{H}_r$ i.e. the relay-destination channel is estimated perfectly during the first phase of channel estimation. The effect of channel estimation error during the first phase of estimation shall be demonstrated in later simulations. Figure 3.7 shows the convergence of the proposed iterative algorithm for the case of $\mathrm{SNR_s} = \{0\mathrm{dB}, 15\mathrm{dB}\}$ and $\overline{\mathrm{SNR}}_\mathrm{r} = 15\mathrm{dB}$. In this example the initial source training matrix $\boldsymbol{S}_0$ and the inital relay precoder matrix $\boldsymbol{G}_0$ are chosen as random matrices that are scaled to satisfy the source and relay power constraints. The destination processor initialisation $\boldsymbol{W}_0$ is selected as in (3.129). For comparison

purposes, Figure 3.7 also shows the achievable MSE of the proposed simplified algorithm. We observe that the proposed iterative algorithm does indeed converge, but that the final solution is poorer than our simplified algorithm. The poorer performance of the iterative algorithm in this case can be explained by the fact that the source training matrix and relay precoder were randomly initialised. A better approach is to initialise these matrices using the solution from the proposed simplified algorithm. Figure 3.8 shows the convergence of the iterative algorithm in this case with $\text{SNR}_\text{s} = -5\text{dB}$ and $\overline{\text{SNR}}_\text{r} = 15\text{dB}$. In this case we see that the iterative algorithm provides improved performance compared to the simplified algorithm. However, the performance improvement is very small, which indicates that the proposed simplified algorithm is close to (at least) a locally optimal solution.

In our next examples we compare the performance of the proposed channel estimation solutions to various benchmarks. The proposed algorithms are compared to MMSE designs where the destination processor is given by (3.138) and a naive amplify forward (NAF) relay precoder is used, which is given by $\boldsymbol{G} = \sqrt{\bar{P}_r/\text{tr}\{\boldsymbol{T}\}}\boldsymbol{I}_{KN_r}$. The source training matrix $\boldsymbol{S}$ for these benchmark algorithms is selected as the previously discussed RTS or EPA training matrices. Figure 3.9 shows the MSE performance of the various algorithms against SNR_s(dB) with $\overline{\text{SNR}}_\text{r} = 30\text{dB}$. We see that both proposed algorithms outperform the benchmark MMSE NAF RTS and MMSE NAF EPA solutions with the iterative algorithm providing a slightly better channel estimate than the proposed simplified solution. Figure 3.10 shows the MSE performance where $\overline{\text{SNR}}_\text{r}$ is now lowered to $\overline{\text{SNR}}_\text{r} = 20\text{dB}$. We see that the proposed iterative algorithm, the proposed simplified algorithm, the MMSE NAF RTS, and MMSE NAF EPA solutions all suffer a loss in performance caused by the lower SNR of the relay-destination channel.

In practice, if the SNR during the first phase of channel estimation is not sufficiently high, then the relay-destination channel will be estimated imperfectly which shall affect the performance of the proposed channel estimation schemes. We now investigate the effects of the source-relay channel estimation errors. In all proceeding examples the suboptimal MMSE algorithm discussed in Section 3.2.4 is used to estimate the relay-destination channel during the first phase of channel estimation. Figure 3.11 shows the MSE results of the proposed and benchmark source-relay channel estimation algorithms against varying SNR_s(dB) with $\overline{\text{SNR}}_\text{r} = 30\text{dB}$ and the SNR during the relay-destination estimation phase set as $\text{SNR}_\text{r} = 20\text{dB}$. Figure 3.11 shows the corresponding MSE results when the SNR during the relay-destination estimation phase is lowered to $\text{SNR}_\text{r} = 10\text{dB}$. We see that all algorithms suffer a loss in performance due to the mismatch between the estimated relay-destination channel and the actual relay-destination channel.

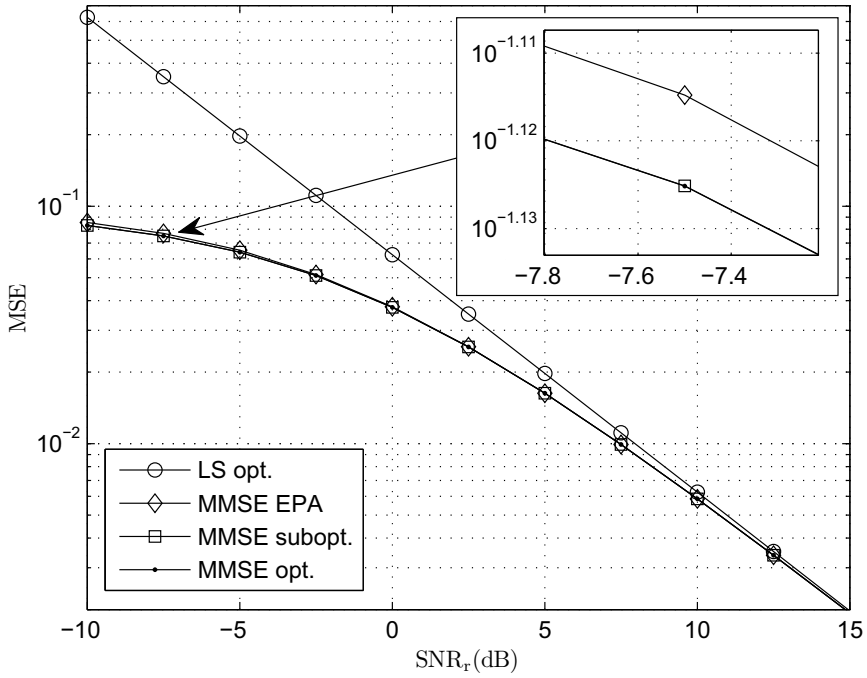

Figure 3.3: MSE against $\mathrm{SNR_r(dB)}$ of LS and MMSE relay-destination channel estimation algorithms for a system with $N_r = N_d = 2$, $L+1 = 10$, $K = 32$, $\sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_r[l] = \varrho_r[l] = 0.2$.

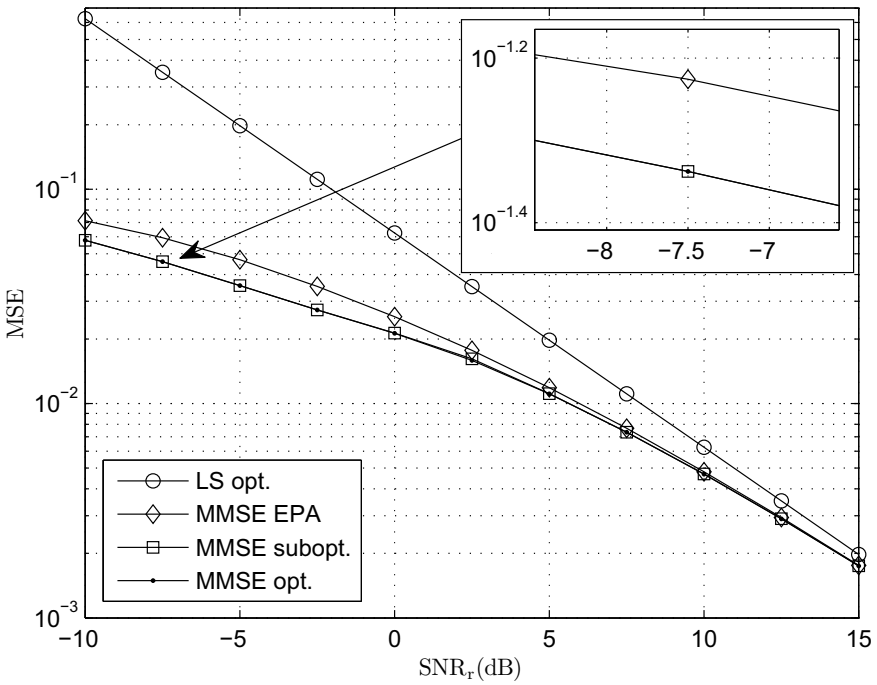

Figure 3.4: MSE against $\mathrm{SNR_r(dB)}$ of LS and MMSE relay-destination channel estimation algorithms for a system with $N_r = N_d = 2$, $L+1 = 10$, $K = 32$, $\sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_r[l] = \varrho_r[l] = 0.8$.

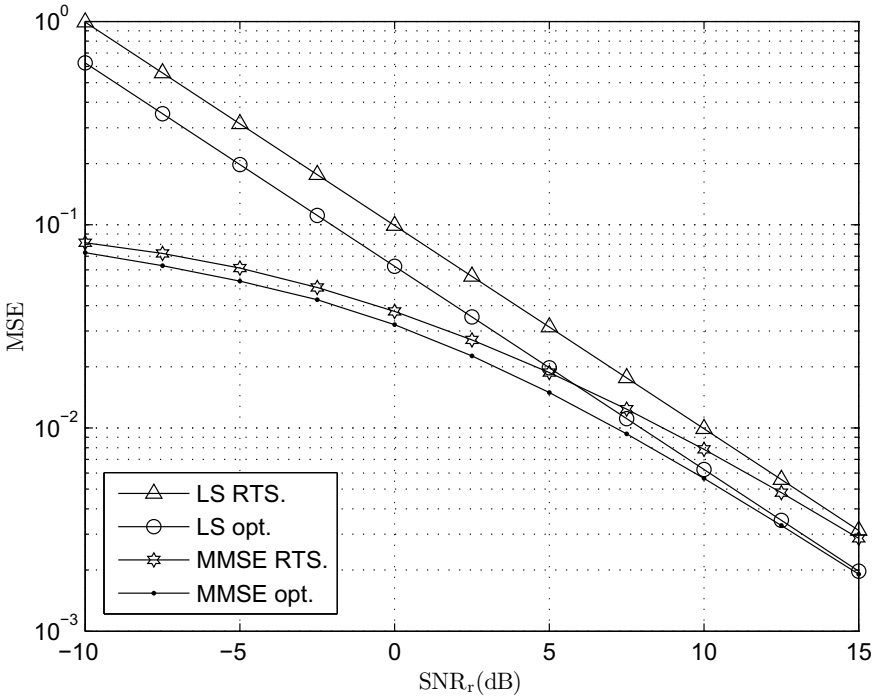

Figure 3.5: MSE against SNR$_r$(dB) of LS and MMSE relay-destination estimation algorithms with optimal or suboptimal training matrix structure for a system with $N_r = N_d = 2$, $L + 1 = 10$, $K = 32$, $\sigma^2_{h_r}[l] = 1/(L + 1)$, and $\rho_r[l] = \varrho_r[l] = 0.5$.

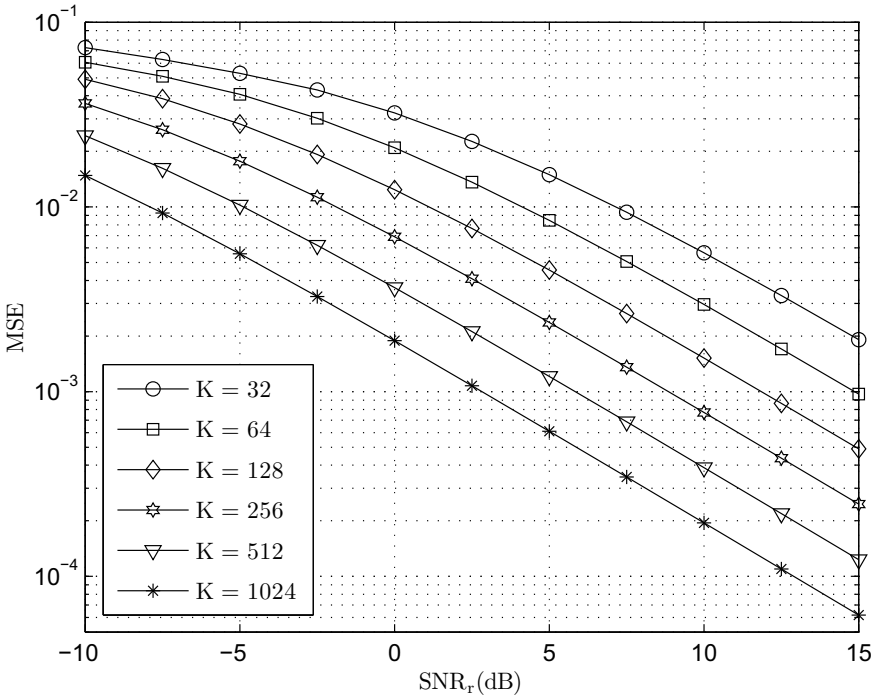

Figure 3.6: MSE against SNR$_r$(dB) of the suboptimal MMSE relay-destination channel estimation algorithm for a system with $N_r = N_d = 2$, $L + 1 = 10$, $\sigma^2_{h_r}[l] = 1/(L + 1)$, $\rho_r[l] = \varrho_r[l] = 0.5$, and $K = \{32, 64, 128, 256, 512, 1024\}$.

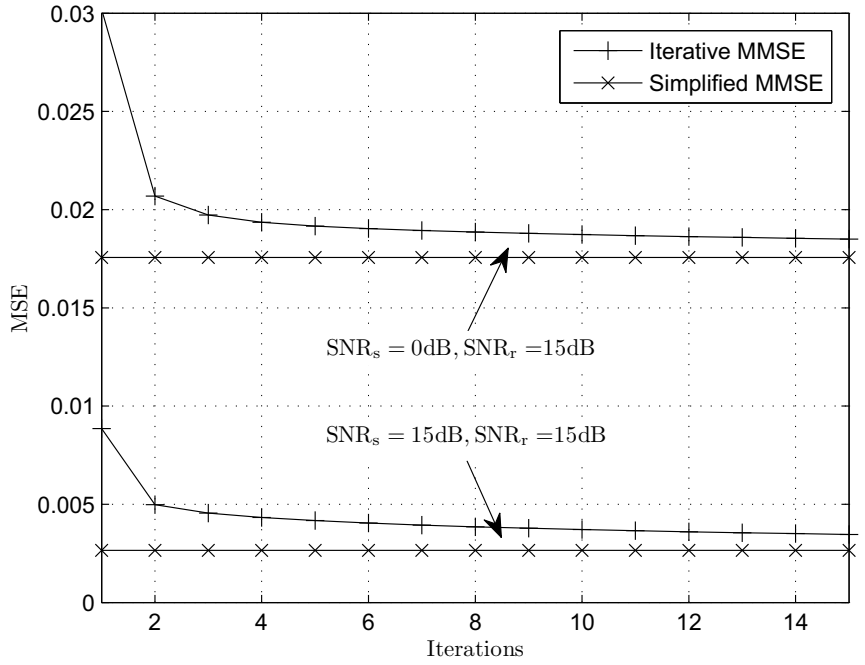

Figure 3.7: Convergence of iterative source-relay estimation algorithm, initialised with random matrices, for a system with $N_s = N_r = N_d = 2$, $L + 1 = 10$, $K = 32$, $\mathrm{SNR_s} = \{0\mathrm{dB}, 15\mathrm{dB}\}$, $\mathrm{SNR_r} = \infty$, $\overline{\mathrm{SNR}}_\mathrm{r} = 15\mathrm{dB}$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$.

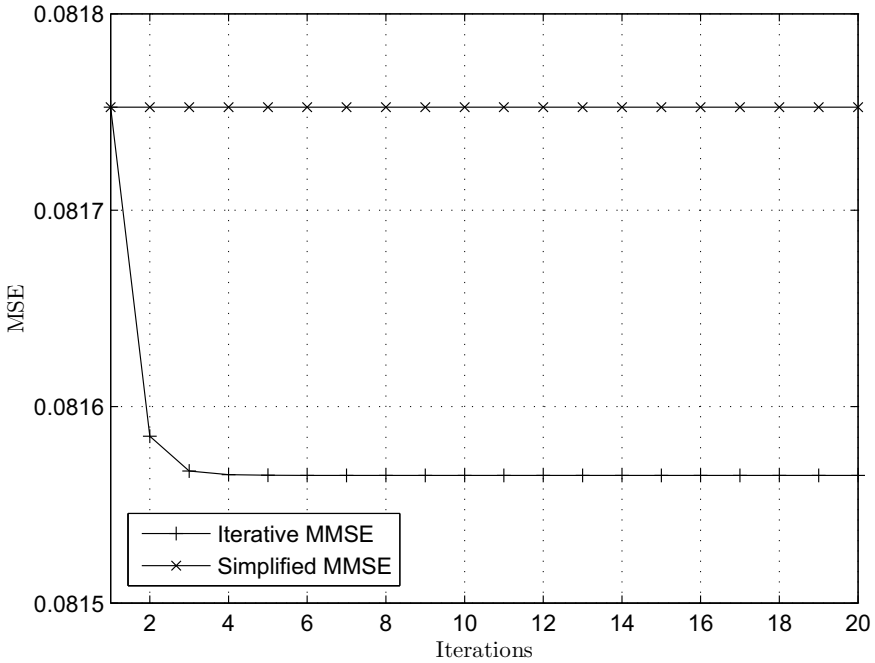

Figure 3.8: Convergence of iterative source-relay estimation algorithm, initialised with simplified MMSE solution, for a system with $N_s = N_r = N_d = 2$, $L + 1 = 10$, $K = 32$, $\mathrm{SNR_s} = -5\mathrm{dB}$, $\mathrm{SNR_r} = \infty$, $\overline{\mathrm{SNR}}_\mathrm{r} = 15\mathrm{dB}$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$.

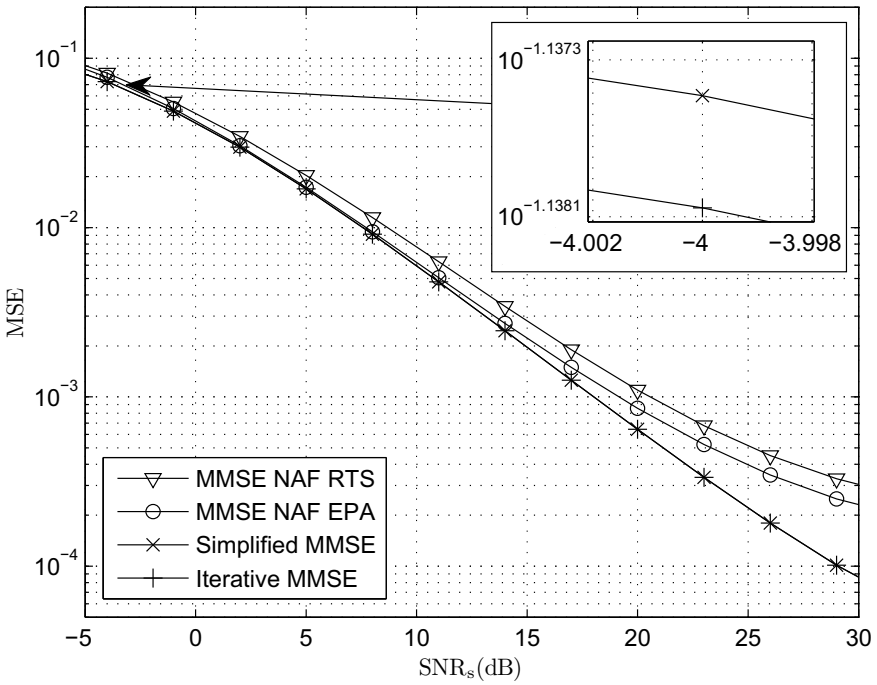

Figure 3.9: MSE against $\mathrm{SNR_s(dB)}$ of MMSE source-relay channel estimation algorithms for a system with $N_s = N_r = N_d = 2$, $L + 1 = 5$, $K = 32$, $\mathrm{SNR_r} = \infty$, $\overline{\mathrm{SNR_r}} = 30\mathrm{dB}$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$.

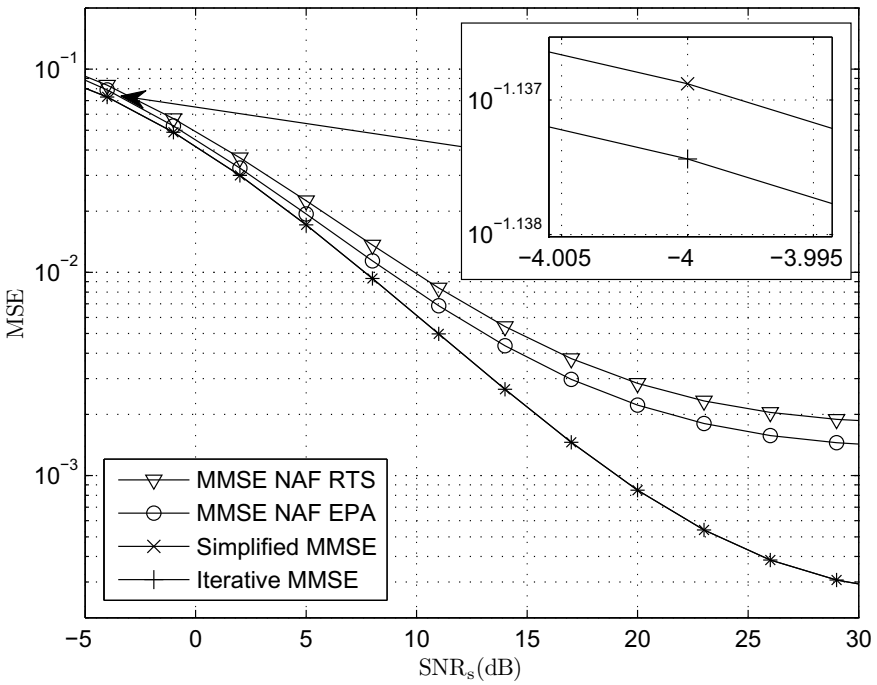

Figure 3.10: MSE against $\mathrm{SNR_s(dB)}$ of MMSE source-relay channel estimation algorithms for a system with $N_s = N_r = N_d = 2$, $L + 1 = 10$, $K = 32$, $\mathrm{SNR_r} = \infty$, $\overline{\mathrm{SNR_r}} = 20\mathrm{dB}$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$.

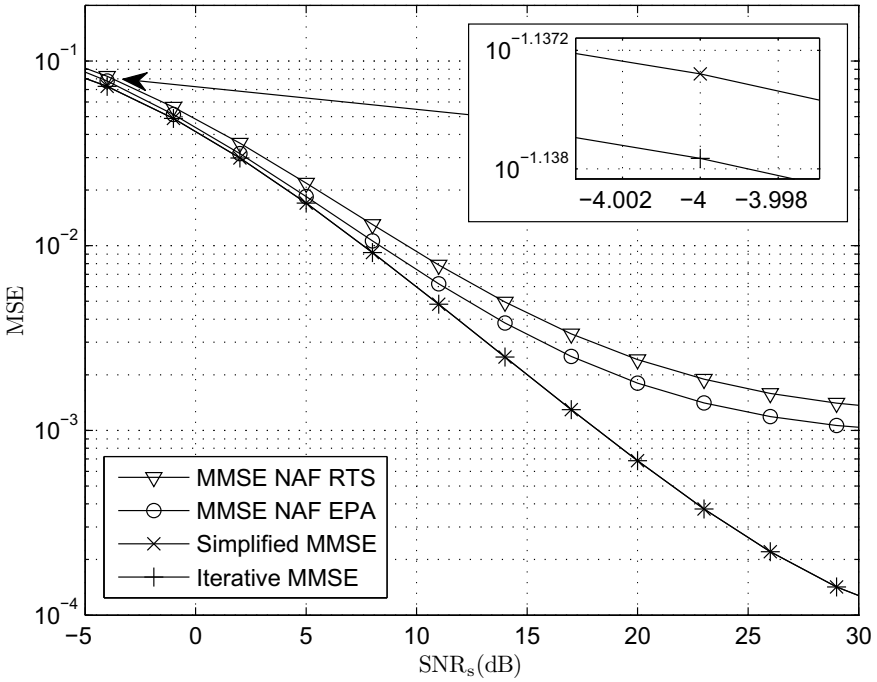

Figure 3.11: MSE against $\mathrm{SNR_s(dB)}$ of MMSE source-relay channel estimation algorithms for a system with $N_s = N_r = N_d = 2$, $L+1 = 10$, $K = 32$, $\mathrm{SNR_r} = 20\mathrm{dB}$, $\overline{\mathrm{SNR_r}} = 30\mathrm{dB}$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$.

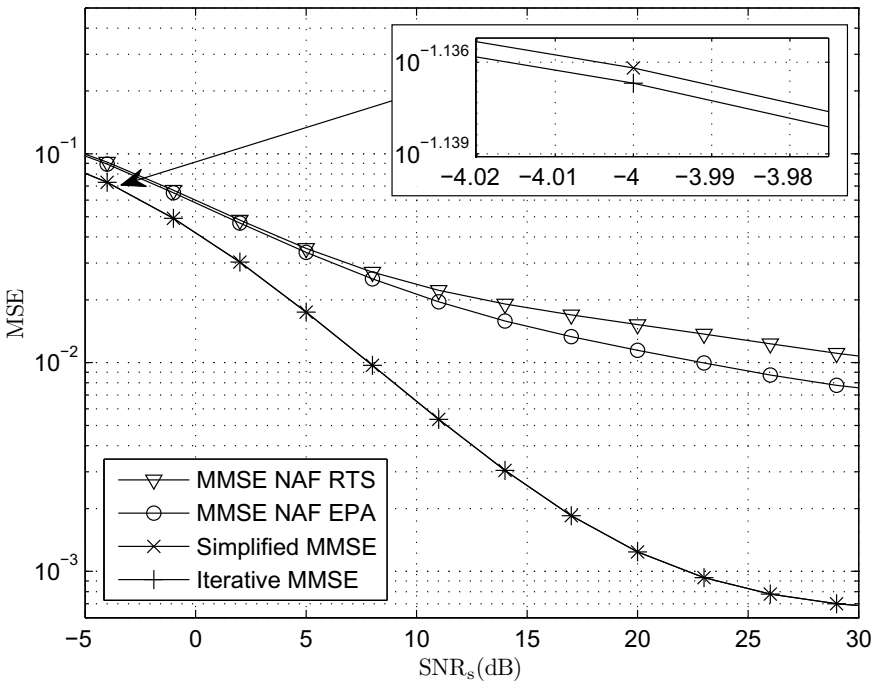

Figure 3.12: MSE against $\mathrm{SNR_s(dB)}$ of MMSE source-relay channel estimation algorithms for a system with $N_s = N_r = N_d = 2$, $L+1 = 10$, $K = 32$, $\mathrm{SNR_r} = 10\mathrm{dB}$, $\overline{\mathrm{SNR_r}} = 30\mathrm{dB}$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, and $\rho_s[l] = \varrho_s[l] = \rho_r[l] = \varrho_r[l] = 0.5$.

3.7 Chapter Summary and Conclusions

In this chapter we considered the task of channel estimation for two-hop MIMO OFDM relaying systems. Specifically we focused on developing algorithms for estimating the frequency selective source-relay and relay-destination channels using training based methods. The task of channel estimation was divided into two phases. In the first phase the relay-destination channel was estimated by sending known training symbols from the relay to the destination device. It was highlighted that this is a standard point-to-point MIMO channel estimation problem and well known LS and MMSE algorithms were discussed in detail. The optimal LS training matrix and both optimal and suboptimal MMSE training matrices were derived. In the second phase of channel estimation the source-relay channel was estimated at the destination using known training symbols from the source-device. In this phase of channel estimation, the source training symbols were transmitted from the source to the relay, which then precoded the received symbols and forwarded them to the destination. Since the source training symbols were forwarded by the relay to the destination, it was highlighted that the destination required knowledge of the relay-destination channel in order to obtain a source-relay channel estimate. For tractability of the source-relay channel estimation problem we assumed that the relay-destination channel was estimated perfectly in the first phase of channel estimation. We then firstly proposed an iterative algorithm to obtain the source-relay channel estimate where the destination processor, the relay precoder, and the source training matrix were updated sequentially. It was shown through theoretical analysis that the proposed iterative algorithm was guaranteed to converge to at least a locally optimal solution. The convergence of the proposed iterative algorithm was also demonstrated through numerical simulations. We also proposed a suboptimal source-relay channel estimation algorithm where all processors could be calculated in closed form solution. The suboptimal solution was derived under a high SNR approximation, which allowed the derivation of the source training matrix to be decoupled from the relay precoder and as such could be derived in closed form. Despite the suboptimality of this approach it was demonstrated through numerical simulations that its performance was comparable to that of the iterative algorithm at a substantially reduced computational complexity. Simulation results also revealed that both the proposed iterative and suboptimal algorithms outperformed competing channel estimation algorithms in terms of achievable MSE.

3.8 Chapter Derivations and Proofs

This section provides some chapter derivations.

3.8.1 Proof of Optimal LS Relay Training Matrix

In this section we prove the structure of the optimal LS training matrix $\boldsymbol{X}$ given in (3.23) of Theorem 1. The proof follows similar lines of argument as those made in e.g. [71, 72, 77]. The proof is heavily reliant on the following lemma:

Lemma 3: [78] For a Hermitian positive definite matrix $\boldsymbol{A} \in \mathbb{H}_{++}^{N\times N}$ we have

$$\mathrm{tr}\{\boldsymbol{A}^{-1}\} \geq \sum_{i=1}^{N} \left[\boldsymbol{A}^{-1}\right]_{ii}. \tag{3.166}$$

where equality holds if $\boldsymbol{A}$ is a diagonal matrix.

Before utilising Lemma 3 we shall find it convenient to decompose the training matrix $\boldsymbol{X}$ w.l.o.g. as

$$\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}, \tag{3.167}$$

where $\boldsymbol{Q} \in \mathbb{C}^{K\times N_r}$ is a semi-unitary matrix and $\bar{\boldsymbol{X}} \in \mathbb{C}^{N_r\times N_r}$ is arbitrary. From Lemma 3 it is straightforward to see that the optimal solution to the problem (3.21)-(3.22) should result in the product $\boldsymbol{M}_x^H\boldsymbol{M}_x$ being diagonal. Using (3.167) as well as the definition of $\boldsymbol{M}_x$ in (3.10) we therefore require that

$$\boldsymbol{M}_x^H\boldsymbol{M}_x = \begin{bmatrix} \bar{\boldsymbol{X}}^H\boldsymbol{Q}^H\boldsymbol{F}_0^H\boldsymbol{F}_0\boldsymbol{Q}\bar{\boldsymbol{X}} & \dots & \bar{\boldsymbol{X}}^H\boldsymbol{Q}^H\boldsymbol{F}_0^H\boldsymbol{F}_L\boldsymbol{Q}\bar{\boldsymbol{X}} \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{X}}^H\boldsymbol{Q}^H\boldsymbol{F}_L^H\boldsymbol{F}_0\boldsymbol{Q}\bar{\boldsymbol{X}} & \dots & \bar{\boldsymbol{X}}^H\boldsymbol{Q}^H\boldsymbol{F}_L^H\boldsymbol{F}_L\boldsymbol{Q}\bar{\boldsymbol{X}} \end{bmatrix}, \tag{3.168}$$

is a diagonal matrix, which clearly holds if $\boldsymbol{Q}$ satisfies the properties stated in (3.24)-(3.26). For the moment we shall assume this to be true (it shall be shown later how such a matrix can be constructed). Assuming $\boldsymbol{Q}$ satisfies the properties stated in (3.24)-(3.26) then $\boldsymbol{M}_x^H\boldsymbol{M}_x$ in (3.168) can be written as

$$\boldsymbol{M}_x^H\boldsymbol{M}_x = \begin{bmatrix} \bar{\boldsymbol{X}}^H\bar{\boldsymbol{X}} & & \boldsymbol{0}_{N_r\times N_r} \\ & \ddots & \\ \boldsymbol{0}_{N_r\times N_r} & & \bar{\boldsymbol{X}}^H\bar{\boldsymbol{X}} \end{bmatrix}, \tag{3.169}$$

Clearly (3.169) is a diagonal matrix only if $\bar{\boldsymbol{X}}^H\bar{\boldsymbol{X}}$ is diagonal, and we can therefore assume w.l.o.g. that $\bar{\boldsymbol{X}} = \boldsymbol{\Lambda}_x$ for some diagonal matrix $\boldsymbol{\Lambda}_x \in \mathbb{R}_+^{N_r\times N_r}$ (note that

with $\bar{\boldsymbol{X}} = \boldsymbol{\Lambda}_x$ for some diagonal $\boldsymbol{\Lambda}_x$, then $\bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} = \boldsymbol{\Lambda}_x^H \boldsymbol{\Lambda}_x \in \mathbb{R}_+^{N_r \times N_r}$ and we can therefore assume w.l.o.g. that $\boldsymbol{\Lambda}_x$ has non-negative real diagonal entries). The remaining task to prove Theorem 1 is then to show that $\boldsymbol{\Lambda}_x$ has equal diagonal elements given by P_r/N_r. To this end we firstly note that with $\bar{\boldsymbol{X}} = \boldsymbol{\Lambda}_x$, the optimisation problem in (3.21)-(3.22) reduces to

$$\min_{\boldsymbol{\lambda}_x} \quad \sum_{i=1}^{N_r} \frac{\sigma_{v_r}^2}{\lambda_{x,i}^2} \tag{3.170}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_r} \lambda_{x,i}^2 \leq P_r, \tag{3.171}$$

$$\lambda_{x,i}^2 \geq 0 \qquad 1 \leq i \leq N_r, \tag{3.172}$$

where $\boldsymbol{\lambda}_x \triangleq [\lambda_{x,1}, ..., \lambda_{x,N_r}]^T$, with $\lambda_{x,i}$ being the ith diagonal element of $\boldsymbol{\Lambda}_x$. Let us now introduce the following lemma:

Lemma 4: For positive scalars $\{a_i\}_{i=1}^N \in \mathbb{R}_+$ we have the inequality

$$\frac{1}{N} \sum_{i=1}^{N} a_i \geq \prod_{i=1}^{N} a_i^{1/N}, \tag{3.173}$$

where equality holds if $\{a_i\}_{i=1}^N$ are all equal. This is the well known arithmetic-geometric mean inequality.

From Lemma 4 the objective function in (3.170) is lower bounded by

$$\sum_{i=1}^{N_r} \frac{\sigma_{v_r}^2}{\lambda_{x,i}^2} \geq N \prod_{i=1}^{N_r} \left(\frac{\sigma_{v_r}^2}{\lambda_{x,i}^2} \right)^{1/N}, \tag{3.174}$$

and we see that the objective function is minimised when $\{\lambda_{x,i}\}_{i=1}^{N_r}$ are all equal. Furthermore, since the objective function in (3.170) is a decreasing function of $\{\lambda_{x,i}\}_{i=1}^{N_r}$ it is straightforward to derive from the constraint in (3.171) that the optimal solution is $\lambda_{x,i} = \sqrt{P_r/N_r}$, $\forall i$. We therefore have $\bar{\boldsymbol{X}} = \sqrt{P_r/N_r}\boldsymbol{I}_{N_r}$, which upon substituting into (3.167) results in the training matrix being given by $\boldsymbol{X} = \sqrt{P_r/N_r}\boldsymbol{Q}$ and completes the proof of Theorem 1.

3.8.2 Proof of Optimal MMSE Relay Training Matrix

In this section we prove the optimal MMSE relay-destination channel training matrix $\boldsymbol{X}$ as stated in Theorem 2. The proof of Theorem 2 is similar to that of the optimal training matrix structure derived in [66].

The proof proceeds similarly to the proof of the optimal LS channel training matrix. Let us firstly write the training matrix w.l.o.g. as

$$\boldsymbol{X} = \boldsymbol{Q}\bar{\boldsymbol{X}}, \tag{3.175}$$

where $\boldsymbol{Q} \in \mathbb{C}^{K\times N_r}$ is a semi-unitary matrix and $\bar{\boldsymbol{X}} \in \mathbb{C}^{N_r\times N_r}$ is arbitrary. To prove Theorem 2 we require to show that $\boldsymbol{Q}$ satisfies the properties in (3.43)-(3.45). To this end let us consider the following lemma:

Lemma 5: For a Hermitian positive definite matrix $\boldsymbol{A} \in \mathbb{H}_{++}^{N\times N}$ we have

$$\mathrm{tr}\big\{\boldsymbol{A}^{-1}\big\} \geq \sum_{i=1}^{P} \mathrm{tr}\big\{\boldsymbol{B}_i^{-1}\big\}. \tag{3.176}$$

where $\boldsymbol{B}_i \in \mathbb{H}_{++}^{M\times M}$ is the ith diagonal submatrix of $\boldsymbol{A}$ and $N = MP$. Equality holds in (3.176) if $\boldsymbol{A}$ is a block diagonal matrix with each diagonal block being Hermitian positive definite.

Before utilising Lemma 5 let us recall that the relay-destination channel covariance matrix is given by

$$\boldsymbol{R}_{h_r} \triangleq \mathbb{E}\{\boldsymbol{h}_r\boldsymbol{h}_r^H\} = \begin{bmatrix} \sigma_{h_r}^2[0]\boldsymbol{\Theta}_r[0]\otimes\boldsymbol{\Upsilon}_r[0] & & \boldsymbol{0}_{N_rN_d\times N_rN_d} \\ & \ddots & \\ \boldsymbol{0}_{N_rN_d\times N_rN_d} & & \sigma_{h_r}^2[L]\boldsymbol{\Theta}_r[L]\otimes\boldsymbol{\Upsilon}_r[L] \end{bmatrix}, \tag{3.177}$$

where we note that the diagonal blocks in (3.177) are of dimension $N_rN_d \times N_rN_d$. Now, applying Lemma 5 to the objective function in (3.40) we have the inequality

$$\begin{aligned} &\mathrm{tr}\left\{\left(\boldsymbol{R}_{h_r}^{-1} + \frac{1}{\sigma_{v_r}^2}\Big(\boldsymbol{M}_x^H\boldsymbol{M}_x\otimes\boldsymbol{I}_{N_d}\Big)\right)^{-1}\right\} \\ &\geq \sum_{l=0}^{L}\mathrm{tr}\left\{\left(\frac{1}{\sigma_{h_r}^2[l]}\boldsymbol{\Theta}_r^{-1}[l]\otimes\boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2}\boldsymbol{N}_l\right)^{-1}\right\}, \end{aligned} \tag{3.178}$$

where we define $\boldsymbol{N} \triangleq \boldsymbol{M}_x^H\boldsymbol{M}_x\otimes\boldsymbol{I}_{N_d}$ with $\boldsymbol{N}_l$ denoting the lth $N_rN_d\times N_rN_d$ diagonal submatrix of $\boldsymbol{N}$. From the inequality in (3.178) we find that the optimal solution to (3.40)-(3.41) should result in the training matrix $\boldsymbol{X}$ having a structure such that $\boldsymbol{N} \triangleq \boldsymbol{M}_x^H\boldsymbol{M}_x\otimes\boldsymbol{I}_{N_d}$ is a block diagonal matrix with $N_rN_d\times N_rN_d$ dimensional diagonal blocks. Using the definition $\boldsymbol{M}_x \triangleq [\boldsymbol{F}_0\boldsymbol{X},...,\boldsymbol{F}_L\boldsymbol{X}]$ in (3.10)

as well as the training matrix decomposition in (3.175) we can expand

$$\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d} = \begin{bmatrix} \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_0^H \boldsymbol{F}_0 \boldsymbol{Q} \bar{\boldsymbol{X}} & \dots & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_0^H \boldsymbol{F}_L \boldsymbol{Q} \bar{\boldsymbol{X}} \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_L^H \boldsymbol{F}_0 \boldsymbol{Q} \bar{\boldsymbol{X}} & \dots & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_L^H \boldsymbol{F}_L \boldsymbol{Q} \bar{\boldsymbol{X}} \end{bmatrix} \otimes \boldsymbol{I}_{N_d} \tag{3.179}$$

$$= \begin{bmatrix} \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_0^H \boldsymbol{F}_0 \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} & \dots & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_0^H \boldsymbol{F}_L \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_L^H \boldsymbol{F}_0 \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} & \dots & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_L^H \boldsymbol{F}_L \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \end{bmatrix}. \tag{3.180}$$

In order for $\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}$ to be a block diagonal matrix with $N_r N_d \times N_r N_d$ dimensional diagonal blocks we clearly require that

$$\bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{F}_m^H \boldsymbol{F}_n \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} = \boldsymbol{0}_{N_r N_d \times N_r N_d} \qquad \forall m \neq n, \tag{3.181}$$

which holds if

$$\boldsymbol{Q}^H \boldsymbol{F}_m^H \boldsymbol{F}_n \boldsymbol{Q} = \boldsymbol{0}_{N_r \times N_r} \qquad \forall m \neq n. \tag{3.182}$$

This is precisely the same condition given in (3.43) of Theorem 2. Assuming that (3.182) holds we can write (3.180) as

$$\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d} = \begin{bmatrix} \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} & & \boldsymbol{0}_{N_r N_d \times N_r N_d} \\ & \ddots & \\ \boldsymbol{0}_{N_r N_d \times N_r N_d} & & \bar{\boldsymbol{X}}^H \boldsymbol{Q}^H \boldsymbol{Q} \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \end{bmatrix}, \tag{3.183}$$

where we have also used the fact that $\boldsymbol{F}_l^H \boldsymbol{F}_l = \boldsymbol{I}_K$, $\forall l$ (see the definition of $\boldsymbol{F}_l$ in (3.8)). We can take $\boldsymbol{Q}^H \boldsymbol{Q} = \boldsymbol{I}_{N_r}$ w.l.o.g. which results in

$$\boldsymbol{Q}^H \boldsymbol{F}_l^H \boldsymbol{F}_l \boldsymbol{Q} = \boldsymbol{I}_{N_r} \qquad \forall m \neq n. \tag{3.184}$$

We have therefore shown that the semi-unitary matrix $\boldsymbol{Q}$ should satisfy the conditions stated in (3.43)-(3.45) which completes the proof of Theorem 2.

3.8.3 Proof of MMSE Objective Function Upper Bound

In this section we show that the MMSE objective function is upper bounded as in (3.65). The upper bound in (3.65) is the same upper bound as derived in [66].

However, in the following, we provide a slightly different proof to that in [66]. To prove the upper bound in (3.65) we shall require the following definition of matrix convexity/concavity:

Definition 2: [39] Given a function $q(\boldsymbol{X}) : \mathbb{X}^{N\times M} \rightarrow \mathbb{R}$ let us define $f(t) \triangleq q(\boldsymbol{U} + t\boldsymbol{V})$, where $\boldsymbol{U} \in \mathbb{C}^{N\times M}$, $\boldsymbol{V} \in \mathbb{C}^{N\times M}$, and t are such that the matrix $\boldsymbol{X} = \boldsymbol{U} + t\boldsymbol{V} \in \mathbb{X}^{N\times M}$. The function $q(\boldsymbol{X})$ is convex w.r.t. $\boldsymbol{X}$ if $f(t)$ satisfies

$$\frac{\partial f^2(t)}{t^2} \geq 0. \tag{3.185}$$

Note that this implies $f(t)$ is convex. Thus to prove the convexity of $q(\boldsymbol{X})$ w.r.t. $\boldsymbol{X}$ we simply require to prove the convexity of $f(t)$ w.r.t. the scalar t. On the other hand, if we have that

$$\frac{\partial f^2(t)}{\partial t^2} \leq 0, \tag{3.186}$$

then the function $q(\boldsymbol{X})$ is concave in $\boldsymbol{X}$.

Based on Definition 2 we can now prove the following lemma:

Lemma 6: Given positive definite matrices $\boldsymbol{X} \in \mathbb{H}_{++}^{N\times M}$ and $\boldsymbol{A} \in \mathbb{H}_{++}^{N\times M}$ the function $q(\boldsymbol{X}) = \mathrm{tr}\{(\boldsymbol{X}^{-1} + \boldsymbol{A})^{-1}\}$ is concave in $\boldsymbol{X}$.

Proof: Using the inverse identity $(\boldsymbol{I}_N + \boldsymbol{P})^{-1} = \boldsymbol{I}_N - (\boldsymbol{I}_N + \boldsymbol{P})^{-1}\boldsymbol{P}$ we have

$$q\left(\boldsymbol{X}\right) = \mathrm{tr}\Big\{\left(\boldsymbol{X}^{-1} + \boldsymbol{A}\right)^{-1}\Big\} = \mathrm{tr}\big\{\boldsymbol{A}^{-1}\big\} - \mathrm{tr}\big\{\left(\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} + \boldsymbol{A}\right)^{-1}\big\}, \tag{3.187}$$

from which we see that proving the concavity of $q(\boldsymbol{X}) = \mathrm{tr}\{(\boldsymbol{X}^{-1} + \boldsymbol{A})^{-1}\}$ w.r.t. $\boldsymbol{X}$ is equivalent to showing that $f(\boldsymbol{X}) \triangleq \mathrm{tr}\{(\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} + \boldsymbol{A})^{-1}\}$ is convex w.r.t. $\boldsymbol{X}$. From Definition 2 we therefore require to show that the function

$$f(t) \triangleq \mathrm{tr}\big\{\left(\boldsymbol{A}\left(\boldsymbol{U} + t\boldsymbol{V}\right)\boldsymbol{A} + \boldsymbol{A}\right)^{-1}\big\}, \tag{3.188}$$

is convex w.r.t. t, where $\boldsymbol{U} \in \mathbb{H}_{++}^{N\times N}$, $\boldsymbol{V} \in \mathbb{H}_{+}^{N\times N}$, and t are such that the matrix $\boldsymbol{U} + t\boldsymbol{V} \in \mathbb{H}_{++}^{N\times N}$. Through straightforward deductions we can write (3.188) as

$$f(t) = \mathrm{tr}\Big\{\boldsymbol{B}^{-1/2}\boldsymbol{A}^{-1}\boldsymbol{B}^{-1/2}\left(\boldsymbol{I}_N + t\boldsymbol{C}\right)^{-1}\Big\} \tag{3.189}$$

where we define the matrices $\boldsymbol{B} \in \mathbb{H}_{++}^{N\times N}$ and $\boldsymbol{C} \in \mathbb{H}_{++}^{N\times N}$ as

$$\boldsymbol{B} \triangleq \boldsymbol{A}^{1/2}\boldsymbol{U}\boldsymbol{A}^{1/2} + \boldsymbol{I}_N \tag{3.190}$$

$$\boldsymbol{C} \triangleq \boldsymbol{B}^{-1/2}\boldsymbol{A}^{1/2}\boldsymbol{V}\boldsymbol{A}^{1/2}\boldsymbol{B}^{-1/2}. \tag{3.191}$$

Introducing the eigendecomposition

$$\boldsymbol{C} = \boldsymbol{U}_c \boldsymbol{\Lambda}_c \boldsymbol{U}_c^H, \tag{3.192}$$

where $\boldsymbol{U}_c \in \mathbb{C}^{N\times N}$ is unitary and $\boldsymbol{\Lambda}_c \in \mathbb{R}_{++}^{N\times N}$ is diagonal with diagonal elements given by $\{\lambda_{c,i}\}_{i=1}^N$, we can further write (3.189) as

$$f(t) = \mathrm{tr}\left\{ \underbrace{\boldsymbol{U}_c^H \boldsymbol{B}^{-1/2} \boldsymbol{A}^{-1} \boldsymbol{B}^{-1/2} \boldsymbol{U}_c}_{\boldsymbol{E}} (\boldsymbol{I}_N + t\boldsymbol{\Lambda}_c)^{-1} \right\} \tag{3.193}$$

$$= \sum_{i=1}^{N} \frac{e_{ii}}{1 + t\lambda_{c,i}}, \tag{3.194}$$

where $e_{ii} \in \mathbb{R}_{++}$ is the ith diagonal elements of the Hermitian positive definite matrix $\boldsymbol{E} \in \mathbb{H}_{++}^{N\times N}$ defined in (3.193). From (3.194) we can now evaluate

$$\frac{\partial f^2(t)}{\partial t^2} = \sum_{i=1}^{N} \frac{\lambda_{c,i} e_{ii}}{\left(1 + t\lambda_{c,i}\right)^4} \geq 0, \tag{3.195}$$

where the inequality results from the fact that $\lambda_{c,i} \geq 0$ and $e_{ii} \geq 0$ since both $\boldsymbol{C}$ and $\boldsymbol{E}$ are Hermitian positive definite matrices. We have thus shown that $f(t)$ is convex w.r.t. t which, from Definition 2, proves that the function $f(\boldsymbol{X}) = \mathrm{tr}\{(\boldsymbol{A}\boldsymbol{X}\boldsymbol{A} + \boldsymbol{A})^{-1}\}$ is convex w.r.t. $\boldsymbol{X}$. From (3.187) this immediately implies that $q(\boldsymbol{X}) = \mathrm{tr}\{\left(\boldsymbol{X}^{-1} + \boldsymbol{A}\right)\}$ is concave in $\boldsymbol{X}$ and completes the proof. □

From Lemma 6 we can now establish that the function

$$\mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\}, \tag{3.196}$$

is concave in $\sigma_{h_r}^2[l]\boldsymbol{\Theta}_r[l] \otimes \boldsymbol{\Upsilon}_r[l]$. Since (3.196) is concave in $\sigma_{h_r}^2[l]\boldsymbol{\Theta}_r[l] \otimes \boldsymbol{\Upsilon}_r[l]$ then from Jensens inequality [39] we obtain

$$\frac{1}{L+1} \sum_{l=0}^{L} \mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\}$$

$$\leq \mathrm{tr}\left\{ \left((L+1) \left(\sum_{l=0}^{L} \sigma_{h_r}^2[l] \boldsymbol{\Theta}_r[l] \otimes \boldsymbol{\Upsilon}_r[l] \right)^{-1} + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\}. \tag{3.197}$$

Since each $\boldsymbol{\Upsilon}_r[l]$ are positive definite, then by defining a matrix $\bar{\boldsymbol{\Upsilon}}_r \in \mathbb{R}_{++}^{N_d \times N_d}$ as

$$\bar{\boldsymbol{\Upsilon}}_r \triangleq \begin{bmatrix} \max\left(\{\upsilon_{r,1}[l]\}_{l=0}^{L}\right) & & 0 \\ & \ddots & \\ 0 & & \max\left(\{\upsilon_{r,N_d}[l]\}_{l=0}^{L}\right) \end{bmatrix}, \tag{3.198}$$

where $\upsilon_{r,i}[l] \in \mathbb{R}_{++}$ is the ith largest eigenvalue of $\boldsymbol{\Upsilon}_r[l]$, we clearly have $\boldsymbol{\Upsilon}_r[l] \preceq \bar{\boldsymbol{\Upsilon}}_r$, $\forall l$. With this observation we can straightforwardly show from (3.197) that

$$\begin{aligned} &\frac{1}{L+1}\sum_{l=0}^{L} \mathrm{tr}\left\{ \left(\frac{1}{\sigma_{h_r}^2[l]} \boldsymbol{\Theta}_r^{-1}[l] \otimes \boldsymbol{\Upsilon}_r^{-1}[l] + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\} \\ &\leq \mathrm{tr}\left\{ \left((L+1) \left(\sum_{l=0}^{L} \sigma_{h_r}^2[l] \boldsymbol{\Theta}_r[l] \right)^{-1} \otimes \bar{\boldsymbol{\Upsilon}}_r^{-1} + \frac{1}{\sigma_{v_r}^2} \bar{\boldsymbol{X}}^H \bar{\boldsymbol{X}} \otimes \boldsymbol{I}_{N_d} \right)^{-1} \right\}. \end{aligned} \tag{3.199}$$

Finally, by multiplying both sides of (3.199) by $L+1$ we prove the upper bound for the MMSE objective function as stated in (3.65).

3.8.4 Proof of Optimal Relay Precoder

In this section we prove the optimal structure of the relay precoding matrix $\boldsymbol{G}$ as given in Theorem 3. In order to prove the optimal relay precoder structure we require the following lemma:

Lemma 7: [78] For Hermitian positive semi-definite matrices $\boldsymbol{A} \in \mathbb{H}_+^{N \times N}$ and $\boldsymbol{B} \in \mathbb{H}_+^{N \times N}$ with non-zero eigenvalues $\{\lambda_{a,i}\}_{i=1}^{N} \in \mathbb{R}_+$ and $\{\lambda_{b,i}\}_{i=1}^{N} \in \mathbb{R}_+$, respectively, which are arranged in the same order, we have the inequality

$$\mathrm{tr}\{\boldsymbol{AB}\} \geq \sum_{i=1}^{N} \lambda_{a,i} \lambda_{b,N-i+1}, \tag{3.200}$$

where equality holds when $\boldsymbol{A}$ is diagonal with elements arranged in descending order and $\boldsymbol{B}$ is diagonal with elements arranged in ascending order.

Before utilising Lemma 7 let us write the objective function in (3.151) as

$$\mathrm{tr}\{\boldsymbol{E}_2\} = \mathrm{tr}\left\{ \boldsymbol{R}_{h_s} \left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r} \right) \left(\frac{1}{\sigma_{\bar{v}_r}^2} \boldsymbol{T}\boldsymbol{G}^H \hat{\boldsymbol{H}}_r^H \hat{\boldsymbol{H}}_r \boldsymbol{G}\boldsymbol{T} + \boldsymbol{T} \right)^{-1} \left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{R}_{h_s} \right\} \tag{3.201}$$

$$= \mathrm{tr}\Bigg\{ \boldsymbol{T}^{-1/2} \left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{R}_{h_s} \boldsymbol{R}_{h_s} \left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{T}^{-1/2} \times \left(\frac{1}{\sigma_{\bar{v}_r}^2} \tilde{\boldsymbol{G}}^H \hat{\boldsymbol{H}}_r^H \hat{\boldsymbol{H}}_r \tilde{\boldsymbol{G}} + \boldsymbol{I}_{KN_r} \right)^{-1} \Bigg\}, \tag{3.202}$$

where we define a new precoding matrix $\tilde{\boldsymbol{G}} \in \mathbb{C}^{KN_r \times KN_r}$ as

$$\tilde{\boldsymbol{G}} \triangleq \boldsymbol{G}\boldsymbol{T}^{1/2}. \tag{3.203}$$

With the newly defined precoder in (3.203), the relay power constraint in (3.152) is equivalent to

$$\mathrm{tr}\left\{ \tilde{\boldsymbol{G}} \tilde{\boldsymbol{G}}^H \right\} \leq \bar{P}_r. \tag{3.204}$$

Finding the optimal relay precoder $\boldsymbol{G}$ as the solution to (3.151)-(3.152) is then clearly equivalent to finding the optimal structure of $\tilde{\boldsymbol{G}}$ that minimises the objective function in (3.202) whilst satisfying the power constraint in (3.204). Let us now recall the SVD in (3.153) which is given by

$$\boldsymbol{T}^{-1/2} \left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{R}_{h_s} = \boldsymbol{U}_t \boldsymbol{\Lambda} \boldsymbol{V}_t^H \tag{3.205}$$

where $\boldsymbol{U}_t \in \mathbb{C}^{KN_r \times KN_r}$ and $\boldsymbol{V}_t \in \mathbb{C}^{N_s N_r (L+1) \times N_s N_r (L+1)}$ are unitary matrices, and the diagonal matrix $\boldsymbol{\Lambda} \in \mathbb{R}_+^{KN_r \times N_s N_r (L+1)}$ contains the non-zero singular values $\{\lambda_i\}_{i=1}^{R_t} \in \mathbb{R}_{++}$, with $R_t \triangleq \mathrm{rank}\{\boldsymbol{\Lambda}\}$, on its upper left main diagonal. The singular values $\{\lambda_i\}_{i=1}^{R_t} \in \mathbb{R}_{++}$ are assumed w.l.o.g. to be arranged in descending order. Furthermore, we introduce the SVD

$$\hat{\boldsymbol{H}}_r \tilde{\boldsymbol{G}} = \boldsymbol{U}_y \boldsymbol{D} \boldsymbol{V}_y^H, \tag{3.206}$$

where $\boldsymbol{U}_y \in \mathbb{C}^{KN_d \times KN_d}$ and $\boldsymbol{V}_y \in \mathbb{C}^{KN_r \times KN_r}$ are unitary matrices, and the diagonal matrix $\boldsymbol{D} \in \mathbb{R}_+^{KN_d \times KN_r}$ contains the non-zero singular values $\{d_i\}_{i=1}^{R_y} \in \mathbb{R}_{++}$, with $R_y \triangleq \mathrm{rank}\{\hat{\boldsymbol{H}}_r \tilde{\boldsymbol{G}}\}$, on its upper left main diagonal. The singular values are assumed to be arranged in decreasing order. By substituting $\hat{\boldsymbol{H}}_r = \boldsymbol{U}_r \boldsymbol{\Delta} \boldsymbol{V}_r^H$ from (3.153) into (3.206), and solving the resulting equation for $\tilde{\boldsymbol{G}}$, we can parameterise

the set of matrices $\tilde{\boldsymbol{G}}$ as

$$\tilde{\boldsymbol{G}} = \boldsymbol{V}_r \boldsymbol{\Delta}^{\dagger} \boldsymbol{U}_r^H \boldsymbol{U}_y \boldsymbol{D} \boldsymbol{V}_y^H. \tag{3.207}$$

From the family of matrices given in (3.207) we wish to find the specific matrix that minimises the objective function in (3.202) whilst satisfying the power constraint in (3.204). To this end we note that by substituting (3.205) and (3.206) into (3.202) we can write

$$\operatorname{tr}\{\boldsymbol{E}_2\} = \operatorname{tr}\left\{ \boldsymbol{U}_t \boldsymbol{\Lambda}^T \boldsymbol{\Lambda} \boldsymbol{U}_t^H \boldsymbol{V}_y \left(\frac{1}{\sigma_{\bar{v}_r}^2} \boldsymbol{D}^T \boldsymbol{D} + \boldsymbol{I}_{KN_r} \right)^{-1} \boldsymbol{V}_y^H \right\} \tag{3.208}$$

$$\geq \sum_{i=1}^{KN_r} \frac{\lambda_i^2 \sigma_{\bar{v}_r}^2}{d_i^2 + \sigma_{\bar{v}_r}^2}, \tag{3.209}$$

where the lower bound is obtained by applying Lemma 7 to (3.208). Also, by substituting (3.207) into the left hand side of the power constraint in (3.204), and making use of Lemma 7, we have

$$\operatorname{tr}\left\{ \tilde{\boldsymbol{G}} \tilde{\boldsymbol{G}}^H \right\} = \operatorname{tr}\left\{ \boldsymbol{U}_r \left(\boldsymbol{\Delta}^T \right)^{\dagger} \boldsymbol{\Delta}^{\dagger} \boldsymbol{U}_r^H \boldsymbol{U}_y \boldsymbol{D} \boldsymbol{D}^T \boldsymbol{U}_y^H \right\} \tag{3.210}$$

$$\geq \sum_{i=1}^{KN_d} \frac{d_i^2}{\delta_i^2}. \tag{3.211}$$

We note that the lower bound in (3.209) is achieved with $\boldsymbol{V}_y = \boldsymbol{U}_t$. Since the relay power constraint in (3.210) is invariant to $\boldsymbol{V}_y$ we can select $\boldsymbol{V}_y = \boldsymbol{U}_t$ to minimise the objective function without affecting the power constraint. We also note that the objective function in (3.209) is invariant to the unitary matrix $\boldsymbol{U}_y$. We will therefore select $\boldsymbol{U}_y$ that results in the least power consumption without affecting the objective function. From (3.210) and (3.211) we see that the relay power consumption is minimised with $\boldsymbol{U}_y = \boldsymbol{U}_r$. By substituting $\boldsymbol{V}_y = \boldsymbol{U}_t$ and $\boldsymbol{U}_y = \boldsymbol{U}_r$ into (3.207) we arrive at the optimal matrix structure

$$\tilde{\boldsymbol{G}} = \boldsymbol{V}_r \boldsymbol{\Delta}^{\dagger} \boldsymbol{D} \boldsymbol{U}_t^H. \tag{3.212}$$

From (3.205) it is straightforward to see that $\boldsymbol{U}_t$ can be equivalently written as

$$\boldsymbol{U}_t = \boldsymbol{T}^{-1/2} \left(\boldsymbol{M}_s \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{R}_{h_s} \boldsymbol{V}_t \boldsymbol{\Lambda}^{\dagger}, \tag{3.213}$$

which upon substituting into (3.212) results in

$$\tilde{\boldsymbol{G}} = \boldsymbol{V}_r \boldsymbol{\Delta}^{\dagger} \boldsymbol{D} \left(\boldsymbol{\Lambda}^T \right)^{\dagger} \boldsymbol{V}_t^H \boldsymbol{R}_{h_s} \left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{T}^{-1/2}. \tag{3.214}$$

Finally, by substituting (3.214) into (3.203) we have the optimal relay precoder $\boldsymbol{G}$ given by

$$\boldsymbol{G} = \boldsymbol{V}_r \boldsymbol{\Delta}^{\dagger} \boldsymbol{D} \left(\boldsymbol{\Lambda}^T \right)^{\dagger} \boldsymbol{V}_t^H \boldsymbol{R}_{h_s} \left(\boldsymbol{M}_s^H \otimes \boldsymbol{I}_{N_r} \right) \boldsymbol{T}^{-1}, \tag{3.215}$$

which, after the definition $\boldsymbol{\Phi} \triangleq \boldsymbol{\Delta}^{\dagger} \boldsymbol{D}$, is precisely the same structure as given in (3.155) of Theorem 3 .

The remaining task to fully prove Theorem 3 is to show that $\text{rank}\{\boldsymbol{\Phi}\} \leq R$ where $R \triangleq \min(R_t, R_r)$, with $R_t = \text{rank}\{\boldsymbol{\Lambda}\}$ and $R_r = \text{rank}\{\boldsymbol{\Delta}\}$. To do so we note that the optimal precoder structure in (3.215) results in (3.209) holding with equality. Substituting $\boldsymbol{\Phi} \triangleq \boldsymbol{\Delta}^{\dagger} \boldsymbol{D}$ (or equivalently $d_i = \phi_i \delta_i$) into (3.209) we can therefore write the objective function as

$$\text{tr}\{\boldsymbol{E}_2\} = \sum_{i=1}^{KN_r} \frac{\lambda_i^2 \sigma_{\bar{v}_r}^2}{\phi_i^2 \delta_i^2 + \sigma_{\bar{v}_r}^2}. \tag{3.216}$$

Similarly, since the optimal precoder structure of (3.215) results in (3.211) holding with equality, then by substituting $d_i = \phi_i \delta_i$ into (3.211) we can write the power consumed by the relay as

$$\text{tr}\Big\{ \tilde{\boldsymbol{G}} \tilde{\boldsymbol{G}}^H \Big\} = \sum_{i=1}^{KN_d} \phi_i^2. \tag{3.217}$$

Noting that $R_t \leq \min(KN_r, N_s N_r (L+1))$ and $R_r \leq \min(KN_d, KN_r)$ then we see from (3.216) that having any $\{\phi_i > 0\}_{i=R+1}^{KN_r}$, where $R \triangleq \min(R_t, R_r)$, will not affect the objective function but from (3.217) will increase the transmission power. It is then obvious that we should have $\text{rank}\{\boldsymbol{\Phi}\} \leq R$ so as not to waste any transmission power. This completes the proof of Theorem 3.

Chapter 4

Non-Linear Transceiver Designs with Perfect CSI

4.1 Introduction

In this chapter we consider transceiver designs for the two-hop MIMO OFDM relaying system introduced in Section 2.1 of Chapter 2. Linear transceiver designs for narrowband MIMO relaying have been studied extensively in [40, 47, 85] for the specific design criteria of minimising the MSE, maximising the mutual information, and minimising the signal-information noise ratio (SINR), respectively. In these works it is assumed that the relay and destination devices can acquire perfect CSI of the source-relay and relay-destination channels and that the source precoder is a scaled identity matrix, which is in fact optimal for the case of no CSI at the source. It is also shown that the optimal linear equaliser is the MMSE equaliser, which is provided by the Wiener-Hopf solution. Interestingly, despite the fact that [40, 47, 85] consider different design criteria, they all derive the same optimal relay precoder structure. Specifically, the optimal relay precoder is composed of the right singular vectors of the relay-destination channel, a real diagonal power allocation matrix with non-negative diagonal elements, and the Hermitian transpose of the source-relay channels left singular vectors. With the MMSE equaliser, and the structure of the relay precoder, the overall transmission process is essentially decomposed into parallel SISO transmissions, which take place over the channel eigenmodes (albeit with a unitary transformation of the transmit symbols as well as the noise signals). Interestingly, this greatly simplifies the transceiver design procedure, since it results in the original matrix valued optimisation problem reducing to a power allocation problem with scalar variables. The solution

to the power allocation problem is specific to the particular design criteria under consideration.

Contrary to the designs in [40, 47, 85], which assume narrowband relaying with no CSI at the source, the linear transceiver designs presented in [45] consider multicarrier transmission and also assume the availability of CSI at the source. A unified framework based on majorisation theory is presented in [45] where the optimal processors are derived for any Schur convex and Schur concave objective function, which are two classes of functions that cover most reasonable communications design criteria. The work in [45] is essentially an extension of the design framework in [16] to the more complicated case of two-hop MIMO relaying. It is shown in [45] that, with the utilisation of the MMSE equaliser at the destination, the optimal linear source and relay precoders completely diagonalise the error covariance matrix for any Schur concave objective function. Such a structure results in the MIMO relay link being simplified to parallel SISO transmissions over the channel eigenmodes. On the other hand, for any Schur convex objective function, a channel diagonalisation only occurs up to a specific unitary transformation of the transmit data symbols. The work of [45] was further extended to the case of MIMO relaying with an arbitrary number of hops in [50], where the optimality of the channel diagonalisation property was shown to hold for any number of successive relayed links.

Besides the study of linear transceiver designs, non-linear transceivers have also received attention in [52–54] under the assumption of narrowband channels and for the case that all nodes can acquire perfect CSI. In [52] and [53] DFE transceiver designs for two-hop MIMO relaying and multi-hop MIMO relaying are studied, respectively. A THP transceiver design is studied in [54] that aims to minimise the arithmetic MSE for two-hop narrowband MIMO relay networks. In this chapter we extend our previous results on DFE and THP transceiver designs in [52] and [54] to the case of transmission over frequency selective channels using OFDM. We consider the optimisation of the DFE and THP processors to minimise the arithmetic MSE subject to transmission power constraints at the source and relay terminals under the assumption that perfect CSI can be acquired at all nodes in the network. We present two different designs depending on whether ZF or MMSE equalisation is used at the receiver.

The remainder of this chapter is organised as follows: The optimal ZF and MMSE DFE/THP processors are derived in Sections 4.2 and 4.3, respectively, under the assumption that all nodes can acquire perfect CSI. In these sections we derive the optimal DFE and THP processors as the solution to the optimisation problem in (2.78)-(2.80) discussed in Section 2.4 of Chapter 2. It is worthwhile

mentioning here that the ZF and MMSE designs presented in sections 4.2 and 4.3 are similar in their problem formulations and indeed their solutions. Nevertheless, the details differ substantially and it is therefore worthwhile treating them independently. Simulation results are discussed in Section 4.4 and demonstrate the advantages of the proposed non-linear techniques compared to linear benchmarks. Finally, conclusions are drawn in Section 4.5.

4.2 ZF DFE/THP Transceiver Design

In this section we derive the processors $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, $\boldsymbol{W}_k$, and $\boldsymbol{U}_k$ as the solution to the optimisation problem (2.78)-(2.80) when ZF equalisation is used at the destination and under the assumption that the source, relay, and destination processors can acquire perfect CSI of all the channels. For the case of ZF equalisation the problem in (2.78)-(2.80) should also include the ZF constraint which shall be established in the following section.

4.2.1 Optimal ZF Equaliser

In solving the optimisation problem in (2.78)-(2.80) we shall firstly derive the optimal equalisers $\boldsymbol{W}_k$ that satisfy the ZF criterion. For given OFDM subcarrier channel matrices, the ZF criterion imposes the following constraint on the relationship between the processors

$$\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k = \boldsymbol{U}_k, \tag{4.1}$$

and is such that, in the absence of noise, the symbols at the receiver can be perfectly reconstructed. Under the assumption that the matrices $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$ are perfectly known at the destination, the optimal ZF equalisers can be obtained by directly solving the constraints in (4.1) resulting in

$$\begin{aligned} \boldsymbol{W}_k &= \boldsymbol{U}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k \right)^{\dagger} && (4.2) \\ &= \boldsymbol{U}_k \left(\boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k \right)^{-1} \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H, && (4.3) \end{aligned}$$

where (4.3) is obtained from the fact that the Moore Penrose pseudo-inverse of a matrix $\boldsymbol{A} \in \mathbb{C}^{M \times N}$, with $M \geq N$, is given by $\boldsymbol{A}^{\dagger} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H$. We note that in order for the inverse in (4.3) to exist, and subsequently for the ZF constraints

in (4.1) to be satisfied, we require that the following conditions are met

$$\text{rank}\left\{\boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k\right\} = N_k. \tag{4.4}$$

For arbitrary matrices $\boldsymbol{A} \in \mathbb{C}^{N\times M}$ and $\boldsymbol{B} \in \mathbb{C}^{M\times N}$ we know that $\text{rank}\{\boldsymbol{AB}\} \leq \min(\text{rank}\{\boldsymbol{A}\}, \text{rank}\{\boldsymbol{B}\})$. Since we can always select processors that satisfy

$$R_{f,k} \triangleq \text{rank}\{\boldsymbol{F}_k\} = N_k \tag{4.5}$$

$$R_{g,k} \triangleq \text{rank}\{\boldsymbol{G}_k\} \geq N_k, \tag{4.6}$$

it is straightforward to show that the conditions in (4.4) can only be met if the channel matrices satisfy the rank constraints

$$R_{s,k} \triangleq \text{rank}\{\boldsymbol{H}_{s,k}\} \geq N_k \tag{4.7}$$

$$R_{r,k} \triangleq \text{rank}\{\boldsymbol{H}_{r,k}\} \geq N_k, \tag{4.8}$$

which in the following we shall assume to be true. Before proceeding it is worthwhile mentioning how these restrictions affect the ZF design. The constraints in (4.7) and (4.8) place a strict limit on the maximum number of data streams N_k which can be transmitted on each subcarrier. This occurs since the number of transmit data streams must satisfy

$$N_k \leq \min(\text{rank}\{\boldsymbol{H}_{s,k}\}, \text{rank}\{\boldsymbol{H}_{r,k}\}). \tag{4.9}$$

For rank deficient channel matrices this can become quite prohibitive. For example in the extreme case where either of the subcarrier channels are rank 1, the ZF design can only support the transmission of a single data stream on that subcarrier. Obviously this will drastically impact on spatial multiplexing gain and consequently the achievable data throughput.

Substituting (4.1) and (4.3) into either (2.60) or (2.75), after some tedious but straightforward deductions, the error covariance matrices for the case of ZF equalisation can be written in a more compact form as

$$\boldsymbol{R}_{e,k} = \boldsymbol{U}_k \Bigg(\boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \Big(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1} \times \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k \Bigg)^{-1} \boldsymbol{U}_k^H, \tag{4.10}$$

where we note that the error covariance matrices are now no longer a function of the equalisers $\boldsymbol{W}_k$ and only depend on the remaining processors $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, and $\boldsymbol{U}_k$.

4.2.2 Optimal ZF Source, Relay, and Feedback Matrices

Having established the optimal ZF equalisers we now focus on deriving the optimal source and relay precoders $\boldsymbol{F}_k$ and $\boldsymbol{G}_k$, as well as the feedback matrices $\boldsymbol{U}_k$ as the solution to (2.78)-(2.80). In order to do so we shall firstly reformulate the optimisation problem. Such a reformulation is based on finding a lower bound to the arithmetic MSE objective function in (2.78) and then showing that appropriate matrices can be constructed to achieve this lower bound. We note that similar approaches can be found in e.g. [24] and [25] for the case of ZF and MMSE DFE transceiver designs in point-to-point MIMO scenarios. Obviously our work differs from these due to the inclusion of the relay terminal and the corresponding additional power constraint this brings.

Before deriving the lower bound to the arithmetic MSE we shall find it convenient in the following analysis to decompose the source precoders w.l.o.g as

$$\boldsymbol{F}_k = \bar{\boldsymbol{F}}_k \boldsymbol{\Psi}_k, \tag{4.11}$$

where $\bar{\boldsymbol{F}}_k \in \mathbb{C}^{N_s \times N_k}$, are arbitrary matrices and $\boldsymbol{\Psi}_k \in \mathbb{C}^{N_k \times N_k}$ are unitary. Substituting (4.11) into (4.10) it is straightforward to write

$$\boldsymbol{R}_{e,k} = \boldsymbol{U}_k \boldsymbol{\Psi}_k^H \boldsymbol{E}_k \boldsymbol{\Psi}_k \boldsymbol{U}_k^H, \tag{4.12}$$

where we have used the fact that $\boldsymbol{\Psi}_k$ are unitary and for notational convenience we define the matrices

$$\boldsymbol{E}_k \triangleq \left(\bar{\boldsymbol{F}}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \right)^{-1} \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \bar{\boldsymbol{F}}_k \right)^{-1}, \tag{4.13}$$

In the following we must assume that the matrices $\boldsymbol{E}_k$ in (4.13) are positive definite (otherwise they do not exist).

We are now ready to derive the lower bound to the arithmetic MSE objective function in (2.78) which can be obtained based on the following lemma:

Lemma 8: For a positive semi-definite matrix $\boldsymbol{A} \in \mathbb{C}^{M \times M}$ we have the inequality $|\boldsymbol{A}|^{1/M} \leq \mathrm{tr}\{\boldsymbol{A}\}/M$ where equality is achieved if and only if $\boldsymbol{A} = \alpha \boldsymbol{I}_M$, for some $\alpha \in \mathbb{R}_+$. In otherwords equality is achieved when $\boldsymbol{A}$ is a diagonal matrix with equal non-negative diagonal entries. This is the well known arithmetic-geometric mean inequality.

Substituting (4.12) into (2.78) and applying Lemma 8 we can show that the arithmetic MSE objective function in (2.78) is lower bounded as follows:

$$\frac{1}{K}\sum_{k=1}^{K}|\boldsymbol{E}_k|^{1/N_k} = \frac{1}{K}\sum_{k=1}^{K}|\boldsymbol{R}_{e,k}|^{1/N_k} \tag{4.14}$$

$$\leq \frac{1}{K}\sum_{k=1}^{K}\frac{\mathrm{tr}\{\boldsymbol{R}_{e,k}\}}{N_k}, \tag{4.15}$$

where to obtain the equality in (4.14) we have used the fact that, for matrices $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{B} \in \mathbb{C}^{N\times N}$, $|\boldsymbol{AB}| = |\boldsymbol{BA}| = |\boldsymbol{B}||\boldsymbol{A}|$, as well as the facts that $\boldsymbol{U}_k$ are unit diagonal triangular and $\boldsymbol{\Psi}_k$ are unitary. We note that the inequality in (4.15) holds for any arbitrary matrices $\boldsymbol{R}_{e,k}$. Thus it is straightforward to deduce that, in order for our transceiver design to minimise the objective function in (2.78), the optimal processors should result in (4.15) holding with equality. From Lemma 8 we know that this can only be achieved when $\boldsymbol{R}_{e,k}$ are diagonal with equal diagonal elements. Noting now that

$$|\boldsymbol{E}_k|^{1/N_k} = \prod_{i=1}^{N_k}\epsilon_{k,i}^{1/N_k}, \tag{4.16}$$

where $\{\epsilon_{k,i}\}_{i=1}^{N_k} \in \mathbb{R}_{++}$ are the non-zero eigenvalues of $\boldsymbol{E}_k$ we can show from Lemma 8 and with the structure of the error covariance matrices in (4.12), that the optimal solution to (2.78)-(2.80) should result in

$$\boldsymbol{R}_{e,k} = \boldsymbol{U}_k\boldsymbol{\Psi}_k^H\boldsymbol{E}_k\boldsymbol{\Psi}_k\boldsymbol{U}_k^H = \prod_{i=1}^{N_k}\epsilon_{k,i}^{1/N_k}\boldsymbol{I}_{N_k}, \tag{4.17}$$

Using the Cholesky factorisation we find that the calculation of $\boldsymbol{U}_k$ and $\boldsymbol{\Psi}_k$ that satisfy (4.17) is equivalent to finding the matrices that satisfy

$$\boldsymbol{U}_k\boldsymbol{\Psi}_k^H\boldsymbol{E}_k^{1/2} = \prod_{i=1}^{N_k}\epsilon_{k,i}^{1/2N_k}\boldsymbol{S}_k^H, \tag{4.18}$$

where $\boldsymbol{S}_k \in \mathbb{C}^{N_k\times N_k}$ are unitary matrices. Upon re-arranging (4.18) we arrive at the matrix decompositions

$$\prod_{i=1}^{N_k}\epsilon_{k,i}^{1/2N_k}\boldsymbol{E}_k^{-1/2} = \boldsymbol{S}_k\boldsymbol{U}_k\boldsymbol{\Psi}_k^H. \tag{4.19}$$

The decompositions in (4.19) are known as the equal diagonal QR decompositions [26] or the Geometric Mean Decompositions (GMD) [28, 86] and result in the matrices $\boldsymbol{U}_k$ being unit diagonal upper right triangular and $\boldsymbol{\Psi}_k$ being unitary as required. The decompositions in (4.19) can be computed using the algorithm provided in [26]. With $\boldsymbol{U}_k$ and $\boldsymbol{\Psi}_k$ being computed according to (4.19) then by construction (4.17) holds with equality. Furthermore, this results in the lower bound in (4.15) holding with equality which, as previously mentioned, is required for our solution to be optimal. With this observation we can simplify the optimisation problem in (2.78)-(2.80). Using the fact that our previous calculation of $\boldsymbol{U}_k$ and $\boldsymbol{\Psi}_k$ result in (4.15) holding with equality we can replace the objective function in (2.78) with the lower bound in (4.15). With this observation, and also using the source precoder decomposition in (4.11), we can reformulate the optimisation problem in (2.78)-(2.80) as

$$\min_{\bar{\boldsymbol{F}}_k, \boldsymbol{G}_k} \quad \sum_{k=1}^{K} |\boldsymbol{E}_k| \tag{4.20}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \text{tr}\Big\{ \bar{\boldsymbol{F}}_k \bar{\boldsymbol{F}}_k^H \Big\} \leq P_s \tag{4.21}$$

$$\sum_{k=1}^{K} \text{tr}\Big\{ \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \bar{\boldsymbol{F}}_k \bar{\boldsymbol{F}}_k^H \boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \right) \boldsymbol{G}_k^H \Big\} \leq P_r, \tag{4.22}$$

Evidently, the remaining task is to find the matrices $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ as the solution to (4.20)-(4.22). Before finding the optimal $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ let us firstly introduce the singular value decompositions (SVDs)

$$\boldsymbol{H}_{s,k} = \boldsymbol{U}_{s,k} \boldsymbol{\Lambda}_k \boldsymbol{V}_{s,k}^H \tag{4.23}$$

$$\boldsymbol{H}_{r,k} = \boldsymbol{U}_{r,k} \boldsymbol{\Delta}_k \boldsymbol{V}_{r,k}^H, \tag{4.24}$$

where $\boldsymbol{U}_{s,k} \in \mathbb{C}^{N_r \times N_r}$ and $\boldsymbol{V}_{s,k} \in \mathbb{C}^{N_s \times N_s}$ are unitary matrices that contain the left and right singular vectors, respectively, of $\boldsymbol{H}_{s,k}$, and similarly $\boldsymbol{U}_{r,k} \in \mathbb{C}^{N_d \times N_d}$ and $\boldsymbol{V}_{r,k} \in \mathbb{C}^{N_r \times N_r}$ are the unitary matrices containing the left and right singular vectors of $\boldsymbol{H}_{r,k}$. The diagonal matrices $\boldsymbol{\Lambda}_k \in \mathbb{R}_+^{N_r \times N_s}$ and $\boldsymbol{\Delta}_k \in \mathbb{R}_+^{N_d \times N_r}$ contain the non-zero singular values $\{\lambda_{k,i}\}_{i=1}^{R_{s,k}} \in \mathbb{R}_{++}$ and $\{\delta_{k,i}\}_{i=1}^{R_{r,k}} \in \mathbb{R}_{++}$, respectively, which are located on their upper left main diagonals and are assumed w.l.o.g. to be in decreasing order. It is worth reiterating here that in order for the ZF equaliser in (4.3) to exist we require, from (4.4), that $N_k \leq \{R_{s,k}, R_{r,k}\}$. Using the decompositions in (4.23) and (4.24) we establish the optimal source and relay precoders as the solution to (4.20)-(4.22) in the following theorem:

Theorem 4: The structure of the optimal processors that minimise the objective function in (4.20) and satisfy the source and relay power constraints in (4.21) and (4.22) are

$$\bar{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Gamma}_k \tag{4.25}$$

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Phi}_k\boldsymbol{U}_{s,k}^H, \tag{4.26}$$

where $\boldsymbol{\Gamma}_k \in \mathbb{R}_+^{N_s \times N_k}$ and $\boldsymbol{\Phi}_k \in \mathbb{R}_+^{N_r \times N_r}$ are diagonal matrices satisfying $\text{rank}\{\boldsymbol{\Gamma}_k\} = \text{rank}\{\boldsymbol{\Phi}_k\} = N_k$ and have elements $\{\gamma_{k,i}\}_{i=1}^{N_k} \in \mathbb{R}_{++}$ and $\{\phi_{k,i}\}_{i=1}^{N_k} \in \mathbb{R}_{++}$, respectively, on their upper left main diagonals, with all other elements being zero.

Proof: The detailed proof of the precoder structures given in (4.25) and (4.26) can be found in Section 4.6.1 on page 110. □

Substituting the optimal precoding matrices (4.25) and (4.26) into (4.20)-(4.22) results in the matrix valued optimisation problem reducing to the simpler scalar valued optimisation problem

$$\min_{\boldsymbol{\gamma},\boldsymbol{\phi}} \quad \sum_{k=1}^{K}\prod_{i=1}^{N_k}\left(\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2\sigma_{v_s}^2+\sigma_{v_r}^2}\right)^{-1} \tag{4.27}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{i=1}^{N_k}\gamma_{k,i}^2 \le P_s \tag{4.28}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{N_k}\phi_{k,i}^2\left(\gamma_{k,i}^2\lambda_{k,i}^2+\sigma_{v_s}^2\right) \le P_r \tag{4.29}$$

$$\gamma_{k,i}^2 > 0, \quad \phi_{k,i}^2 > 0, \quad 1 \le k \le K, \quad 1 \le i \le N_k, \tag{4.30}$$

where we define vectors $\boldsymbol{\gamma} \triangleq [\gamma_{k,i}]_{k,i}^{K,N_k} \in \mathbb{R}_{++}^{\bar{N}}$ and $\boldsymbol{\phi} \triangleq [\phi_{k,i}]_{k,i}^{K,N_k} \in \mathbb{R}_{++}^{\bar{N}}$. In the following section we show that the optimisation problem in (4.27)-(4.30) can be reformulated as a convex optimisation problem and thus the optimal solution can be derived.

4.2.3 ZF Joint Source and Relay Power Allocation Algorithm Using Geometric Programming

In order to simplify our following analysis let us firstly introduce the new variables

$$\vartheta_{k,i} \triangleq \gamma_{k,i}^2 \tag{4.31}$$

$$\varphi_{k,i} \triangleq \phi_{k,i}^2\left(\gamma_{k,i}^2\lambda_{k,i}^2+\sigma_{v_s}^2\right), \tag{4.32}$$

and restate the problem in (4.27)-(4.30) as

$$\min_{\boldsymbol{\vartheta},\boldsymbol{\varphi}} \quad \sum_{k=1}^{K}\prod_{i=1}^{N_k}\left(\frac{\vartheta_{k,i}\lambda_{k,i}^2\sigma_{v_r}^2+\varphi_{k,i}\delta_{k,i}^2\sigma_{v_s}^2+\sigma_{v_s}^2\sigma_{v_r}^2}{\vartheta_{k,i}\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}\right) \tag{4.33}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{i=1}^{N_k}\vartheta_{k,i}\leq P_s \tag{4.34}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{N_k}\varphi_{k,i}\leq P_r \tag{4.35}$$

$$\vartheta_{k,i}>0, \quad \varphi_{k,i}>0, \quad 1\leq k\leq K, \quad 1\leq i\leq N_k, \tag{4.36}$$

where we define $\boldsymbol{\vartheta}\triangleq[\vartheta_{k,i}]_{k,i}^{K,N_k}\in\mathbb{R}_{++}^{\bar{N}}$ and $\boldsymbol{\varphi}\triangleq[\varphi_{k,i}]_{k,i}^{K,N_k}\in\mathbb{R}_{++}^{\bar{N}}$. We note that the power constraints in (4.34) and (4.35) are now decoupled which will simplify our following analysis. The optimal solution to the optimisation problem in (4.33)-(4.36) can be found using geometric programming. To this end let us firstly introduce auxiliary variables $t_{k,i}$, with the associated vector $\boldsymbol{t}\triangleq[t_{k,i}]_{k,i}^{K,N_k}\in\mathbb{R}_{++}^{\bar{N}}$, and equivalently state the optimisation problem in (4.33)-(4.36) as

$$\min_{\boldsymbol{\vartheta},\boldsymbol{\varphi},\boldsymbol{t}} \quad \sum_{k=1}^{K}\prod_{i=1}^{N_k}t_{k,i} \tag{4.37}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{i=1}^{N_k}\vartheta_{k,i}\leq P_s \tag{4.38}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{N_k}\varphi_{k,i}\leq P_r \tag{4.39}$$

$$\frac{\vartheta_{k,i}\lambda_{k,i}^2\sigma_{v_r}^2+\varphi_{k,i}\delta_{k,i}^2\sigma_{v_s}^2+\sigma_{v_s}^2\sigma_{v_r}^2}{t_{k,i}\vartheta_{k,i}\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}\leq 1 \tag{4.40}$$

$$\vartheta_{k,i}>0, \quad \varphi_{k,i}>0, \quad 1\leq k\leq K, \quad 1\leq i\leq N_k. \tag{4.41}$$

We note firstly that the objective function in (4.37) is a posynomial function[1] [39]. Secondly, since the numerator of the constraint in (4.40) is a posynomial and the denominator is a monomial[2] [39], the constraint in (4.40) is also a posynomial. Finally, the constraints in (4.38), (4.39), and (4.41), are trivially posynomial functions. Thus, the problem in (4.37)-(4.41) is a geometric programming problem and the global optimal solution can be found [39]. Geometric programs are in, general, not naturally convex optimisation problems but may be transformed to geometric programs in convex form with a change of variables and a reformulation

[1] A posynomial function is one of the form $f(\boldsymbol{x})=\sum_{k=1}^{K}c_k x_1^{a_1 k}x_2^{a_2 k}...x_K^{a_n K}$.

[2] A monomial function is one of the form $f(\boldsymbol{x})=cx_1^{a_1}x_2^{a_2}...x_n^{a_n}$.

of the objective function and constraints [39]. In order to transform the problem into convex form we introduce the new variables

$$\tilde{t}_{k,i} \triangleq \log\left(t_{k,i}\right) \tag{4.42}$$

$$\tilde{\vartheta}_{k,i} \triangleq \log\left(\vartheta_{k,i}\right) \tag{4.43}$$

$$\tilde{\varphi}_{k,i} \triangleq \log\left(\varphi_{k,i}\right), \tag{4.44}$$

along with the vectors $\tilde{\boldsymbol{t}} \triangleq [\tilde{t}_{k,i}]_{k,i}^{K,N_k} \in \mathbb{R}_{++}^{\bar{N}}$, $\tilde{\boldsymbol{\vartheta}} \triangleq [\tilde{\vartheta}_{k,i}]_{k,i}^{K,N_k} \in \mathbb{R}_{++}^{\bar{N}}$ and $\tilde{\boldsymbol{\varphi}} \triangleq [\tilde{\varphi}_{k,i}]_{k,i}^{K,N_k} \in \mathbb{R}_{++}^{\bar{N}}$. Substituting (4.42)-(4.44) into (4.37)-(4.41) allows us to restate the optimisation problem as

$$\min_{\tilde{\boldsymbol{\vartheta}},\tilde{\boldsymbol{\varphi}},\tilde{\boldsymbol{t}}} \quad \sum_{k=1}^{K} \exp\left(\sum_{i=1}^{N_k} \tilde{t}_{k,i}\right) \tag{4.45}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{i=1}^{N_k} \exp\left(\tilde{\vartheta}_{k,i}\right) \leq P_s \tag{4.46}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{N_k} \exp\left(\tilde{\varphi}_{k,i}\right) \leq P_r \tag{4.47}$$

$$\frac{\exp\left(\tilde{\vartheta}_{k,i}\right)\lambda_{k,i}^2\sigma_{v_r}^2 + \exp\left(\tilde{\varphi}_{k,i}\right)\delta_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_s}^2\sigma_{v_r}^2}{\exp\left(\tilde{t}_{k,i} + \tilde{\vartheta}_{k,i} + \tilde{\varphi}_{k,i}\right)\lambda_{k,i}^2\delta_{k,i}^2} \leq 1,$$

$$1 \leq k \leq K, \quad 1 \leq i \leq N_k. \tag{4.48}$$

Before proceeding we note that with (4.43) and (4.44), the inequality constraints in (4.41) are equivalent to $\exp(\tilde{\vartheta}_{k,i}) > 0$ and $\exp(\tilde{\varphi}_{k,i}) > 0$. Obviously such constraints are redundant since they do not affect the feasible sets for the new variables $\tilde{\boldsymbol{\vartheta}}$ and $\tilde{\boldsymbol{\varphi}}$, and have therefore been removed from the optimisation problem. We now observe that the change of variable in (4.42) has resulted in the objective function (4.37) being converted from a posynomial function into the logarithmically convex function in (4.45). Similarly, through the change of variables (4.43), (4.44), and (4.42), the constraints in (4.46)-(4.48) are also logarithmically convex. Taking the logarithm of (4.45)-(4.48) we can write the problem as a geometric

problem in convex form as

$$\min_{\tilde{\boldsymbol{\vartheta}},\tilde{\boldsymbol{\varphi}},\tilde{\boldsymbol{t}}} \quad \log\left(\sum_{k=1}^{K}\exp\left(\sum_{i=1}^{N_k}\tilde{t}_{k,i}\right)\right) \tag{4.49}$$

$$\text{s.t.} \quad \log\left(\sum_{k=1}^{K}\sum_{i=1}^{N_k}\exp\left(\tilde{\vartheta}_{k,i}\right)\right) \leq \log\left(P_s\right) \tag{4.50}$$

$$\log\left(\sum_{k=1}^{K}\sum_{i=1}^{N_k}\exp\left(\tilde{\varphi}_{k,i}\right)\right) \leq \log\left(P_r\right) \tag{4.51}$$

$$\log\left(\exp\left(\tilde{\vartheta}_{k,i}\right)\lambda_{k,i}^2\sigma_{v_r}^2 + \exp\left(\tilde{\varphi}_{k,i}\right)\delta_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_s}^2\sigma_{v_r}^2\right)$$
$$\leq \left(\tilde{t}_{k,i} + \tilde{\vartheta}_{k,i} + \tilde{\varphi}_{k,i}\right) + \log\left(\lambda_{k,i}^2\delta_{k,i}^2\right), \quad 1 \leq k \leq K, \quad 1 \leq i \leq N_k. \tag{4.52}$$

Since the problem in (4.49)-(4.52) is a convex optimisation problem the optimal solution can be found using interior point methods [39]. Once the optimal $\tilde{\boldsymbol{\vartheta}}$ and $\tilde{\boldsymbol{\varphi}}$ as the solution to (4.49)-(4.52) have been computed, the optimal $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$ as the solution to the original optimisation problem in (4.27)-(4.30) can be recovered using (4.43), (4.44), (4.31) and (4.32).

4.3 MMSE DFE/THP Transceiver Design

In this section we derive the optimal DFE and THP processors as the solution to the optimisation problem in (2.78)-(2.80) for the case when perfect CSI is available to all nodes and the MMSE equaliser is used at the receiver. We begin the optimal MMSE design by firstly deriving the optimal MMSE equalisers $\boldsymbol{W}_k$.

4.3.1 Optimal MMSE Equaliser

Although the ZF solution derived in (4.3) has the often desirable property that it can avoid crosstalk between the different data streams, it is well known to suffer a penalty from noise amplification. The noise amplification phenomenon may manifest itself in poor BER and MSE performance, particularly at low SNR values. The optimal MMSE equaliser, which provides a tradeoff between complete interference cancellation and the noise enhancement exhibited by the ZF receiver, is obtained by setting the derivative of the objective function in (2.78) w.r.t. $\boldsymbol{W}_k^*$

to zero, and then solving for $\boldsymbol{W}_k$. This results in the optimal solution

$$\boldsymbol{W}_k = \boldsymbol{U}_k \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \Big(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H + \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1}, \tag{4.53}$$

which is the well known Wiener Hopf solution. Contrary to the ZF solution in (4.3) we observe that the MMSE solution in (4.53) takes into account the noise covariance matrices $\sigma_{v_s}^2 \boldsymbol{I}_{N_r}$ and $\sigma_{v_r}^2 \boldsymbol{I}_{N_d}$. Therefore in order to implement the MMSE receiver it is required that the noise statistics be known at the destination device. It is also worthwhile mentioning that the MMSE equaliser does not require any strict assumptions on the rank of $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$, which were needed for the ZF equaliser (see the rank constraint in (4.4)). This is due to the fact that the matrix inverse in (4.53) is regularised by the positive definite matrix $\sigma_{v_r}^2 \boldsymbol{I}_{N_d}$. Upon substituting (4.53) into (2.60) or (2.75) we can write the concentrated error covariance matrix when the MMSE equaliser is used as

$$\begin{aligned} \boldsymbol{R}_{e,k} &= \boldsymbol{U}_k \Big(\boldsymbol{I}_{N_k} - \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \Big(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \Big(\boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \Big) \\ &\qquad \times \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big) \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k \Big) \boldsymbol{U}_k^H \qquad (4.54) \\ &= \boldsymbol{U}_k \Big(\boldsymbol{I}_{N_k} + \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \Big(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1} \\ &\qquad \times \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k \Big)^{-1} \boldsymbol{U}_k^H, \qquad (4.55) \end{aligned}$$

where in order to obtain (4.55) we have applied the Woodbury identity[3] to (4.54). We see that the error covariance matrix in (4.55) for the case of MMSE equalisation is always guaranteed to be positive definite irrespective of the specific properties of $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, $\boldsymbol{H}_{s,k}$, and $\boldsymbol{H}_{r,k}$. This is in contrast to the case of the ZF solution (c.f. (4.13)), where for a given N_k the assumptions of $\text{rank}\{\boldsymbol{H}_{s,k}\} \geq N_k$ and $\text{rank}\{\boldsymbol{H}_{r,k}\} \geq N_k$ were required for the ZF solution to exist. For the ZF design, the rank constraints placed a restrictive assumption on the number of data streams that could be transmitted given by $N_k \leq \min(\text{rank}\{\boldsymbol{H}_{s,k}\}, \text{rank}\{\boldsymbol{H}_{r,k}\})$. On the contrary, the MMSE equaliser allows N_k independent data streams to be transmitted irrespective of the channel rank and thus the MMSE solution is capable of dealing with more general channel matrices than that of the ZF solution. Furthermore, it was shown that the ZF precoders given in Theorem 4 satisfied $\text{rank}\{\boldsymbol{F}_k\} = N_k$ and $\text{rank}\{\boldsymbol{G}_k\} = N_k$ (conditions which were also necessary for the

[3]For matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$, the Woodbury identity states that $\boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1} = (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1}$.

existence of the ZF solution). As will be seen in the following sections, the precoders for the MMSE solution do not necessarily have to satisfy these constraints. This provides the MMSE solution more degrees of freedom in which it allocates the available power budgets.

4.3.2 Optimal MMSE Source, Relay, and Feedback Matrices

Similar to the ZF designs derived in the previous section, we shall derive a lower bound to the arithmetic MSE and then find the appropriate matrices such that the lower bound holds with equality. Substituting the source precoder decomposition $\boldsymbol{F}_k = \bar{\boldsymbol{F}}_k \boldsymbol{\Psi}_k$ from (4.11), where we recall that $\bar{\boldsymbol{F}}_k$ are arbitrary matrices and $\boldsymbol{\Psi}_k$ are unitary, into (4.55) we can decompose the error covariance matrices as

$$\boldsymbol{R}_{e,k} = \boldsymbol{U}_k \boldsymbol{\Psi}_k^H \boldsymbol{E}_k \boldsymbol{\Psi}_k \boldsymbol{U}_k^H, \tag{4.56}$$

where we define the matrices $\boldsymbol{E}_k \in \mathbb{C}^{N_k \times N_k}$ as

$$\boldsymbol{E}_k \triangleq \Big(\boldsymbol{I}_{N_k} + \bar{\boldsymbol{F}}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \Big(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1} \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \bar{\boldsymbol{F}}_k \Big)^{-1}. \tag{4.57}$$

We note that, irrelevant of the structure of $\bar{\boldsymbol{F}}_k$, $\boldsymbol{G}_k$, $\boldsymbol{H}_{s,k}$, and $\boldsymbol{H}_{r,k}$, the matrices $\boldsymbol{E}_k$ in (4.57) are always guaranteed to be positive definite. Comparing (4.56) and (4.57) to the corresponding results for the ZF design in (4.12) and (4.13), it is interesting to see that the only difference in the error covariance matrices is the addition of the matrix $\boldsymbol{I}_{N_k}$ in (4.57).

We recall that by applying Lemma 8 to the objective function in (2.78) the lower bound to the arithmetic MSE is given by

$$\frac{1}{K} \sum_{k=1}^{K} |\boldsymbol{E}_k|^{1/N_k} = \frac{1}{K} \sum_{k=1}^{K} |\boldsymbol{R}_{e,k}|^{1/N_k} \tag{4.58}$$

$$\leq \frac{1}{K} \sum_{k=1}^{K} \frac{\mathrm{tr}\{\boldsymbol{R}_{e,k}\}}{N_k}, \tag{4.59}$$

We also recall from Lemma 8 that our optimal solution should result in (4.59) holding with equality which holds provided that $\boldsymbol{R}_{e,k}$ is diagonal with equal diagonal elements given by $\prod_{i=1}^{N_k} \epsilon_{k,i}^{1/N_k}$, where $\{\epsilon_{k,i}\}_{i=1}^{N_k} \in \mathbb{R}_+$ are the eigenvalues of $\boldsymbol{E}_k$ in (4.57). Following the same procedure that was used for the ZF design we find

that this results when the matrices $\boldsymbol{U}_k$ and $\boldsymbol{\Psi}_k$ are selected from the GMD

$$\prod_{i=1}^{N_k} \epsilon_{k,i}^{1/2N_k} \boldsymbol{E}_k^{-1/2} = \boldsymbol{S}_k \boldsymbol{U}_k \boldsymbol{\Psi}_k^H. \tag{4.60}$$

Using the fact that the matrices $\boldsymbol{U}_k$ and $\boldsymbol{\Psi}_k$ from (4.60) result in (4.59) holding with equality we can reformulate the optimisation problem in (2.78)-(2.80) as

$$\min_{\bar{\boldsymbol{F}}_k, \boldsymbol{G}_k} \quad \sum_{k=1}^{K} |\boldsymbol{E}_k| \tag{4.61}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \text{tr}\left\{ \bar{\boldsymbol{F}}_k \bar{\boldsymbol{F}}_k^H \right\} \leq P_s \tag{4.62}$$

$$\sum_{k=1}^{K} \text{tr}\left\{ \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \bar{\boldsymbol{F}}_k \bar{\boldsymbol{F}}_k^H \boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \right) \boldsymbol{G}_k^H \right\} \leq P_r, \tag{4.63}$$

Theorem 5: The optimal MMSE precoding matrices $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ as the solution to the optimisation problem in (4.61)-(4.63) are given by

$$\bar{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k} \boldsymbol{\Gamma}_k \tag{4.64}$$

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k} \boldsymbol{\Phi}_k \boldsymbol{U}_{s,k}^H, \tag{4.65}$$

where $\boldsymbol{\Gamma}_k \in \mathbb{R}_+^{N_s \times N_k}$ and $\boldsymbol{\Phi}_k \in \mathbb{R}_+^{N_r \times N_r}$ are at most rank $\bar{N}_k$ diagonal matrices with elements $\{\gamma_{k,i}\}_{i=1}^{\bar{N}_k} \in \mathbb{R}_+$ and $\{\phi_{k,i}\}_{i=1}^{\bar{N}_k} \in \mathbb{R}_+$ on their upper left main diagonals. Here we define the variable $\bar{N}_k \triangleq \min(N_k, R_{s,k}, R_{r,k})$. The matrices $\boldsymbol{V}_{s,k}$, $\boldsymbol{V}_{r,k}$, and $\boldsymbol{U}_{s,k}$ are given from the channel decompositions in (4.23) and (4.24).

Proof: The proof of the precoder structures in (4.64) and (4.65) are given in Section 4.6.2 on page 114. □

The structure of the source and relay precoders given in (4.64) and (4.65) result in the original matrix valued optimisation problem in (4.61)-(4.63) reducing to the

scalar valued problem

$$\min_{\boldsymbol{\gamma},\boldsymbol{\phi}} \quad \sum_{k=1}^{K}\prod_{i=1}^{\bar{N}_k}\left(1+\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2\sigma_{v_s}^2+\sigma_{v_r}^2}\right)^{-1} \tag{4.66}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\gamma_{k,i}^2 \leq P_s \tag{4.67}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\phi_{k,i}^2\left(\gamma_{k,i}^2\lambda_{k,i}^2+\sigma_{v_s}^2\right) \leq P_r \tag{4.68}$$

$$\gamma_{k,i}^2 \geq 0, \quad \phi_{k,i}^2 \geq 0, \quad 1 \leq k \leq K, \quad 1 \leq i \leq \bar{N}_k. \tag{4.69}$$

Unfortunately, unlike the power allocation problem for the ZF transceiver design, (4.66)-(4.69) cannot be formulated as a convex optimisation problem. However we note that for high received SNR we have the inequality

$$\sum_{k=1}^{K}\prod_{i=1}^{\bar{N}_k}\left(1+\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2\sigma_{v_s}^2+\sigma_{v_r}^2}\right)^{-1} \leq \sum_{k=1}^{K}\prod_{i=1}^{\bar{N}_k}\left(\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2\sigma_{v_s}^2+\sigma_{v_r}^2}\right)^{-1}. \tag{4.70}$$

By replacing the objective function in (4.66) with the upper bound in (4.70), the resulting optimisation problem is precisely the same as that given in (4.27)-(4.30). Thus we can derive $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$ by using the geometric programming approach discussed in section 4.2.3. However, such an approach is obviously only asymptotically optimal with increasing received SNR. As suggested in e.g. [45, 48, 54] an alternating power allocation algorithm can be used to find at least a locally optimal solution to (4.66)-(4.69) and is discussed in the following section.

4.3.3 MMSE Alternating Power Allocation Algorithm

In order to derive the alternating algorithm we firstly decouple the source and relay power constraints using an appropriate change of variable, and then reformulate the optimisation problem. To this end let us define the new variables

$$\vartheta_{k,i} \triangleq \gamma_{k,i}^2, \tag{4.71}$$

$$\varphi_{k,i} \triangleq \phi_{k,i}^2(\gamma_{k,i}^2\lambda_{k,i}^2+\sigma_{v_s}^2), \tag{4.72}$$

and rewrite the optimisation problem in (4.66)-(4.69) as

$$\min_{\boldsymbol{\vartheta},\boldsymbol{\varphi}} \quad \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \log\left(\frac{\vartheta_{k,i}\lambda_{k,i}^2\sigma_{v_r}^2 + \varphi_{k,i}\delta_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_s}^2\sigma_{v_r}^2}{\left(\vartheta_{k,i}\lambda_{k,i}^2 + \sigma_{v_s}^2\right)\left(\varphi_{k,i}\delta_{k,i}^2 + \sigma_{v_r}^2\right)}\right) \tag{4.73}$$

$$\text{s.t.} \quad \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \vartheta_{k,i} \leq P_s \tag{4.74}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \varphi_{k,i} \leq P_r \tag{4.75}$$

$$\vartheta_{k,i} \geq 0,\ \varphi_{k,i} \geq 0, \quad k = 1, ..., K,\ i = 1, ..., N_k, \tag{4.76}$$

We observe now that, with the definitions in (4.71) and (4.72), the power constraints stated in (4.74) and (4.75) are now decoupled making the problem easier to handle. Furthermore, for a given $\boldsymbol{\varphi}$ it can be shown that the problem in (4.73)-(4.76) is a standard convex optimisation problem w.r.t. $\boldsymbol{\vartheta}$. Thus the optimal solution can be derived from the KKT conditions [39] resulting in

$$\vartheta_{k,i} = \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left[\sqrt{\frac{\delta_{k,i}^4\varphi_{k,i}^2}{\sigma_{v_r}^4} + \frac{4\mu_s\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}} - \frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2} - 2\right]^+, \tag{4.77}$$

where μ_s must be calculated to satisfy the power constraint in (4.74) and can be calculated from the non linear equation

$$\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left[\sqrt{\frac{\delta_{k,i}^4\varphi_{k,i}^2}{\sigma_{v_r}^4} + \frac{4\mu_s\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}} - \frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2} - 2\right]^+ \leq P_s. \tag{4.78}$$

We observe that (4.78) is a monotonically increasing function w.r.t. μ_s and therefore μ_s can be computed using the method of bisection with upper and lower bounds on μ_s given by

$$0 \leq \mu_s \leq \left(\frac{P_s + \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2} + 2\right)}{\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\sqrt{\frac{4\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}}\right)}\right)^2. \tag{4.79}$$

Detailed derivations of the solutions for $\vartheta_{k,i}$ are given in (4.77) and the upper and lower limits for μ_s in (4.79) are given in Section 4.6.3 on page 119.

For a given $\boldsymbol{\vartheta}$ the problem in (4.73)-(4.76) is also a standard convex optimisation problem w.r.t. $\boldsymbol{\varphi}$ and the optimal solution can therefore be calculated from the

KKT conditions resulting in

$$\varphi_{k,i} = \frac{\sigma_{v_r}^2}{2\delta_{k,i}^2}\left[\sqrt{\frac{\lambda_{k,i}^4\vartheta_{k,i}^2}{\sigma_{v_s}^4} + \frac{4\mu_r\delta_{k,i}^2\vartheta_{k,i}\lambda_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}} - \frac{\lambda_{k,i}^2\vartheta_{k,i}}{\sigma_{v_s}^2} - 2\right]^+, \tag{4.80}$$

with the Lagrangian multiplier μ_r required to be computed to satisfy

$$\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\frac{\sigma_{v_r}^2}{2\delta_{k,i}^2}\left[\sqrt{\frac{\lambda_{k,i}^4\vartheta_{k,i}^2}{\sigma_{v_s}^4} + \frac{4\mu_r\delta_{k,i}^2\vartheta_{k,i}\lambda_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}} - \frac{\lambda_{k,i}^2\vartheta_{k,i}}{\sigma_{v_s}^2} - 2\right]^+ \leq P_r, \tag{4.81}$$

and is such that the power constraint in (4.75) is satisfied with equality. The value of μ_r can be calculated using the method of bisection with μ_r being bounded by

$$0 \leq \mu_r \leq \left(\frac{P_r + \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\frac{\sigma_{v_r}^2}{2\delta_{k,i}^2}\left(\frac{\lambda_{k,i}^2\vartheta_{k,i}}{\sigma_{v_s}^2} + 2\right)}{\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\frac{\sigma_{v_r}^2}{2\delta_{k,i}^2}\left(\sqrt{\frac{4\delta_{k,i}^2\vartheta_{k,i}\lambda_{k,i}^2}{\sigma_{v_r}^2\sigma_{v_s}^2}}\right)}\right)^2. \tag{4.82}$$

Since the problem in (4.73)-(4.76) is symmetric in the variables $\boldsymbol{\vartheta}$ and $\boldsymbol{\varphi}$, the solutions in (4.80) and (4.82) can be derived in exactly the same manner as those in (4.77) and (4.79), the details of which are provided in Section 4.6.3 on page 119.

The alternating power algorithm now proceeds as follows: The variables $\boldsymbol{\vartheta}$ and $\boldsymbol{\varphi}$ are initialised to satisfy the constraints in (4.74) and (4.75), respectively. Appropriate initialisations are uniform power allocations i.e. $\vartheta_{k,i} = P_s/KN_k$, and $\varphi_{k,i} = P_r/KN_k$. The source power allocation variable $\boldsymbol{\vartheta}$ is then updated with elements given in (4.77) and μ_s being computed from the method of bisection using the bounds in (4.79). With the given $\boldsymbol{\vartheta}$, the elements of $\boldsymbol{\varphi}$ can then be updated using (4.80) with μ_r being calculated from the method of bisection using the bounds in (4.82). This procedure is repeatedly carried out until an appropriate termination criterion is met. Since $\boldsymbol{\vartheta}$ and $\boldsymbol{\varphi}$ are calculated from solving standard convex optimisation problems, the updating of these variables can only decrease or preserve the objective function in (4.73) and therefore convergence of the alternating algorithm is guaranteed [45, 48, 54]. Finally, with $\boldsymbol{\vartheta}$ and $\boldsymbol{\varphi}$, the elements of $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$ can be calculated from (4.71) and (4.72), respectively.

4.4 Simulation Results

In this section we evaluate the performance of the proposed DFE and THP transceiver designs through numerical simulations. In all simulations we consider a

system utilising $N_s = N_r = N_d = 3$ antennas at each node in the network, with the frequency selective MIMO channel impulse responses being modelled according to (2.8) and (2.9) of Section 2.1.3. In all simulations we set $L+1 = 5$ and the spatial correlation matrices are set to identity matrices i.e. we assume no spatial correlation. The elements of $\boldsymbol{\mathcal{H}}_{sw}[l]$ and $\boldsymbol{\mathcal{H}}_{rw}[l]$ in (2.8) and (2.9) are drawn from zero mean Gaussian random distributions with variances $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $0 \leq l \leq L$. OFDM is employed with $K = 32$ subcarriers with each subcarrier being used to transmit $N_k = 3$ 16-QAM data symbols. We define the SNR of the source-relay and relay-destination channels as $\text{SNR}_\text{s} = P_s/K\sigma^2_{v_s}$ and $\text{SNR}_\text{r} = P_r/K\sigma^2_{v_r}$ respectively. All simulation results are obtained from averageing over 500 channel realisations.

4.4.1 Comparison of ZF Transceivers

We firstly compare the performance of the proposed ZF DFE and ZF THP transceiver designs to linear transceivers utilising ZF equalisation. We compare the performance of the proposed solutions to the naive amplify forward (NAF), maximum mutual information (MMI), maximum geometric signal-interference noise ratio (GSINR), minimum arithmetic MSE (AMSE), and minimax-MSE (MM-MSE), designs which are all derived in[4] [45]. Figures 4.1 and 4.2 compare the BER and MSE performances, respectively, of the proposed and benchmark designs. The results are shown against varying SNR_s(dB) with $\text{SNR}_\text{r} = 25$dB. The dashed curves show the performance of the DFE transceivers in the absence of error propagation (referred to as "genie" DFE algorithms). It is evident from Figures 4.1 and 4.2 that the proposed non-linear DFE and THP transceivers with optimal power allocation (OPA) provide a substantial improvement in performance in terms of both BER and MSE when compared to the linear transceiver designs. For completeness we also provide simulation results showing the BER and MSE performance of the proposed and benchmark ZF algorithms against varying SNR_r(dB) with $\text{SNR}_\text{s} = 25$dB. Figures 4.3 and 4.4 show the BER and MSE results, respectively, for this scenario. We again observe that the proposed ZF DFE and ZF THP algorithms provided substantially improved performance compared to the linear benchmarks.

[4]The transceivers in [45] are all derived with a linear MMSE receiver at the destination. These algorithms have been appropriately modified for the case of ZF equalisation at the destination

4.4.2 Comparison of MMSE Transceivers

We now compare the performance of the proposed MMSE DFE and MMSE THP transceivers derived in Section 4.3 against the transceivers in [45] when the linear MMSE receiver is used at the destination. We also compare the performance of the proposed MMSE DFE and THP transceivers that utilise the alternating power allocation (APA) algorithm discussed in Section 4.3.3 with suboptimal MMSE DFE and THP solutions. The suboptimal solutions are similar to the ones derived in Section 4.3 but utilise an equal power allocation (EPA) as opposed to the APA algorithm provided in Section 4.3.3. For the EPA algorithm, the source and relay power allocation matrices are selected to allocate power uniformly across all data streams. Figures 4.5 and 4.6 show the performance of the proposed and benchmark designs against varying SNR_s(dB) with $\text{SNR}_\text{r} = 25$dB. The dashed curves again show the performance of the DFE transceivers in the absence of error propagation (referred to as "genie" DFE algorithms). From Figures 4.5 and 4.6 we see that the non-linear transceivers provide improved performance compared to the linear benchmark designs. Furthermore, the non-linear algorithms that utilise the APA algorithm provide better performance in terms of BER and MSE compared to the non-linear techniques that use the suboptimal EPA algorithm. Figures 4.7 and 4.8 show the corresponding results for varying SNR_r(dB) with $\text{SNR}_\text{s} = 25$dB. Similar trends to the ones demonstrated in Figures 4.5-4.6 can be observed in Figures 4.7 and 4.8 with the proposed algorithms offering improved performance to the linear benchmarks.

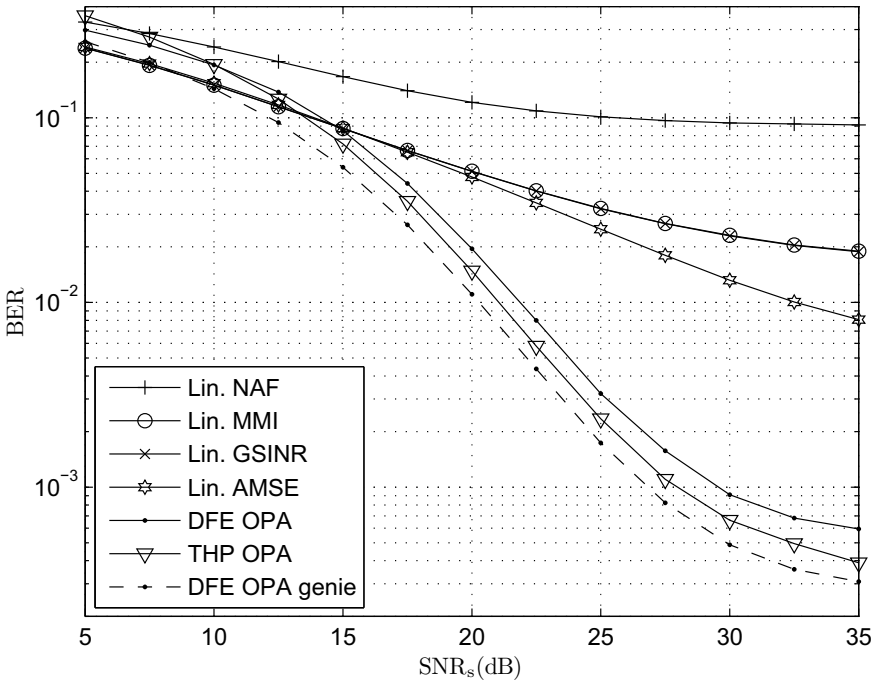

Figure 4.1: BER against $\mathrm{SNR_s(dB)}$ of ZF linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_r} = 25\mathrm{dB}$.

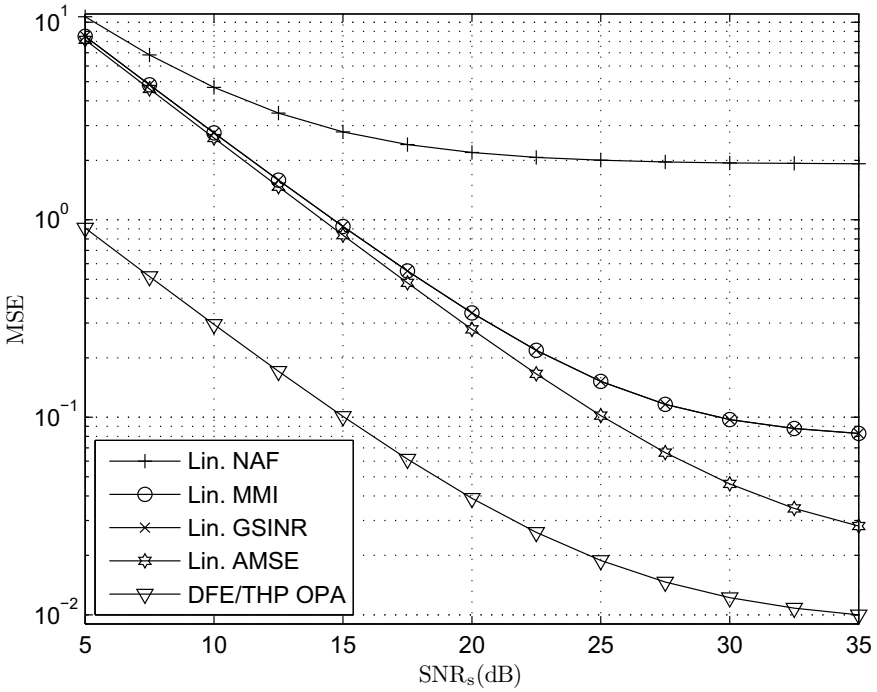

Figure 4.2: MSE against $\mathrm{SNR_s(dB)}$ of ZF linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_r} = 25\mathrm{dB}$.

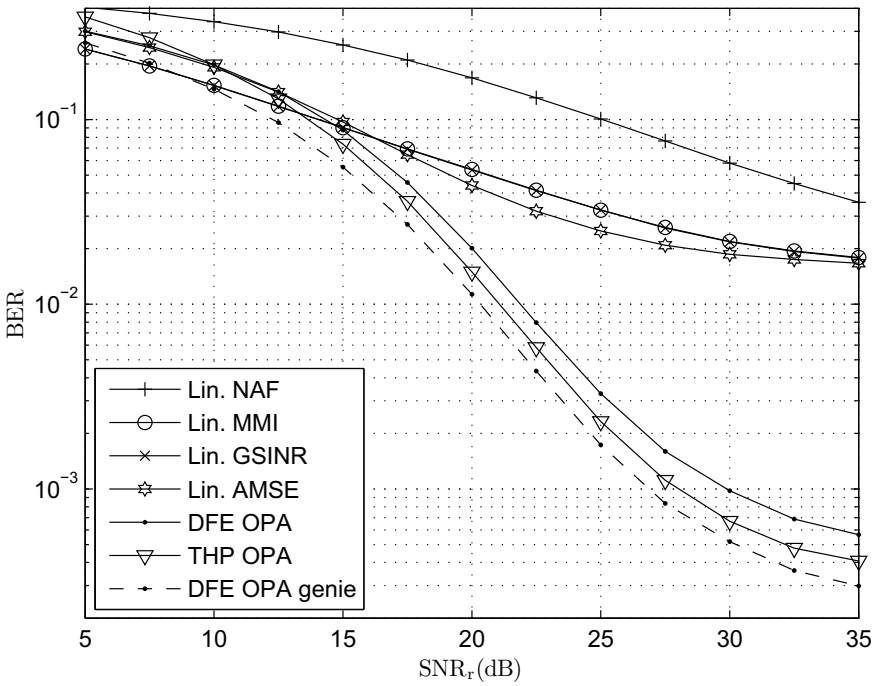

Figure 4.3: BER against $\mathrm{SNR_r(dB)}$ of ZF linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_s} = 25$dB.

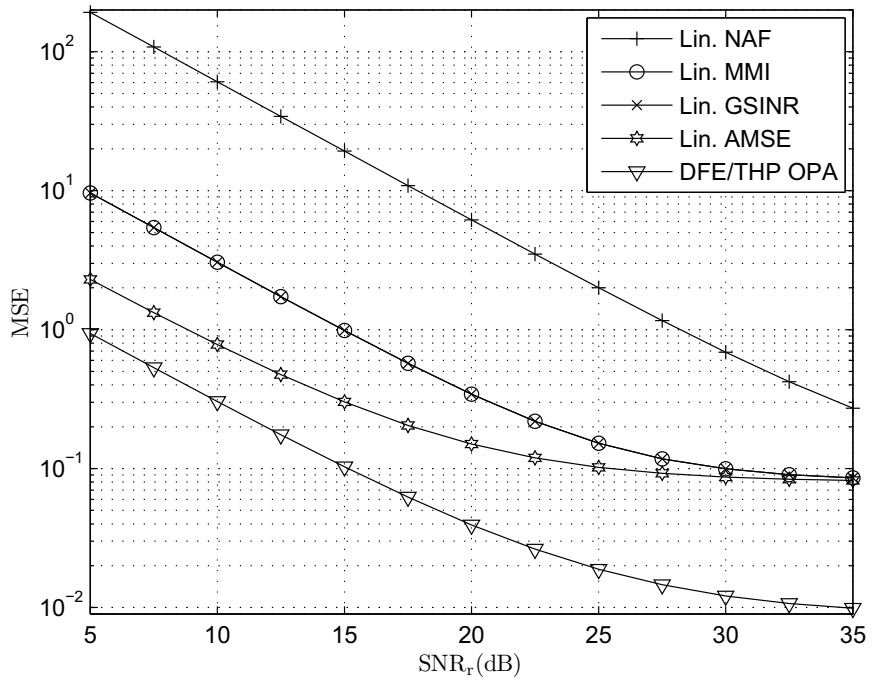

Figure 4.4: MSE against $\mathrm{SNR_r(dB)}$ of ZF linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_s} = 25$dB.

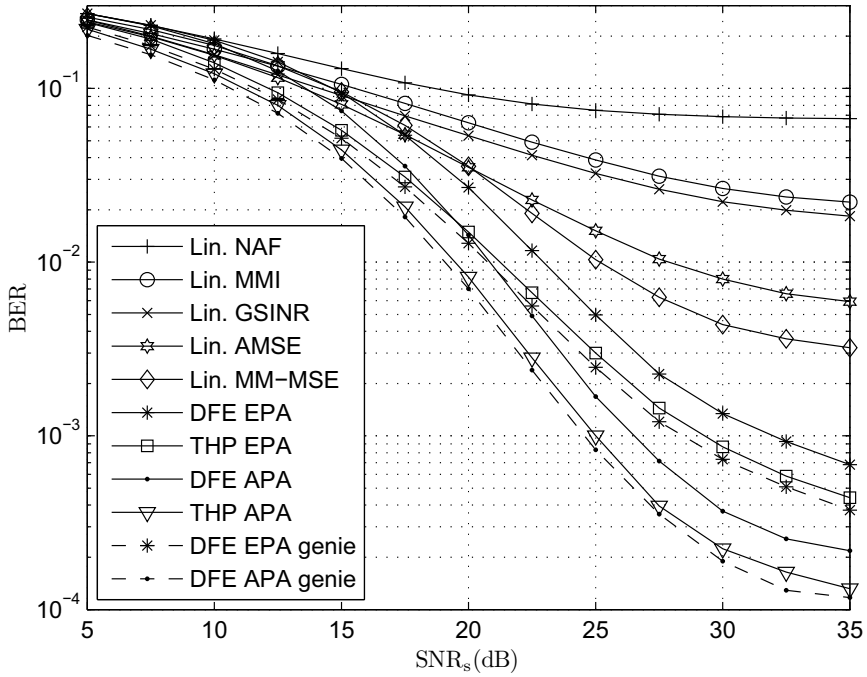

Figure 4.5: BER against $\mathrm{SNR_s(dB)}$ of MMSE linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_r} = 25\mathrm{dB}$.

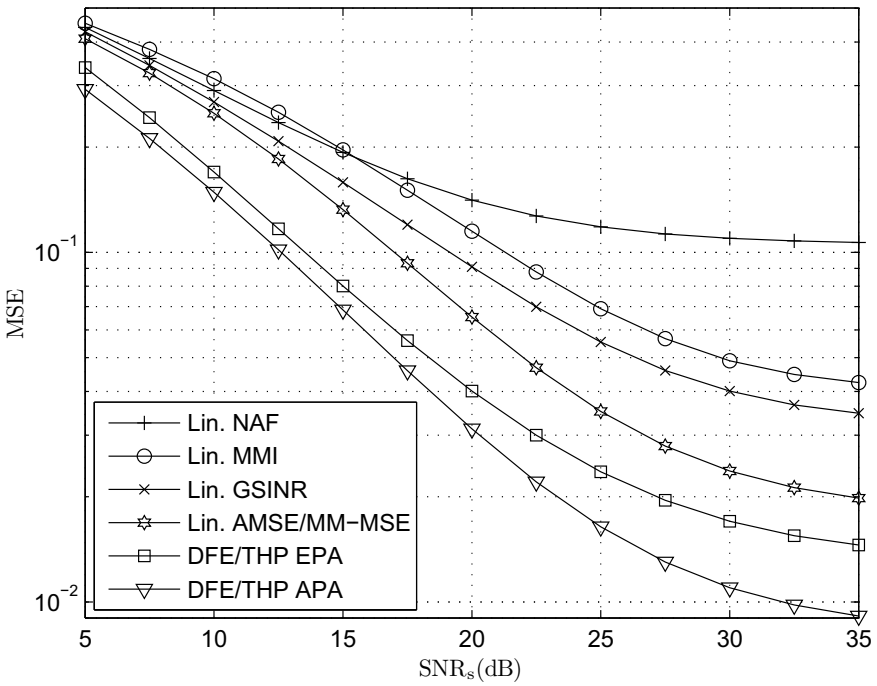

Figure 4.6: MSE against $\mathrm{SNR_s(dB)}$ of MMSE linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_r} = 25\mathrm{dB}$.

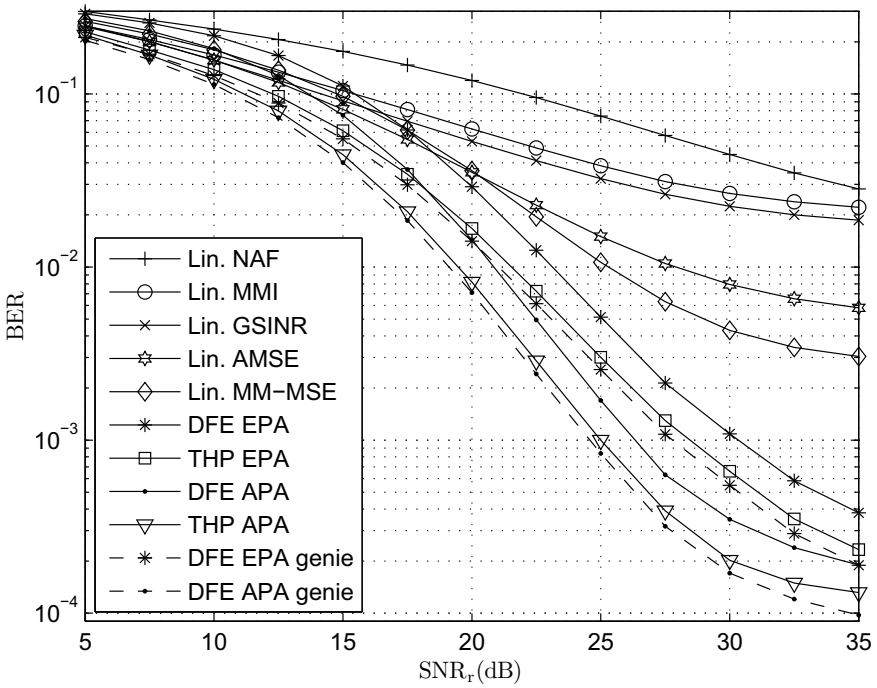

Figure 4.7: BER against $\mathrm{SNR_r(dB)}$ of MMSE linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_s} = 25\mathrm{dB}$.

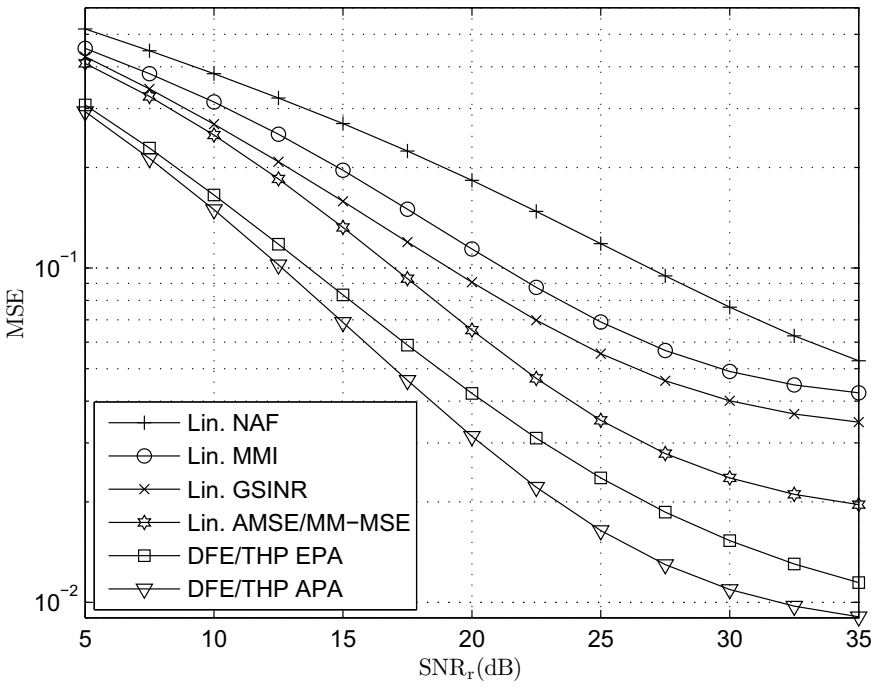

Figure 4.8: MSE against $\mathrm{SNR_r(dB)}$ of MMSE linear and non-linear transceivers for a system with uncorrelated delay paths, $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, and $\mathrm{SNR_s} = 25\mathrm{dB}$.

4.5 Chapter Summary and Conclusions

In this chapter we considered the derivation of the optimal processing matrices for MIMO OFDM relaying systems with DFE or THP. The optimal processors were derived under the assumption that perfect CSI was available to the source, relaying, and destination devices. Such an assumption is reasonable when the SNR during the training/channel estimation phase is high enough to obtain accurate channel estimates. We considered two seperate transceiver designs depending on whether ZF or MMSE equalisation was employed at the destination. For both cases we considered the optimisation of the remaining processors to minimise the arithmetic MSE subject to transmission constraints at the source and relay terminals. The optimal processors were shown to convert the original matrix valued optimisation problem into simpler scalar valued power allocation problems. For the case of ZF equalisation the optimal source and relay power allocation matrices were derived using GP. For the case of MMSE equalisation an alternating power allocation algorithm must be used to optimise the source and relay power allocation matrices. Simulation results demonstrated that the proposed ZF/MMSE DFE and THP algorithms provide improved performance in terms of BER and MSE compared to linear benchmark designs. Although perfect CSI of all channels may be a reasonable assumption in some cases, if the channels are estimated imperfectly then the algorithms proposed in this chapter will suffer a loss in performance. In the next chapter we therefore consider statistically robust DFE and THP transceiver designs that account for imperfect CSI.

4.6 Chapter Derivations and Proofs

This section provides some proofs and derivations which have been omitted from the main text.

4.6.1 Proof of Optimal ZF Source and Relay Precoders

In this section we provide the detailed derivation of the optimal ZF precoders given in (4.25) and (4.26). In order to prove $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ in (4.25) and (4.26) we shall require the following tools from majorisation theory [78]:

Definition 3: For a vector $\boldsymbol{a} \in \mathbb{R}^M$ with elements $\{a_i\}_{i=1}^M \in \mathbb{R}$ then let $\{a_{[i]}\}_{i=1}^M$ denote the elements of $\boldsymbol{a}$ arranged in descending order i.e.

$$a_{[1]} \geq a_{[2]}, ..., \geq a_{[M]}. \tag{4.83}$$

Definition 4: A vector $\boldsymbol{a} \in \mathbb{R}_+^M$ is said to be weakly multiplicatively submajorised by $\boldsymbol{b}$, denoted as $\boldsymbol{a} \preceq_w^{\times} \boldsymbol{b}$, if the following conditions are met

$$\prod_{i=1}^{m} a_{[i]} \leq \prod_{i=1}^{m} b_{[i]}, \quad 1 \leq m \leq M. \tag{4.84}$$

Lemma 9: For a Hermitian positive semi-definite matrix $\boldsymbol{B}$ that is given by $\boldsymbol{B} \triangleq (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)^H (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)$ we have the multiplicative majorisation

$$\left[\lambda_{B,1}, ..., \lambda_{B,M}\right]^T \preceq_w^{\times} \left[\prod_{n=1}^{N} \sigma_{A_n,1}^2, ..., \prod_{n=1}^{N} \sigma_{A_n,M}^2\right]^T \tag{4.85}$$

where $\{\lambda_{B,i}\}_{i=1}^{M} \in \mathbb{R}_{++}$ are the M largest eigenvalues of $\boldsymbol{B}$ (assumed w.l.o.g. to be in decreasing order), $\{\sigma_{A_n,i}\}_{i=1}^{M} \in \mathbb{R}_{++}$ are the M largest singular values of $\boldsymbol{A}_n$ (assumed to be in decreasing order), and we define $M \triangleq \min(\text{rank}\{\boldsymbol{A}_1\}, ..., \text{rank}\{\boldsymbol{A}_N\})$. In otherwords we have

$$\prod_{i=1}^{m} \lambda_{B,i} \leq \prod_{i=1}^{m} \prod_{n=1}^{N} \sigma_{A_n,i}^2, \quad 1 \leq m \leq M \tag{4.86}$$

$$\Leftrightarrow \prod_{i=1}^{m} \lambda_{B,i}^{-1} \geq \left(\prod_{i=1}^{m} \prod_{n=1}^{N} \sigma_{A_n,i}^2\right)^{-1}, \quad 1 \leq m \leq M, \tag{4.87}$$

where the equivalence in (4.87) comes simply by inverting both sides of (4.86).

Lemma 10: For positive semi-definite matrices $\boldsymbol{A} \in \mathbb{C}^{N \times N}$ and $\boldsymbol{B} \in \mathbb{C}^{N \times N}$, with eigenvalues $\{\lambda_{A,i}\}_{i=1}^{N} \in \mathbb{R}_+$ and $\{\lambda_{B,i}\}_{i=1}^{N} \in \mathbb{R}_+$ arranged in descending order, we have the inequality

$$\text{tr}\{\boldsymbol{AB}\} \geq \sum_{i=1}^{N} \lambda_{A,i} \lambda_{B,N+1-i}. \tag{4.88}$$

Note that since $\{\lambda_{B,i}\}_{i=1}^{N}$ are arranged in descending order then $\{\lambda_{B,N+1-i}\}_{i=1}^{N}$ corresponds to the eigenvalues of $\boldsymbol{B}$ arranged in ascending order.

We are now ready to prove the optimal structure of $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ given in Theorem 4. We firstly focus on deriving the structure of $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ that minimise $|\boldsymbol{E}_k|$, and therefore minimise the objective function in (4.20). For convenience we recall that $|\boldsymbol{E}_k|$ is given by

$$|\boldsymbol{E}_k| = \left|\bar{\boldsymbol{F}}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \Big(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d}\Big)^{-1} \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \bar{\boldsymbol{F}}_k\right|^{-1}, \tag{4.89}$$

where we have used the matrices $\boldsymbol{E}_k$ in (4.13) as well as the fact that $|\boldsymbol{A}^{-1}| = |\boldsymbol{A}|^{-1}$. We now introduce the decompositions

$$\boldsymbol{Y}_k \triangleq \boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k = \boldsymbol{U}_{y,k}\boldsymbol{D}_k\boldsymbol{V}_{y,k}^H \tag{4.90}$$

$$\boldsymbol{Z}_k \triangleq \boldsymbol{H}_{r,k}\boldsymbol{G}_k = \boldsymbol{U}_{z,k}\boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H, \tag{4.91}$$

where $\boldsymbol{U}_{y,k} \in \mathbb{C}^{N_r \times N_r}$, $\boldsymbol{V}_{y,k} \in \mathbb{C}^{N_k \times N_k}$, $\boldsymbol{U}_{z,k} \in \mathbb{C}^{N_d \times N_d}$, and $\boldsymbol{V}_{z,k} \in \mathbb{C}^{N_r \times N_r}$ are unitary and the diagonal matrices $\boldsymbol{D}_k \in \mathbb{R}^{N_r \times N_k}$ and $\boldsymbol{\Sigma}_k \in \mathbb{R}^{N_d \times N_r}$ contain the singular values $\{d_{k,i}\}_{i=1}^{N_k}$ and $\{\sigma_{k,i}\}_{i=1}^{\bar{R}_{r,k}}$ of $\boldsymbol{Y}_k$ and $\boldsymbol{Z}_k$ respectively, which are assumed to be arranged in decreasing order. Here we define the variables[5] $\bar{R}_{r,k} \triangleq \min(R_{r,k}, R_{g,k})$. Substituting (4.90) and (4.91) into (4.89) we have the lower bounds

$$|\boldsymbol{E}_k| = \left| \boldsymbol{D}_k^T\boldsymbol{U}_{y,k}^H\boldsymbol{V}_{z,k}\boldsymbol{\Sigma}_k^T \left(\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^T\sigma_{v_s}^2 + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)^{-1} \boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H\boldsymbol{U}_{y,k}\boldsymbol{D}_k \right|^{-1} \tag{4.92}$$

$$\geq \prod_{i=1}^{N_k} \left(\frac{d_{k,i}^2\sigma_{k,i}^2}{\sigma_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_r}^2} \right)^{-1}, \tag{4.93}$$

where Lemma 14 is applied to to (4.92) to obtain the lower bounds in (4.93). We note here that $|\boldsymbol{E}_k|$ in (4.92) is independent of the unitary matrices $\boldsymbol{V}_{y,k}$ and $\boldsymbol{U}_{z,k}$ and that the lower bounds of (4.93) are achieved for any $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. Substituting (4.23) and (4.24) into (4.90) and (4.91), respectively, and solving for $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ we find the precoders that achieve the lower bound in (4.93) are given by

$$\bar{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{U}_{s,k}^H\boldsymbol{U}_{y,k}\boldsymbol{D}_k\boldsymbol{V}_{y,k}^H \tag{4.94}$$

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Delta}_k^{\dagger}\boldsymbol{U}_{r,k}^H\boldsymbol{U}_{z,k}\boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H. \tag{4.95}$$

Any precoders that have the structures given in (4.94) and (4.95) achieve the lower bound in (4.93) provided that $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. We shall select a specific choice of $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ from the family of precoders given by (4.94) and (4.95) that also minimise the source and relay transmission power. Substituting (4.94) into $\mathrm{tr}\{\bar{\boldsymbol{F}}_k\bar{\boldsymbol{F}}_k^H\}$ the power consumed by the source for the kth subcarrier is lower bounded by

$$\mathrm{tr}\left\{\bar{\boldsymbol{F}}_k\bar{\boldsymbol{F}}_k^H\right\} = \mathrm{tr}\left\{\boldsymbol{U}_{y,k}\boldsymbol{U}_{s,k}^H(\boldsymbol{\Lambda}_k^{\dagger})^T\boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{U}_{s,k}^H\boldsymbol{U}_{y,k}\boldsymbol{D}_k\boldsymbol{D}_k^T\right\} \tag{4.96}$$

$$\geq \sum_{i=1}^{N_k} \lambda_{k,i}^{-2}d_{k,i}^2, \tag{4.97}$$

[5]Since we assume that $R_{s,k} \triangleq \mathrm{rank}\{\boldsymbol{H}_{s,k}\} \geq N_k$ and $R_{f,k} \triangleq \mathrm{rank}\{\bar{\boldsymbol{F}}_k\} = N_k$ then we have that $\mathrm{rank}\{\boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k\} = \mathrm{rank}\{\boldsymbol{D}_k\} \leq N_k$. Similarly, since $R_{r,k} \triangleq \mathrm{rank}\{\boldsymbol{H}_{r,k}\}$ and $R_{g,k} \triangleq \mathrm{rank}\{\boldsymbol{G}_k\}$ then $\mathrm{rank}\{\boldsymbol{H}_{r,k}\boldsymbol{G}_k\} = \mathrm{rank}\{\boldsymbol{\Sigma}_k\} \leq \min(R_{r,k}, R_{g,k}) \triangleq \bar{R}_{r,k}$.

where (4.97) is obtained by applying Lemma 10 to (4.96). We see from (4.96) that the source transmit power is independent of $\boldsymbol{V}_{y,k}$. Since (4.92) and (4.93) are also independent of $\boldsymbol{V}_{y,k}$ we can assume w.l.o.g. that $\boldsymbol{V}_{y,k} = \boldsymbol{I}_{N_k}$. We also see that the source transmission power is minimised when $\boldsymbol{U}_{y,k} = \boldsymbol{U}_{s,k}$. Substituting $\boldsymbol{V}_{y,k} = \boldsymbol{I}_{N_k}$ and $\boldsymbol{U}_{y,k} = \boldsymbol{U}_{s,k}$ into (4.94) we have the source precoder structure

$$\bar{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Gamma}_k, \tag{4.98}$$

where we define $\boldsymbol{\Gamma}_k \triangleq \boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{D}_k$. It shall later be shown that $\boldsymbol{\Gamma}_k$ (and therefore $\bar{\boldsymbol{F}}_k$) should be rank N_k.

We now consider the relay transmission power. Using (4.90) and (4.95) we can show from Lemma 10 that the power consumed by the relay is lower bounded by

$$\begin{aligned}
&\mathrm{tr}\Big\{\boldsymbol{G}_k\left(\boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k\bar{\boldsymbol{F}}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)\boldsymbol{G}_k^H\Big\} \\
&= \mathrm{tr}\Big\{\boldsymbol{U}_{z,k}^H\boldsymbol{U}_{r,k}(\boldsymbol{\Delta}_k^{\dagger})^T\boldsymbol{\Delta}_k^{\dagger}\boldsymbol{U}_{r,k}^H\boldsymbol{U}_{z,k}\boldsymbol{\Sigma}_k\left(\boldsymbol{D}_k\boldsymbol{D}_k^T + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)\boldsymbol{\Sigma}_k^T\Big\} \qquad (4.99) \\
&\geq \sum_{i=1}^{N_r}\delta_{k,i}^{-2}\sigma_{k,i}^2\left(d_{k,i}^2 + \sigma_{v_s}^2\right). \qquad (4.100)
\end{aligned}$$

To obtain (4.99) we have used $\boldsymbol{V}_{z,k} = \boldsymbol{U}_{y,k} = \boldsymbol{U}_{s,k}$ and we see from (4.100) that the relay power consumption is minimised with $\boldsymbol{U}_{r,k} = \boldsymbol{U}_{z,k}$. Substituting $\boldsymbol{V}_{z,k} = \boldsymbol{U}_{s,k}$ and $\boldsymbol{U}_{r,k} = \boldsymbol{U}_{z,k}$, into (4.95) and defining $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\Delta}_k^{\dagger}\boldsymbol{\Sigma}_k$, we have

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Phi}_k\boldsymbol{U}_{s,k}^H, \tag{4.101}$$

which is exactly the precoder structure given in (4.25) of Theorem 4.

The last remaining tasks to fully prove Theorem 4 are to show that $\mathrm{rank}\{\boldsymbol{\Gamma}_k\} = N_k$ and $\mathrm{rank}\{\boldsymbol{\Phi}_k\} = N_k$. To do so we firstly note that due to the fact that $\boldsymbol{\Gamma}_k$ are $N_s \times N_k$ dimensional matrices they must immediately satisfy $\mathrm{rank}\{\boldsymbol{\Gamma}_k\} \leq \min(N_s, N_k) \leq N_k$. However in order for the ZF equaliser in (4.3) to exist (see the rank constraint in (4.4)) we must have $\mathrm{rank}\{\bar{\boldsymbol{F}}_k\} = \mathrm{rank}\{\boldsymbol{\Gamma}_k\} = N_k$.

We now consider the matrices $\boldsymbol{\Phi}_k$ and note that since they are of dimensions $N_r \times N_r$ they instantly satisfy $\mathrm{rank}\{\boldsymbol{\Phi}_k\} \leq N_r$. Noting now that the source and relay structures in (4.98) and (4.101) result in (4.93) holding with equality then by substituting $\boldsymbol{\Gamma}_k = \boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{D}_k$ and $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\Delta}_k^{\dagger}\boldsymbol{\Sigma}_k$ into (4.92) we obtain

$$|\boldsymbol{E}_k| = \prod_{i=1}^{N_k}\left(\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_r}^2}\right)^{-1}. \tag{4.102}$$

We observe from (4.102) that $|\boldsymbol{E}_k|$ only depends on $\{\phi_{k,i}\}_{i=1}^{N_k}$ and therefore $\boldsymbol{\Phi}_k$ should be at most rank N_k. This is due to the fact that if we have any $\{\phi_{k,i} > 0\}_{i=N_k+1}^{N_r}$ then we can find a new relay precoder that achieves the same value of the objective function but with reduced transmission power simply by setting $\{\phi_{k,i} = 0\}_{i=N_k+1}^{N_r}$. If $\text{rank}\{\boldsymbol{G}_k\} = \text{rank}\{\boldsymbol{\Phi}_k\} < N_k$ then the equaliser in (4.3) does not exist and we therefore must have $\text{rank}\{\boldsymbol{G}_k\} = \text{rank}\{\boldsymbol{\Phi}_k\} = N_k$.

4.6.2 Proof of Optimal MMSE Source and Relay Precoders

In this section we provide the derivation of the optimal MMSE precoders given in (4.64) and (4.65) of Theorem 5. In order to prove $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ in (4.64) and (4.65) we shall require the following tools from majorisation theory [78] (we note that some of the following tools were introduced in Section 4.6.1 but are repeated here for convenience):

Definition 5: For a vector $\boldsymbol{a} \in \mathbb{R}^M$ with elements $\{a_i\}_{i=1}^M \in \mathbb{R}$, then let $\{a_{[i]}\}_{i=1}^M$ denote the elements of $\boldsymbol{a}$ arranged in descending order i.e.

$$a_{[1]} \geq a_{[2]}, ..., \geq a_{[M]}. \tag{4.103}$$

Definition 6: A vector $\boldsymbol{a} \in \mathbb{R}^M$ is said to be weakly additively submajorised by $\boldsymbol{b} \in \mathbb{R}^M$, denoted as $\boldsymbol{a} \preceq_w^+ \boldsymbol{b}$, if

$$\sum_{i=1}^{m} a_{[i]} \leq \sum_{i=1}^{m} b_{[i]}, \quad 1 \leq m \leq M. \tag{4.104}$$

Definition 7: A vector $\boldsymbol{a} \in \mathbb{R}_+^M$ is said to be weakly multiplicatively submajorised by $\boldsymbol{b}$, denoted as $\boldsymbol{a} \preceq_w^\times \boldsymbol{b}$, if the following conditions are met

$$\prod_{i=1}^{m} a_{[i]} \leq \prod_{i=1}^{m} b_{[i]}, \quad 1 \leq m \leq M. \tag{4.105}$$

Lemma 11: For any convex increasing function $g(.)$ then given any vectors satisfying $\boldsymbol{a} \preceq_w^+ \boldsymbol{b}$ we have $[g(a_1), ..., g(a_M)]^T \preceq_w^+ [g(b_1), ..., g(b_M)]^T$ i.e.

$$\sum_{i=1}^{m} g\left(a_{[i]}\right) \leq \sum_{i=1}^{m} g\left(b_{[i]}\right), \quad 1 \leq m \leq M. \tag{4.106}$$

Lemma 12: For vectors $\boldsymbol{a} \in \mathbb{R}_{++}^{M}$ and $\boldsymbol{b} \in \mathbb{R}_{++}^{M}$ then $\boldsymbol{a} \preceq_{w}^{\times} \boldsymbol{b}$ is equivalent to $[\log(a_1), ..., \log(a_M)]^T \preceq_{w}^{+} [\log(b_1), ..., \log(b_M)]^T$ i.e. $\boldsymbol{a}$ and $\boldsymbol{b}$ satisfy

$$\sum_{i=1}^{m} \log\left(a_{[i]}\right) \leq \sum_{i=1}^{m} \log\left(b_{[i]}\right), \quad 1 \leq m \leq M. \tag{4.107}$$

Lemma 13: For a Hermitian positive semi-definite matrix $\boldsymbol{B}$ that is given by $\boldsymbol{B} \triangleq (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)^H (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)$ we have the multiplicative majorisation

$$\left[\lambda_{B,1}, ..., \lambda_{B,M}\right]^T \preceq_{w}^{\times} \left[\prod_{n=1}^{N} \sigma_{A_n,1}^2, ..., \prod_{n=1}^{N} \sigma_{A_n,M}^2\right]^T \tag{4.108}$$

where $\{\lambda_{B,i}\}_{i=1}^{M} \in \mathbb{R}_{+}$ are the M largest eigenvalue of $\boldsymbol{B}$, $\{\sigma_{A_n,i}\}_{i=1}^{M} \in \mathbb{R}_{++}$ are the M largest singular value of $\boldsymbol{A}_n$, and we define $M \triangleq \min(\text{rank}\{\boldsymbol{A}_1\}, ..., \text{rank}\{\boldsymbol{A}_N\})$. In other words we have

$$\prod_{i=1}^{m} \lambda_{B,i} \leq \prod_{i=1}^{m} \prod_{n=1}^{N} \sigma_{A_n,i}^2, \quad 1 \leq m \leq M. \tag{4.109}$$

Lemma 14: For $\boldsymbol{B} \triangleq (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)^H (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)$ we have the inequality

$$|\boldsymbol{I} + \boldsymbol{B}|^{-1} \geq \prod_{i=1}^{M} \left(1 + \prod_{n=1}^{N} \sigma_{A_n,i}^2\right)^{-1}. \tag{4.110}$$

Proof: To prove the inequality in Lemma 14 we firstly note that $\boldsymbol{B}$ is Hermitian and at most rank M. It can then be straightforwardly shown that proving (4.110) is equivalent to proving that

$$\prod_{i=1}^{M} \left(1 + \lambda_{B,i}\right) \leq \prod_{i=1}^{M} \left(1 + \prod_{n=1}^{N} \sigma_{A_n,i}^2\right). \tag{4.111}$$

Furthermore, since $f(x) = \log(x)$ is a strictly increasing function for $x \in \mathbb{R}_{++}$ then proving (4.111) is equivalent to showing that

$$\sum_{i=1}^{M} \log\left(1 + e^{b_i}\right) \leq \sum_{i=1}^{M} \log\left(1 + e^{a_i}\right), \tag{4.112}$$

where we define $a_i \triangleq \log(\prod_{n=1}^{N} \sigma_{A_n,i}^2)$ and $b_i \triangleq \log(\lambda_{B,i})$. Thus proving (4.110) is equivalent to proving (4.112). Since $f(x) = \log(1 + e^x)$ is an increasing convex function, then from Lemma 11 we know the inequality in (4.112) holds provided

that

$$\left[\log\left(\lambda_{B,1}\right),...,\log\left(\lambda_{B,M}\right)\right]^T \preceq_w^+ \left[\log\left(\prod_{n=1}^{N}\sigma_{A_n,1}^2\right),...,\log\left(\prod_{n=1}^{N}\sigma_{A_n,M}^2\right)\right]^T, \tag{4.113}$$

which from Lemma 12 is equivalent to

$$\left[\lambda_{B,1},...,\lambda_{B,M}\right]^T \preceq_w^{\times} \left[\prod_{n=1}^{N}\sigma_{A_n,1}^2,...,\prod_{n=1}^{N}\sigma_{A_n,M}^2\right]^T. \tag{4.114}$$

We know that the previous weak multiplicative majorisation is true from Lemma 13 and we have therefore proven the inequality in (4.112), which as previously mentioned, is equivalent to proving (4.110)[6]. □

Lemma 15: For positive semi-definite matrices $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{B} \in \mathbb{C}^{N\times N}$, with eigenvalues $\{\lambda_{A,i}\}_{i=1}^N \in \mathbb{R}_+$ and $\{\lambda_{B,i}\}_{i=1}^N \in \mathbb{R}_+$ arranged in descending order, we have the inequality

$$\text{tr}\{\boldsymbol{AB}\} \geq \sum_{i=1}^{N}\lambda_{A,i}\lambda_{B,N+1-i}. \tag{4.115}$$

We now prove the optimal structure of the source and relay precoding matrices $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ given in Theorem 5. We firstly focus on deriving the structure of $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ that minimise $|\boldsymbol{E}_k|$, with $\boldsymbol{E}_k$ given in (4.57), and therefore minimise the objective function in (4.61). For convenience we recall that

$$|\boldsymbol{E}_k| \triangleq \left|\boldsymbol{I}_{N_k} + \bar{\boldsymbol{F}}_k^H\boldsymbol{H}_{s,k}^H\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H\Big(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H\sigma_{v_s}^2 + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\Big)^{-1}\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k\right|^{-1}. \tag{4.116}$$

Let us now introduce the decompositions

$$\boldsymbol{Y}_k \triangleq \boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k = \boldsymbol{U}_{y,k}\boldsymbol{D}_k\boldsymbol{V}_{y,k}^H \tag{4.117}$$

$$\boldsymbol{Z}_k \triangleq \boldsymbol{H}_{r,k}\boldsymbol{G}_k = \boldsymbol{U}_{z,k}\boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H, \tag{4.118}$$

where $\boldsymbol{U}_{y,k} \in \mathbb{C}^{N_r\times N_r}$, $\boldsymbol{V}_{y,k} \in \mathbb{C}^{N_k\times N_k}$, $\boldsymbol{U}_{z,k} \in \mathbb{C}^{N_d\times N_d}$, and $\boldsymbol{V}_{z,k} \in \mathbb{C}^{N_r\times N_r}$ are unitary and the upper left portions of the diagonal matrices $\boldsymbol{D}_k \in \mathbb{R}^{N_r\times N_k}$ and $\boldsymbol{\Sigma}_k \in \mathbb{R}^{N_d\times N_r}$ contain the singular values $\{d_{k,i}\}_{i=1}^{\bar{R}_s}$ and $\{\sigma_{k,i}\}_{i=1}^{R_r}$, respectively,

[6] In actual fact we have proven that $\prod_{i=1}^{m}\left(1+\lambda_{B,i}\right) \leq \prod_{i=1}^{m}\left(1+\prod_{n=1}^{N}\sigma_{A_n,i}^2\right)$, $1 \leq m \leq M$. However for our purposes we need only consider the case $m = M$

which are assumed to be in decreasing order. Here we define the variable[7] $\bar{R}_{s,k} \triangleq \min(N_k, R_{s,k})$. Substituting (4.117) and (4.118) into (4.116) we have

$$\begin{aligned}|\boldsymbol{E}_k| &= \left|\boldsymbol{I}_{N_k} + \boldsymbol{Y}_k^H \boldsymbol{Z}_k^H \left(\boldsymbol{Z}_k \boldsymbol{Z}_k^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_r}\right)^{-1} \boldsymbol{Z}_k \boldsymbol{Y}_k\right|^{-1} \\ &= \left|\boldsymbol{I}_{N_k} + \boldsymbol{D}_k^T \boldsymbol{U}_{y,k}^H \boldsymbol{V}_{z,k} \boldsymbol{\Sigma}_k^T \left(\boldsymbol{\Sigma}_k \boldsymbol{\Sigma}_k^T \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_r}\right)^{-1} \boldsymbol{\Sigma}_k \boldsymbol{V}_{z,k}^H \boldsymbol{U}_{y,k} \boldsymbol{D}_k\right|^{-1},\end{aligned} \tag{4.119}$$

We note here that $|\boldsymbol{E}_k|$ in (4.119) is independent of the unitary matrices $\boldsymbol{V}_{y,k}$ and $\boldsymbol{U}_{z,k}$. It is also easily verified that

$$\operatorname{rank}\left\{\boldsymbol{\Sigma}_k^T \left(\boldsymbol{\Sigma}_k \boldsymbol{\Sigma}_k^T \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_r}\right)^{-1} \boldsymbol{\Sigma}_k\right\} \leq \min(N_k, R_{s,k}, R_{r,k}) \triangleq \bar{N}_k. \tag{4.120}$$

With these observations then applying Lemma 14 to (4.119) we obtain

$$|\boldsymbol{E}_k| \geq \prod_{i=1}^{\bar{N}_k} \left(1 + \frac{d_{k,i}^2 \sigma_{k,i}^2}{\sigma_{k,i}^2 \sigma_{v_s}^2 + \sigma_{v_r}^2}\right)^{-1}. \tag{4.121}$$

The lower bound of (4.121) is achieved for any $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. Substituting (4.23) and (4.24) into (4.117) and (4.118), respectively, and solving for $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$, we find the precoders that achieve the lower bound in (4.121) are given by

$$\bar{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k} \boldsymbol{\Lambda}_k^{\dagger} \boldsymbol{U}_{s,k}^H \boldsymbol{U}_{y,k} \boldsymbol{D}_k \boldsymbol{V}_{y,k}^H \tag{4.122}$$

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k} \boldsymbol{\Delta}_k^{\dagger} \boldsymbol{U}_{r,k}^H \boldsymbol{U}_{z,k} \boldsymbol{\Sigma}_k \boldsymbol{V}_{z,k}^H, \tag{4.123}$$

Any precoders that have the structures given in (4.122) and (4.123) achieve the lower bound in (4.121) provided that $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. We shall select a specific choice of $\bar{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ from the family of precoders given by (4.122) and (4.123) that also minimise the source and relay transmission power. Substituting (4.122) into $\operatorname{tr}\{\bar{\boldsymbol{F}}_k \bar{\boldsymbol{F}}_k^H\}$, the power consumed by the source is lower bounded by

$$\operatorname{tr}\left\{\bar{\boldsymbol{F}}_k \bar{\boldsymbol{F}}_k^H\right\} = \operatorname{tr}\left\{\boldsymbol{U}_{y,k}^H \boldsymbol{U}_{s,k} (\boldsymbol{\Lambda}_k^{\dagger})^T \boldsymbol{\Lambda}_k^{\dagger} \boldsymbol{U}_{s,k}^H \boldsymbol{U}_{y,k} \boldsymbol{D}_k \boldsymbol{D}_k^T\right\} \tag{4.124}$$

$$\geq \sum_{i=1}^{\bar{R}_s} \lambda_{k,i}^{-2} d_{k,i}^2, \tag{4.125}$$

[7] Since $\operatorname{rank}\{\boldsymbol{H}_{s,k}\} = R_{s,k} \leq \min(N_s, N_r)$ and $\operatorname{rank}\{\bar{\boldsymbol{F}}_k\} \leq \min(N_k, N_s)$ then we have that $\operatorname{rank}\{\boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k\} = \operatorname{rank}\{\boldsymbol{D}_k\} \leq \min(N_k, R_{s,k}) \triangleq \bar{R}_{s,k}$. Similarly, since $\operatorname{rank}\{\boldsymbol{H}_{r,k}\} = R_{r,k} \leq \min(N_d, N_r)$ and $\operatorname{rank}\{\boldsymbol{G}_k\} \leq N_r$ then $\operatorname{rank}\{\boldsymbol{H}_{r,k}\boldsymbol{G}_k\} = \operatorname{rank}\{\boldsymbol{\Sigma}_k\} \leq R_{r,k}$.

where (4.125) is obtained by applying Lemma 15 to (4.124). We see from (4.124) that the source transmit power is independent of $\boldsymbol{V}_{y,k}$. Since (4.119) is also independent of $\boldsymbol{V}_{y,k}$ we can assume w.l.o.g. that $\boldsymbol{V}_{y,k} = \boldsymbol{I}_{N_k}$. We also see that the source transmission power is minimised when $\boldsymbol{U}_{s,k} = \boldsymbol{U}_{y,k}$. Substituting $\boldsymbol{V}_{y,k} = \boldsymbol{I}_{N_k}$ and $\boldsymbol{U}_{s,k} = \boldsymbol{U}_{y,k}$ into (4.122) we have the source precoder structures

$$\bar{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Gamma}_k, \tag{4.126}$$

where we define $\boldsymbol{\Gamma}_k \triangleq \boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{D}_k$. This is exactly the precoder structures given in (4.64) of Theorem 5. It shall be shown later that $\boldsymbol{\Gamma}_k$ should be at most rank $\bar{N}_k$. We now consider the relay transmission power. Using (4.117) and (4.123) it can be shown from Lemma 15 that the power consumed by the relay is lower bounded by

$$\begin{aligned} &\mathrm{tr}\Big\{\boldsymbol{G}_k\left(\boldsymbol{H}_{s,k}\bar{\boldsymbol{F}}_k\bar{\boldsymbol{F}}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)\boldsymbol{G}_k^H\Big\} \\ =\ &\mathrm{tr}\Big\{\boldsymbol{U}_{z,k}^H\boldsymbol{U}_{r,k}(\boldsymbol{\Delta}_k^{\dagger})^T\boldsymbol{\Delta}_k^{\dagger}\boldsymbol{U}_{r,k}^H\boldsymbol{U}_{z,k}\boldsymbol{\Sigma}_k\left(\boldsymbol{D}_k\boldsymbol{D}_k^T + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)\boldsymbol{\Sigma}_k^T\Big\} \qquad (4.127) \\ \geq\ &\sum_{i=1}^{N_r}\delta_{k,i}^{-2}\sigma_{k,i}^2\left(d_{k,i}^2 + \sigma_{v_s}^2\right). \qquad (4.128) \end{aligned}$$

To obtain (4.127) we have used $\boldsymbol{V}_{z,k} = \boldsymbol{U}_{y,k} = \boldsymbol{U}_{s,k}$ and we see from (4.128) that the relay power consumption is minimised with $\boldsymbol{U}_{z,k} = \boldsymbol{U}_{r,k}$. Substituting $\boldsymbol{V}_{z,k} = \boldsymbol{U}_{s,k}$ and $\boldsymbol{U}_{z,k} = \boldsymbol{U}_{r,k}$ into (4.123) and defining $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\Delta}_k^{\dagger}\boldsymbol{\Sigma}_k$ we have

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Phi}_k\boldsymbol{U}_{s,k}^H, \tag{4.129}$$

as given in (4.64) of Theorem 5. The last remaining tasks to fully prove Theorem 5 are to show that $\boldsymbol{\Gamma}_k$ and $\boldsymbol{\Phi}_k$ are at most rank $\bar{N}_k$. By substituting $\boldsymbol{\Gamma}_k = \boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{D}$ and $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\Delta}_k^{\dagger}\boldsymbol{\Sigma}_k$ into (4.121) we can obtain

$$|\boldsymbol{E}_k| = \prod_{i=1}^{\bar{N}_k}\left(1 + \frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_r}^2}\right)^{-1}. \tag{4.130}$$

To find the source and relay precoders that minimise (4.130) whilst utilising the minimum transmit power we can assume that $\{\gamma_{k,i} = 0\}_{i=\bar{N}_k+1}^{N_k}$ and $\{\phi_{k,i} = 0\}_{i=\bar{N}_k+1}^{N_r}$ since from (4.130) it is easily seen that these values do not affect the objective function. This shows that $\mathrm{rank}\{\boldsymbol{\Gamma}_k\} = \bar{N}_k$ and $\mathrm{rank}\{\boldsymbol{\Phi}_k\} = \bar{N}_k$ and concludes the proof of Theorem 5.

4.6.3 Proof of MMSE Power Allocation

In this section we prove the MMSE power allocation matrices given in (4.77) and (4.77) before showing the upper and lower limits for μ_s and μ_r given in (4.79) and (4.82). We firstly note that the problem in (4.73)-(4.76) is symmetric in $\boldsymbol{\vartheta}$ and $\boldsymbol{\varphi}$, as indeed are the solutions. Therefore, for brevity, we only provide the proofs of (4.77) and (4.79) as an example. The following analysis can be straightforwardly modified to prove (4.80) and (4.82).

For set $\boldsymbol{\varphi}$ we consider the Lagrangian function associated with (4.73)-(4.76) which is given by

$$\begin{aligned}\mathcal{L} = &\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \log\left(\frac{\vartheta_{k,i}\lambda_{k,i}^2\sigma_{v_r}^2 + \varphi_{k,i}\delta_{k,i}^2\sigma_{v_s}^2 + \sigma_{v_s}^2\sigma_{v_r}^2}{\left(\vartheta_{k,i}\lambda_{k,i}^2 + \sigma_{v_s}^2\right)\left(\varphi_{k,i}\delta_{k,i}^2 + \sigma_{v_r}^2\right)}\right) \\ &+ \mu_s\left(\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\vartheta_{k,i} - P_s\right) - \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\upsilon_{k,i}\vartheta_{k,i},\end{aligned} \tag{4.131}$$

The following KKT conditions are sufficient and necessary for optimality

$$-\frac{\varphi_{k,i}\delta_{k,i}^2\lambda_{k,i}^2\sigma_{v_s}^2(\varphi_{k,i}\delta_{k,i}^2 + \sigma_{v_r}^2)}{\alpha_{k,i}\beta_{k,i}} + \mu_s - \upsilon_{k,i} = 0, \tag{4.132}$$

$$\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\vartheta_{k,i} - P_s \leq 0 \tag{4.133}$$

$$\vartheta_{k,i} \geq 0,\ \mu_s \geq 0,\ \upsilon_{k,i} \geq 0, \tag{4.134}$$

$$\mu_s\left(\sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k}\vartheta_{k,i} - P_s\right) = 0 \tag{4.135}$$

$$\upsilon_{k,i}\vartheta_{k,i} = 0, \tag{4.136}$$

where for notational convenience we define

$$\alpha_{k,i} \triangleq \varphi_{k,i}\delta_{k,i}^2\vartheta_{k,i}\lambda_{k,i}^2 + \varphi_{k,i}\delta_{k,i}^2\sigma_{v_s}^2 + \vartheta_{k,i}\lambda_{k,i}^2\sigma_{v_r}^2 + \sigma_{v_s}^2\sigma_{v_r}^2, \tag{4.137}$$

$$\beta_{k,i} \triangleq \varphi_{k,i}\delta_{k,i}^2\sigma_{v_s}^2 + \vartheta_{k,i}\lambda_{k,i}^2\sigma_{v_r}^2 + \sigma_{v_s}^2\sigma_{v_r}^2. \tag{4.138}$$

Solving (4.132) for $\upsilon_{k,i}$, and substituting the resulting expressions into (4.136) and $\upsilon_{k,i} \geq 0$, from (4.134) we have the conditions

$$\left(-\frac{\varphi_{k,i}\delta_{k,i}^2\lambda_{k,i}^2\sigma_{v_s}^2(\varphi_{k,i}\delta_{k,i}^2+\sigma_{v_r}^2)}{\alpha_{k,i}\beta_{k,i}} + \mu_s \right) \vartheta_{k,i} = 0, \tag{4.139}$$

$$\mu_s \geq \frac{\varphi_{k,i}\delta_{k,i}^2\lambda_{k,i}^2\sigma_{v_s}^2(\varphi_{k,i}\delta_{k,i}^2+\sigma_{v_r}^2)}{\alpha_{k,i}\beta_{k,i}}. \tag{4.140}$$

The solutions for $\varphi_{k,i}$ can be found based on the conditions in (4.139) and (4.140) and can be partitioned into two cases. We firstly notice that if

$$\mu_s \geq \frac{\varphi_{k,i}\delta_{k,i}^2\lambda_{k,i}^2}{\sigma_{v_s}^2(\varphi_{k,i}\delta_i^2+\sigma_{v_r}^2)} \tag{4.141}$$

then we find that (4.139) only holds when we have $\vartheta_{k,i} = 0$ and we can write

$$\vartheta_{k,i} = 0, \quad \forall k, \forall i \mid \mu_s \geq \frac{\varphi_{k,i}\delta_{k,i}^2\lambda_{k,i}^2}{\sigma_{v_s}^2(\varphi_{k,i}\delta_{k,i}^2+\sigma_{v_r}^2)} \tag{4.142}$$

If on the other hand

$$\mu_s < \frac{\vartheta_{k,i}\delta_{k,i}^2\lambda_{k,i}^2}{\sigma_{v_s}^2(\varphi_{k,i}\delta_{k,i}^2+\sigma_{v_r}^2)} \tag{4.143}$$

then we find that (4.140) only holds when $\vartheta_{k,i} > 0$. In this case, in order for (4.139) to hold, the bracketed term must be zero. Setting the bracketed term in (4.139) to zero and solving for $\vartheta_{k,i}$ results in

$$\begin{aligned} \vartheta_{k,i} &= \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\sqrt{\frac{\delta_{k,i}^4\varphi_{k,i}^2}{\sigma_{v_r}^4} + \frac{4\mu_s\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}} - \frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2} - 2 \right), \\ &\forall k, \forall i \mid \mu_s < \frac{\vartheta_{k,i}\delta_{k,i}^2\lambda_{k,i}^2}{\sigma_{v_s}^2(\varphi_{k,i}\delta_{k,i}^2+\sigma_{v_r}^2)}. \end{aligned} \tag{4.144}$$

The two cases specified in (4.142) and (4.144) can be written more compactly as in the solution in (4.77).

We now prove the bounds for μ_s in (4.79) that should be used for the method of bisection. The lower bound in (4.79) comes directly from the KKT condition in (4.134). To prove the upper bound in (4.79) it is straightforward to show from

(4.78) that we have the following inequalities

$$P_s \geq \sum_{i=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left[\sqrt{\frac{\delta_{k,i}^4\varphi_{k,i}^2}{\sigma_{v_r}^4}+\frac{4\mu_s\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}}-\frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2}-2\right]^+ \tag{4.145}$$

$$\geq \sum_{i=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\sqrt{\frac{\delta_{k,i}^4\varphi_{k,i}^2}{\sigma_{v_r}^4}+\frac{4\mu_s\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}}-\frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2}-2\right) \tag{4.146}$$

$$\geq \sum_{i=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\sqrt{\frac{4\mu_s\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}}-\frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2}-2\right) \tag{4.147}$$

$$\geq \sqrt{\mu_s}\sum_{i=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\sqrt{\frac{4\lambda_{k,i}^2\varphi_{k,i}\delta_{k,i}^2}{\sigma_{v_s}^2\sigma_{v_r}^2}}\right)-\sum_{i=1}^{K}\sum_{i=1}^{\bar{N}_k} \frac{\sigma_{v_s}^2}{2\lambda_{k,i}^2}\left(\frac{\delta_{k,i}^2\varphi_{k,i}}{\sigma_{v_r}^2}+2\right) \tag{4.148}$$

Solving (4.148) for μ_s results in the upper bound given in (4.79).

Chapter 5

Robust Non-linear Transceiver Designs with Imperfect CSI

5.1 Introduction

In the previous chapter we considered DFE and THP transceiver designs for two-hop MIMO OFDM relaying systems under the assumption that perfect CSI of the source-relay and relay-destination channels could be acquired at all nodes. In practical systems perfect estimation of the communication channel is generally very difficult to achieve and channel estimation errors inevitably occur. Sources of channel estimation errors include, for example, limited length training sequences, noisy pilot/training symbols used for channel estimation, or an insufficient number of training intervals in a time varying wireless environment. A simple yet possibly naive approach in dealing with such imperfect CSI is to simply regard the estimated channel as being the actual channel and design the precoding and equalisation matrices according to algorithms that have been developed under the assumption of perfect CSI. For example, the full CSI algorithms discussed in Chapter 4 could be used to optimise the processors but with the actual channel matrices being replaced with their estimated counterparts. If the channel estimation errors are sufficiently small, i.e. the estimated channel is a good approximation of the actual channel, then this approach may work well in practice. However, for larger channel estimation errors such a naive approach can exhibit a serious degradation in performance caused by the mismatch between the true and estimated channels. In such a scenario it is more appropriate to consider robust transceiver designs that take into account the channel estimation errors in the design of the transmit and receive processors.

In general there are two main philosophies to the approach of robust transceiver designs that depend on the manner in which the channel estimation errors are modelled. Generally speaking, robust transceivers can be classified as being either worst-case designs which are also commonly referred to as maximin approaches (see e.g.[87–89]), or stochastic designs which are otherwise known as Bayesian techniques. In worst-case approaches the channel estimation errors are considered to be bounded in some manner (usually according to some norm constraint on the estimation error) such that any specific channel realisation falls within a known and bounded uncertainty region. The transceiver processors are then designed to guarantee a minimum level of system performance for any specific channel realisation that falls within the considered uncertainty region. In other words the transceiver is optimised for the worst-case channel realisation. Worst-case approaches are useful for quality of service (QoS) constrained designs since for any channel condition a guaranteed QoS can always be achieved. However such approaches tend to lead to rather pessimistic designs. On the other hand, Bayesian approaches consider the optimisation of some objective function that is averaged over the channel estimation error and therefore generally lead to improved performance.

The Bayesian philosophy has been extensively adopted for the design of robust transceivers in narrowband MIMO relaying systems in [61, 62, 81–83, 90–94]. The robust joint linear design of the destination equaliser and relay precoder is considered in [81–83, 90] for minimising the arithmetic MSE subject to the transmission power constraint at the relay terminal. As discussed in these works the optimal solution is difficult to obtain for general channel estimation errors due to the non-convexity of the optimisation problem. In [81–83] iterative algorithms are derived to compute the equaliser and relay precoder for the case of general estimation errors. It is shown that the subproblems for updating the equaliser and the relay processor are convex and that the iterative algorithms are guaranteed to converge to at least a locally optimal solution. Suboptimal solutions are discussed in [83, 90] where the destination equaliser and relay precoder are calculated in closed form solution. It is also shown that, for the special case of uncorrelated estimation errors, these solutions are in fact the optimal ones. In [91] an iterative algorithm is derived to minimise the arithmetic MSE when a linear precoder is also introduce at the source and is shown to offer improved performance compared to the iterative algorithms in [81–83]. The study of robust transceivers for maximising the mutual information and minimising the weighted MSE can also be found in [92] and [93], respectively. Besides the study of robust linear designs, robust non-linear transceivers utilising THP have also been considered in [61, 62, 94] for narrowband MIMO relaying.

In this chapter we extend our previous results in [61], and consider the robust design of DFE and THP transceivers for the MIMO OFDM relaying model introduced in Section 2.2 of Chapter 2. To incorporate the effects of imperfect CSI we discuss a general channel estimation error model and show that the estimation errors arising from specific channel estimation algorithms fits into this general model. Based on the imperfect CSI model we consider the design of the DFE and THP processors that minimise the arithmetic MSE subject to transmission power constraints. Unfortunately the optimal solution is difficult to obtain and we relax the problem by minimising an upper bound to the objective function subject to relaxed constraints. The source and relay precoder matrices can then be derived as the solution to this relaxed optimisation problem. Simulation results are presented demonstrating the advantages of the proposed algorithms compared to benchmark designs proposed in the literature.

The rest of this chapter is organised as follows: In Section 5.2 we formulate the robust optimisation problem and derive the DFE and THP processors in Section 5.3. Simulation results are presented in Section 5.4 and finally conclusions are drawn in Section 5.5.

5.2 Problem Formulation

In this section we formulate the robust optimisation problem for minimising the arithmetic MSE subject to transmission power constraints at the source and relay terminals. For convenience we recall that the optimisation problem for the considered DFE and THP transceivers is (see Section 2.4 of Chapter 2)

$$\min_{\boldsymbol{F}_k, \boldsymbol{G_k}, \boldsymbol{W}_k, \boldsymbol{U}_k} \quad \frac{1}{K} \sum_{k=1}^{K} \frac{\mathrm{tr}\{\boldsymbol{R}_{e,k}\}}{N_k} \tag{5.1}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \mathrm{tr}\Big\{ \boldsymbol{F}_k \boldsymbol{F}_k^H \Big\} \leq P_s \tag{5.2}$$

$$\sum_{k=1}^{K} \mathrm{tr}\Big\{ \boldsymbol{G}_k \left(\boldsymbol{H}_{s,k} \boldsymbol{F}_k \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \right) \boldsymbol{G}_k^H \Big\} \leq P_r. \tag{5.3}$$

where $\boldsymbol{R}_{e,k}$ is the error covariance matrix for the kth subcarrier given by

$$\begin{aligned} \boldsymbol{R}_{e,k} = & \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right) \left(\boldsymbol{W}_k \boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{H}_{s,k} \boldsymbol{F}_k - \boldsymbol{U}_k \right)^H \\ & + \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k} \boldsymbol{G}_k \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \sigma_{v_s}^2 + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \right) \boldsymbol{W}_k^H. \end{aligned} \tag{5.4}$$

From (5.4) we see that the error covariance matrix for the kth subcarrier depends directly on the subcarrier channel matrices $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$. Furthermore, the relay power constraint in (5.3) also depends on the source-relay subcarrier matrix $\boldsymbol{H}_{s,k}$. When the subcarrier channel matrices are subject to channel estimation errors, i.e. they are not completely known to all nodes, then the optimisation problem in (5.1)-(5.3) is intractable and cannot be directly solved. In order to reformulate the optimisation problem into a tractable form, and to derive robust transceiver solutions, we must incorporate the effect of channel estimation errors. In the following we firstly consider a general channel estimation error model and then reformulate (5.1)-(5.3) as a statistically robust optimisation problem.

5.2.1 Channel Error Model

When the source-relay and relay-destination subcarrier channel matrices $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$ are subject to channel estimation errors then they can be written as

$$\boldsymbol{H}_{s,k} = \bar{\boldsymbol{H}}_{s,k} + \boldsymbol{E}_{s,k} \tag{5.5}$$

$$\boldsymbol{H}_{r,k} = \bar{\boldsymbol{H}}_{r,k} + \boldsymbol{E}_{r,k}, \tag{5.6}$$

where $\bar{\boldsymbol{H}}_{s,k} \in \mathbb{C}^{N_r \times N_s}$ and $\bar{\boldsymbol{H}}_{r,k} \in \mathbb{C}^{N_d \times N_r}$ are the estimated source-relay and relay-destination subcarrier matrices, respectively, with $\boldsymbol{E}_{s,k} \in \mathbb{C}^{N_r \times N_s}$ and $\boldsymbol{E}_{r,k} \in \mathbb{C}^{N_d \times N_r}$ representing the corresponding channel estimation errors. The specific structure of $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$ depend largely on the source of channel estimation error [62, 81–83, 90–95]. We assume that the error matrices can be modelled according to the following definition for matrix variate Gaussian distributions:

Definition 8: [67] A random matrix $\boldsymbol{A} \in \mathbb{C}^{N \times M}$ is said to have a matrix variate Gaussian distribution with mean $\bar{\boldsymbol{A}} \in \mathbb{C}^{N \times M}$ and covariance matrix $\boldsymbol{B} \otimes \boldsymbol{C}$, where $\boldsymbol{B} \in \mathbb{C}^{M \times M}$ and $\boldsymbol{C} \in \mathbb{C}^{N \times N}$, if it satisfies

$$\mathrm{vec}\,[\boldsymbol{A}] \sim \mathcal{CN}\left(\mathrm{vec}[\bar{\boldsymbol{A}}], \boldsymbol{B} \otimes \boldsymbol{C}\right), \tag{5.7}$$

where

$$\mathbb{E}\{\boldsymbol{A}\} = \bar{\boldsymbol{A}} \tag{5.8}$$

$$\mathbb{E}\Big\{\left(\mathrm{vec}\,[\boldsymbol{A}] - \mathrm{vec}\left[\bar{\boldsymbol{A}}\right]\right)\left(\mathrm{vec}\,[\boldsymbol{A}] - \mathrm{vec}\left[\bar{\boldsymbol{A}}\right]\right)^H\Big\} = \boldsymbol{B} \otimes \boldsymbol{C}. \tag{5.9}$$

The matrix variate Gaussian distribution of the random matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A} \sim \mathcal{CN}(\bar{\boldsymbol{A}}, \boldsymbol{B} \otimes \boldsymbol{C})$.

As a general model we shall assume that the channel estimation error matrices have the following independent matrix variate Gaussian distributions [67]

$$\boldsymbol{E}_{s,k} \sim \mathcal{CN}\left(\boldsymbol{0}_{N_r \times N_s}\ ,\ \boldsymbol{\Xi}_{s,k} \otimes \boldsymbol{\Omega}_{s,k}\right) \tag{5.10}$$

$$\boldsymbol{E}_{r,k} \sim \mathcal{CN}\left(\boldsymbol{0}_{N_d \times N_r}\ ,\ \boldsymbol{\Xi}_{r,k} \otimes \boldsymbol{\Omega}_{r,k}\right), \tag{5.11}$$

where $\boldsymbol{\Xi}_{s,k} \in \mathbb{C}^{N_s \times N_s}$ and $\boldsymbol{\Omega}_{s,k} \in \mathbb{C}^{N_r \times N_r}$ are positive definite matrices that represent the covariance of the source-relay subcarrier channel estimation error $\boldsymbol{E}_{s,k}$. Similarly, $\boldsymbol{\Xi}_{r,k} \in \mathbb{C}^{N_r \times N_r}$ and $\boldsymbol{\Omega}_{r,k} \in \mathbb{C}^{N_d \times N_d}$ are positive definite and represent the covariance of the relay-destination subcarrier channel estimation error $\boldsymbol{E}_{r,k}$. Before formulating a robust optimisation problem based on the preceeding channel error model, it is worthwhile highlighting a specific case where the subcarrier channel estimation error matrices can be modelled according to (5.10) and (5.11).

5.2.2 LS Channel Estimation Error

We consider the case that the source-relay and relay-destination frequency selective channels are estimated independently, where the source-relay channel is estimated at the relay and the relay-destination channel is estimated at the destination. We show that if the frequency selective channels are estimated using the LS channel estimation algorithm discussed in Section 3.2.2 of Chapter 3, then the resulting subcarrier estimation matrices $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$ can be modelled as a special case of (5.10) and (5.11). Since we consider the case that the source-relay channel is estimated at the relay and the relay-destination channel is estimated at the destination, the channel estimation for both channels is similar. We therefore take the estimation of the relay-destination channel as an example. We recall that the LS channel estimation error is given by (see (3.17) and (3.84))

$$\boldsymbol{e}_r \triangleq \mathrm{vec}\left[\boldsymbol{\mathcal{E}}_r[0], ..., \boldsymbol{\mathcal{E}}_r[L]\right] \tag{5.12}$$

$$= \left(\boldsymbol{M}_x^H \boldsymbol{M}_x \otimes \boldsymbol{I}_{N_d}\right)^{-1} \left(\boldsymbol{M}_x^H \otimes \boldsymbol{I}_{N_d}\right) \boldsymbol{v}_r, \tag{5.13}$$

where $\boldsymbol{\mathcal{E}}_r[l]$ is the channel estimation error matrix of $\boldsymbol{\mathcal{H}}_r[l]$. In (5.13) $\boldsymbol{M}_x$ is defined as (see (3.10))

$$\boldsymbol{M}_x \triangleq \left[\boldsymbol{F}_0 \boldsymbol{X}, ..., \boldsymbol{F}_L \boldsymbol{X}\right], \tag{5.14}$$

where $\boldsymbol{X}$ is the relay training matrix and

$$\boldsymbol{F}_l \triangleq \begin{bmatrix} e^{-j2\pi 0l/K} & & 0 \\ & \ddots & \\ 0 & & e^{-j2\pi(K-1)l/K} \end{bmatrix}. \tag{5.15}$$

It was shown in Theorem 1 of Chapter 3 that the optimal LS training matrix $\boldsymbol{X}$ is given

$$\boldsymbol{X} = \sqrt{\frac{P_r}{N_r}}\boldsymbol{Q}, \tag{5.16}$$

where $\boldsymbol{Q}$ is a semi-unitary matrix that satisfies the properties

$$\boldsymbol{Q}^H\boldsymbol{F}_m^H\boldsymbol{F}_n\boldsymbol{Q} = \boldsymbol{0}_{N_r\times N_r} \qquad \forall m \neq n \tag{5.17}$$

$$\boldsymbol{Q}^H\boldsymbol{F}_l^H\boldsymbol{F}_l\boldsymbol{Q} = \boldsymbol{I}_{N_r} \qquad \forall l \tag{5.18}$$

$$\boldsymbol{Q}^H\boldsymbol{Q} = \boldsymbol{I}_{N_r}. \tag{5.19}$$

We now note that the subcarrier estimation error matrix $\boldsymbol{E}_{r,k}$ is given by

$$\boldsymbol{E}_{r,k} = [\boldsymbol{\mathcal{E}}_r[0], ..., \boldsymbol{\mathcal{E}}_r[L]] \underbrace{\begin{bmatrix} e^{-j2\pi(k-1)0/K}\boldsymbol{I}_{N_r} \\ \vdots \\ e^{-j2\pi(k-1)L/K}\boldsymbol{I}_{N_r} \end{bmatrix}}_{\boldsymbol{D}_k}. \tag{5.20}$$

Using (5.13), and the fact that $\boldsymbol{v}_r$ is a zero mean Gaussian random variable, it is straightforward to show that $\mathbb{E}\{\boldsymbol{E}_{r,k}\} = \boldsymbol{0}_{N_d\times N_r}$. Furthermore, it is also straightforwardly shown that

$$\mathbb{E}\Big\{\text{vec}\left[\boldsymbol{E}_{r,k}\right]\text{vec}^H\left[\boldsymbol{E}_{r,k}\right]\Big\} = \sigma_{v_r}^2\left(\boldsymbol{D}_k^T\otimes\boldsymbol{I}_{N_d}\right)\left(\boldsymbol{M}_x^H\boldsymbol{M}_x\otimes\boldsymbol{I}_{N_d}\right)^{-1}\left(\boldsymbol{D}_k^*\otimes\boldsymbol{I}_{N_d}\right) \tag{5.21}$$

$$= \frac{\sigma_{v_r}^2 N_r(L+1)}{P_r}\boldsymbol{I}_{N_r}\otimes\boldsymbol{I}_{N_d}, \tag{5.22}$$

where we have used the fact that the optimal training matrix is given by (5.16) with the matrix $\boldsymbol{Q}$ satisfying the conditions in (5.17)-(5.19). Using the definition of matrix variate Gaussian distributions in Definition 8 we finally have that the

kth relay-destination subcarrier has the matrix variate distribution

$$\boldsymbol{E}_{r,k} \sim \mathcal{CN}\left(\boldsymbol{0}_{N_d \times N_r}\ ,\ \sigma_{e_r}^2 \boldsymbol{I}_{N_r} \otimes \boldsymbol{I}_{N_d}\right), \tag{5.23}$$

where for convenience we define $\sigma_{e_r}^2 \triangleq \sigma_{v_r}^2 N_r(L+1)/P_r$, which can be viewed as the inverse SNR during the channel estimation phase. A similar analysis shows that $\boldsymbol{E}_{s,k}$ is distributed according to

$$\boldsymbol{E}_{s,k} \sim \mathcal{CN}\left(\boldsymbol{0}_{N_r \times N_s}\ ,\ \sigma_{e_s}^2 \boldsymbol{I}_{N_s} \otimes \boldsymbol{I}_{N_r}\right), \tag{5.24}$$

where $\sigma_{e_s}^2 \triangleq \sigma_{v_s}^2 N_s(L+1)/P_s$. Clearly (5.23) and (5.24) are special cases of the more general distributions in (5.11) and (5.10), respectively.

5.2.3 Robust Constrained Optimisation Problem

Having derived a general channel error model we now formulate the robust optimisation problem for deriving the DFE and THP processors $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, $\boldsymbol{W}_k$, and $\boldsymbol{U}_k$ to minimise the arithmetic MSE subject to transmission power constraints at both the source and relay terminals. As previously remarked the error covariance matrix for each subcarrier in (5.4) depends on the subcarrier channel matrices $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$, which are not completely known, and therefore an optimisation problem using $\boldsymbol{R}_{e,k}$ cannot be conducted. Rather than considering the instantaneous error covariance matrix $\boldsymbol{R}_{e,k}$ we shall reformulate the optimisation problem in (5.1)-(5.3) by considering an objective function based on $\mathbb{E}\{\boldsymbol{R}_{e,k}\}$, where the expectation is taken w.r.t. the random matrices $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$. Furthermore, since the relay power constraint in (5.3) depends on the unknown source-relay subcarrier matrix $\boldsymbol{H}_{s,k}$ we shall consider constraining the expected transmission power of the relay node, with the expectation being taken w.r.t. $\boldsymbol{E}_{s,k}$. In order to compute the necessary expectations in our following analysis we firstly consider the following lemma:

Lemma 16: [67] For a random matrix $\boldsymbol{A} \in \mathbb{C}^{M \times N}$ with mean $\bar{\boldsymbol{A}} \in \mathbb{C}^{M \times N}$ and covariance $\boldsymbol{C} \otimes \boldsymbol{D}$, i.e. $\boldsymbol{A}$ has the matrix variate complex Gaussian distribution $\boldsymbol{A} \sim \mathcal{CN}\left(\bar{\boldsymbol{A}}\ ,\ \boldsymbol{C} \otimes \boldsymbol{D}\right)$, we have for any $\boldsymbol{Q} \in \mathbb{C}^{N \times N}$ that

$$\mathbb{E}\left\{\boldsymbol{A}\boldsymbol{Q}\boldsymbol{A}^H\right\} = \bar{\boldsymbol{A}}\boldsymbol{Q}\bar{\boldsymbol{A}}^H + \mathrm{tr}\left\{\boldsymbol{Q}\boldsymbol{C}^T\right\}\boldsymbol{D}. \tag{5.25}$$

With the use of Definition 8, Lemma 16, and the matrix variate Gaussian distributions for $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$ given in (5.10) and (5.11), respectively, we can

establish the following results

$$\mathbb{E}\{\boldsymbol{E}_{s,k}\} = \mathbf{0}_{N_r \times N_s} \tag{5.26}$$

$$\mathbb{E}\{\boldsymbol{E}_{r,k}\} = \mathbf{0}_{N_d \times N_r} \tag{5.27}$$

$$\mathbb{E}\Big\{\boldsymbol{E}_{s,k}\boldsymbol{Y}_{s,k}\boldsymbol{E}_{s,k}^H\Big\} = \mathrm{tr}\Big\{\boldsymbol{Y}_{s,k}\boldsymbol{\Xi}_{s,k}^T\Big\}\,\boldsymbol{\Omega}_{s,k} \tag{5.28}$$

$$\mathbb{E}\Big\{\boldsymbol{E}_{r,k}\boldsymbol{Y}_{r,k}\boldsymbol{E}_{r,k}^H\Big\} = \mathrm{tr}\Big\{\boldsymbol{Y}_{r,k}\boldsymbol{\Xi}_{r,k}^T\Big\}\,\boldsymbol{\Omega}_{r,k}, \tag{5.29}$$

where $\boldsymbol{Y}_{s,k} \in \mathbb{C}^{N_s \times N_s}$ and $\boldsymbol{Y}_{r,k} \in \mathbb{C}^{N_r \times N_r}$ are arbitrary matrices.

We are now ready to reformulate (5.1)-(5.3) as a statistically robust optimisation problem for deriving the DFE and THP processors $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, $\boldsymbol{W}_k$, and $\boldsymbol{U}_k$. Substituting (5.5) and (5.6) into (5.4) and taking the expectation w.r.t. the random matrices $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$, we can compute that $\mathbb{E}\{\boldsymbol{R}_{e,k}\}$ is given by

$$\begin{aligned}\mathbb{E}\{\boldsymbol{R}_{e,k}\} = \boldsymbol{W}_k\Big(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \mathrm{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{\Xi}_{r,k}^T\Big\}\,\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\Big)\boldsymbol{W}_k^H \\ - \boldsymbol{W}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H - \boldsymbol{U}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\boldsymbol{W}_k^H + \boldsymbol{U}_k\boldsymbol{U}_k^H,\end{aligned} \tag{5.30}$$

where for notational convenience we define the matrix $\bar{\boldsymbol{X}}_k \in \mathbb{C}^{N_r \times N_r}$ as

$$\bar{\boldsymbol{X}}_k \triangleq \bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathrm{tr}\Big\{\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{\Xi}_{s,k}^T\Big\}\,\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r}. \tag{5.31}$$

A detailed derivation of (5.30) is provided in Section 5.6.1 on page 147. Since the power consumed by the relay depends on $\boldsymbol{H}_{s,k}$, we must also average the relay transmission power w.r.t. $\boldsymbol{E}_{s,k}$. The power consumed by the relay on the kth subcarrier is $\mathrm{tr}\{\boldsymbol{G}_k(\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r})\boldsymbol{G}_k^H\}$. Substituting (5.5) into this expression and taking the expectation w.r.t $\boldsymbol{E}_{s,k}$ we can compute

$$\begin{aligned}&\mathrm{tr}\Big\{\boldsymbol{G}_k\mathbb{E}\Big\{\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\Big\}\boldsymbol{G}_k^H\Big\} \\ &= \mathrm{tr}\Big\{\boldsymbol{G}_k\Big(\bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\mathbb{E}\Big\{\bar{\boldsymbol{E}}_{s,k}^H\Big\} \\ &\quad + \mathbb{E}\{\bar{\boldsymbol{E}}_{s,k}\}\,\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathbb{E}\Big\{\bar{\boldsymbol{E}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{E}}_{s,k}^H\Big\} + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\Big)\boldsymbol{G}_k\Big\} &(5.32)\\ &= \mathrm{tr}\Big\{\boldsymbol{G}_k\Big(\bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathrm{tr}\Big\{\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{\Xi}_{s,k}^T\Big\}\,\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r}\Big)\Big\} &(5.33)\\ &= \mathrm{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\Big\}, &(5.34)\end{aligned}$$

where $\bar{\boldsymbol{X}}_k$ is defined in (5.31). To obtain (5.33) from (5.32) we have used (5.26), (5.27), as well as the result in (5.29). With the expected error covariance matrix in (5.30) and the expected relay power constraint for the kth subcarrier given in

(5.34), we can reformulate the optimisation problem in (5.1)-(5.3) as

$$\min_{\boldsymbol{F}_k,\boldsymbol{G}_k,\boldsymbol{W}_k,\boldsymbol{U}_k} \quad \frac{1}{K}\sum_{k=1}^{K}\frac{\mathrm{tr}\{\mathbb{E}\{\boldsymbol{R}_{e,k}\}\}}{N_k} \tag{5.35}$$

$$\text{s.t.} \quad \sum_{k=1}^{N_c}\mathrm{tr}\Big\{\boldsymbol{F}_k\boldsymbol{F}_k^H\Big\} \leq P_s \tag{5.36}$$

$$\sum_{k=1}^{N_c}\mathrm{tr}\left\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\right\} \leq P_r. \tag{5.37}$$

It is worthwhile noting here that for the case of perfect channel estimations, i.e. when $\boldsymbol{E}_{s,k} = \boldsymbol{0}_{N_r\times N_s}$ and $\boldsymbol{E}_{r,k} = \boldsymbol{0}_{N_d\times N_r}$, then the optimisation problem in (5.35)-(5.37) reduces to that in (5.1)-(5.3). In other words (5.35)-(5.37) includes the scenario of perfect channel estimation as a specific case and therefore represents a more general optimisation problem.

5.3 Robust Transceiver Design

In this section we focus on deriving the processors $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, $\boldsymbol{W}_k$, and $\boldsymbol{U}_k$, as the solution to the problem in (5.35)-(5.37).

5.3.1 Optimal Equaliser

We begin by deriving the optimal MMSE equaliser. Since the power constraints in (5.36) and (5.37) do not depend on $\boldsymbol{W}_k$, the optimal equaliser is the solution to the unconstrained problem of minimising (5.35). Since $\mathrm{tr}\{\mathbb{E}\{\boldsymbol{R}_{e,k}\}\}$ is convex quadratic [39] in $\boldsymbol{W}_k$, the equaliser is obtained by setting the derivative of $\mathrm{tr}\{\mathbb{E}\{\boldsymbol{R}_{e,k}\}\}$ w.r.t. $\boldsymbol{W}_k^*$ to zero and solving for $\boldsymbol{W}_k$, resulting in

$$\begin{aligned}\boldsymbol{W}_k = \boldsymbol{U}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\Big(&\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H \\ &+ \mathrm{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{\Xi}_{r,k}^T\Big\}\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\Big)^{-1}.\end{aligned} \tag{5.38}$$

Substituting (5.38) into (5.30) and using (5.31) we can write the expected error covariance matrix in the more concentrated manner as

$$\mathbb{E}\{\boldsymbol{R}_{e,k}\} = \boldsymbol{U}_k \Bigg(\boldsymbol{I}_{N_k} + \boldsymbol{F}_k^H \bar{\boldsymbol{H}}_{s,k}^H \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H \Big(\bar{\boldsymbol{H}}_{r,k} \boldsymbol{G}_k \Big(\mathrm{tr}\Big\{ \boldsymbol{F}_k \boldsymbol{F}_k^H \boldsymbol{\Xi}_{s,k}^T \Big\} \boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \Big) \times \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H + \mathrm{tr}\Big\{ \boldsymbol{G}_k \bar{\boldsymbol{X}}_k \boldsymbol{G}_k^H \boldsymbol{\Xi}_{r,k}^T \Big\} \boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1} \bar{\boldsymbol{H}}_{r,k} \boldsymbol{G}_k \bar{\boldsymbol{H}}_{s,k} \boldsymbol{F}_k \Bigg)^{-1} \boldsymbol{U}_k^H, \tag{5.39}$$

where we have also used the matrix inversion lemma.

5.3.2 Source, Relay, and Feedback Matrices

In a similar fashion to the derivation of the optimal source, relay, and feedback matrices for the case of perfect channel estimation derived in Chapter 4, we proceed by firstly obtaining a lower bound to the objective function in (5.35) and showing that a specific choice of processors can achieve the lower bound. To obtain such a lower bound we recall the following lemma:

Lemma 17: For a positive semi-definite matrix $\boldsymbol{A} \in \mathbb{C}^{M \times M}$ we have the inequality $|\boldsymbol{A}|^{1/M} \leq \mathrm{tr}\{\boldsymbol{A}\}/M$ where equality is achieved if and only if $\boldsymbol{A} = \alpha \boldsymbol{I}_M$, for some $\alpha \in \mathbb{R}_+$. In otherwords equality is achieved when $\boldsymbol{A}$ is a diagonal matrix with equal non-negative diagonal entries. This is the well known arithmetic-geometric mean inequality.

Based on Lemma 17 the objective function in (5.35) is lower bounded by

$$\frac{1}{K} \sum_{k=1}^{K} |\mathbb{E}\{\boldsymbol{R}_{e,k}\}|^{1/N_k} \leq \frac{1}{K} \sum_{k=1}^{K} \frac{\mathrm{tr}\{\mathbb{E}\{\boldsymbol{R}_{e,k}\}\}}{N_k} \tag{5.40}$$

where equality holds if and only if $\mathbb{E}\{\boldsymbol{R}_{e,k}\}$ is diagonal with equal diagonal elements. To simplify the lower bound in (5.40) we shall find it convenient to parameterise the linear source precoding matrix as

$$\boldsymbol{F}_k = \tilde{\boldsymbol{F}}_k \boldsymbol{\Psi}_k, \tag{5.41}$$

where $\tilde{\boldsymbol{F}}_k \in \mathbb{C}^{N_s \times N_k}$ is an arbitrary matrix and $\boldsymbol{\Psi}_k \in \mathbb{C}^{N_k \times N_k}$ is a unitary matrix. To find the solution to (5.35)-(5.37) we shall firstly derive the matrices $\boldsymbol{W}_k$, $\boldsymbol{U}_k$, and $\boldsymbol{\Psi}_k$, before then calculating $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$. Upon substituting (5.41) into (5.39)

we can decompose the expected error covariance matrix as

$$\mathbb{E}\{\boldsymbol{R}_{e,k}\} = \boldsymbol{U}_k \boldsymbol{\Psi}_k^H \boldsymbol{E}_k \boldsymbol{\Psi}_k \boldsymbol{U}_k^H, \tag{5.42}$$

where the positive definite matrix $\boldsymbol{E}_k \in \mathbb{C}^{N_k \times N_k}$ is defined as

$$\boldsymbol{E}_k \triangleq \Bigg(\boldsymbol{I}_{N_k} + \tilde{\boldsymbol{F}}_k^H \bar{\boldsymbol{H}}_{s,k}^H \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H \Big(\bar{\boldsymbol{H}}_{r,k} \boldsymbol{G}_k \Big(\mathrm{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \boldsymbol{\Xi}_{s,k}^T \Big\} \boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \Big) \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H + \mathrm{tr}\Big\{ \boldsymbol{G}_k \bar{\boldsymbol{X}}_k \boldsymbol{G}_k^H \boldsymbol{\Xi}_{r,k}^T \Big\} \boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1} \bar{\boldsymbol{H}}_{r,k} \boldsymbol{G}_k \bar{\boldsymbol{H}}_{s,k} \tilde{\boldsymbol{F}}_k \Bigg)^{-1}. \tag{5.43}$$

We also note that with the structure of the source precoder in (5.41) the matrix $\bar{\boldsymbol{X}}_k$ in (5.31) is independent of $\boldsymbol{\Psi}_k$ i.e.

$$\bar{\boldsymbol{X}}_k = \bar{\boldsymbol{H}}_{s,k} \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \bar{\boldsymbol{H}}_{s,k}^H + \mathrm{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \boldsymbol{\Xi}_{s,k}^T \Big\} \boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2 \boldsymbol{I}_{N_r}. \tag{5.44}$$

With the structure of the expected error covariance matrix in (5.42) we can write the inequality in (5.40) equivalently as

$$\frac{1}{K} \sum_{k=1}^{K} |\boldsymbol{E}_k|^{1/N_k} \leq \frac{1}{K} \sum_{k=1}^{K} \frac{\mathrm{tr}\{\mathbb{E}\{\boldsymbol{R}_{e,k}\}\}}{N_k}, \tag{5.45}$$

where to obtain the lower bound we have used the fact that $\boldsymbol{U}_k$ is unit diagonal triangular and $\boldsymbol{\Psi}_k$ is unitary. In the following we shall see that, given any $\boldsymbol{E}_k$, an appropriate unit diagonal upper right triangular matrix $\boldsymbol{U}_k$ and unitary $\boldsymbol{\Psi}_k$ can be found to ensure that (5.45) holds with equality. From the arithmetic-geometric mean inequality we find that (5.45) holds with equality when

$$\mathbb{E}\{\boldsymbol{R}_{e,k}\} = \boldsymbol{U}_k \boldsymbol{\Psi}_k^H \boldsymbol{E}_k \boldsymbol{\Psi}_k \boldsymbol{U}_k^H = \prod_{i=1}^{N_k} \varepsilon_{k,i}^{1/N_k} \boldsymbol{I}_{N_k}, \tag{5.46}$$

where $\{\varepsilon_{k,i}\}_{i=1}^{N_k} \in \mathbb{R}_{++}$ are the non-zero eigenvalues of $\boldsymbol{E}_k$. Since both sides of (5.46) are Hermitian, solving (5.46) is equivalent to finding matrices that satisfy the equality $\boldsymbol{U}_k \boldsymbol{\Psi}_k^H \boldsymbol{E}_k^{1/2} = \prod_{i=1}^{N_k} \varepsilon_{k,i}^{1/2N_k} \boldsymbol{S}_k^H$, which upon rearranging leads to the GMD [26]

$$\prod_{i=1}^{N_k} \varepsilon_{k,i}^{1/2N_k} \boldsymbol{E}_k^{-1/2} = \boldsymbol{S}_k \boldsymbol{U}_k \boldsymbol{\Psi}_k^H. \tag{5.47}$$

When the matrices $\boldsymbol{U}_k$ and $\boldsymbol{\Psi}_k$ are computed according to (5.47) then (5.45) holds

with equality. As such we can replace the objective function in (5.35) with the lower bound in (5.45) and state the optimisation problem in (5.35)-(5.37) as

$$\min_{\tilde{\boldsymbol{F}}_k, \boldsymbol{G}_k} \quad \sum_{k=1}^{K} |\boldsymbol{E}_k| \tag{5.48}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \text{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \Big\} \leq P_s \tag{5.49}$$

$$\sum_{k=1}^{K} \text{tr}\left\{ \boldsymbol{G}_k \bar{\boldsymbol{X}}_k \boldsymbol{G}_k^H \right\} \leq P_r. \tag{5.50}$$

The remaining task is to compute the source precoding matrix $\tilde{\boldsymbol{F}}_k$ and relay precoding matrix $\boldsymbol{G}_k$ as the solution to (5.48)-(5.50). Unfortunately from the definition of $\boldsymbol{E}_k$ in (5.43) we see that (5.48) is a complicated function of $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ and the optimisation problem is difficult to handle. In order to proceed, we relax the problem by minimising an upper bound of the objective function in (5.48) subject to relaxed power constraints. To this end, it can be shown that (5.48) and (5.50) are upper bounded by

$$\sum_{k=1}^{K} |\boldsymbol{E}_k| \leq \sum_{k=1}^{K} |\tilde{\boldsymbol{E}}_k| \tag{5.51}$$

$$\sum_{k=1}^{K} \text{tr}\left\{ \boldsymbol{G}_k \bar{\boldsymbol{X}}_k \boldsymbol{G}_k^H \right\} \leq \sum_{k=1}^{K} \text{tr}\left\{ \boldsymbol{G}_k \tilde{\boldsymbol{X}}_k \boldsymbol{G}_k^H \right\}, \tag{5.52}$$

where the matrices $\tilde{\boldsymbol{E}}_k$ and $\tilde{\boldsymbol{X}}_k$ are defined as

$$\tilde{\boldsymbol{E}}_k \triangleq \Bigg(\boldsymbol{I}_{N_k} + \tilde{\boldsymbol{F}}_k^H \bar{\boldsymbol{H}}_{s,k}^H \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H \Big(\bar{\boldsymbol{H}}_{r,k} \boldsymbol{G}_k \Big(\text{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \Big\} \xi_{s,k} \boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2 \boldsymbol{I}_{N_r} \Big) \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H + \text{tr}\Big\{ \boldsymbol{G}_k \tilde{\boldsymbol{X}}_k \boldsymbol{G}_k^H \Big\} \xi_{r,k} \boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2 \boldsymbol{I}_{N_d} \Big)^{-1} \bar{\boldsymbol{H}}_{r,k} \boldsymbol{G}_k \bar{\boldsymbol{H}}_{s,k} \tilde{\boldsymbol{F}}_k \Bigg)^{-1} \tag{5.53}$$

$$\tilde{\boldsymbol{X}}_k \triangleq \bar{\boldsymbol{H}}_{s,k} \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \bar{\boldsymbol{H}}_{s,k}^H + \xi_{s,k} \text{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \Big\} \boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2 \boldsymbol{I}_{N_r}. \tag{5.54}$$

In (5.53) and (5.54) the scalars $\xi_{s,k}$ and $\xi_{r,k}$ are the largest eigenvalues of $\boldsymbol{\Xi}_{s,k}$ and $\boldsymbol{\Xi}_{r,k}$, respectively. The detailed derivations of the bounds in (5.51) and (5.52) are given in Section 5.6.2 on page 148.

The optimisation problem in (5.48)-(5.50) can be relaxed by replacing the objective function and the relay power constraint with the upper bounds in (5.51)

and (5.52), respectively, resulting in the relaxed problem

$$\min_{\tilde{\boldsymbol{F}}_k, \boldsymbol{G}_k} \quad \sum_{k=1}^{K} |\tilde{\boldsymbol{E}}_k| \tag{5.55}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \text{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \Big\} \leq P_s \tag{5.56}$$

$$\sum_{k=1}^{K} \text{tr}\left\{ \boldsymbol{G}_k \tilde{\boldsymbol{X}}_k \boldsymbol{G}_k^H \right\} \leq P_r. \tag{5.57}$$

It is worth noting that the relaxed optimisation problem in (5.55)-(5.57) is equivalent to the original problem in (5.48)-(5.50) when $\boldsymbol{\Xi}_{s,k} = \xi_{s,k} \boldsymbol{I}_{N_s}$ and $\boldsymbol{\Xi}_{r,k} = \xi_{r,k} \boldsymbol{I}_{N_r}$. It is shown in Section 5.6.3 on page 150 that the optimal processors $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ as the solution to (5.55)-(5.57) should satisfy the following power constraint equalities

$$\text{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \Big\} = P_{s,k} \tag{5.58}$$

$$\sum_{k=1}^{K} P_{s,k} = P_s \tag{5.59}$$

$$\text{tr}\Big\{ \boldsymbol{G}_k \tilde{\boldsymbol{X}}_k \boldsymbol{G}_k^H \Big\} = P_{r,k} \tag{5.60}$$

$$\sum_{k=1}^{K} P_{r,k} = P_r, \tag{5.61}$$

where $P_{s,k} \geq 0$ and $P_{r,k} \geq 0$ are the maximum power budgets available to the source and relay on the kth subcarrier. Incorporating the constraints (5.58)-(5.61) into (5.55)-(5.57), the problem can be equivalently stated as

$$\min_{\tilde{\boldsymbol{F}}_k, \boldsymbol{G}_k, P_{s,k}, P_{r,k}} \quad \sum_{k=1}^{K} |\tilde{\boldsymbol{E}}_k| \tag{5.62}$$

$$\text{s.t.} \quad \text{tr}\Big\{ \tilde{\boldsymbol{F}}_k \tilde{\boldsymbol{F}}_k^H \Big\} = P_{s,k} \tag{5.63}$$

$$\sum_{k=1}^{K} P_{s,k} = P_s \tag{5.64}$$

$$\text{tr}\Big\{ \boldsymbol{G}_k \tilde{\boldsymbol{X}}_k \boldsymbol{G}_k^H \Big\} = P_{r,k} \tag{5.65}$$

$$\sum_{k=1}^{K} P_{r,k} = P_r \tag{5.66}$$

$$P_{s,k} \geq 0, \; P_{r,k} \geq 0, \quad 1 \leq k \leq K. \tag{5.67}$$

We emphasise that the objective function in (5.62) is now easier to handle since

from (5.58) and (5.60) we can replace $\mathrm{tr}\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\}$ and $\mathrm{tr}\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\}$ in (5.53) with $P_{s,k}$ and $P_{r,k}$, respectively. Thus the matrix $\tilde{\boldsymbol{E}}_k$ for the objective function is

$$\tilde{\boldsymbol{E}}_k \triangleq \Bigg(\boldsymbol{I}_{N_k} + \tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\Big(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\Big(P_{s,k}\xi_{s,k}\boldsymbol{\Omega}_{s,k}+\sigma_{v_s}^2\boldsymbol{I}_{N_r}\Big)\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + P_{r,k}\xi_{r,k}\boldsymbol{\Omega}_{r,k}+\sigma_{v_r}^2\boldsymbol{I}_{N_d}\Big)^{-1}\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\Bigg)^{-1}. \tag{5.68}$$

Before deriving the optimal processors $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ let us now introduce the singular value decompositions

$$\left(P_{s,k}\xi_{s,k}\boldsymbol{\Omega}_{s,k}+\sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)^{-1/2}\bar{\boldsymbol{H}}_{s,k} = \boldsymbol{U}_{s,k}\boldsymbol{\Lambda}_k\boldsymbol{V}_{s,k}^H, \tag{5.69}$$

$$\left(P_{r,k}\xi_{r,k}\boldsymbol{\Omega}_{r,k}+\sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)^{-1/2}\bar{\boldsymbol{H}}_{r,k} = \boldsymbol{U}_{r,k}\boldsymbol{\Delta}_k\boldsymbol{V}_{r,k}^H, \tag{5.70}$$

where $\boldsymbol{U}_{s,k}\in\mathbb{C}^{N_r\times N_r}$, $\boldsymbol{V}_{s,k}\in\mathbb{C}^{N_s\times N_s}$, $\boldsymbol{U}_{r,k}\in\mathbb{C}^{N_d\times N_d}$, and $\boldsymbol{V}_{r,k}\in\mathbb{C}^{N_r\times N_r}$ are unitary matrices. The upper left submatrices of $\boldsymbol{\Lambda}_k\in\mathbb{R}^{N_r\times N_s}$ and $\boldsymbol{\Delta}_k\in\mathbb{R}^{N_d\times N_r}$ contain the non-zero singular values $\{\lambda_{k,i}\}_{i=1}^{R_{s,k}}\in\mathbb{R}_+$ and $\{\delta_{k,i}\}_{i=1}^{R_{r,k}}\in\mathbb{R}_+$ respectively, which are assumed to be arranged in decreasing order. Here we define $R_{s,k}\triangleq\mathrm{rank}\{\boldsymbol{\Lambda}_k\}$ and $R_{r,k}\triangleq\mathrm{rank}\{\boldsymbol{\Delta}_k\}$.

Theorem 6: The structure of the optimal processors $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ as the solution to (5.62)-(5.67) are given by

$$\tilde{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Gamma}_k \tag{5.71}$$

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Phi}_k\boldsymbol{U}_{s,k}^H\left(P_{s,k}\xi_{s,k}\boldsymbol{\Omega}_{s,k}+\sigma_{v_s}^2\boldsymbol{I}_{N_r}\right)^{-1/2}, \tag{5.72}$$

where the diagonal matrices $\boldsymbol{\Gamma}_k\in\mathbb{R}^{N_s\times N_k}$ and $\boldsymbol{\Phi}_k\in\mathbb{R}^{N_r\times N_r}$ contain the elements $\{\gamma_{k,i}\}_{i=1}^{\bar{N}_k}\in\mathbb{R}_+$ and $\{\phi_{k,i}\}_{i=1}^{\bar{N}_k}\in\mathbb{R}_+$ on their upper left main diagonals, and we define the variable $\bar{N}_k\triangleq\min(N_k,R_{s,k},R_{r,k})$.

Proof: See Section 5.6.4 on page 152. □

The structure of the precoders in (5.71) and (5.72) result in the problem (5.62)-(5.67) reducing to

$$\min_{\gamma_{k,i},\phi_{k,i},P_{s,k},P_{r,k}} \quad \sum_{k=1}^{K}\sum_{i=1}^{\bar{N}_k} \log\left(1+\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2+1}\right)^{-1} \tag{5.73}$$

$$\text{s.t.} \quad \sum_{i=1}^{\bar{N}_k}\gamma_{k,i}^2 = P_{s,k} \tag{5.74}$$

$$\sum_{k=1}^{K} P_{s,k} = P_s \tag{5.75}$$

$$\sum_{i=1}^{\bar{N}_k}\phi_{k,i}^2\left(\gamma_{k,i}^2\lambda_{k,i}^2+1\right) = P_{r,k} \tag{5.76}$$

$$\sum_{k=1}^{K} P_{r,k} = P_r \tag{5.77}$$

$$\gamma_{k,i}\geq 0,\ \phi_{k,i}\geq 0,\ P_{s,k}\geq 0,\ P_{r,k}\geq 0,$$
$$1\leq k\leq K\ , 1\leq i\leq \bar{N}_k. \tag{5.78}$$

We see that the structure of $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ in (5.71) and (5.72) results in (5.62)-(5.67) being reduced to a scalar power allocation problem.

5.3.3 Power Allocation Algorithm

It is evident that the problem in (5.73)-(5.78) is intractable, since from the decompositions in (5.69) and (5.70) we see that $\lambda_{k,i}$ and $\delta_{k,i}$ depend on the source and relay subcarrier power allocation parameters $P_{s,k}$ and $P_{r,k}$ respectively, which are currently unknown. In order to achieve a simple analytic solution we assume that power is allocated uniformly among the subcarriers i.e. $P_{s,k} = P_s/K$ and $P_{r,k} = P_r/K$, $\forall k$. With the given $P_{s,k}$ and $P_{r,k}$ the decompositions in (5.69) and (5.70) are now known and the constraints in (5.75) and (5.77) are both satisfied with equality. The problem then reduces to K parallel optimisation problems

given by

$$\min_{\gamma_{k,i},\phi_{k,i}} \quad \sum_{i=1}^{\bar{N}_k} \log\left(1+\frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2+1}\right)^{-1} \tag{5.79}$$

$$\text{s.t.} \quad \sum_{i=1}^{\bar{N}_k} \gamma_{k,i}^2 = P_s/K \tag{5.80}$$

$$\sum_{i=1}^{\bar{N}_k} \phi_{k,i}^2\left(\gamma_{k,i}^2\lambda_{k,i}^2+1\right) = P_r/K \tag{5.81}$$

$$\gamma_{k,i} \geq 0,\ \phi_{k,i} \geq 0,\ 1 \leq i \leq \bar{N}_k, \tag{5.82}$$

which have to be solved for each subcarrier k. The problem stated in (5.79)-(5.82) can be solved using the iterative power allocation algorithms utilised in e.g. [45, 50, 54, 96] to achieve a locally optimal solution. We firstly introduce the new variables

$$\vartheta_{k,i} = \gamma_{k,i}^2 \tag{5.83}$$

$$\varphi_{k,i} = \phi_{k,i}^2\left(\gamma_{k,i}^2\lambda_{k,i}^2+1\right), \tag{5.84}$$

and rewrite the problem in (5.79)-(5.82) as

$$\min_{\vartheta_{k,i},\varphi_{k,i}} \quad \sum_{i=1}^{\bar{N}_k} \log\left(\frac{\left(\vartheta_{k,i}\lambda_{k,i}^2+1\right)\left(\varphi_{k,i}\delta_{k,i}^2+1\right)}{\vartheta_{k,i}\lambda_{k,i}^2+\varphi_{k,i}\delta_{k,i}^2+1}\right) \tag{5.85}$$

$$\text{s.t.} \quad \sum_{i=1}^{\bar{N}_k} \vartheta_{k,i} = P_s/K \tag{5.86}$$

$$\sum_{i=1}^{\bar{N}_k} \varphi_{k,i} = P_r/K \tag{5.87}$$

$$\vartheta_{k,i} \geq 0,\ \varphi_{k,i} \geq 0,\ 1 \leq i \leq \bar{N}_k. \tag{5.88}$$

It can be seen that the power constraints in (5.86) and (5.87) are now decoupled. For a set $\varphi_{k,i}$ satisfying (5.87), the problem for finding $\vartheta_{k,i}$ is given by minimising (5.85) subject to (5.86). Similarly, for a given $\vartheta_{k,i}$ satisfying (5.86), $\varphi_{k,i}$ can be computed by minimising (5.85) subject to the constraint (5.87). Both problems for finding $\vartheta_{k,i}$ and $\varphi_{k,i}$ are standard convex optimisation problems which can be

solved using the Karush Kuhn Tucker (KKT) conditions [39], resulting in

$$\vartheta_{k,i} = \frac{1}{2\lambda_{k,i}^2}\left[\sqrt{\varphi_{k,i}^2\delta_{k,i}^4 + \frac{4\varphi_{k,i}\delta_{k,i}^2\lambda_{k,i}^2}{\mu_s}} - \varphi_{k,i}\delta_{k,i}^2 - 2\right]^+ \tag{5.89}$$

$$\varphi_{k,i} = \frac{1}{2\delta_{k,i}^2}\left[\sqrt{\vartheta_{k,i}^2\lambda_{k,i}^4 + \frac{4\vartheta_{k,i}\lambda_{k,i}^2\delta_{k,i}^2}{\mu_r}} - \vartheta_{k,i}\lambda_{k,i}^2 - 2\right]^+, \tag{5.90}$$

where μ_s and μ_r are the Lagrangian multipliers that are required to be calculated to satisfy the constraints in (5.86) and (5.87) respectively. Starting with any feasible $\vartheta_{k,i}$, the algorithm updates $\varphi_{k,i}$ and $\vartheta_{k,i}$ according to (5.90) and (5.89) in an alternating fashion until a locally optimal solution is reached, which as noted in [45, 50, 96] is guaranteed. Once the alternating algorithm has converged, the variables $\gamma_{k,i}$ and $\phi_{k,i}$ are computed from (5.83) and (5.84).

5.4 Simulation Results

This section provides simulation results to evaluate the performance of the proposed robust DFE and THP transceiver solutions. In all simulations we consider a system employing $N_s = N_r = N_d = 3$ antennas at the source, relay , and destination devices. The MIMO time domain channel responses are modelled according to (2.8) and (2.9) and in all simulations we set $L + 1 = 5$. The elements of $\boldsymbol{\mathcal{H}}_{sw}[l]$ and $\boldsymbol{\mathcal{H}}_{rw}[l]$ in (2.8) and (2.9) are drawn from zero mean Gaussian random distributions with variances $\sigma_{h_s}^2[l] = \sigma_{h_r}^2[l] = 1/(L+1)$. The transmit and receive spatial correlation matrices in (2.8) and (2.9) are generated from the exponential correlation model (see e.g. [81–84]) and have elements given by

$$[\boldsymbol{\Theta}_s[l]]_{mn} = \rho_s[l]^{|m-n|} \tag{5.91}$$

$$[\boldsymbol{\Theta}_r[l]]_{mn} = \rho_r[l]^{|m-n|} \tag{5.92}$$

$$[\boldsymbol{\Upsilon}_s[l]]_{mn} = \varrho_s[l]^{|m-n|} \tag{5.93}$$

$$[\boldsymbol{\Upsilon}_r[l]]_{mn} = \varrho_r[l]^{|m-n|}, \tag{5.94}$$

where the correlation co-efficients $\rho_s[l]$, $\rho_r[l]$, $\varrho_s[l]$, and $\varrho_r[l]$ are selected from the interval $[0, 1]$ and define the level of spatial correlation. OFDM is employed with $K = 32$ subcarriers with each subcarrier being used to transmit $N_k = 3$ 16-QAM data symbols. We define the SNR of the source-relay and relay-destination channels as $\mathrm{SNR_s} = P_s/K\sigma_{v_s}^2$ and $\mathrm{SNR_r} = P_r/K\sigma_{v_r}^2$, respectively. For channel

estimation we consider that the source-relay and relay-destination channels are estimated independently, with the source-relay channel being estimated at the relay device whilst the relay-destination channel is estimated at the destination. Furthermore, both channels are estimated using LS channel estimation as discussed in Section 5.2.2 and thus the channel estimation error matrices $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$ in (5.5) and (5.6) have the matrix variate Gaussian distributions in (5.24) and (5.23). All results are obtained from averaging measurements over 500 channel realisations.

5.4.1 Comparison of Robust Transceiver Designs

In our first simulation example we compare the performance of the proposed robust DFE and THP transceivers with robust linear transceiver designs proposed in the literature. In these simulations the spatial correlation coefficients are set as $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.3$, $\forall l$, and the channel estimation error variances in (5.24) and (5.23) are set as $\sigma_{e_s}^2 = \sigma_{e_r}^2 = 0.0025$. We compare the proposed transceivers with the optimal relay precoded (ORP) and optimal source and relay precoded (OSRP) algorithms in [83], as well as the linear naive amplify forward (NAF) algorithm [45]. As a further benchmark we also include the performance of the proposed algorithms when the alternating power allocation (APA) algorithm discussed in Section 5.3.3 is replaced by an equal power allocation (EPA) algorithm. Figures 5.1 and 5.2 show the BER and MSE[1] performances of the proposed and benchmark designs against varying SNR_s(dB) with $\text{SNR}_\text{r} = 30$dB, whilst Figures 5.3 and 5.4 show the BER and MSE results for varying SNR_r(dB) with $\text{SNR}_\text{s} = 30$dB. The dashed curves in Figures 5.1 and 5.3 represent the theoretical performances of the DFE transceivers in the absence of error propagation. We observe that the proposed DFE and THP transceivers significantly outperform the linear benchmark designs in terms of BER and MSE, with the THP design providing the best BER performance. Although the DFE EPA and THP EPA algorithms utilise a suboptimal power allocation they also provide improved performance compared to the linear designs. We also observe that the linear ORP algorithm is outperformed by the linear OSRP algorithm which results from the fact that it does not utilise a source precoding matrix. It is worth mentioning that, of all the algorithms, the OSRP in [83] is the most computationally expensive since it is an iterative algorithm and in each iteration a matrix valued convex optimisation problem must be solved to update the source precoder.

[1]The MSE performance is the theoretical MSE given by $(1/K)\sum_{k=1}^{K} \text{tr}\{\boldsymbol{R}_{e,k}\}/N_k$. We note that since the error covariance matrices for DFE and THP transceivers is the same, they result in the same theoretical MSE performance.

5.4.2 Effect of Channel Estimation Error

In this simulation example we demonstrate the effect of channel estimation error and compare the performance of the proposed robust DFE and THP algorithms to non-robust DFE and THP solutions. The non-robust techniques assume that the estimated channels are accurate and therefore do not take into account the channel estimation errors. The non-robust designs can be obtained from the MMSE DFE and THP algorithm provided in Section 4.3 of Chapter 4 simply by replacing the actual subcarrier channels $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$ with their estimates $\bar{\boldsymbol{H}}_{s,k}$ and $\bar{\boldsymbol{H}}_{r,k}$. In these simulations the spatial correlation coefficients are set as $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, $\forall l$, and we show results for $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$. Figure 5.5 compares the BER performance of the robust DFE and non-robust DFE algorithms when varying $\mathrm{SNR_s}$(dB) with $\mathrm{SNR_r} = 30$dB. The corresponding BER results for the robust and non-robust THP algorithms are shown in Figure 5.6. The MSE results for these algorithms in this situation are shown in Figure 5.7. From these results we see that the proposed robust algorithms of this chapter outperform their non-robust counterparts for different levels of channel estimation error. For completeness we also include the results for varying $\mathrm{SNR_r}$(dB) with $\mathrm{SNR_s} = 30$dB, which are displayed in Figures 5.8-5.10.

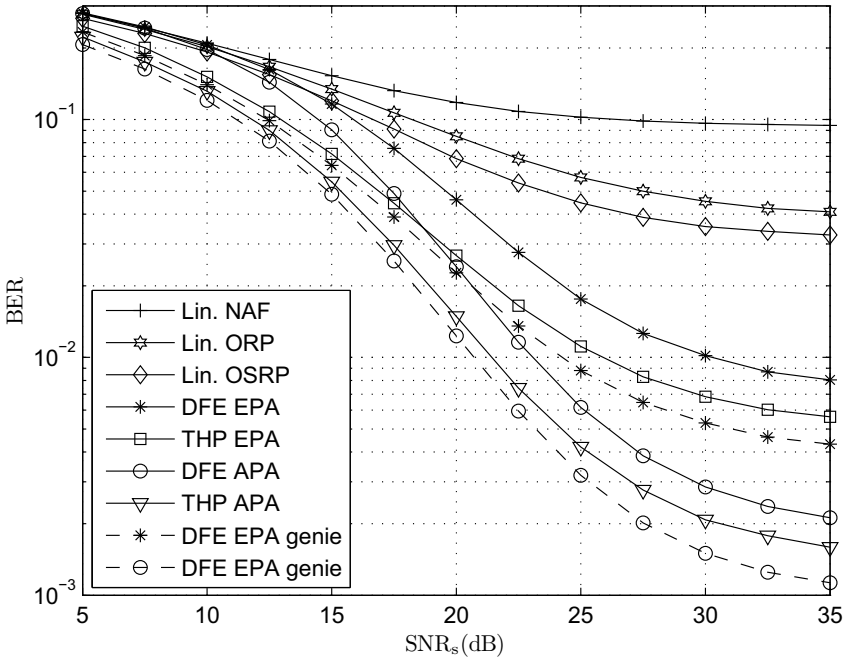

Figure 5.1: BER against varying $\mathrm{SNR_s(dB)}$ of robust linear and non-linear transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_r} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.3$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = 0.0025$.

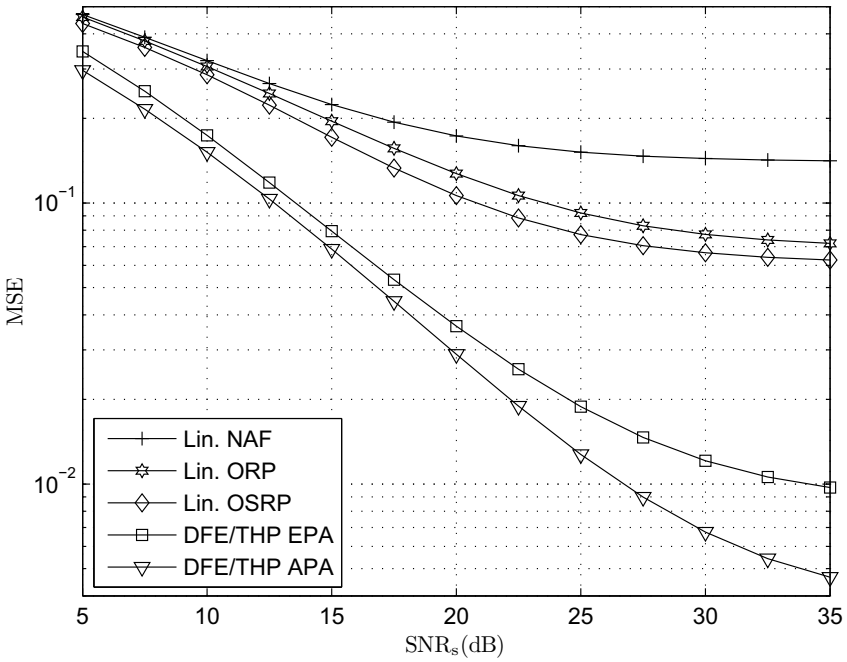

Figure 5.2: MSE against varying $\mathrm{SNR_s(dB)}$ of robust linear and non-linear transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_r} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.3$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = 0.0025$.

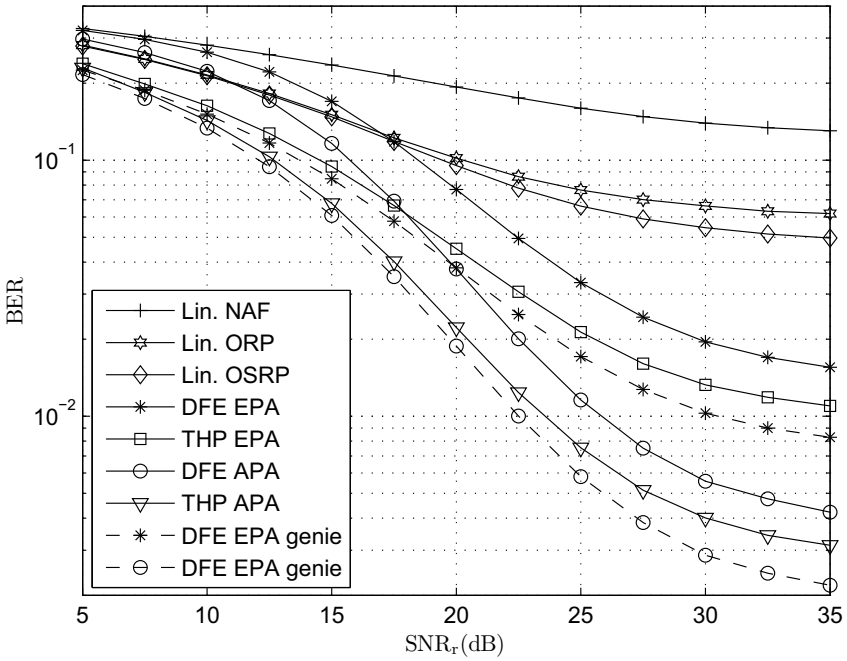

Figure 5.3: BER against varying $\mathrm{SNR_r(dB)}$ of robust linear and non-linear transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_s} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.3$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = 0.0025$.

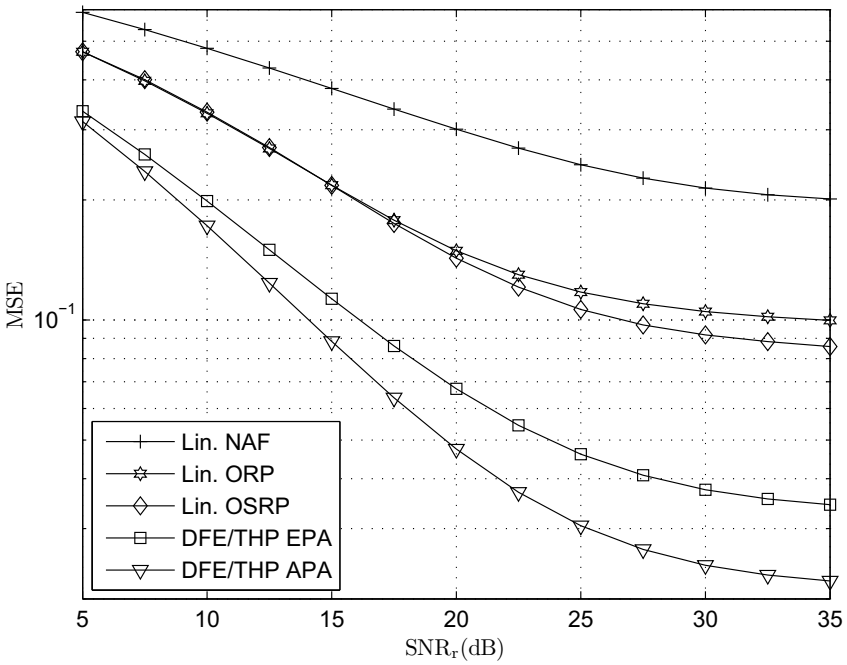

Figure 5.4: MSE against varying $\mathrm{SNR_r(dB)}$ of robust linear and non-linear transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_s} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.3$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = 0.0025$.

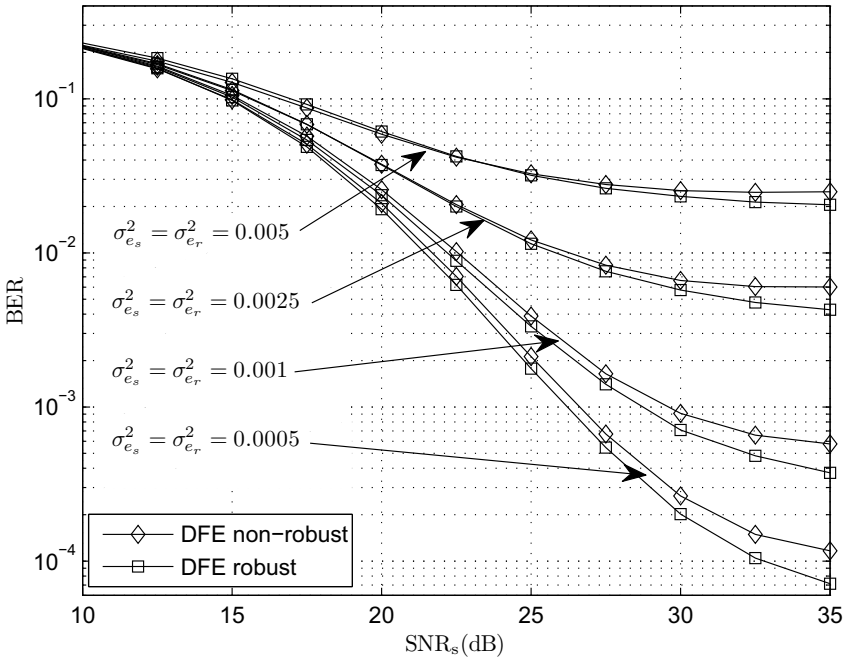

Figure 5.5: BER against varying SNR_s(dB) of robust and non-robust DFE transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\text{SNR}_\text{r} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$.

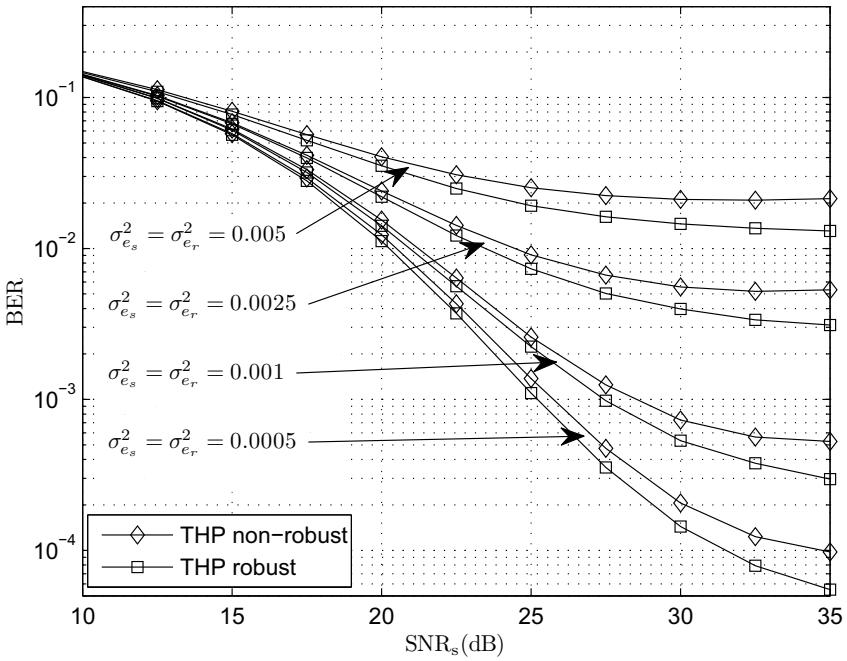

Figure 5.6: BER against varying SNR_s(dB) of robust and non-robust THP transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\text{SNR}_\text{r} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$.

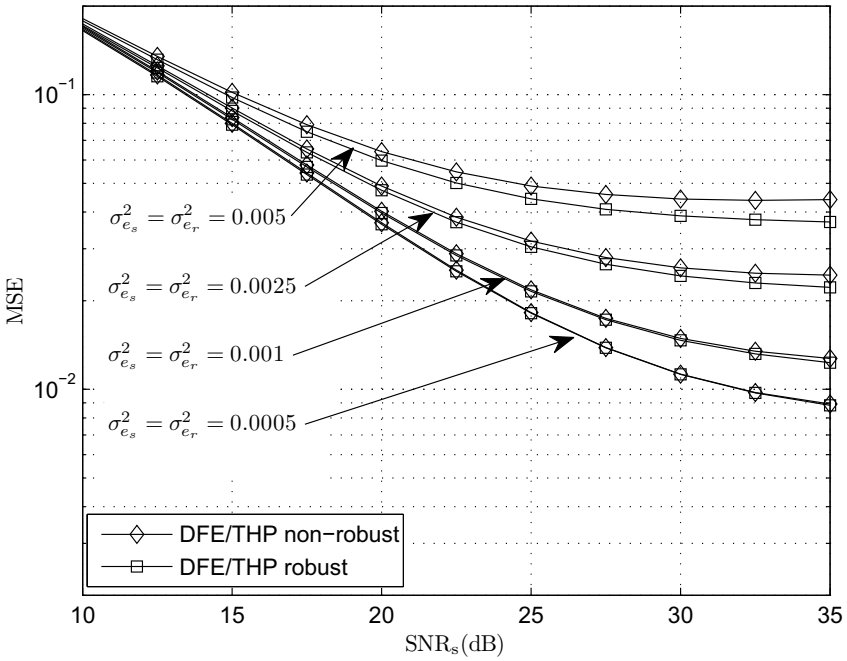

Figure 5.7: MSE against varying $\mathrm{SNR_s(dB)}$ of robust and non-robust DFE/THP transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_r} = 30\mathrm{dB}$, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$.

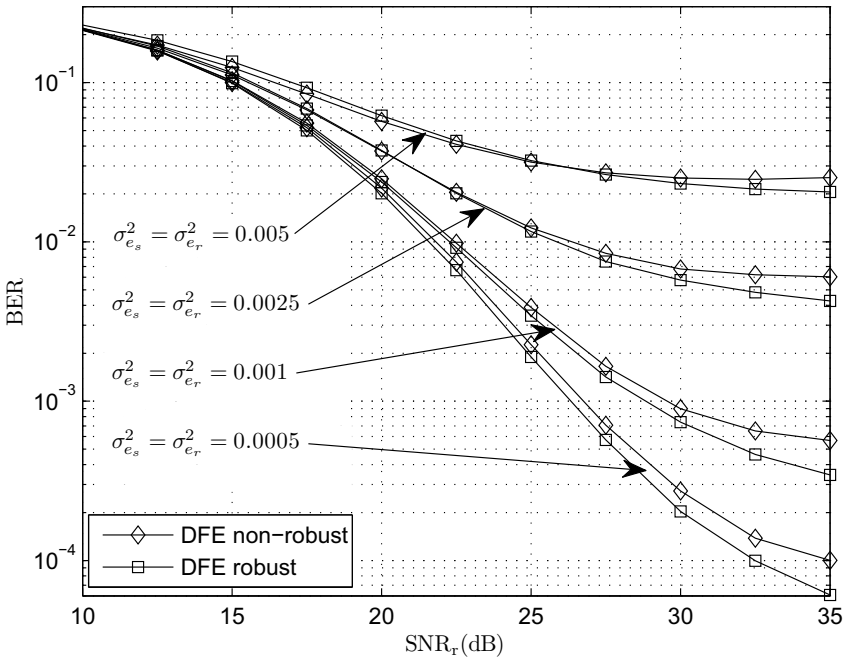

Figure 5.8: BER against varying $\mathrm{SNR_r(dB)}$ of robust and non-robust DFE transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_s} = 30\mathrm{dB}$, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$.

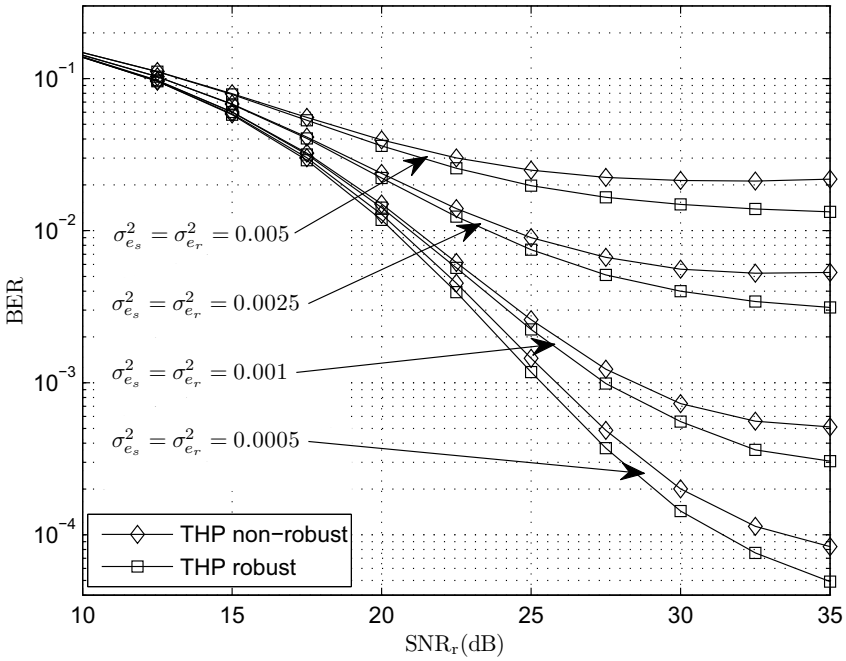

Figure 5.9: BER against varying $\mathrm{SNR_r(dB)}$ of robust and non-robust THP transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_s} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$.

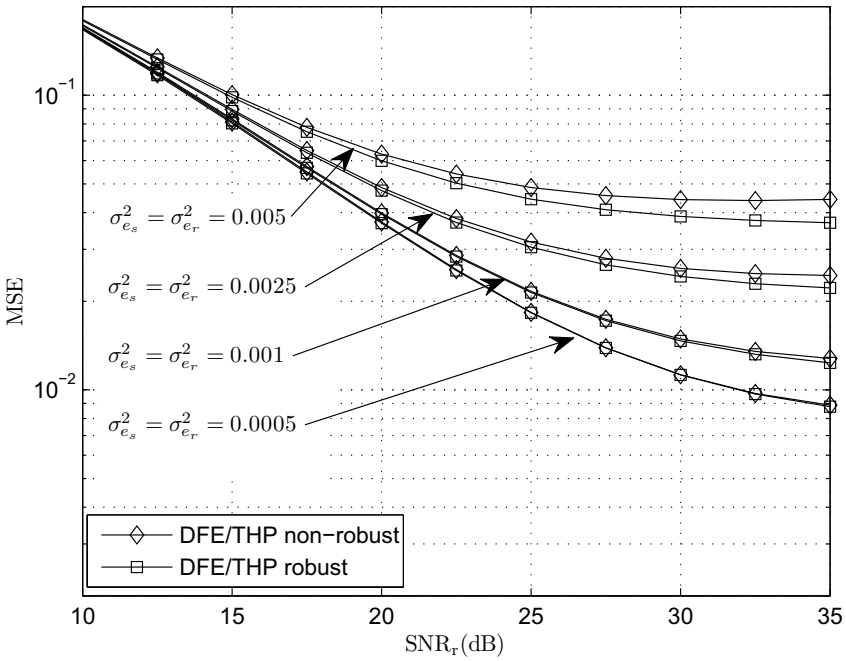

Figure 5.10: MSE against varying $\mathrm{SNR_r(dB)}$ of robust and non-robust DFE/THP transceivers for a system with $N_s = N_r = N_d = 3$, $L+1 = 5$, $\sigma^2_{h_s}[l] = \sigma^2_{h_r}[l] = 1/(L+1)$, $K = 32$, $N_k = 3$ 16-QAM symbols, $\mathrm{SNR_s} = 30$dB, $\rho_s[l] = \rho_r[l] = \varrho_s[l] = \varrho_r[l] = 0.5$, and $\sigma^2_{e_s} = \sigma^2_{e_r} = \{0.005, 0.0025, 0.001, 0.0005\}$.

5.5 Chapter Summary and Conclusions

In this chapter we considered the effects of CSI mismatch on the design of DFE and THP transceivers for MIMO OFDM relaying systems. We introduced a general error model where the actual source-relay and relay-destination subcarrier channels were modelled as the summation of the estimated channels and error matrices, with the additive error matrices assumed to have zero mean complex Gaussian distributions with known covariances. It was shown that two specific channel estimation algorithms resulted in the the channel estimation error matrices having such a structure. For the general channel estimation error model we derived the robust optimisation problem for minimising the arithmetic MSE subject to transmission power constraints. Unfortunately, for the general case, the resulting optimisation problem was extremely difficult to solve. In order to find an analytic solution we therefore considered a relaxation of the problem for which the source and relay precoder matrices could be more easily derived. The relaxation of the original optimisation problem resulted from deriving an upper bound to the original objective function and minimising the upper bound subject to the source power constraint and a relaxed relay power constraint. The source and relay precoding matrices were then derived for the relaxed optimisation problem. It was shown that for special cases of channel estimation error the relaxed problem was in fact equivalent to the original problem. Simulation results showed that the proposed robust non-linear solutions offered improved performance compared to robust linear techniques proposed in the literature. It was also demonstrated that the proposed algorithms were more robust to CSI mismatch compared to non-robust non-linear solutions that did not take into account the channel estimation errors, and simply considered the estimated channel to be the actual channel.

5.6 Chapter Derivations and Proofs

In this section we provide detailed derivations and proofs for this chapter that have been omitted from the main text.

5.6.1 Derivation of Expected Error Covariance Matrix

Here we prove the expression for $\mathbb{E}\{\boldsymbol{R}_{e,k}\}$ given in (5.30). Let us firstly rewrite the instantaneous error covariance matrix $\boldsymbol{R}_{e,k}$ in (5.4) as

$$\begin{aligned}\boldsymbol{R}_{e,k} = \boldsymbol{W}_k \left(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{X}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\boldsymbol{W}_k^H - \boldsymbol{W}_k\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H \\ - \boldsymbol{U}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H\boldsymbol{W}_k^H + \boldsymbol{U}_k\boldsymbol{U}_k^H,\end{aligned} \tag{5.95}$$

where for convenience we define the matrix $\boldsymbol{X}_k \in \mathbb{C}^{N_r\times N_r}$ as

$$\boldsymbol{X}_k \triangleq \boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H + \sigma_{v_s}^2\boldsymbol{I}_{N_r}. \tag{5.96}$$

Taking the expectation of (5.95) we require to compute

$$\begin{aligned}\mathbb{E}\{\boldsymbol{R}_{e,k}\} = \mathbb{E}\Big\{\boldsymbol{W}_k\left(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{X}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\boldsymbol{W}_k^H\Big\} \\ - \mathbb{E}\Big\{\boldsymbol{W}_k\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H\Big\} \\ - \mathbb{E}\Big\{\boldsymbol{U}_k\boldsymbol{F}_k^H\boldsymbol{H}_{s,k}^H\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H\boldsymbol{W}_k^H\Big\} + \boldsymbol{U}_k\boldsymbol{U}_k^H,\end{aligned} \tag{5.97}$$

where it should be understood that the expectations shall be taken w.r.t. the random matrices $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$, which are related to the subcarrier channels $\boldsymbol{H}_{s,k}$ and $\boldsymbol{H}_{r,k}$ through (5.5) and (5.6), respectively[2]. We focus firstly on calculating the term $\mathbb{E}\{\boldsymbol{W}_k\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H\}$ in (5.97). Substituting (5.5) and (5.6) into this expression it is straightforward to show that

$$\begin{aligned}&\mathbb{E}\Big\{\boldsymbol{W}_k\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{H}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H\Big\} = \boldsymbol{W}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H + \boldsymbol{W}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\mathbb{E}\{\boldsymbol{E}_{s,k}\}\boldsymbol{F}_k\boldsymbol{U}_k^H \\ &+ \boldsymbol{W}_k\mathbb{E}\{\boldsymbol{E}_{r,k}\}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H + \boldsymbol{W}_k\mathbb{E}\{\boldsymbol{E}_{r,k}\}\boldsymbol{G}_k\mathbb{E}\{\boldsymbol{E}_{s,k}\}\boldsymbol{F}_k\boldsymbol{U}_k^H. \qquad (5.98) \\ &= \boldsymbol{W}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{U}_k^H \qquad (5.99)\end{aligned}$$

where to obtain the right hand side of (5.98) we have used the fact that $\bar{\boldsymbol{H}}_{s,k}$ and $\bar{\boldsymbol{H}}_{r,k}$ are deterministic quantities, and the fact that $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$ are independent. To obtain (5.99) from (5.98) we have used the results $\mathbb{E}\{\boldsymbol{E}_{s,k}\} = \boldsymbol{0}_{N_r\times N_s}$ and

[2]Note that in the expression (5.97) an expectation is not taken w.r.t. the term $\boldsymbol{U}_k\boldsymbol{U}_k^H$ since, as it is not a function of the random matrices $\boldsymbol{E}_{s,k}$ and/or $\boldsymbol{E}_{r,k}$, it is considered a deterministic quantity.

$\mathbb{E}\{\boldsymbol{E}_{r,k}\} = \boldsymbol{0}_{N_d \times N_r}$ from (5.26) and (5.27), respectively. Taking the Hermitian transpose of the left hand side of (5.98) and (5.99) we also have

$$\mathbb{E}\Big\{\boldsymbol{U}_k \boldsymbol{F}_k^H \boldsymbol{H}_{s,k}^H \boldsymbol{G}_k^H \boldsymbol{H}_{r,k}^H \boldsymbol{W}_k^H\Big\} = \boldsymbol{U}_k \boldsymbol{F}_k^H \bar{\boldsymbol{H}}_{s,k}^H \boldsymbol{G}_k^H \bar{\boldsymbol{H}}_{r,k}^H \boldsymbol{W}_k^H, \tag{5.100}$$

which provides the third expectation term in (5.97). We lastly require to compute the term $\mathbb{E}\{\boldsymbol{W}_k(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{X}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d})\boldsymbol{W}_k^H\}$ in (5.97). Since only $\boldsymbol{X}_k$ and $\boldsymbol{H}_{r,k}$ contain the random matrices $\boldsymbol{E}_{s,k}$ and $\boldsymbol{E}_{r,k}$, this expectation is given by

$$\begin{aligned}&\mathbb{E}\Big\{\boldsymbol{W}_k\left(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\boldsymbol{X}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\boldsymbol{W}_k^H\Big\}\\ &= \boldsymbol{W}_k\mathbb{E}\Big\{\left(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\mathbb{E}\{\boldsymbol{X}_k\}\,\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\Big\}\boldsymbol{W}_k^H.\end{aligned} \tag{5.101}$$

With regards to the inner expectation $\mathbb{E}\{\boldsymbol{X}_k\}$ in (5.101), by substituting (5.5) into (5.96) and taking the expectation w.r.t. $\boldsymbol{E}_{s,k}$ we have

$$\mathbb{E}\{\boldsymbol{X}_k\} = \bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathbb{E}\Big\{\boldsymbol{E}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{E}_{s,k}^H\Big\} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \tag{5.102}$$

$$= \bar{\boldsymbol{H}}_{s,k}\boldsymbol{F}_k\boldsymbol{F}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathrm{tr}\Big\{\boldsymbol{F}_k\boldsymbol{F}_k^H\boldsymbol{\Xi}_{s,k}^T\Big\}\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \triangleq \bar{\boldsymbol{X}}_k. \tag{5.103}$$

To obtain (5.103) we have used the result in (5.28), which was obtained based on the matrix variate Gaussian distribution of $\boldsymbol{E}_{s,k}$ given in (5.10) as well as Lemma 16. Substituting (5.103) into (5.101), the remaining task now is to compute $\boldsymbol{W}_k\mathbb{E}\{(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d})\}\boldsymbol{W}_k^H$. Using (5.6) as well as the result in (5.29) we can expand this term as

$$\begin{aligned}&\boldsymbol{W}_k\mathbb{E}\Big\{\left(\boldsymbol{H}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{H}_{r,k}^H + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\Big\}\boldsymbol{W}_k^H\\ &= \boldsymbol{W}_k\left(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \mathbb{E}\Big\{\bar{\boldsymbol{E}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{E}}_{r,k}^H\Big\} + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\boldsymbol{W}_k^H \qquad (5.104)\\ &= \boldsymbol{W}_k\left(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \mathrm{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{\Xi}_{r,k}^T\Big\}\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d}\right)\boldsymbol{W}_k^H. \qquad (5.105)\end{aligned}$$

Finally, by substituting (5.99), (5.100), and (5.105) into (5.97) we prove $\mathbb{E}\{\boldsymbol{R}_{e,k}\}$ as given in (5.30).

5.6.2 Proof of Upper Bounds for Objective Function and Relay Power Constraint

In this section we prove the upper bounds in (5.51) and (5.52). To do so we require the following lemmas:

Lemma 18: [82] For Hermitian positive semi-definite matrices $\boldsymbol{A} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{B} \in \mathbb{C}^{N\times N}$, with eigenvalues $\{\lambda_{a,i}\}_{i=1}^{N} \in \mathbb{R}_+$ and $\{\lambda_{b,i}\}_{i=1}^{N} \in \mathbb{R}_+$ arranged in descending order we have the inequality $\mathrm{tr}\{\boldsymbol{AB}\} \leq \mathrm{tr}\{\boldsymbol{A}\}\,\lambda_{b,1}$.

Lemma 19: [97] For Hermitian positive semi-definite matrices $\boldsymbol{A}_1 \in \mathbb{C}^{N\times N}$ and $\boldsymbol{A}_2 \in \mathbb{C}^{N\times N}$, where $\boldsymbol{A}_1 \preceq \boldsymbol{A}_2$, then given an arbitrary matrix $\boldsymbol{B} \in \mathbb{C}^{M\times N}$ the following inequality holds $\boldsymbol{BA}_1\boldsymbol{B}^H \preceq \boldsymbol{BA}_2\boldsymbol{B}^H$.

Lemma 20: [97] For Hermitian positive semi-definite matrices $\boldsymbol{A}_1 \in \mathbb{C}^{N\times N}$ and $\boldsymbol{A}_2 \in \mathbb{C}^{N\times N}$, if $\boldsymbol{A}_1 \preceq \boldsymbol{A}_2$, then we also have $\boldsymbol{A}_1^{-1} \succeq \boldsymbol{A}_2^{-1}$.

We now set out to prove (5.51). Using Lemma 18, and the fact that $\boldsymbol{\Xi}_{s,k}$ is a Hermitian positive semi-definite matrix, we can state the following inequality

$$\boldsymbol{J}_k \triangleq \mathrm{tr}\Big\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\boldsymbol{\Xi}_{s,k}^T\Big\}\,\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \tag{5.106}$$

$$\preceq \mathrm{tr}\Big\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\Big\}\,\xi_{s,k}\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \triangleq \bar{\boldsymbol{J}}_k, \tag{5.107}$$

where $\xi_{s,k}$ is the largest eigenvalue of $\boldsymbol{\Xi}_{s,k}$. Substituting (5.107) into (5.31) it is straightforward to show that

$$\bar{\boldsymbol{X}}_k = \bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathrm{tr}\Big\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\boldsymbol{\Xi}_{s,k}^T\Big\}\,\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \tag{5.108}$$

$$\preceq \bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \mathrm{tr}\Big\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\Big\}\,\xi_{s,k}\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \triangleq \tilde{\boldsymbol{X}}_k. \tag{5.109}$$

Using (5.109), along with the properties of the trace operator and Lemma 19, it can be shown that

$$\boldsymbol{K}_k \triangleq \mathrm{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{\Xi}_{r,k}^T\Big\}\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d} \tag{5.110}$$

$$\preceq \mathrm{tr}\Big\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\boldsymbol{\Xi}_{r,k}^T\Big\}\,\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d} \tag{5.111}$$

$$\preceq \mathrm{tr}\Big\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\Big\}\,\xi_{r,k}\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d} \triangleq \bar{\boldsymbol{K}}_k, \tag{5.112}$$

where $\xi_{r,k}$ is the largest eigenvalue of $\boldsymbol{\Xi}_{r,k}$ and we have applied Lemma 18 to (5.111) to obtain (5.112). Using the definitions of the positive semi-definite Hermitian matrices $\boldsymbol{J}_k$, $\bar{\boldsymbol{J}}_k$, $\boldsymbol{K}_k$, and $\bar{\boldsymbol{K}}_k$, in (5.106), (5.107), (5.110), and (5.112), we can state that

$$\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\boldsymbol{J}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \boldsymbol{K}_k \preceq \bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{J}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \bar{\boldsymbol{K}}_k, \tag{5.113}$$

where we have made use of Lemma 19. Furthermore, inverting both sides of (5.113), and applying Lemma 20 we have

$$\begin{aligned}\boldsymbol{L}_k &\triangleq \left(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\boldsymbol{J}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H+\boldsymbol{K}_k\right)^{-1}\\ &\succeq \left(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{J}}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H+\bar{\boldsymbol{K}}_k\right)^{-1}\triangleq\bar{\boldsymbol{L}}_k. \end{aligned}\tag{5.114}$$

Using (5.114) as well as Lemma 19 we have

$$\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\boldsymbol{L}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\succeq\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\bar{\boldsymbol{L}}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k.\tag{5.115}$$

Adding $\boldsymbol{I}_{N_k}$ to both sides of (5.115) and taking the inverse of both sides we finally arrive at

$$\begin{aligned}&\left(\boldsymbol{I}_{N_k}+\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\boldsymbol{L}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\right)^{-1}\\ \preceq&\left(\boldsymbol{I}_{N_k}+\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\bar{\boldsymbol{L}}_k\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\right)^{-1},\end{aligned}\tag{5.116}$$

where again we have utilised Lemma 20. With $\boldsymbol{E}_k$ and $\tilde{\boldsymbol{E}}_k$ given in (5.43) and (5.53) we see that (5.116) is equivalent to $\boldsymbol{E}_k\preceq\tilde{\boldsymbol{E}}_k$. Since $|\boldsymbol{A}|\leq|\boldsymbol{B}|$, for any positive semi-definite Hermitian matrices satisfying $\boldsymbol{A}\preceq\boldsymbol{B}$, we find that taking the determinant of both sides of (5.116) we have $|\boldsymbol{E}_k|\leq|\tilde{\boldsymbol{E}}_k|$. The inequality in (5.51) then directly follows.

We now prove the inequality in (5.52). Using the relationship $\bar{\boldsymbol{X}}_k\preceq\tilde{\boldsymbol{X}}_k$ from (5.109) as well as Lemma 19 we have

$$\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\preceq\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H.\tag{5.117}$$

Using the fact that $\mathrm{tr}\{\boldsymbol{A}\}\leq\mathrm{tr}\{\boldsymbol{B}\}$ for positive semi-definite matrices $\boldsymbol{A}\preceq\boldsymbol{B}$ we have from (5.117) that

$$\mathrm{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{X}}_k\boldsymbol{G}_k^H\Big\}\leq\mathrm{tr}\Big\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\Big\}.\tag{5.118}$$

Finally by taking the summation of both sides of (5.118) over $k=1...K$ we prove the inequality in (5.52).

5.6.3 Proof of Power Constraint Equalities

In this section we prove that the source and relay precoding processors $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ should satisfy the power constraint equalities in (5.58)-(5.61). Proving the constraints in (5.58) and (5.59) relies on showing that $|\tilde{\boldsymbol{E}}_k|$ in (5.53) is a decreasing

function of $\mathrm{tr}\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\}$. To this end let us write

$$\tilde{\boldsymbol{F}}_k = \sqrt{\mathrm{tr}\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\}}\breve{\boldsymbol{F}}_k, \tag{5.119}$$

where $\breve{\boldsymbol{F}}_k$ satisfies $\mathrm{tr}\{\breve{\boldsymbol{F}}_k\breve{\boldsymbol{F}}_k^H\} = 1$. Substituting (5.119) into (5.53) we can write

$$|\tilde{\boldsymbol{E}}_k| = \left|\boldsymbol{I}_{N_k} + \breve{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H\left(\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\boldsymbol{M}_k\boldsymbol{G}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \boldsymbol{N}_k\right)^{-1}\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{H}}_{s,k}\breve{\boldsymbol{F}}_k\right|^{-1}, \tag{5.120}$$

where we define the new variables

$$\boldsymbol{M}_k \triangleq \xi_{s,k}\boldsymbol{\Omega}_{s,k} + \frac{\sigma_{v_s}^2}{\mathrm{tr}\left\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\right\}}\boldsymbol{I}_{N_r} \tag{5.121}$$

$$\boldsymbol{N}_k \triangleq \mathrm{tr}\left\{\boldsymbol{G}_k\breve{\boldsymbol{X}}_k\boldsymbol{G}_k^H\right\}\xi_{r,k}\boldsymbol{\Omega}_{r,k} + \frac{\sigma_{v_r}^2}{\mathrm{tr}\left\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\right\}}\boldsymbol{I}_{N_d} \tag{5.122}$$

$$\breve{\boldsymbol{X}}_k \triangleq \bar{\boldsymbol{H}}_{s,k}\breve{\boldsymbol{F}}_k\breve{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H + \boldsymbol{M}_k. \tag{5.123}$$

For a given $\breve{\boldsymbol{F}}_k$ it can now be seen that (5.120) is a decreasing function of $\mathrm{tr}\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\}$. Thus, given a maximum available power budget $P_{s,k}$ available to the kth subcarrier, it is straightforward to deduce that the optimal $\tilde{\boldsymbol{F}}_k$ that minimises (5.120) should satisfy $\mathrm{tr}\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\} = P_{s,k}$. This proves the condition in (5.58). To prove the condition in (5.59) we note that if the total power budget $P_s \neq \sum_{k=1}^{N_c} P_{s,k}$ then allocating the unused power from P_s to the subcarriers can only further minimise (5.120). We should thus have $\sum_{k=1}^{N_c} P_{s,k} = P_s$ as in (5.59).

In a similar fashion, to show (5.60) and (5.61) we write

$$\boldsymbol{G}_k = \sqrt{\mathrm{tr}\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\}}\breve{\boldsymbol{G}}_k, \tag{5.124}$$

where $\breve{\boldsymbol{G}}_k$ satisfies $\mathrm{tr}\{\breve{\boldsymbol{G}}_k\tilde{\boldsymbol{X}}_k\breve{\boldsymbol{G}}_k^H\} = 1$. Substituting (5.124) into (5.53) we have

$$|\tilde{\boldsymbol{E}}_k| = \left|\boldsymbol{I}_{N_k} + \tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\breve{\boldsymbol{G}}_k^H\bar{\boldsymbol{H}}_{r,k}^H\left(\bar{\boldsymbol{H}}_{r,k}\breve{\boldsymbol{G}}_k\breve{\boldsymbol{M}}_k\breve{\boldsymbol{G}}_k^H\bar{\boldsymbol{H}}_{r,k}^H + \breve{\boldsymbol{N}}_k\right)^{-1}\bar{\boldsymbol{H}}_{r,k}\breve{\boldsymbol{G}}_k\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\right|^{-1}, \tag{5.125}$$

where we have defined the new variables

$$\breve{\boldsymbol{M}}_k \triangleq \mathrm{tr}\Big\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\Big\}\,\xi_{s,k}\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r} \tag{5.126}$$

$$\breve{\boldsymbol{N}}_k \triangleq \xi_{r,k}\boldsymbol{\Omega}_{r,k} + \frac{\sigma_{v_r}^2}{\mathrm{tr}\Big\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\Big\}}\boldsymbol{I}_{N_d}. \tag{5.127}$$

It can be seen that (5.125) is a decreasing function of $\mathrm{tr}\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\}$. Following the same arguments made previously for $\tilde{\boldsymbol{F}}_k$, we can straightforwardly show that $\boldsymbol{G}_k$ should satisfy the conditions in (5.60) and (5.61).

5.6.4 Proof of Source and Relay Precoder Structures

In this section we provide the proof of Theorem 6. We begin by showing that the optimal source and relay processing matrices $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ are given by (5.71) and (5.72). In order to do so we shall require the following tools:

Definition 9: For a vector $\boldsymbol{a} \in \mathbb{R}^M$ with elements $\{a_i\}_{i=1}^M \in \mathbb{R}$, let us denote $\{a_{[i]}\}_{i=1}^M$ as being the elements of $\boldsymbol{a}$ arranged in descending order i.e.

$$a_{[i]} \geq a_{[i+1]}, \qquad 1 \leq i \leq M-1. \tag{5.128}$$

Definition 10: [39] For vectors $\boldsymbol{a} \in \mathbb{R}^M$ and $\boldsymbol{b} \in \mathbb{R}^M$ we say that $\boldsymbol{a}$ is weakly additively submajorised by $\boldsymbol{b}$, denoted as $\boldsymbol{a} \preceq_w^+ \boldsymbol{b}$, if

$$\sum_{i=1}^{m} a_{[i]} \leq \sum_{i=1}^{m} b_{[i]}, \qquad 1 \leq m \leq M. \tag{5.129}$$

Definition 11: [39] For vectors $\boldsymbol{a} \in \mathbb{R}_+^M$ and $\boldsymbol{b} \in \mathbb{R}_+^M$ we say that $\boldsymbol{a}$ is weakly multiplicatively submajorised by $\boldsymbol{b}$, denoted as $\boldsymbol{a} \preceq_w^\times \boldsymbol{b}$, if

$$\prod_{i=1}^{m} a_{[i]} \leq \prod_{i=1}^{m} b_{[i]}, \qquad 1 \leq m \leq M. \tag{5.130}$$

Lemma 21: [39] For $\boldsymbol{a} \in \mathbb{R}^M$ and $\boldsymbol{b} \in \mathbb{R}^M$ satisfying $\boldsymbol{a} \preceq_w^+ \boldsymbol{b}$, then for any convex increasing function $g(.)$ we have $[g(a_1), ..., g(a_M)]^T \preceq_w^+ [g(b_1), ..., g(b_M)]^T$ i.e.

$$\sum_{i=1}^{m} g\left(a_{[i]}\right) \leq \sum_{i=1}^{m} g\left(b_{[i]}\right), \qquad 1 \leq m \leq M. \tag{5.131}$$

Lemma 22: [39] For $\boldsymbol{a} \in \mathbb{R}_{++}^{M}$ and $\boldsymbol{b} \in \mathbb{R}_{++}^{M}$ then $\boldsymbol{a} \preceq_{w}^{\times} \boldsymbol{b}$ is equivalent to saying that $[\log(a_1), ..., \log(a_M)]^T \preceq_{w}^{+} [\log(b_1), ..., \log(b_M)]^T$ and vice versa. In other words $\boldsymbol{a}$ and $\boldsymbol{b}$ satisfy

$$\sum_{i=1}^{m} \log\left(a_{[i]}\right) \leq \sum_{i=1}^{m} \log\left(b_{[i]}\right), \qquad 1 \leq m \leq M. \tag{5.132}$$

Lemma 23: [39] For $\boldsymbol{B} \triangleq (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)^H (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)$, we have the weak multiplicative submajorisation $[\lambda_{b,1}, ..., \lambda_{b,M}]^T \preceq_{w}^{\times} [\prod_{n=1}^{N} \sigma_{A_n,1}^2, ..., \prod_{n=1}^{N} \sigma_{A_n,M}^2]^T$, where $\{\lambda_{b,i}\}_{i=1}^{M} \in \mathbb{R}_{++}$ is the ith largest eigenvalue of $\boldsymbol{B}$, $\{\sigma_{A_n,i}\}_{i=1}^{M} \in \mathbb{R}_{++}$ is the ith largest singular value of $\boldsymbol{A}_n$, and $M \triangleq \min(\text{rank}\{\boldsymbol{A}_1\}, ..., \text{rank}\{\boldsymbol{A}_N\})$. This means that

$$\prod_{i=1}^{m} \lambda_{b,i} \leq \prod_{i=1}^{m} \prod_{n=1}^{N} \sigma_{A_n,i}^2, \qquad 1 \leq m \leq M. \tag{5.133}$$

Lemma 24: For $\boldsymbol{B} \triangleq (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)^H (\bigotimes_{n=1}^{N} \boldsymbol{A}_n)$ we have that

$$|\boldsymbol{I} + \boldsymbol{B}|^{-1} \geq \prod_{i=1}^{M} \left(1 + \prod_{n=1}^{N} \sigma_{A_n,i}^2\right)^{-1}, \tag{5.134}$$

where again, $\{\sigma_{A_n,i}\}_{i=1}^{M} \in \mathbb{R}_{++}$ is the ith largest singular value of $\boldsymbol{A}_n$, and we define the scalar $M \triangleq \min(\text{rank}\{\boldsymbol{A}_1\}, ..., \text{rank}\{\boldsymbol{A}_N\})$.

Lemma 25: [78] For positive semi-definite Hermitian matrices $\boldsymbol{A} \in \mathbb{C}^{N \times N}$ and $\boldsymbol{B} \in \mathbb{C}^{N \times N}$, with eigenvalues $\{\lambda_{a,i}\}_{i=1}^{N} \in \mathbb{R}_{+}$ and $\{\lambda_{b,i}\}_{i=1}^{N} \in \mathbb{R}_{+}$ arranged in descending order, we have the inequality

$$\text{tr}\{\boldsymbol{AB}\} \geq \sum_{i=0}^{N} \lambda_{a,i} \lambda_{b,N+1-i}. \tag{5.135}$$

We now set out to prove the structure of $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ given in Theorem 6. We start by finding a family of processors that minimise the objective function in (5.62). To this end it is sufficient to find $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ that minimise $|\tilde{\boldsymbol{E}}_k|$. From the family of precoders that minimise $|\tilde{\boldsymbol{E}}_k|$ we select specific processors that also minimise the source and relay power consumption.

Substituting (5.58) into $\bar{\boldsymbol{J}}_k$ in (5.107) and substituting (5.60) into the definition of $\bar{\boldsymbol{K}}_k$ in (5.112) we can write

$$\bar{\boldsymbol{J}}_k = P_{s,k}\xi_{s,k}\boldsymbol{\Omega}_{s,k} + \sigma_{v_s}^2\boldsymbol{I}_{N_r}, \tag{5.136}$$

$$\bar{\boldsymbol{K}}_k = P_{r,k}\xi_{r,k}\boldsymbol{\Omega}_{r,k} + \sigma_{v_r}^2\boldsymbol{I}_{N_d}. \tag{5.137}$$

By further substituting (5.136) and (5.137) into (5.68), after some straightforward deductions, we can write $|\tilde{\boldsymbol{E}}_k|$ as

$$|\tilde{\boldsymbol{E}}_k| = \left|\boldsymbol{I}_{N_k} + \boldsymbol{Y}_k^H\boldsymbol{Z}_k^H\left(\boldsymbol{Z}_k\boldsymbol{Z}_k^H + \boldsymbol{I}_{N_d}\right)^{-1}\boldsymbol{Z}_k\boldsymbol{Y}_k\right|^{-1}, \tag{5.138}$$

where $\boldsymbol{Y}_k \triangleq \bar{\boldsymbol{J}}_k^{-1/2}\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{Z}_k \triangleq \bar{\boldsymbol{K}}_k^{-1/2}\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{J}}_k^{1/2}$. We now consider the decompositions

$$\boldsymbol{Y}_k = \bar{\boldsymbol{J}}_k^{-1/2}\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k = \boldsymbol{U}_{y,k}\boldsymbol{D}_k\boldsymbol{V}_{y,k}^H \tag{5.139}$$

$$\boldsymbol{Z}_k = \bar{\boldsymbol{K}}_k^{-1/2}\bar{\boldsymbol{H}}_{r,k}\boldsymbol{G}_k\bar{\boldsymbol{J}}_k^{1/2} = \boldsymbol{U}_{z,k}\boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H, \tag{5.140}$$

where $\boldsymbol{U}_{y,k} \in \mathbb{C}^{N_r\times N_r}$, $\boldsymbol{V}_{y,k} \in \mathbb{C}^{N_k\times N_k}$, $\boldsymbol{U}_{z,k} \in \mathbb{C}^{N_d\times N_d}$, and $\boldsymbol{V}_{z,k} \in \mathbb{C}^{N_r\times N_r}$, are unitary. The matrices $\boldsymbol{D}_k \in \mathbb{C}^{N_r\times N_k}$ and $\boldsymbol{\Sigma}_k \in \mathbb{C}^{N_d\times N_k}$ are diagonal and contain the singular values of $\boldsymbol{Y}_k$ and $\boldsymbol{Z}_k$, respectively, which are assumed to be in decreasing order. Let us now define the variable $R_{s,k} \triangleq \text{rank}\{\bar{\boldsymbol{J}}_k^{-1/2}\bar{\boldsymbol{H}}_{s,k}\}$. Noting that $\text{rank}\{\tilde{\boldsymbol{F}}_k\} \leq \min(N_s, N_k)$ and $R_{s,k} \leq N_s$, we can straightforwardly show from (5.139) that $\text{rank}\{\boldsymbol{Y}_k\} = \text{rank}\{\boldsymbol{D}_k\} \leq \min(R_{s,k}, N_k) \triangleq R_{y,k}$. Similarly, by defining $R_{r,k} \triangleq \text{rank}\{\bar{\boldsymbol{K}}_k^{-1/2}\bar{\boldsymbol{H}}_{r,k}\}$, we see from (5.140) that $\text{rank}\{\boldsymbol{Z}_k\} = \text{rank}\{\boldsymbol{\Sigma}_k\} \leq R_{r,k}$. With these observations and substituting (5.139) and (5.140) into (5.138) we can state that

$$|\tilde{\boldsymbol{E}}_k| = \left|\boldsymbol{I}_{N_k} + \boldsymbol{D}_k^T\boldsymbol{U}_{y,k}^H\boldsymbol{V}_{z,k}\boldsymbol{\Sigma}_k^T\left(\boldsymbol{\Sigma}_k\boldsymbol{\Sigma}_k^T + \boldsymbol{I}_{N_d}\right)^{-1}\boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H\boldsymbol{U}_{y,k}\boldsymbol{D}_k\right|^{-1} \tag{5.141}$$

$$\geq \prod_{i=1}^{\bar{N}_k}\left(1 + \frac{d_{k,i}^2\sigma_{k,i}^2}{\sigma_{k,i}^2 + 1}\right)^{-1}, \tag{5.142}$$

where we have used Lemma 24 to obtain the lower bound in (5.142) where $\{d_{k,i}\}_{i=1}^{R_{y,k}}$ and $\{\sigma_{k,i}\}_{i=1}^{R_{r,k}}$ are the singular values in $\boldsymbol{D}_k$ and $\boldsymbol{\Sigma}_k$ respectively. In (5.142) we also define $\bar{N}_k \triangleq \min\{N_k, R_{s,k}, R_{r,k}\}$. We see that (5.141) is invariant to the unitary matrix $\boldsymbol{V}_{y,k}$. It shall also be seen later that the source and relay power consumptions are also independent of $\boldsymbol{V}_{y,k}$ and we can assume w.l.o.g. that $\boldsymbol{V}_{y,k} = \boldsymbol{I}_{N_k}$. We also note that (5.142) holds with equality when $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. Substituting the decompositions of (5.69) and (5.70) into (5.139) and (5.140) (c.f. (5.136) and

(5.137)) and solving the resulting equations for $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ we have

$$\tilde{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{T}_{1,k}\boldsymbol{D}_k\boldsymbol{V}_{y,k}^H \tag{5.143}$$

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Delta}_k^{\dagger}\boldsymbol{T}_{2,k}\boldsymbol{\Sigma}_k\boldsymbol{V}_{z,k}^H\bar{\boldsymbol{J}}_k^{-1/2}, \tag{5.144}$$

where $\boldsymbol{T}_{1,k} \triangleq \boldsymbol{U}_{s,k}^H\boldsymbol{U}_{y,k}$ and $\boldsymbol{T}_{2,k} \triangleq \boldsymbol{U}_{r,k}^H\boldsymbol{U}_{z,k}$. The family of precoders given in (5.143) and (5.144) achieve the lower bound of (5.142), and therefore minimise $|\tilde{\boldsymbol{E}}_k|$, provided that $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. We note that the lower bound in (5.142) depends only on the largest $\bar{N}_k$ singular values of $\boldsymbol{D}_k$ and $\boldsymbol{\Sigma}_k$ and we can assume w.l.o.g. that $\text{rank}\{\boldsymbol{D}_k\} \leq \bar{N}_k$ and $\text{rank}\{\boldsymbol{\Sigma}_k\} \leq \bar{N}_k$. From (5.143) and (5.144) it is then obvious that we should have $\text{rank}\{\tilde{\boldsymbol{F}}_k\} \leq \bar{N}_k$ and $\text{rank}\{\boldsymbol{G}_k\} \leq \bar{N}_k$. With this observation and using (5.143), the power consumed by the source on the kth subcarrier is

$$\text{tr}\Big\{\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\Big\} = \text{tr}\Big\{\boldsymbol{T}_{1,k}^H(\boldsymbol{\Lambda}_k^{\dagger})^T\boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{T}_{1,k}\boldsymbol{D}_k\boldsymbol{D}_k^T\Big\} \tag{5.145}$$

$$\geq \sum_{i=1}^{\bar{N}_k}\lambda_{k,i}^{-2}d_{k,i}^2, \tag{5.146}$$

where the lower bound in (5.146) is obtained by applying Lemma 25 to (5.145). From (5.146) and (5.145) we see that the power consumed by the source on the kth subcarrier is minimised with $\boldsymbol{T}_{1,k} = \boldsymbol{I}_{N_r}$, which holds for $\boldsymbol{U}_{s,k} = \boldsymbol{U}_{y,k}$. Substituting this result and $\boldsymbol{V}_{y,k} = \boldsymbol{I}_{N_k}$ into (5.143) we arrive at the source precoder structure

$$\tilde{\boldsymbol{F}}_k = \boldsymbol{V}_{s,k}\boldsymbol{\Gamma}_k, \tag{5.147}$$

where we define the diagonal matrix $\boldsymbol{\Gamma}_k \triangleq \boldsymbol{\Lambda}_k^{\dagger}\boldsymbol{D}_k$. This proves the source precoder given in Theorem 6.

We now turn our attention to $\text{tr}\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\}$. Using $\bar{\boldsymbol{J}}_k$ in (5.136) and $\tilde{\boldsymbol{X}}_k$ in (5.54) we can show that

$$\text{tr}\Big\{\boldsymbol{G}_k\tilde{\boldsymbol{X}}_k\boldsymbol{G}_k^H\Big\} = \text{tr}\Big\{\boldsymbol{G}_k\bar{\boldsymbol{J}}_k^{1/2}\Big(\bar{\boldsymbol{J}}_k^{-1/2}\bar{\boldsymbol{H}}_{s,k}\tilde{\boldsymbol{F}}_k\tilde{\boldsymbol{F}}_k^H\bar{\boldsymbol{H}}_{s,k}^H\bar{\boldsymbol{J}}_k^{-\text{H}/2} + \boldsymbol{I}_{N_r}\Big)\bar{\boldsymbol{J}}_k^{\text{H}/2}\boldsymbol{G}_k^H\Big\} \tag{5.148}$$

$$= \text{tr}\Big\{\boldsymbol{T}_{2,k}^H(\boldsymbol{\Delta}_k^{\dagger})^T\boldsymbol{\Delta}_k^{\dagger}\boldsymbol{T}_{2,k}\boldsymbol{\Sigma}_k\Big(\boldsymbol{D}_k\boldsymbol{D}_k^T + \boldsymbol{I}_{N_r}\Big)\boldsymbol{\Sigma}_k^T\Big\} \tag{5.149}$$

$$\geq \sum_{i=1}^{\bar{N}_k}\frac{\sigma_{k,i}^2(d_{k,i}^2+1)}{\delta_{k,i}^2}. \tag{5.150}$$

To obtain (5.148) we have used the decomposition in (5.139), the structure of $\boldsymbol{G}_k$ in (5.144), and the result $\boldsymbol{U}_{y,k} = \boldsymbol{V}_{z,k}$. The lower bound in (5.149) follows from the

application of Lemma 25 and holds with equality when $\boldsymbol{T}_{2,k} = \boldsymbol{I}_{N_d}$. Substituting $\boldsymbol{T}_{2,k} = \boldsymbol{I}_{N_d}$ and $\boldsymbol{V}_{z,k} = \boldsymbol{U}_{s,k}$ into (5.144) we have

$$\boldsymbol{G}_k = \boldsymbol{V}_{r,k}\boldsymbol{\Phi}_k\boldsymbol{U}_{s,k}^H\bar{\boldsymbol{J}}_k^{-1/2}, \tag{5.151}$$

where $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\Delta}_k^\dagger\boldsymbol{\Sigma}_k$. Finally by substituting (5.136) into (5.151) we prove the relay precoder structure given in Theorem 6.

Having proved the particular structure of the source and relay precoder matrices $\tilde{\boldsymbol{F}}_k$ and $\boldsymbol{G}_k$ given in Theorem 6, we now show that the diagonal matrices $\boldsymbol{\Gamma}_k$ and $\boldsymbol{\Phi}_k$ in (5.71) and (5.72), respectively, should have at most $\bar{N}_k$ positive diagonal elements. We firstly recall the definitions $\boldsymbol{\Gamma}_k \triangleq \boldsymbol{\Lambda}_k^\dagger\boldsymbol{D}_k$ and $\boldsymbol{\Phi}_k \triangleq \boldsymbol{\Delta}_k^\dagger\boldsymbol{\Sigma}_k$ made previously. Consequently we have the relationships

$$d_{k,i} = \gamma_{k,i}\lambda_{k,i} \tag{5.152}$$

$$\sigma_{k,i} = \phi_{k,i}\delta_{k,i}, \tag{5.153}$$

Noting that the optimal source and relay precoders result in (5.141) and (5.142) holding with equality, then substituting (5.152) and (5.153) into (5.142) we can write

$$|\boldsymbol{E}_k| = \prod_{i=1}^{\bar{N}_k}\left(1 + \frac{\gamma_{k,i}^2\lambda_{k,i}^2\phi_{k,i}^2\delta_{k,i}^2}{\phi_{k,i}^2\delta_{k,i}^2 + 1}\right)^{-1}. \tag{5.154}$$

From (5.154) we observe that only the first $\bar{N}_k$ diagonal elements of $\boldsymbol{\Gamma}_k$ and $\boldsymbol{\Phi}_k$ affect the objective function. It is therefore straightforward to show that, in order to conserve the source and relay transmission power, the remaining elements of $\boldsymbol{\Gamma}_k$ and $\boldsymbol{\Phi}_k$ should be zero.

Chapter 6

Conclusions and Future Work

In this concluding chapter we firstly summarise the main contributions of the thesis before discussing some potential lines of future research based on the work proposed in the previous chapters.

6.1 Thesis Summary

This thesis has been concerned with the topic of channel estimation and non-linear transceiver designs for MIMO relaying networks. We introduced the signal model for such a system in Chapter 2 considering the transmission between a single source and destination device, with the communication process being aided by a single relaying node. The source, relay, and destination devices were each equipped with multiple antennas and we assumed the source-relay and relay-destination channels to be frequency selective. To deal with the frequency selectivity of these channels we considered the use of OFDM which allowed each transmission stage to be decoupled into a number of parallel narrowband subcarriers. We then considered non-linear transceivers for the MIMO OFDM relaying system. Specifically, we considered two transceiver models which either utilised linear precoding at the source and relay with a non-linear DFE at the destination, or non-linear THP at the source with linear precoding and equalisation at the destination. It was shown that for both transceivers the same optimisation problem could be formulated to solve for source, relay, and destination processors.

In Chapter 3 we discussed channel estimation algorithms for the considered MIMO OFDM relaying system. We considered the estimation of the frequency selective time domain channels as opposed to the direct estimation of the frequency domain subcarrier channels. The estimation process for the MIMO OFDM relaying system was divided into two seperate phases. In the first phase the estimation

of the relay-destination channel was conducted through transmitting known training symbols from the relay to the destination. This was equivalent to the channel estimation problem in point-to-point MIMO systems for which known algorithms have been studied in the literature. We reviewed well known LS and MMSE solutions to this problem. In the second phase the source-relay channel was estimated. Known pilot symbols were firstly transmitted from the source to the relay device. It was noted that if the relay performed channel estimation then the previously discussed LS and MMSE point-to-point MIMO channel estimation algorithms could be utilised to obtain the estimate of the source-relay channel at the relay.

We further considered the case that the destination was tasked with channel estimation. In this scenario the relay precoded and forwarded the pilot symbols to the destination device and we then suggested several channel estimation algorithms for this problem. To estimate the source-relay channel the relay-destination channel was required to be known. For simplicity it was assumed that the relay-destination channel estimate was sufficiently accurate and channel estimation errors could be neglected. In our first algorithm an iterative procedure was used where we updated the MMSE channel estimate at the destination, the relay precoding matrix, and the source pilot matrix in an alternating fashion. It was shown that each subproblem was a standard convex optimisation problem and as such the proposed iterative algorithm had guaranteed convergence. Due to the computational complexity of this algorithm we then suggested suboptimal but simplifed solutions. It was shown that, under a high SNR approximation, the source-relay channel estimation using destination measurements could be greatly simplified. Simulation results showed that the suboptimal solutions had comparable performance to the iterative algorithm whilst offering substantially reduced complexity and furthermore all proposed algorithms offered better channel estimates compared to orthogonal source and relay pilot matrix designs.

In Chapter 4 we considered the problem of deriving the processors for the DFE and THP MIMO OFDM relay transceivers introduced in Chapter 2. We assumed that the source-relay and relay-destination channels were perfectly known to all nodes in the network. The assumption of perfect CSI was validated in Chapter 3 when the SNR during the channel estimation phases was sufficiently large. Two different design solutions were derived depending on whether ZF or MMSE equalisation was performed at the receiver. For both cases of ZF and MMSE equalisation we considered the problem of deriving the source and relay precoding matrices to minimise the arithmetic MSE subject to transmission power constraints at both the source and relay terminals. The optimal source and relay precoder structures were derived independently firstly for the use of a ZF equaliser

and secondly for the use of MMSE equalisation. It was shown that in both cases the precoder structures resulted in the original matrix valued optimisation problem reducing to simpler power allocation problems that involved only scalar variables. When ZF equalisation was used, it was shown that the optimal source and power allocation variables could be calculated in closed form with the source allocating power uniformly across all data streams. The optimal relay power allocation was derived using the KKT conditions. On the other hand, when MMSE equalisation was employed the optimal power allocation could not be derived in closed form. An iterative power allocation algorithm was utilised to derive the source and relay power allocation matrices. Simulation results showed that the proposed ZF and MMSE DFE/THP transceivers provided improved BER and MSE performance compared to several linear transceiver designs proposed in the literature.

The algorithms proposed in 4 suffer a degradation in performance when the channels are estimated imperfectly. In Chapter 5 we therefore considered the problem of designing the DFE and THP processors in a statistically robust manner to derive transceivers that could deal with channel estimation errors. The subcarrier channel estimation errors were modelled as being additive matrices having zero mean Gaussian distributions with known covariance. The vailidity of this model was verified by showing that some specific channel estimation algorithms resulted in the channel estimation errors having such distributions. A Bayesian approach was adopted to formulate a robust optimisation problem where the objective function was obtained by averaging the DFE and THP error covariance matrices over the channel estimation errors using their statistical properties. It was shown that for general channel estimation error covariance matrices the optimal solution to the optimisation problem was difficult to obtain due to the objective function being an extremely complicated non-convex function of the source and relay precoders. We simplified the procedure by relaxing the optimisation problem and derived the source and relay precoders as the solution to the relaxed problem. Similar to the designs in Chapter 4, for the case of perfect CSI, the robust source and relay precoders result in a scalar valued optimisation problem and an iterative power allocation procedure was utilised to solve this problem. Simulation results demonstrated that the proposed robust non-linear transceivers outperformed robust linear transceivers in both BER and MSE. Simulation results also showed that the robust designs proposed in this chapter outperformed the non-robust algorithms derived in Chapter 4.

6.2 Future Work

Throughout this thesis we have discussed various algorithms for two-hop MIMO OFDM relay systems. From this work a number of very interesting future lines of research have arisen, some of which we shall now briefly discuss:

- **Robust MIMO OFDM Relay Channel Estimation**
  In Chapter 3 we developed MIMO OFDM relay channel estimation algorithms where the channel estimates could only be acquired using measurements made at the destination. In the first phase of channel estimation the relay-destination channel was estimated by transmitting known pilot symbols from the relay to destination. In the second phase known pilot symbols were transmitted from the source to relay, which were then forwarded to the destination allowing for an estimation of the source-relay channel. The source-relay channel estimate depended directly on the relay-destination channel estimate obtained in the first phase. In our proposed algorithms, for simplicity, we assumed that the relay-destination channel estimate was sufficiently accurate that channel estimation errors could be ignored. However, simulation results demonstrated that an inaccurate relay-destination channel estimate adversely impacted the accuracy of the source-relay estimate. A possible approach to dealing with this problem is to develop robust algorithms that take into account the error from the first phase of channel estimation. Such an approach has been developed in [57] for narrowband MIMO relaying and should be extended to MIMO OFDM relay channel estimation in future research work.

- **Superimposed MIMO OFDM Relay Channel Estimation**
  The main drawback of conventional training based channel estimation algorithms, where known pilot symbols are used for channel estimation, is that dedicated training periods must be used. Since no information bearing symbols are transmitted during training the overall system spectral efficiency is reduced. Another class of channel training algorithms known as superimposed channel estimation have therefore been developed to circumvent this problem. Superimposed training for narrowband point-to-point MIMO channel estimation has been studied in e.g. [98], and for point-to-point MIMO OFDM systems in e.g. [99]. The advantage of superimposed training compared to conventional training based algorithms is that it allows data transmission and channel estimation to be carried out simultaneously, which makes it a very spectrally efficient technique. This is achieved

through superimposing training symbols over data carrying symbols. The idea of superimposed training has recently been extended to two-hop narrowband MIMO relaying systems in [100]. In this algorithm the source node sends a training signal to the relay, which then amplifies the received signal, superimposes another training signal, and forwards the resulting signal to the destination. At the destination, the source training signal and the superimposed relay training signal are used to simultaneously estimate the source-relay and relay-destination channels. Superimposed channel training for two-hop MIMO relaying is therefore a more spectrally efficient technique compared to conventional training based algorithms since it does not require the source-relay and relay-destination channels to be estimated in seperate phases. Whilst superimposed training has been considered in [100] for narrowband MIMO relaying, it does not appear to have been studied for MIMO OFDM relaying and is an interesting future research topic.

- **Vector Precoding for MIMO Relaying**
  In this thesis we investigated non-linear DFE and THP transceivers for MIMO OFDM relay systems. Another non-linear technique which is very similar to THP is vector precoding (VP) which is also commonly referred to as vector perturbation. The main difference between THP and VP techniques mainly lies in the choice of the perturbation vector. In fact, THP is actually a particular case of the more general VP techniques. It is shown in [101] for the case of point-point MIMO systems that VP can offer improved performance when compared to various THP designs. VP has also been studied in [102–104] for multi-user narrowband MIMO relaying systems under the assumption of perfect CSI but appears to be largely unstudied for single user MIMO OFDM relay communications. It is therefore of interest to further study the use of VP for MIMO OFDM relaying networks and to compare such algorithms to the proposed DFE and THP transceivers.

- **Optimal Robust DFE and THP Transceivers**
  In Chapter 5 we studied robust DFE and THP transceiver designs for MIMO OFDM relaying systems with imperfect CSI. The source and relay precoder structures were derived by solving a relaxed optimisation problem. For a specific class of channel estimation error covariance matrices it was shown that the derived source and relay precoders were in fact optimal. However for general channel estimation errors the proposed approach is suboptimal. The optimal solution still remains an open problem and should be further

investigated. For the case of robust linear transceiver designs for narrowband MIMO relay systems the optimal solution for general estimation errors has been derived in e.g. [82] and relies on an iterative procedure. It is possible that a similar approach could be used to design robust DFE and THP transceiver designs in MIMO OFDM relay systems.

- **Optimal DFE and THP Transceivers with Direct Link**
  Throughout this thesis we have restricted our attention to MIMO relay systems in the absence of a direct link between the source and destination terminals. In practice, the inclusion of this link can provide a valuable source of spatial diversity that can be utilised to improve performance. Optimal linear precoders with the inclusion of a direct link can be found in [43, 105] where iterative algorithms are required in order to update the source and/or relay precoder matrices. Suboptimal non-linear approaches can also be found in [51, 55, 63, 106, 107]. Although the inclusion of the direct source-destination link has been considered by some authors, the optimal DFE and THP source and relay precoders when a direct link is included appears to still remain an open problem and requires further investigation.

References

[1] C. Shannon, "A mathematical theory of communications," *Bell Systems Technical Journal*, vol. 27, pp. 623–656, Jul. 1948.

[2] H. Bolcskei, D. Gesbert, C. Papadias, A. J. Van Der Veen, and A. M. Cipriano, *Space-Time Wireless Systems: From Array Processing to MIMO Communications*. Cambridge University Press, 2006.

[3] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2008.

[4] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. Poor, *MIMO Wireless Communications*. Cambridge University Press, 2007.

[5] E. Larsson, P. Stoica, and G. Ganesan, *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2003.

[6] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Information Theory*, vol. 44, pp. 744 – 765, Mar. 1998.

[7] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Information Theory*, vol. 45, pp. 1456 – 1467, Jul. 1999.

[8] H. Wang, X.-G. Xia, Q. Yin, and B. Li, "A family of space-time block codes achieving full diversity with linear receivers," *IEEE Trans. Communications*, vol. 57, pp. 3607 – 3617, Dec. 2009.

[9] D. Gesbert, M. Shafi, D. Shiu, P. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE Journ. Selected Areas in Communications*, vol. 21, pp. 281 – 302, Apr. 2003.

[10] I. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Telecommunications*, vol. 10, pp. 585 – 595, Nov./Dec. 1999.

[11] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311 – 335, Mar. 1998.

[12] H. Bolcskei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Communications*, vol. 50, pp. 225 – 234, Feb. 2002.

[13] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Information Theory*, vol. 49, pp. 1073 – 1096, May. 2003.

[14] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Information Theory*, vol. 51, pp. 4152 – 4172, Dec. 2005.

[15] L. Ordonez, D. Palomar, and J. Fonollosa, "On the diversity, multiplexing, and array gain tradeoff in mimo channels," *IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jun. 2010.

[16] D. Palomar, J. Cioffi, and M. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Processing*, vol. 51, pp. 2381 – 2401, Sept. 2003.

[17] D. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends in Communications and Information Theory*, vol. 3, Nov. 2006.

[18] A. Scaglione, G. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. i. unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2006, Jul. 1999.

[19] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Communications*, vol. 49, pp. 2198 – 2206, Dec. 2001.

[20] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May. 2002.

[21] A. Scaglione, S. Barbarossa, and G. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Information Theory*, vol. 45, pp. 1019 – 1032, Apr. 1999.

[22] D. Palomar, M. Bengtsson, and B. Ottersten, "Minimum BER linear transceivers for MIMO channels via primal decomposition," *IEEE Trans. Signal Processing*, vol. 53, pp. 2866 – 2882, Aug. 2005.

[23] L. Ordonez, D. Palomar, A. Pages-Zamora, and J. Fonollosa, "On equal constellation minimum ber linear mimo transceivers," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2007.

[24] F. Xu, T. Davidson, J.-K. Zhang, and K. Wong, "Design of block transceivers with decision feedback detection," *IEEE Trans. Signal Processing*, vol. 54, pp. 965 – 978, Mar. 2006.

[25] F. Xu, T. Davidson, J.-K. Zhang, S. Chan, and K. Wong, "Design of block transceivers with MMSE decision feedback detection," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 2005*, Mar. 2005.

[26] 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.

[27] A. Stamoulis, G. Giannakis, and A. Scaglione, "Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Trans. Communications*, vol. 49, pp. 69 –83, Jan. 2001.

[28] Y. Jiang, L. J., and W. W. Hager, "MIMO transceiver design using geometric mean decomposition," *IEEE Information Theory Workshop*, Oct. 2004.

[29] Y. Jiang, J. Li, and W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Processing*, vol. 53, pp. 3791 – 3803, Oct. 2005.

[30] Y. Jiang, D. Palomar, and M. Varanasi, "Precoder optimization for nonlinear MIMO transceiver based on arbitrary cost function," *41st Annual Conference on Information Sciences and Systems, 2007. CISS '07.*, Mar. 2007.

[31] K. Kusume, M. Joham, W. Utschick, and G. Bauch, "Cholesky factorization with symmetric permutation applied to detecting and precoding spatially multiplexed data streams," *IEEE Trans. Signal Processing*, vol. 55, pp. 3089–3103, Jun. 2007.

[32] M. Shenouda and T. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding," *IEEE Journ. Selected Areas in Communications*, vol. 26, pp. 401 – 411, Feb. 2008.

[33] A. D'Amico and M. Morelli, "Joint Tx-Rx MMSE design for MIMO multicarrier systems with Tomlinson-Harashima precoding," *IEEE Trans. Wireless Communications*, vol. 7, pp. 3118 – 3127, Aug. 2008.

[34] A. D'Amico, "Tomlinson-harashima precoding in MIMO systems: A unified approach to transceiver optimization based on multiplicative schur-convexity," *IEEE Trans. Signal Processing*, vol. 56, pp. 3662 – 3677, Aug. 2008.

[35] R. Fischer, C. Windpassinger, A. Lampe, and J. H. Huber, "Space-time transmission using Tomlinson-Harashima precoding," *Proc. 4th ITC Conf. on Source and Channel Coding,*, Feb. 2008.

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

[37] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. Wireless Communications*, vol. 6, pp. 1774 – 1786, May. 2007.

[38] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journ. Selected Areas in Communications*, vol. 25, pp. 379 – 3892, Feb. 2007.

[39] S. Boyd and L. Vandenberghe, *Convex Optimisation*. Cambridge University Press, 1985.

[40] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Communications Letters*, vol. 12, pp. 517 – 519, Jul. 2008.

[41] C. Song, K. Lee, and I. Lee, "MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems," *IEEE Trans. Wireless Communications*, vol. 9, pp. 2310 – 2319, Jul. 2010.

[42] A. Behbahani, R. Merched, and A. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Processing*, vol. 56, pp. 5062 – 5073, Oct. 2008.

[43] Y. Rong, "Optimal joint source and relay beamforming for MIMO relays with direct link," *IEEE Communications Letters*, vol. 14, pp. 390 – 392, May. 2010.

[44] F. Tseng and W. Wu, "Linear MMSE transceiver design in amplify-and-forward MIMO relay systems," *IEEE Trans. Vehicular Technology*, vol. 59, pp. 754 – 765, Feb. 2010.

[45] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Processing*, vol. 57, pp. 4837 – 4851, Dec. 2009.

[46] X. Tang and Y. Hua, "Optimal design of non-regenerative mimo wireless relays," *IEEE Trans. Wireless Communications*, vol. 6, pp. 1398 – 1407, Apr. 2007.

[47] O. Munoz-Medina, J. Vidal, and A. Agustin, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Processing*, vol. 55, pp. 2593 – 2604, Jun. 2007.

[48] Z. Fang, Y. Hua, and J. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," *Fourth IEEE Workshop on Sensor Array and Multichannel Processing*, Jul. 2006.

[49] R. Mo and Y. Chew, "Precoder design for non-regenerative MIMO relay systems," *IEEE Trans. Wireless Communications*, vol. 8, pp. 5041 – 5049, Oct. 2009.

[50] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Communications*, vol. 8, pp. 6068 – 6077, Dec. 2009.

[51] A. Millar, S. Weiss, and R. Stewart, "Precoder design for MIMO relay networks with direct link and decision feedback equalisation," *IEEE Communications Letters*, vol. 15, pp. 1044 – 1046, Oct. 2011.

[52] A. Millar and S. Weiss, "Transceiver design for non-regenerative MIMO relay systems with decision feedback detection," *18th European Signal Processing Conference (EUSIPCO), 2010*, Aug. 2010.

[53] Y. Rong, "Optimal linear non-regenerative multi-hop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Communications*, vol. 9, pp. 2268 – 2279, Jul. 2010.

[54] A. Millar, S. Weiss, and R. Stewart, "Tomlinson harashima precoding design for non-regenerative MIMO relay networks," *73rd IEEE Vehicular Technology Conference (VTC Spring)*, May. 2011.

[55] F.-S. Tseng, M.-Y. Chang, and W.-R. Wu, "Joint tomlinson-harashima source and linear relay precoder design in amplify-and-forward MIMO relay systems via MMSE criterion," *IEEE Trans. Vehicular Technology*, vol. 60, pp. 1687 –1698, May. 2011.

[56] P. Lioliou, M. Viberg, and M. Matthaiou, "Bayesian approach to channel estimation for AF MIMO relaying systems," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 1440–1451, Sep. 2012.

[57] C. Chiong, Y. Rong, and Y. Xiang, "Robust channel estimation algorithm for dual-hop MIMO relay channels," *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2012.

[58] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Processing*, vol. 59, pp. 4438–4446, Sep. 2011.

[59] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Communications*, vol. 7, pp. 1907–1916, May. 2008.

[60] T. Cui, F. Gao, and A. Nallanathan, "Optimal training design for channel estimation in amplify and forward relay networks," *Global Telecommunications Conference (GLOBECOM)*, Nov. 2007.

[61] A. Millar, S. Weiss, and R. Stewart, "Robust transceiver design for MIMO relay systems with tomlinson harashima precoding," *20th European Signal Processing Conference (EUSIPCO), 2012*, Aug. 2012.

[62] C. Xing, M. Xia, F. Gao, and Y. Wu, "Robust transceiver with tomlinson-harashima precoding for amplify-and-forward MIMO relaying systems," *IEEE Journ. Selected Areas in Communications*, vol. 30, pp. 1370 – 1382, Sep. 2012.

[63] F.-S. Tseng, M.-Y. Chang, and W.-R. Wu, "Robust MMSE transceiver design in amplify-and-forward MIMO relay system with tomlinson-harashima source precoding," *IEEE Wireless Communications and Networking Conference (WCNC), 2012.*, Apr. 2012.

[64] H. Bolcskei, M. Borgmann, and A. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE Journ. Selected Areas in Communications*, vol. 21, pp. 427 – 439, Apr. 2003.

[65] C.-N. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Information Theory*, vol. 48, pp. 637 –650, Mar. 2002.

[66] H. Tuan, H. Kha, H. Nguyen, and V.-J. Luong, "Optimized training sequences for spatially correlated MIMO-OFDM," *IEEE Trans. Wireless Communications*, vol. 9, pp. 2768–2778, Sep. 2010.

[67] A. Gupta and D. Nagar, *Matrix Variate Distributions*. London, UK,: Chapman & Hall/CRC, 2000.

[68] H. Bahrami and T. Le-Ngoc, "MIMO precoder designs for frequency-selective fading channels using spatial and path correlation," *IEEE Trans. Vehicular Technology*, vol. 57, pp. 3441 –3452, Nov. 2008.

[69] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, pp. 138–139, Mar. 1971.

[70] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Communications*, vol. 20, pp. 774–780, Aug. 1972.

[71] M. Biguesh and A. Gershman, "MIMO channel estimation: optimal training and tradeoffs between estimation techniques," *IEEE International Conference on Communications*, Jun. 2004.

[72] M. Biguesh and A. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, pp. 884–893, Mar. 2006.

[73] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Communications*, vol. 5, pp. 1158 – 1168, May. 2006.

[74] H. Zhang, Y. G. Li, A. Reid, and J. Terry, "Optimum training symbol design for mimo ofdm in correlated fading channels," *IEEE Trans. Wireless Communications*, vol. 5, pp. 2343–2347, Sep. 2006.

[75] P. Lioliou and M. Viberg, "Least square based channel estimation for mimo relays," *International ITG Workshop on Smart Antennas (WSA)*, Feb. 2008.

[76] J. Ma, P. Orlik, J. Zhang, and G. Y. Li, "Pilot matrix design for interim channel estimation in two hop mimo af relay systems," *IEEE International Conference on Communications (ICC)*, Jun. 2009.

[77] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1615–1624, Jun. 2003.

[78] A. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*. Ney York, Academic Press, 1979.

[79] D. Palomar and J. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Processing*, vol. 53, pp. 686–695, Feb. 2005.

[80] Y. Rong, "Simplified algorithms for optimizing multiuser multi-hop MIMO relay systems," *IEEE Trans. Communications*, vol. 59, pp. 2896–2904, Oct. 2011.

[81] C. Xing, S. Ma, and Y. Wu, "Iterative LMMSE transceiver design for dual-hop AF MIMO relay systems under channel uncertainties," *Proceedings of IEEE Personal, Indoor and Mobile Radio Communications Symposium (PIMRC)*, Sep. 2009.

[82] C. Xing, S. Ma, and C. Wu, Y, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Processing*, vol. 58, pp. 2273 –2283, Apr. 2010.

[83] C. Xing, S. Ma, C. Wu, Y, and T. Ng, "Transceiver design for dual-hop nonregenerative MIMO-OFDM relay systems under channel uncertainties," *IEEE Trans. Signal Processing*, vol. 58, pp. 6325 –6339, Dec. 2010.

[84] X. Zhang, D. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Processing*, vol. 56, pp. 3678 –3689, Aug. 2008.

[85] T. Zeng, Q. Chen, P. Xiao, and J. Wu, "MSINR based precoding design for non-regenerative MIMO relay system," *1st IEEE International Conference on Communications in China (ICCC)*, Aug. 2012.

[86] Y. Jiang, L. J., and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition," *IEEE Trans. Signal Processing*, vol. 53, pp. 3791–3803, 2005.

[87] A. Pascual-Iserte, D. Palomar, A. Perez-Neira, and M.-A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Processing*, vol. 54, pp. 346–360, Jan. 2006.

[88] J. Wang and D. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Processing*, vol. 57, pp. 3086–3100, Aug. 2009.

[89] J. Wang and M. Bengtsson, "Joint optimization of the worst-case robust mmse mimo transceiver," *IEEE Signal Processing Letters*, vol. 18, pp. 295–298, May. 2011.

[90] C. Xing, S. Ma, Y. Wu, T. Ng, and H. Poor, "Linear transceiver design for amplify-and-forward mimo relay systems under channel uncertainties," *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2010.

[91] C. Xing, S. Ma, Y. Wu, and T. Ng, "Robust beamforming for amplify-and-forward mimo relay systems based on quadratic matrix programming," *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Mar. 2010.

[92] C. Xing, Z. Fei, S. Ma, J. Kuang, and Y. Wu, "Maximum mutual information design for amplify-and-forward multi-hop mimo relaying systems under channel uncertainties," *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2012.

[93] C. Xing, S. Ma, Z. Fei, Y. Wu, and J. Kuang, "Joint robust weighted LMMSE transceiver design for dual-hop AF multiple-antenna relay systems," *Proceedings of the Global Communications Conference, (GLOBECOM)*, Dec. 2011.

[94] C. Xing, M. Xia, F. Gao, and Y. Wu, "Robust tomlinson-harashima precoding for non-regenerative multi-antenna relaying systems," *IEEE Wireless Communications and Networking Conference (WCNC), 2012.*, Apr. 2012.

[95] B. Chalise and L. Vandendorpe, "Joint linear processing for an amplify-and-forward MIMO relay channel with imperfect channel state information," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, Article ID 640186, 13 pages.

[96] Y. Rong, "Robust design for linear non-regenerative MIMO relays with imperfect channel state information," *IEEE Trans. Signal Processing*, vol. 59, pp. 2455 –2460, May. 2011.

[97] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[98] V. Nguyen, H. Tuan, H. Nguyen, and N. Tran, "Optimal superimposed training design for spatially correlated fading MIMO channels," *IEEE Trans. Wireless Communications*, vol. 7, pp. 3206 – 3217, Aug. 2008.

[99] N. Tran, H. Tuan, and H. Nguyen, "Superimposed training designs for spatially correlated MIMO-OFDM systems," *IEEE Trans. Wireless Communications*, vol. 9, pp. 876 – 880, Mar. 2010.

[100] Y. Rong, "Superimposed channel training for MIMO relay systems," *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2012.

[101] M. Joham, H. Brunner, R. Hunger, D. Schmidt, and W. Utschick, "Point-to-point MIMO MMSE vector precoding and thp achieving capacity," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2008.

[102] I. Jimenez, M. Barrenechea, M. Mendicute, and E. Arruti, "Iterative joint MMSE design of relaying MIMO downlink schemes with non-linearly precoded transmission," *IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2012.

[103] I. Jimenez, S. Weiss, M. Mendicute, and E. Arruti, "Multiuser MIMO amplify-and-forward relaying schemes with vector precoding," *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Dec. 2011.

[104] I. Jimenez, M. Barrenechea, M. Mendicute, and E. Arruti, "Non-linear precoding approaches for non-regenerative multiuser MIMO relay systems," *Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Aug. 2012.

[105] Y. Rong and F. Gao, "Optimal beamforming for non-regenerative mimo relays with direct link," *IEEE Communications Letters*, vol. 13, pp. 926–928, Dec. 2009.

[106] A. Millar, S. Weiss, and R. Stewart, "ZF DFE transceiver design for MIMO relay systems with direct source-destination link," *19th European Signal Processing Conference (EUSIPCO), 2011*, Aug. 2011.

[107] A. Millar, S. Weiss, and R. Stewart, "THP transceiver design for MIMO relaying with direct link and partial CSI," *IEEE Communications Letters*, vol. 17, pp. 1204–1207, Jun. 2013.