THE UNIVERSITY *of* EDINBURGH

Advanced Interference Management Techniques for Future Generation Cellular Networks

Paula Aquilina

Doctor of Philosophy
School of Engineering
University of Edinburgh
2017

Declaration

I declare that this thesis was composed by myself, that the work contained therein is my own except where explicitly stated otherwise, and that none of the work has been submitted for any other degree or professional qualification.

Paula Aquilina

Acknowledgements

Firstly, my deepest gratitude goes to my supervisor Prof. Tharmalingam Ratnarajah for his constant guidance and help throughout this entire process. His breadth of knowledge, coupled with his boundless enthusiasm for the subject, have been invaluable to my work, and have made this a far more pleasurable and interesting process than I could have ever hoped for. I also greatly appreciate his patience and confidence in me, right from the very beginning when I was still an MSc. student by providing the necessary funding[1] to enable me to pursue this PhD programme.

I would also like to thank my colleagues at IDCOM for their company and assistance throughout these years. Particularly, I am thankful to Dr Ali Cagatay Cirik for providing me with advice and sharing his knowledge on optimisation techniques, which have been used for Chapter 5 of this thesis. I am also grateful to Sudip, Yuki and Lin for their friendship and excellent company.

I thank my friends from back home, particularly Deborah, Leanne, Lucianne and Maria, who have helped bring the sunshine of Malta to Edinburgh. Finally, this thesis would not have been possible without the support and love of the people nearest to my heart - my father George, my mother Mary, my brothers Daniel and Samuel, my grandparents, the rest of my extended family, and Jonathan.

[1]This PhD project was supported by the Seventh Framework Programme for Research of the European Commission under grant numbers: ADEL-619647, HARP-318489 and HIATUS-265578.

Abstract

The demand for mobile wireless network resources is constantly on the rise, pushing for new communication technologies that are able to support unprecedented rates. In this thesis we address the issue by considering advanced interference management techniques to exploit the available resources more efficiently under relaxed channel state information (CSI) assumptions. While the initial studies focus on current half-duplex (HD) technology, we then move on to full-duplex (FD) communication due to its inherent potential to improve spectral efficiency. Work in this thesis is divided into four main parts as follows.

In the first part, we focus on the two-cell two-user-per-cell interference broadcast channel (IBC) and consider the use of topological interference management (TIM) to manage inter-cell interference in an alternating connectivity scenario. Within this context we derive novel outer bounds on the achievable degrees of freedom (DoF) for different system configurations, namely, single-input single-output (SISO), multiple-input single-output (MISO) and multiple-input multiple-output (MIMO) systems. Additionally, we propose new transmission schemes based on joint coding across states that exploit global topological information at the transmitter to increase achievable DoF. Results show that when a single state has a probability of occurrence equal to one, the derived bounds are tight with up to a twofold increase in achievable DoF for the best case scenario. Additionally, when all alternating connectivity states are equiprobable: the SISO system gains $\frac{11}{16}$ DoF, achieving 96.4% of the derived outer bound; while the MISO/MIMO scenario has a gain of $\frac{1}{2}$ DoF, achieving the outer bound itself.

In the second part, we consider a general G-cell K-user-per-cell MIMO IBC and analyse the performance of linear interference alignment (IA) under imperfect CSI. Having imperfect channel knowledge impacts the effectiveness of the IA beamformers, and leads to a significant amount of residual leakage interference. Understanding the extent of this impact is a fundamental step towards obtaining a performance characterisation that is more relevant to practical scenarios. The CSI error model used is highly versatile, allowing the error to be treated either as a function of the signal-to-noise ratio (SNR) or as independent of it. Based on this error model, we derive a novel upper bound on the asymptotic mean sum rate loss and quantify the DoF loss due to imperfect CSI. Furthermore, we propose a new version of the maximum signal-to-interference plus noise ratio (Max-SINR) algorithm which takes into account statistical knowledge of the CSI

error in order to improve performance over the naive counterpart in the presence of CSI mismatch.

In the third part, we shift our attention to FD systems and consider weighted sum rate (WSR) maximisation for multi-user multi-cell networks where FD base-stations (BSs) communicate with HD downlink (DL) and uplink (UL) users. Since WSR problems are non-convex we transform them into weighted minimum mean squared error (WMMSE) ones that are proven to converge. Our analysis is first carried out for perfect CSI and then expanded to cater for imperfect CSI under two types of error models, namely, a norm-bounded error model and a stochastic error model. Additionally, we propose an algorithm that maximises the total DL rate subject to each UL user achieving a desired target rate. Results show that the use of FD BSs provides significant gains in achievable rate over the use of HD BSs, with a gain of 1.92 for the best case scenario under perfect CSI. They also demonstrate the robust performance of the imperfect CSI designs, and confirm that FD outperforms HD even under CSI mismatch conditions.

Finally, the fourth part considers the use of linear IA to manage interference in a multi-user multi-cell network with FD BSs and HD users under imperfect CSI. The number of interference links present in such a system is considerably greater than that present in the HD network counterpart; thus, understanding the impact of residual leakage interference on performance is even more important for FD enabled networks. Using the same generalised CSI error model from the second part, we study the performance of IA by characterising the sum rate and DoF losses incurred due to imperfect CSI. Additionally, we propose two novel IA algorithms applicable to this network; the first one is based on minimising the mean squared error (MMSE), while the second is based on Max-SINR. The proposed algorithms exploit statistical knowledge of the CSI error variance in order to improve performance. Moreover, they are shown to be equivalent under certain conditions, even though the MMSE based one has lower computational complexity. Furthermore for the multi-cell case, we also derive the proper condition for IA feasibility.

Lay Summary

Having constant access to mobile data is gaining significant importance in our day-to-day lives, with mobile devices and the services they provide us with changing many aspects, from the way we contact each other, to the way we consume video or even process payments. However, wireless network resources are scarce and need to be allocated conservatively. This has pushed the wireless communications research community to look into new technologies that will be able to support the increasing rate demands of future generation cellular networks. In this thesis we address the issue by focusing on ways that allow us to exploit the available resources in a more efficient manner.

One major concern in wireless networks is interference. Mobile devices communicating in the same time/frequency resource, can conflict with each other resulting in unreliable communication. The traditional way to manage interference is to simply avoid it by allocating orthogonal resources. However, recent information theory results show that by exploiting knowledge of the channel state information (CSI), more advanced interference management techniques can be applied in order to boost wireless network capacity. Here, we focus on advanced interference management solutions that allow us to serve multiple users across multiple cells concurrently. This is done under relaxed CSI conditions, where the channel information is not assumed to be perfect, in order to study scenarios that are more easily relatable to practical situations.

The first half of the thesis deals with half-duplex (HD) systems where uplink and downlink communication are separate. Within this context we consider a two-cell two-user-per-cell network with alternating connectivity where transmitters only have topological knowledge of the global channel links; we characterise the potential gains and devise new transmission schemes to achieve them. Next, we consider a more general G-cell K-user-per-cell network and apply a new interference management technique, called interference alignment (IA), to manage interference under imperfect CSI. We characterise the capacity losses incurred due to the CSI mismatch, and propose a novel IA algorithm that exploits statistical knowledge of the error in order to improve performance over the standard version.

In the second half of the thesis we consider full-duplex (FD) communication. In FD systems, uplink and downlink communication take place in the same time/frequency resource, thereby providing us with potential to double spectral

efficiency. While this technology is not used in current networks, it is a highly promising candidate for future generation ones. Firstly, we consider weighted sum rate maximisation problems for a multi-cell multi-user network with FD base-stations (BSs) and HD users, under both perfect and imperfect CSI. Next, we consider the use of linear IA within the same setting, and quantify the losses incurred due to imperfect CSI, while also proposing novel IA algorithms applicable to this scenario.

Contents

List of Figures

List of Tables

List of Abbreviations

ADC	Analogue-to-digital converter
AWGN	Additive white Gaussian noise
BER	Bit error rate
BS	Base-station
CCI	Co-channel interference
CPS	Central processing site
CSI	Channel state information
CSIT	Channel state information at the transmitter
D2D	Device-to-device
DAC	Digital-to-analogue converter
dB	Decibel
dBm	Decibel-milliwatts
DL	Downlink
DoF	Degrees of freedom
EB	Exabytes
FD	Full-duplex
FDD	Frequency division duplex
HD	Half-duplex
Hz	Hertz
IA	Interference alignment
IBC	Interference broadcast channel
IC	Interference channel
i.i.d.	Independent and identically distributed
IMAC	Interference multiple access channel
ITU	International telecommunication union
km	Kilometres
LOS	Line of sight
LTE	Long-term evolution

m	Metres
Max-Det	Determinant maximisation
Max-SINR	Maximum signal-to-interference-plus-noise ratio
MHz	Megahertz
MIMO	Multiple-input multiple-output
Min-LI	Minimum leakage interference
MISO	Multiple-input single-output
MMSE	Minimum mean squared error
MSE	Mean squared error
NLOS	Non line of sight
OFDM	Orthogonal frequency division multiplexing
RF	Radio frequency
R-SI	Residual self-interference
SKCE	Statistical knowledge of the CSI error
SDP	Semi-definite programming
SI	Self-interference
SINR	Signal-to-interference-plus-noise ratio
SISO	Single-input single-output
SNR	Signal-to-noise ratio
SOCP	Second-order cone programming
SRS	Sound reference signals
s.t.	Such that
TDD	Time division duplex
TDMA	Time division multiple access
TIA	Topological interference alignment
TIM	Topological interference management
UL	Uplink
VOIP	Voice over internet protocol
WMMSE	Weighted minimum mean squared error
WSR	Weighted sum rate
QPSK	Quadrature phase shift keying
ZF	Zero-forcing

List of Notations

$\lvert \cdot \rvert$	Absolute value
$\lVert \cdot \rVert$	Euclidean norm
$\lVert \cdot \rVert_F$	Frobenius norm
$(\cdot)^{-1}$	Inverse
$(\cdot)^{\dagger}$	Pseudo-inverse
$(\cdot)^{H}$	Hermitian transpose
$\otimes$	Kronecker product
$\succeq$	Positive semi-definite
a	Scalar
$\mathbf{a}$	Vector
$\mathbf{A}$	Matrix
$[\mathbf{A}]_m$	The mth element along the diagonal of $\mathbf{A}$
$\mathbb{C}$	Set of complex numbers
$\mathcal{CN}(a, b)$	Complex Gaussian random variable with mean a and variance b
$\mathrm{cov}(\cdot)$	Covariance
$\det$	Determinant
$\mathrm{diag}(\mathbf{A})$	Diagonal matrix containing the elements along the diagonal of $\mathbf{A}$
$\mathbb{E}\{\cdot\}$	Expectation operator
$\mathrm{eigvec}(\mathbf{A})$	Eigenvectors of $\mathbf{A}$
$\exp(a)$	Returns e^a, where e is Euler's number
$\mathbb{GF}$	Galois Field
$h(\cdot)$	Differential entropy
$\mathbf{I}$	Identity matrix
$\lim$	Limit
$\ln$	Base e logarithm
$\log_2$	Base 2 logarithm
$\log_{10}$	Base 10 logarithm
$\max$	Maximise

min	Minimise
$\mathcal{N}(a, b)$	Real Gaussian random variable with mean a and variance b
$\text{not}(b)$	Complement operation for binary term b
$o(\cdot)$	For $f(x) = o(g(x))$, $\lim_{x\to\infty} f(x)/g(x) = 0$
$\text{Tr}(\cdot)$	Trace
$\mathbb{QR}(\mathbf{A})$	Unitary part of QR-decomposition of $\mathbf{A}$
$\mathcal{V}_n[\mathbf{A}]$	Set of eigenvectors corresponding to the n smallest eigenvalues of $\mathbf{A}$
$\text{vec}(\mathbf{A})$	Vector obtained by stacking columns of $\mathbf{A}$
χ	Chi-squared distributed random variable

List of Symbols

α	SNR exponent of CSI error
β	Scalar component of CSI error
b_d/b_u	Number of streams per DL/UL user
ΔD	DoF loss
ΔR	Rate loss
D	DoF with perfect CSI
$\widehat{D}$	DoF with imperfect CSI
η	CSI error variance
G	Number of cells
$\mathbf{H}$	Perfect CSI
$\widehat{\mathbf{H}}$	Imperfect CSI
ι	Scalar component of receiver distortion variance
$\widehat{J}$	Interference leakage
κ	Scalar component of transmitter distortion variance
K	Total number of users per cell
K_g^d/K_g^u	Number of DL/UL users in cell g
K_d/K_u	Number of DL/UL users per cell in symmetric system
λ_i	Probability of occurrence of state i
μ	Weight used for WSR and WMMSE problems
M_B	Number of antennas at BS
N_d/N_u	Number of antennas at DL/UL user
P	Transmit power
R	Rate
$\widehat{R}$	Rate with imperfect CSI
ρ	SNR
σ^2	Noise variance
$\mathbf{s}$	Transmitted signal

Θ / $\bar{\Theta}$	Binary term to reflect presence of additional residual SI component due to imperfect CSI/ complement of Θ
ϖ, ϱ	Lagrange multipliers
ε	Upper bound for F-norm of CSI error
$\mathbf{U}$	Receive beamformer calculated with perfect CSI
$\widehat{\mathbf{U}}$	Receive beamformer calculated with imperfect CSI
$\mathbf{V}$	Precoder calculated with perfect CSI
$\widehat{\mathbf{V}}$	Precoder calculated with imperfect CSI
$\mathbf{y}$	Received signal vector
$\mathbf{z}$	Noise vector

Chapter 1

Introduction

1.1 Background

Cellular networks and the constant connectivity they provide us with, have entered a period of ever increasing importance in our daily lives, having a positive ripple effect on the economy and society. This has sparked a huge growth in the demand for mobile data services, causing a significant shift from the original voice-centric, circuit-switched, centrally optimised networks to data-centric, packet-switched, organically deployed ones in an effort to cope with the higher rate requirements.

Cisco's capacity forecast, depicted in Fig. 1.1, estimates an eightfold increase in mobile data traffic between 2015 and 2020. The expected compound annual growth rate over this period is 53%, with mobile data traffic increasing from 3.7 exabytes per month in 2015 to 30.6 exabytes per month in 2020. By 2020 projections estimate there will be 11.6 billion mobile connected devices for a global population of 7.8 billion, corresponding to approximately 1.5 devices per person. The increase in the number of mobile devices, along with the rise in mobile data requirements is mainly due to new trends in device usage. Fig. 1.2 shows the percentage of mobile data traffic per year distributed over four main categories. In particular it can be noticed that mobile video traffic, which for 2015 corresponded to 55% (2 exabytes per month) of the total mobile traffic, will grow to contribute 75% (20 exabytes per month) by 2020. The proliferation of high-end mobile devices has increased the propensity for users to consume higher bandwidth content and use data hungry applications. Therefore, the wireless

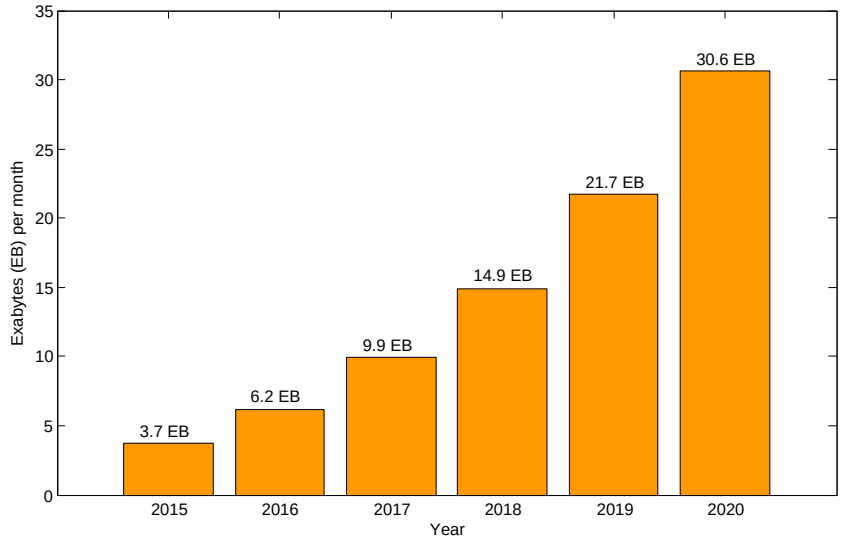


Figure 1.1: Cisco global mobile data traffic forecast [1].

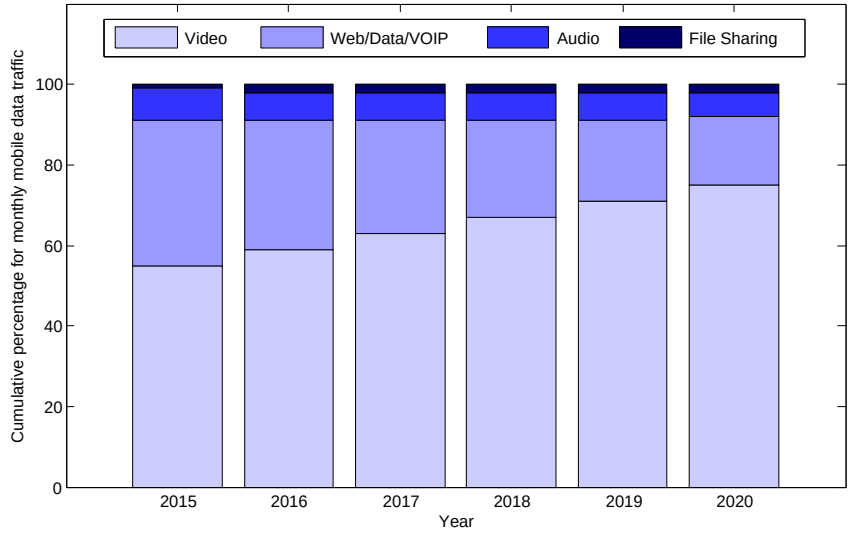


Figure 1.2: Distribution of mobile data traffic over different categories [1].

communications research community is focused on finding new ways to cater for the ever increasing data rate demands.

Cellular networks transmit information over radio frequency (RF) bands that can be characterised by the carrier frequency, bandwidth, propagation properties and interference conditions. The RF spectrum is a scarce and highly valuable resource, with usage that is strictly regulated worldwide by national laws coordinated by the International Telecommunication Union (ITU). Such regulation is

necessary in order to ensure that different systems can coexist without interfering with each other, however it introduces significantly rigid usage constraints. Moreover, opportunities to obtain exclusive rights to suitable frequency bands occur very rarely and require high expenditure. Therefore, utilising the available spectrum as efficiently as possible is a crucial step towards being able to meet the increasing data rate requirements in future generation cellular networks. This thesis aims to contribute to this effort by proposing novel solutions for advanced interference management. While the first half considers current half-duplex (HD) systems, in the second half attention shifts to full-duplex (FD) ones, since the use of FD nodes has inherent potential to significantly increase spectral efficiency.

1.2 Contributions

Recent information-theoretic results show that it is possible to use the available spectrum in a more efficient manner than the currently employed orthogonalisation based approaches. Such results indicate that the maximum achievable capacity is significantly higher than conventionally presumed, however they have generally been obtained under the assumption of perfect channel state information (CSI) at the transmitter and the receiver. The CSI acquirement process requires dedicated system resources, causing overheads and rendering it unable to provide perfect and instantaneous CSI delivery [2]. Recognising the significance of this issue has driven the wireless communications research community to focus on scenarios where the CSI assumptions are relaxed, in order to obtain results that are more easily relatable to practical settings. In this thesis we focus on the relaxed CSI context, and consider interference management solutions under such conditions. The main contributions are listed as follows.

- We study the degrees of freedom (DoF) of a two-cell two-user-per-cell interference broadcast channel (IBC) with alternating inter-cell connectivity and global topological CSI at the transmitter (CSIT), deriving novel DoF outer bounds for a variety of system configurations, namely, single-input single-output (SISO), multiple-input single-output (MISO) and multiple-input multiple-output (MIMO) systems. Additionally, we propose new transmission schemes based on joint coding across states in order to approach the derived bounds. Results show that DoF higher than those conventionally obtained without global topological information are achievable,

with up to twofold increase for the best case scenario, indicating that even such a minimal level of global CSIT is still highly useful.

- We consider the performance of linear interference alignment (IA) in the general G-cell K-user-per-cell MIMO IBC under imperfect CSI, using a generalised CSI msimatch model that allows us to treat the CSI error variance either as a function of the signal-to-noise ratio (SNR) or as independent of it. Given this error model, we derive an upper bound on the asymptotic mean loss in sum rate and quantify the achievable DoF with imperfect CSI. Results show that when the error variance and SNR are inversely proportional, full DoF can be achieved and the rate loss is finite and upper bounded by a derived value dependent on the system configuration and the CSI error parameters. When the error variance scales with SNR to the power of a negative proper fraction, the DoF loss is quantified in terms of the error parameters and the asymptotic sum rate loss is unbounded. Moreover, we propose a novel version of the maximum signal-to-interference plus noise ratio (Max-SINR) algorithm that exploits statistical knowledge of the CSI error. This algorithm provides significant performance improvements over the naive version, without incurring additional computational costs.

- We design beamformers for weighted sum rate (WSR) maximisation in a multi-user multi-cell MIMO scenario with FD base-stations (BSs) and HD downlink (DL) and uplink (UL) users. Since WSR problems are non-convex, we exploit the relationship between rate and mean squared error (MSE) to propose low complexity weighted minimum mean squared error (WMMSE) alternating optimisation algorithms. While the initial design assumes perfect CSI, we also cater for imperfect CSI under two different models, namely, a norm-bounded error model and a stochastic error model. Results show that rates achieved in FD mode are significantly higher than those achieved by the baseline HD schemes for low to intermediate distortion levels, even under imperfect CSI conditions. Additionally, we extend our original WSR problem to one which maximises the total DL rate subject to each UL user achieving a desired target rate.

- We consider the use of linear IA to manage interference in a multi-user multi-cell MIMO network with FD BSs and HD users, under imperfect CSI. Within this context, we derive an upper bound on the asymptotic

mean sum rate loss and quantify the DoF loss due to CSI mismatch. Results show that the way the error scales with SNR affects the general performance trend significantly, with both losses going to zero under certain conditions. Additionally, we propose two linear IA algorithms applicable to the FD scenario under consideration; the first one is based on minimising the mean squared error (MMSE), while the second is a Max-SINR based solution. Both algorithms exploit statistical knowledge of the CSI error to improve performance under imperfect CSI and are shown to result in identical beamformers under certain conditions, even though the MMSE one is less computationally complex. Furthermore, for the multi-cell case we also derive the proper condition for IA feasibility.

1.3 Thesis overview

The rest of this thesis is organised as follows.

- Chapter 2 provides a brief overview of the fundamental wireless communications concepts that underpin the work presented in this thesis, including: MIMO systems, different interference management approaches and FD communication.

- In Chapter 3 we focus on the two-cell two-user-per-cell IBC and consider the use of topological interference management (TIM) in order to manage inter-cell interference under an alternating connectivity scenario. Bounds for the achievable DoF are derived for SISO, MISO and MIMO system configurations. Additionally, we propose new transmission schemes based on joint coding across states and demonstrate the DoF gains that can be achieved by applying them.

- Chapter 4 focuses on the use of linear IA in the general HD MIMO IBC under imperfect CSI. It provides a bound on the sum rate loss and quantifies the DoF loss experienced due to the CSI mismatch. It also presents a novel Max-SINR algorithm that takes into account statistical knowledge of the CSI error to improve performance under imperfect CSI.

- Chapter 5 considers WSR problems for a multi-user multi-cell MIMO scenario with FD BSs and HD users. It uses the rate to MSE relationship to

propose solutions under: (a) perfect CSI, (b) a norm-bounded CSI error model, and (c) a stochastic CSI error model. Additionally, it extends the perfect CSI design to one which maximises the DL rate subject to a per UL user rate constraint.

- Chapter 6 studies the use of linear IA in a multi-user multi-cell MIMO scenario with FD BSs and HD users, and characterises the sum rate and DoF losses incurred due to imperfect CSI. It also presents MMSE and Max-SINR based IA algorithms applicable to this type of network; these take in consideration the effect of CSI mismatch for added robustness.

- Finally, Chapter 7 provides a summary of the contributions and directions for future work.

Chapter 2

Overview of Wireless Communications Concepts

2.1 Introduction

In this chapter, we provide a brief overview of the fundamental wireless communications concepts that underpin the work presented in the rest of this thesis. First, we consider point-to-point MIMO systems, outlining their advantages and characterising their capacity. Then, we move on to multi-user systems, introducing a variety of theoretical models that are used to capture different interference aspects of practical communication scenarios. Next, our focus shifts to interference management techniques, where we introduce the concepts of interference alignment (IA) and topological interference management (TIM) which play an important role in Chapters 3, 4 and 6. Finally, we consider FD communication, which is relevant to Chapters 5 and 6, highlighting its characteristics, challenges and the resulting interference scenario.

2.2 MIMO systems

The most basic form for a wireless communication link is a point-to-point one where a single antenna transmitter communicates with a single antenna receiver. Before reaching the receiver, the transmitted signal undergoes attenuation due to fading. This can be classified into two general categories:

- *small-scale fading* - due to the presence of reflectors and scatterers that cause multiple versions of the transmitted signal to arrive at the receiver, each one distorted in amplitude, phase and angle of arrival;
- *large-scale fading* - due to effects such as distance related attenuation and shadowing by obstacles.

The effects of fading, along with interference (owing to the use of a shared medium), are the most fundamental aspects of wireless communications [3]. By relying on channel knowledge, transmit and receive beamforming can be applied to mitigate their negative impact on the overall received signal. Additionally, in recent years focus has shifted from single antenna nodes to ones equipped with multiple antennas, as a means of further improving reliability and spectral efficiency, due to the inherent ability of MIMO systems to address the issues of fading and interference. Fig. 2.1 provides an illustration of a point-to-point MIMO wireless link with M transmit antennas and N receive antennas.

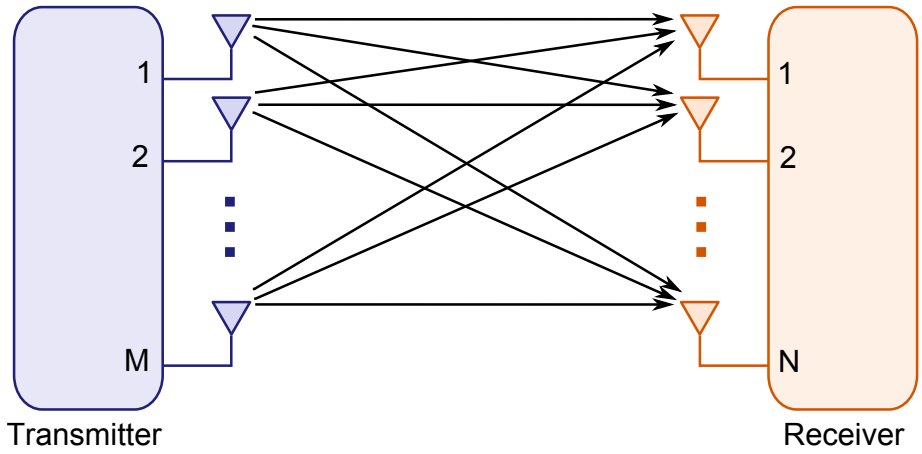


Figure 2.1: MIMO point-to-point link with N transmit antennas and M receive antennas.

The signal seen at the receiver in Fig. 2.1 can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{z}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,M} \\ h_{2,1} & h_{2,2} & \dots & h_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N,1} & h_{N,2} & \dots & h_{N,M} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_M \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \tag{2.1}$$

where y_i for $i = 1, \ldots, N$ is the signal received at the ith antenna, s_j for $j = 1, \ldots, M$ is the signal transmitted by the jth antenna, $h_{i,j}$ is the channel from the jth antenna at the transmitter to the ith antenna at the receiver, and z_i for $i = 1, \ldots, N$ is the noise seen at the ith receive antenna.

For a rich scattering environment, where there are a significant number of reflectors and no dominant propagation along the line of sight path, the channel matrix, $\mathbf{H}$, can be represented using Rayleigh fading. A Rayleigh fading channel is modelled as a statistical process with independent and identically distributed (i.i.d) entries with zero mean and phase evenly distributed between 0 and 2π radians, i.e.

$$h_{i,j} \sim \mathcal{N}\left(0, \frac{1}{2}\right) + \sqrt{-1}\,\mathcal{N}\left(0, \frac{1}{2}\right) . \tag{2.2}$$

The noise vector, $\mathbf{z}$, can be modelled as additive white Gaussian noise (AWGN), with i.i.d. Gaussian random variable entries with zero mean and variance σ^2, i.e.

$$z_i \sim \mathcal{N}\left(0, \frac{\sigma^2}{2}\right) + \sqrt{-1}\,\mathcal{N}\left(0, \frac{\sigma^2}{2}\right) . \tag{2.3}$$

2.2.1 Advantages of MIMO

Compared to SISO systems, MIMO provides a number of advantages as follows [3].

- *Power gain:* Via processing at the transmitter and the receiver, the average received SNR can be increased. At the receiver side this can be done by coherent combining of the received signals, while at the transmitter side it can be achieved by allocating the transmit power to favour higher quality links. These methods require knowledge of the CSI at the respective ends. While CSI at the receiver is more easily realisable, CSI at the transmitter is generally more difficult to obtain.

- *Spatial diversity gain:* Channel quality in wireless systems is subject to random fluctuations because of fading; diversity can be exploited to combat this issue. For SISO systems, only time or frequency diversity can be exploited, owing to the one-to-one nature of the link. However, MIMO systems offer a new type of diversity known as spatial diversity, due to the

presence of multiple channel links. Spatial diversity can be exploited at both transmit and receive ends, with a maximal gain of $M \times N$.

- *Interference suppression gain:* In wireless networks co-channel interference (CCI) arises due to frequency reuse. When multiple antenna nodes are deployed, there is increased opportunity to differentiate between the desired signal and CCI. This is done via beamforming, either at the transmitter, the receiver or a combination of both, depending on the type of inference network and the type of beamforming solution applied. Such methods require knowledge of the desired signal's path, and ideally also of the interferers' paths.

- *Spatial multiplexing gain:* Having multiple transmit and receive antennas provides extra spatial dimensions for communication, allowing for the transmission of multiple independent data streams without using additional power or bandwidth, i.e. it provides us with a spatial multiplexing gain, also known as a DoF gain. At high SNR capacity scales linearly with DoF, with the maximum achievable DoF being equivalent to $\min(M, N)$ for an $M \times N$ MIMO link.

The spatial multiplexing gain, along with the increased opportunity to suppress interference, are the main defining features of MIMO systems, putting them at the forefront for technologies that can help increase spectral efficiency. These aspects underpin most of the work presented in this thesis, where the networks studied generally involve communication between multiple antenna nodes in multi-user networks.

2.2.2 Capacity of MIMO links

For a time-varying channel, the capacity of a MIMO link with N transmit and M receive antennas with CSI at the receiver is given by

$$C_{\text{MIMO}} = \max_{\mathbf{K}:\text{Tr}(\mathbf{K})\leq P} \mathbb{E}\left\{\log_2 \det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\mathbf{H}\mathbf{K}\mathbf{H}^H\right)\right\} \tag{2.4}$$

where $\mathbf{H}$ is the channel, $\mathbf{K}$ is the covariance of the transmitted signal and P is the maximum transmit power.

Under a Rayleigh fading channel assumption, it can be shown that the optimal covariance matrix is $\mathbf{K} = \frac{P}{M}\mathbf{I}_M$ [3], i.e. equal power allocation is optimal. In such cases, capacity can be represented as

$$C_{\text{MIMO}} = \mathbb{E}\left\{\log_2 \det\left(\mathbf{I}_N + \frac{\text{SNR}}{M}\mathbf{H}\mathbf{H}^H\right)\right\}. \tag{2.5}$$

Defining $d = \min(M, N)$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ as the ordered singular values of $\mathbf{H}$, the capacity can be further expressed as

$$C_{\text{MIMO}} = \sum_{i=1}^{d} \mathbb{E}\left\{\log_2\left(1 + \frac{\text{SNR}}{M}\lambda_i^2\right)\right\}. \tag{2.6}$$

Under a high SNR assumption this can be approximated as

$$C_{\text{MIMO}} \approx d\log_2\frac{\text{SNR}}{M} + \sum_{i=1}^{d} \mathbb{E}\left\{\log_2\lambda_i^2\right\}. \tag{2.7}$$

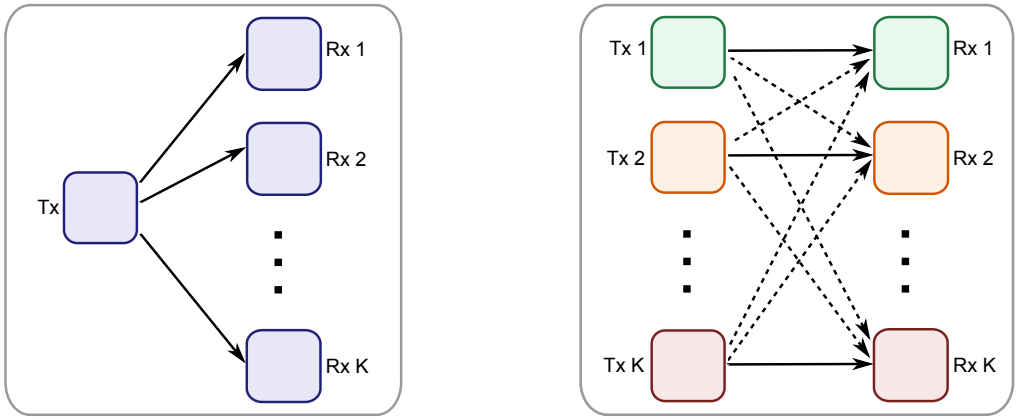
Additionally, $\sum_{i=1}^{d} \mathbb{E}\left\{\log_2\lambda_i^2\right\} = \sum_{m=|M-N|+1}^{\max\{M,N\}} \mathbb{E}\left\{\log_2\chi_{2m}^2\right\}$ where χ_{2m}^2 is a χ-square distributed random variable with $2m$ degrees of freedom. This allows us to represent the sum capacity as

$$C_{\text{MIMO}} \approx d\log_2\text{SNR} + o(\log_2\text{SNR}) \tag{2.8}$$

which shows a linear DoF gain of d in the rate pre-log factor, implying that full DoF are attained. For a SISO scenario, the corresponding capacity expression is $C_{\text{SISO}} \approx \log_2\text{SNR} + o(\log_2\text{SNR})$, where the rate pre-log factor of 1 represents the maximum achievable DoF for SISO links.

2.3 Multi-user systems

Moving beyond the point-to-point MIMO link illustrated in Fig. 2.1, real-world networks involve multiple receivers and transmitters that all need access to a limited amount of resources, giving rise to interference. There are several theoretical models that capture separate aspects of such communication systems. Two commonly studied models are the broadcast channel (BC) and the interference channel (IC), both depicted in Fig. 2.2. For the BC, a single transmitter wants to communicate with multiple users delivering a unique message to each.



(a) Broadcast channel. **(b)** Interference channel.

Figure 2.2: Broadcast channel and interference channel.

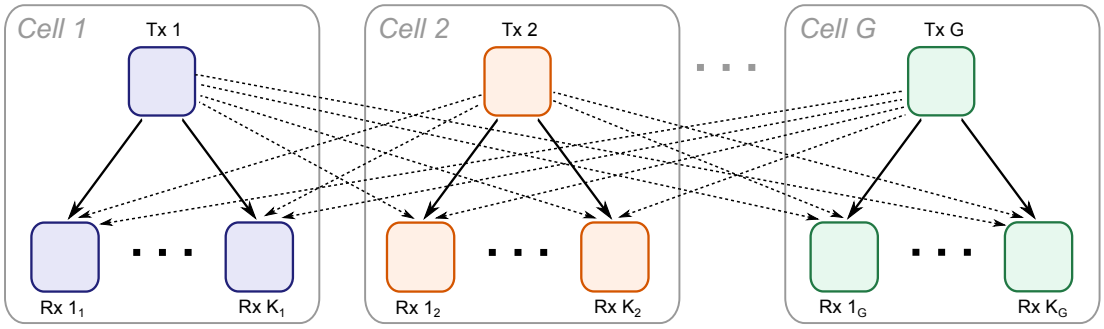


Figure 2.3: Interference broadcast channel.

Therefore, both the desired signal and interference are received over the same link, represented by the solid black arrows in Fig. 2.2(a). The IC consists of multiple point-to-point links interfering with each other, therefore desired signals and interference ones are received over separate links. This is reflected in the IC depicted in Fig. 2.2(b), where solid black arrows represent desired links and dashed black ones represent interference links.

A more complex network model is the interference broadcast channel (IBC), which captures both types of interference experienced in the BC and the IC. This consists of multiple cells that contain one transmitter and several receivers each, where every receiver requires a unique message from the corresponding transmitter. The IBC is depicted in Fig. 2.3, where the dashed black arrows represent *inter-cell interference* from other cell transmitters, and the solid black arrows represent links over which both the desired message and *intra-cell interference* are

received. The IBC models DL communication, its UL counterpart is the interference multiple access channel (IMAC) and consists of the scenario depicted in Fig. 2.3 with the direction of communication (represented by the arrows) reversed.

The IBC having all nodes equipped with multiple antennas is used as a starting point for all the HD scenarios considered throughout this thesis, since it provides a solid representation of real-world DL cellular systems where transmitters take the role of the base-stations (BSs) and receivers represent mobile users.

2.4 Interference management

Having seen the complex interference scenario posed by trying to serve multiple users simultaneously, it is clear that successful interference management strategies need to be devised in order to enable communication. From an information-theoretic perspective there are three traditional approaches to handling interference [4] as listed below.

- *Decode:* In cases where interference is strong, the interfering part of the signal can be decoded along with the portion of interest. The decodability of the interfering signals limits the users' rates and also the applicability of this technique in practice.
- *Treat as noise:* When interference is sufficiently weak, interference signals can be treated as noise. In such cases, single user encoding and decoding procedures can be applied.
- *Orthogonalise:* When both the signal of interest and the interference signals are of comparable strength, orthogonalisation techniques solve the issue by allocating totally separate channel access, thereby avoiding the occurrence of interference in the first place.

Practical approaches generally involve a combination of the last two techniques. Orthogonal time/frequency resources are allocated to neighbouring links of comparable strength, such that any resultant interference is weak enough to be treated as noise. While such solutions are successful in mitigating the effect of interference, they are wasteful with respect to the use of time and frequency resources. Therefore, the wireless communications research community is shifting towards trying to find innovative interference management solutions that enable

multiple users to be served in the same time/frequency resource; with the ultimate aim of developing systems that can meet the more demanding data rate requirements of future generation cellular networks. From an information-theoretic perspective these efforts have resulted in a much deeper understanding of the fundamental capacity limits of interference limited wireless networks. There have been numerous studies focusing on the determination of the number of achievable DoF as a first order characterisation of network capacity due to the high SNR relationship between these two metrics (see (2.8)). Results within this context show that the maximum achievable capacity is significantly higher than what is currently obtained via the use of conventional techniques. These have been obtained by exploiting the availability of CSIT [5], giving rise to a number of innovative techniques to manage interference via the design of novel transmission schemes and beamforming solutions.

The rest of this section will focus on two classes of novel interference management techniques, namely, interference alignment (IA) and topological interference management (TIM), due to their relevance to Chapters 3, 4 and 6.

2.4.1 Interference alignment

IA exploits the numerous DoF available from the time, frequency and spatial domains, in order to ensure that the interfering signals seen at each receiver occupy a low-dimensional subspace such that the desired signal can be decoded. By coding over multiple dimensions and carefully constructing the transmission and receiving strategies, IA maximises the number of non-interfering signals that can be simultaneously communicated over the interference network as a whole.

In their 2008 landmark paper [4] Cadambe and Jafar proposed an IA technique for the IC that aligns an arbitrarily large number of interferers, leading to the fundamental conclusion that wireless networks are not necessarily interference limited. The authors show that their proposed scheme can restrict interference at each receiver to approximately half of the received signal space, leaving the other half interference free for the desired signal. With this approach, a K-user IC where each node is equipped with a single antenna can achieve total DoF of $\frac{K}{2}$ rather than 1 as conventionally assumed. Starting from these initial highly promising IC results, which show that DoF can scale linearly with the number of users, IA has now evolved to address numerous types of interference scenarios

across a wide range of applications. Fig. 2.4 illustrates an IA example for a 3-user IC where each node is equipped with two antennas [5]. As can be noticed, by carefully constructing the precoders $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$ according to (2.9), interference can be made to overlap at the receivers. Each receiver sees two linear equations with three unknowns - since only one unknown is desired and the other two are aligned along the same dimension, the desired symbol can be decoded.

$$\begin{aligned}
\mathbf{V}_1 &= (\mathbf{H}_{1,2})^{-1}\mathbf{H}_{1,3}\mathbf{V}_3 \\
\mathbf{V}_2 &= (\mathbf{H}_{2,1})^{-1}\mathbf{H}_{2,3}\mathbf{V}_3 \\
\mathbf{V}_3 &= \text{eigvec}\big((\mathbf{H}_{2,3})^{-1}\mathbf{H}_{2,1}(\mathbf{H}_{3,1})^{-1}\mathbf{H}_{3,2}(\mathbf{H}_{1,2})^{-1}\mathbf{H}_{1,3}\mathbf{V}_3\big)
\end{aligned} \tag{2.9}$$

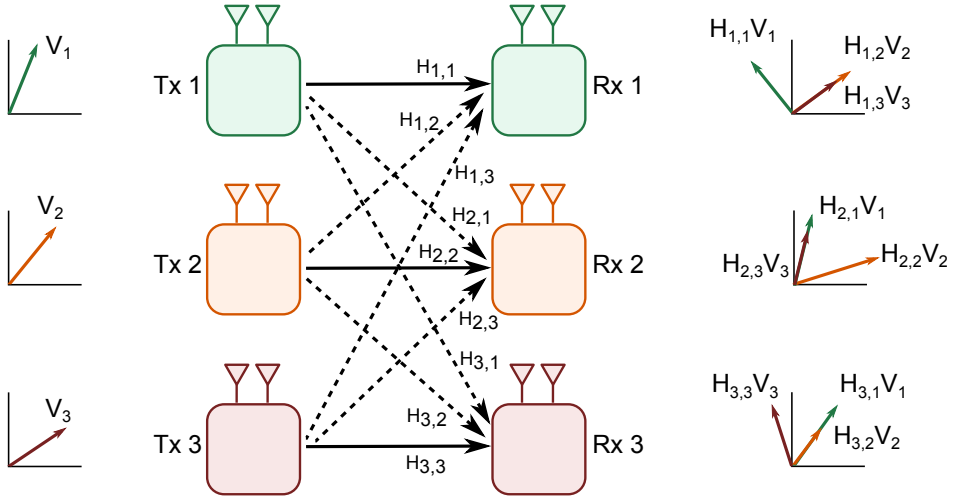


Figure 2.4: IA example for 3-user IC with two antennas at each node [5].

2.4.1.1 Types of IA and implementation challenges

As mentioned earlier IA can exploit a variety of domains (time, frequency and spatial) in order to increase achievable DoF, this has led to the development of a wide range of different IA techniques

The example depicted in Fig. 2.4 applies *linear IA*. This aligns signal spaces via the use of specifically designed transmit and receive beamformers, and provides a one-shot solution to the IA problem, making it highly accessible from a practical perspective [5]. However, as the number of interfering nodes increases, the amount of signals that need to be aligned also grows rapidly. This can cause linear IA problems to become infeasible, unless the number of antennas at the

various nodes is chosen appropriately to sustain the required DoF. Therefore, significant research effort has been targeted to understanding the feasibility of linear IA for a variety of network configurations [6–8]. Such studies derive relationships between the maximum number of streams that can be delivered simultaneously for a specific number of nodes, and the required number of transmit and receive antennas. Additionally, since closed-form solutions such as the one in (2.9) can only be obtained for a very restricted set of networks (e.g. [4, 9]), a parallel body of work has focused on developing iterative IA algorithms. These algorithms are used to test the IA feasibility conditions derived, and more importantly, to provide numerical IA beamforming solutions for various network types. Starting with initial studies on the IC [10, 11], several iterative approaches have been proposed, for example, minimising the leakage interference (Min-LI), Max-SINR and MMSE.

Linear IA schemes are based on spatial beamforming, and exploit the gains provided by having multiple antennas at the nodes. However, when the number of antennas is insufficient other solutions may be applied. One option is to use *symbol extensions*, which implies beamforming across multiple channel uses [12]. Such techniques require the channel to be unique for each channel use to obtain full rank matrices, and often result in schemes that only reach the promised DoF as the number of channel uses goes to infinity, for example, the *asymptotic IA* scheme proposed in [4]. Therefore, the concept of IA via symbol extensions is generally more useful from an information-theoretic perspective rather than a practical one. Other types of IA, such as *ergodic IA* and *blind IA*, rely on specific properties of the channels themselves. Ergodic IA pairs complementary channel states to naturally cancel interference by repetition coding across the identified state pairs [13]. Blind IA requires the channel states to be part of a known set of channel fluctuation patterns that are either naturally occurring [14] or enforced via the use of reconfigurable antennas [15].

The majority of IA techniques mentioned so far, except blind IA, require the availability of perfect CSI. In truth, dedicated system resources are needed to acquire this information, implying that perfect and instantaneous delivery of CSI to all network nodes is not possible in practice. Recognising the highly idealistic nature of the initial studies, research is moving into interference management for scenarios where the CSI assumptions are more relaxed. This is fundamental in order to obtain results that are more closely relatable to practical systems and

has been considered in numerous works, for example by recognising the effect of imperfect, delayed [16], mixed [17, 18] or partial [19] CSIT.

An in depth review of the types of IA available in literature and the challenges involved in their implementation may be found in [5]. For the purpose of this thesis linear IA is of particular interest, since it is the most easily accessible form of IA from a practical perspective [5]. Additionally, imperfect CSI also plays a central role. In Chapters 4 and 6, we focus on the performance of linear IA under imperfect CSI, and develop iterative algorithms that are applicable to the HD MIMO IBC (Chapter 4) and the FD MIMO multi-user multi-cell network (Chapter 6). The availability of imperfect CSI is also considered in Chapter 5 which focuses on WSR maximisation in FD MIMO multi-user multi-cell networks.

2.4.2 Topological interference management

IA studies generally start from a perfect and abundant CSIT assumption and then try to move into more relaxed CSIT scenarios that are more relevant to practical situations. Topological interference management (TIM) introduced in [20] offers a new but complementary perspective. Rather than starting with an idealistic CSIT assumption, it considers the issue from the opposite end of the spectrum as shown in Fig. 2.5; with the ultimate aim of reaching a compromise where higher rates can be achieved in practical CSIT settings.

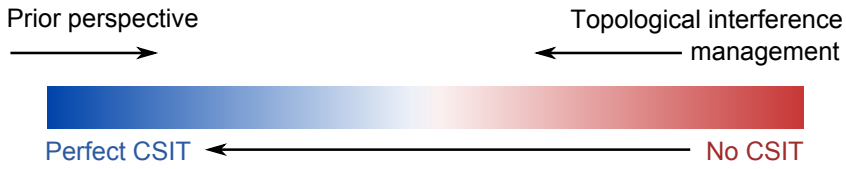


Figure 2.5: Motivation behind TIM.

For wireless networks, having no CSIT at all is a degenerate setting where no reliable communication can be guaranteed. Thus, to ensure that the resultant problem is non-degenerate, the TIM framework assumes that the desired channel links are of sufficient quality to sustain the required SNR for reliable communication in the absence of interference. Using the K-user SISO IC as an example, this interference free SNR guarantee can be expressed as

$$\frac{|h_{i,i}|^2 P_i}{\sigma^2} \geq \text{SNR} \quad \forall\, i \in \{1, \ldots, K\} \tag{2.10}$$

where $h_{i,i}$ represents the desired channel link, P_i is the transmit power constraint and σ^2 is the noise power. Therefore, if no interference is present, each user can achieve a rate of $\log_2(1 + \text{SNR})$. The capacity of the K-user SISO IC under such conditions can be achieved via the use of orthogonal schemes like time division multiple access (TDMA), and corresponds to $\frac{1}{K}\log_2(1 + K\text{SNR})$ per user, equivalent to $\frac{1}{K}$ DoF each.

The application of orthogonal techniques leads to highly conservative values for achievable capacity and DoF. Such results are obtained under two rather strong and pessimistic assumptions: (a) all possible interference links exist and are all equally significantly strong (this is a worst case scenario assumption since fading is an intrinsic characteristic of wireless networks, and it is therefore natural for some links to be weak enough to be considered non-existent), and (b) no information is available with respect to the interfering links. Moreover, literature proposing innovative interference management alternatives to the traditional orthogonal based techniques, tends to combine the overly pessimistic fully connected scenario hypothesis, with a highly idealistic full CSIT availability assumption, rendering the solutions proposed highly difficult to implement in practice.

The TIM framework aims to find a middle ground that moves beyond both, (a) the overly cautious fully connected scenario assumption, and also (b) the extreme either full or no CSIT assumptions with respect to the interfering links. This is achieved by classifying all interference links into two main categories as follows.

- *Weak interference links*: links over which the nominal received power is below a pre-established threshold value equivalent to the noise floor.
- *Strong interference links*: links over which the nominal received power is above this threshold.

Therefore, just one bit of CSIT per interfering link is required to provide topological information to all the transmitters present in the network; weak links are assigned a '0', while strong links are assigned a '1'. This ensures that feedback related overheads are significantly lower, and that the CSIT requirements can be met in practice. Thus, for the general wireless TIM framework, apart from the interference free SNR guarantee for the desired channel (analogous to (2.10)), transmitters only need know the network structure, i.e. they are aware of the existence of interference links but have no knowledge of their exact channel

value.

To understand the potential benefits offered by TIM more clearly, consider for example the partially connected 5-user IC network in Fig. 2.6. If we only know the value of the desired channels, then we have the interference free SNR guarantee in (2.10) and can apply TDMA to achieve $\frac{1}{5}$ DoF per user. However, if topological information is available higher DoF can be achieved. Using orthogonal techniques that schedule non-interfering users groups simultaneously results in achievable DoF per user of $\frac{1}{3}$. Additionally, [20] shows that with fixed channels under the TIM framework, even higher DoF can be obtained by constructing the transmitted signals intelligently such that interference is aligned. This is achieved by activating transmitters $\{1, 2, 5\}$ in time slot 1, and transmitters $\{2, 3, 4\}$ in time slot 2. Receivers 1 and 2 obtain their desired symbols directly in time slot 1, similarly receivers 3 and 4 obtain their symbols directly in time slot 2. Lastly, receiver 5 can subtract the signal received in time slot 2 from the signal received in time slot 1 to obtain its desired symbol. This results in a symbol being delivered to each user over two time slots, equivalent to $\frac{1}{2}$ DoF per user.

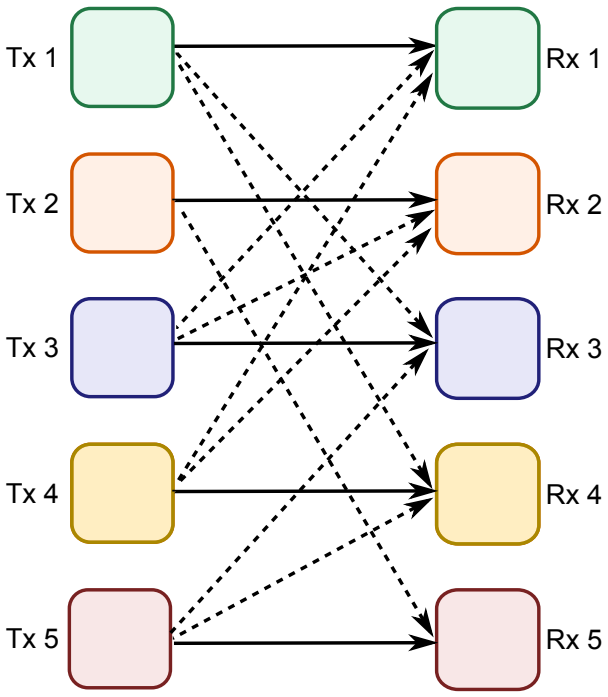


Figure 2.6: Partially connected 5-user IC; solid arrows represent desired links, dashed arrows represent interference links [20].

Furthermore, [20] draws a highly interesting parallel between wireless and wired networks, showing that with the use of linear network coding in wired

networks, the distinction between these two types of networks is virtually non-existent under the TIM framework. While traditionally interference is not an issue in standard wired networks, with the application of linear network coding at intermediate nodes, the resulting network experiences interference between different flows, such that the underlying structure is equivalent to that of the corresponding wireless network. The work in [20] shows how the innovative interference management solutions derived under the TIM framework for wireless networks can be transferred to the wired network equivalents, with the emphasis being placed on the underlying network structure, rather than the value of the channel links. Therefore, TIM allows us to obtain a unified view of wired and wireless networks, such that the normalised network capacity for both networks is identical. In this case, DoF can be viewed as a representation of the capacity of the underlying noiseless linear communication network, where the received signals are a linear combination of the transmitted ones. This generally translates to a direct capacity result for wired networks, since noise is not an issue, and corresponds to a first order high SNR capacity approximation for wireless ones, where noise is unavoidable.

The study in [20] considers partially connected networks where connectivity is fixed throughout the duration of communication. Within this thesis, in Chapter 3, we move beyond this assumption and consider the use of TIM to manage inter-cell interference in a two-cell two-user per cell network with alternating inter-cell connectivity. For the alternating connectivity scenario, the network may range from a fully connected one in the worst case scenario, to one without any inter-cell interference in the best case scenario. While our main focus is on wireless networks, we also show how the results obtained results can be translated into capacity results for the corresponding wired network equivalents.

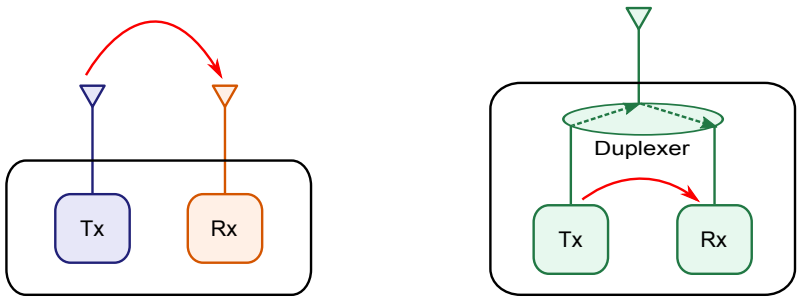
2.5 Full-duplex communication

Traditionally RF nodes were considered to be unable to transmit and receive in the same frequency band simultaneously due to self-interference (SI) between the transmitted and received signals [21]. Effects like path loss and fading cause signals received over the air to be much weaker than those transmitted, rendering it difficult to detect the desired signal. Current communication systems avoid creating SI by operating in HD mode. Bi-directional communication is enabled

via the use of orthogonal channels in the time and frequency domains; this is typically achieved through the use of techniques like time division duplex (TDD) or frequency division duplex (FDD) that provide complete separation between the transmit and receive signals. However, such methods lead to an inefficient use of the limited RF spectrum. This waste, combined with the ever increasing data rate demands of wireless networks, has driven the wireless communications research community to aspire to the realisation of FD nodes that can concurrently transmit and receive information in the same time/frequency resource. The realisation of FD technology promises to improve spectral efficiency by a factor of two compared to conventional HD systems.

2.5.1 FD node architecture

FD nodes can generally be implemented in two different ways as depicted in Fig. 2.7. One method is *separate antenna architecture* [22], where each transmit chain uses a dedicated radiating antenna and each receive chain uses a separate sensing antenna. The second method is *shared antenna architecture* [23] where a duplexer (usually a ferrite device that exploits non-linear propagation in magnetic materials, known as a circulator) is attached to the antenna, and used to direct the received signal to the receiver and route the transmit signal from the transmitter to the antenna, while also isolating the two chains.



(a) Separate antenna. **(b)** Shared antenna.

Figure 2.7: FD node architecture, red solid arrows represent the SI path.

2.5.2 SI suppression

In order to enable the practical realistion of FD systems a lot of research effort has been directed towards identifying ways to suppress SI. To understand the impact of SI more clearly, let us focus on a femto cell cellular system example

from [24], depicted in Fig. 2.8. Here, femto BSs and mobile handsets transmit at a power of 21 dBm, with a receiver noise floor of 100 dBm. Assuming that the FD node can naturally isolate the transmit and receive signal paths by 15 dB[1], then the SI interference will be 106 dB above the noise floor; implying that SI suppression of at least 106 dB must occur in order to achieve a link SNR equal to that of the HD counterpart. Note that for larger cells, where transmit powers are higher, even more SI would need to be suppressed.

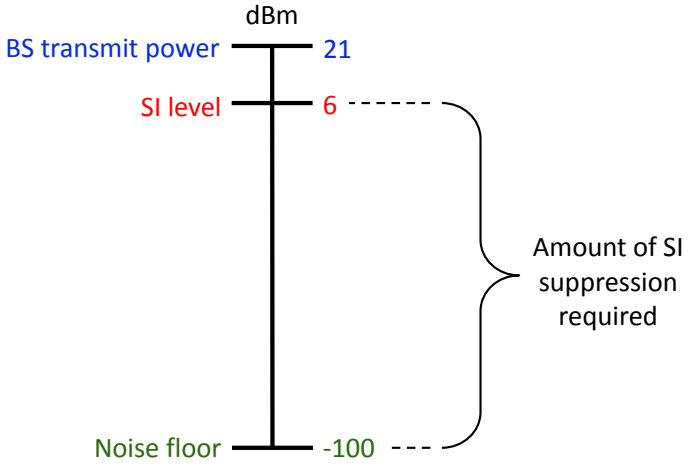


Figure 2.8: Amount of SI suppression required for femto cell example from [24].

Since the transmitted signal is known, one might erroneously think that perfect SI cancellation is possible by simply subtracting the transmitted signal from the received one; however, this cannot be done in practice. For digital domain only cancellation, issues like the dynamic range of the analogue-to-digital converter (ADC), quantisation noise, oscillator phase noise and non-linearities in the amplifiers and mixers, act as a bottleneck for the effectiveness of digital SI cancellation and lead to some residual SI. This has motivated a more holistic approach to SI suppression, where techniques that can be applied both before and after the ADC are combined together, in an effort to achieve acceptable levels of SI suppression. The different methods can be classified into three main categories as follows [24].

- *Propagation domain SI isolation:* Propagation domain techniques aim to

[1]An SI isolation figure of 15 dB is quite conservative. Larger values have been reported in literature for different antenna architectures, for example up to 45 dB of SI isolation is achieved in [25]. See also [22, 26, 27] for further information.

suppress SI before it manifests itself at the receiver chain circuitry, by isolating the transmit and receive chains. For separate antenna systems, propagation domain SI suppression is achieved using a combination of: (a) path loss - by increasing spacing or by placing absorptive shielding between the transmit and receive chains, (b) cross-polarisation - for example by having a node that only transmits horizontally polarised signals and only receives vertically polarised ones or vice-versa, and (c) antenna directionality - by ensuring that the main radiation lobes of the transmit and receive antennas have minimal intersection. For shared antenna systems propagation domain SI is achieved by the duplexer.

- *Analogue circuit domain SI cancellation:* Analogue circuit domain methods aim to suppress SI in the receive chain circuitry before the ADC. Such techniques include classic time domain training based methods that estimate the SI leakage and then apply signal inversion to cancel it. Additionally, for MIMO nodes the increased spatial DoF offered by having multiple antennas may also be exploited to mitigate SI. In each case, the main principle behind analogue SI suppression is the idea of introducing a cancelling signal to diminish the amount of SI experienced at the receive side.

- *Digital circuit domain SI cancellation:* Digital circuit domain methods aim to suppress SI after the ADC, by using knowledge of the SI signal and applying digital signal processing techniques to mitigate its effects. Operating in the digital domain has the advantage of rendering sophisticated signal processing techniques relatively easy to implement compared to the analogue domain. However, hardware impairments like for example the ADC's dynamic range, and noise and non-linearities in various other components, limit the amount of SI reduction possible. Thus, digital domain SI suppression techniques are generally considered as the final resort to cancel residual SI left over from the previously applied propagation domain and analogue circuit domain SI cancellation techniques.

Fig. 2.9 shows how the various SI suppression techniques are combined together in order to obtain the final received signal that can be used for decoding. Recent advances in antenna design and RF circuitry have demonstrated highly promising results in terms of SI suppression levels. For example, [28] uses a combination of signal inversion and digital cancellation to achieve 73 dB of SI suppression for a 10 MHz orthogonal frequency division multiplexed (OFDM) signal.

In [29] all three classes of SI cancellation methods are combined to achieve an average cancellation of 85 dB, with a minimum of 70 dB and a maximum of 100 dB, over a 20 MHz signal. Additionally, [25] proposes a single-antenna FD node prototype that can achieve $40-45$ dB of SI isolation before analogue and digital cancellation. Another single-antenna design is proposed in [23]; this achieves up to 110 dB of SI suppression in total, over an 80 MHz bandwidth.

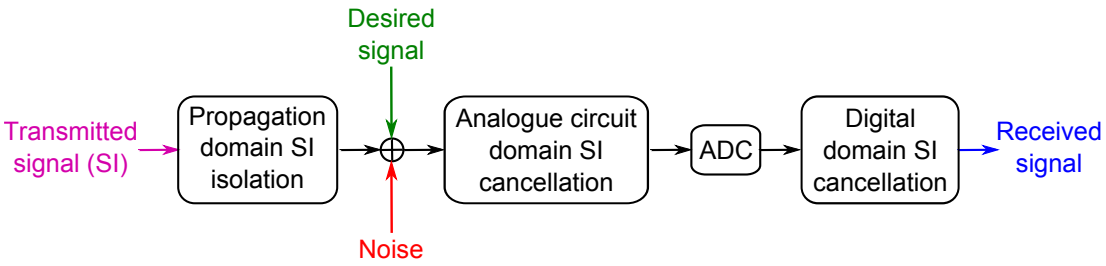


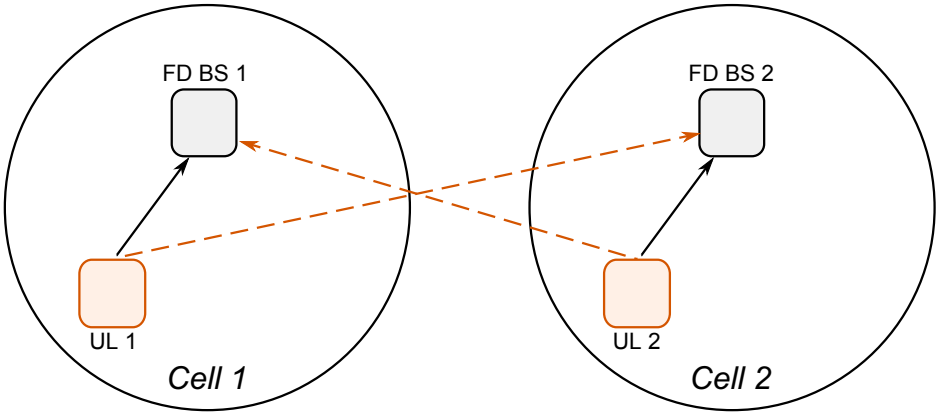
Figure 2.9: Implementation of different SI suppression techniques [30].

2.5.3 Interference in FD networks

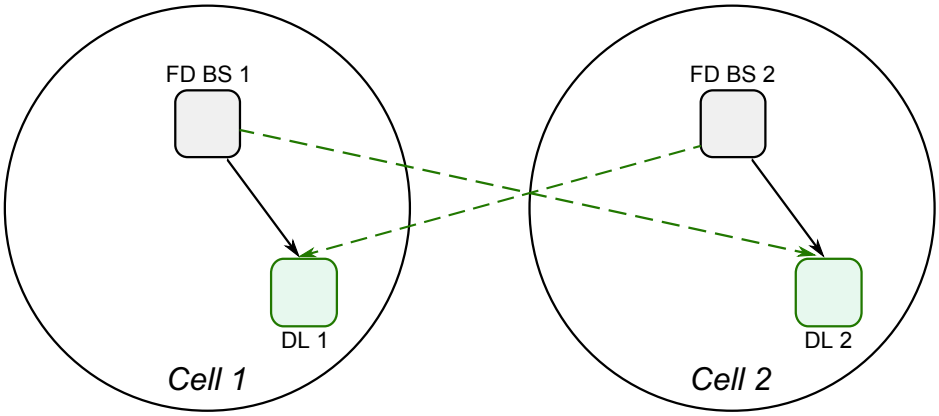
Given the fact that sufficient SI cancellation is well on its way to becoming practically feasible, the next stage is to consider wirelesses networks containing FD enabled nodes. While it is possible to construct networks where all nodes operate in FD mode, since FD operation requires significant hardware changes with higher costs and power usage, it is more practical to initially consider scenarios where only the infrastructure elements (BSs) are upgraded to FD, with user devices still operating in HD mode [31].

The FD capability at the BSs allows both UL and DL users to be served simultaneously, this leads to a surge in the amount of interference experienced across the network compared to its HD counterpart. Consider for example a two-cell network with one BS, one UL user and one DL user per cell, and compare HD operation in Fig. 2.10 with FD operation in Fig. 2.11. For the HD network there are two inter-cell interference links in both UL and DL scenarios. For FD operation, the UL and DL interference links seen in each HD scenario appear simultaneously, in addition to a number of new interference components depicted in blue in Fig. 2.11. The novel interference components that occur when replacing HD BSs with FD ones are:

- *residual SI* - as discussed earlier in Section 2.5.2 there are numerous ways to mitigate SI, however none of them can cancel it completely;



(a) Serve only UL users.



(b) Serve only DL users.

Figure 2.10: HD two-cell scenario with one UL and one DL user per cell. UL and DL users are scheduled in separate time/frequency resources.

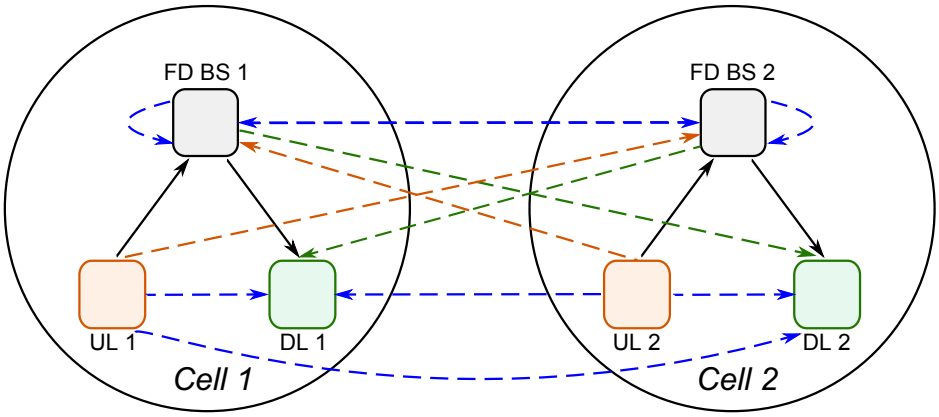


Figure 2.11: FD two-cell scenario with one UL and one DL user per cell.

- *BS-to-BS interference* - since BSs are now transmitting and receiving at the same time, DL data from one BS interferes with UL information desired at the neighbouring BS and vice-versa;
- *UL user to DL user interference* - transmission from UL users, both in the same cell and in neighbouring ones, interferes with the received signal at DL users.

Considering the complex interference scenario that results by replacing HD BSs with FD ones, it is clear that having appropriate techniques to handle them is fundamental for the practical realisation of FD networks. In this thesis we focus on multi-cell multi-user networks with FD BSs and HD users. In Chapter 5, we propose solutions for WSR maximisation both under perfect and imperfect CSI. While, in Chapter 6, we consider the use of linear IA for the same system in the presence of CSI mismatch.

Chapter 3

Topological Interference Management for Interference Broadcast Channels

3.1 Introduction

As highlighted earlier in Chapter 2, recent years have seen major advances in terms of understanding the information-theoretic capacity limits of interference limited networks, mainly under the assumption of abundant CSIT. While this has given rise to a number of innovative ways on how to exploit different aspects of CSIT, the theoretical gains have been difficult to translate into practical ones due to the idealistic CSIT requirements.

Therefore, moving on from the initial perfect CSIT studies, the current research direction is to focus on more relaxed assumptions in order to reach a middle ground where higher rates can be achieved with realistic CSIT requirements. Some works rely on specific properties of the channel links themselves, such as ergodic IA [13] (outlined earlier in Section 2.4.1.1), or the technique proposed in [35] for finite state compound wireless networks, which relies on the fact that the channel realisations come from a finite set of possibilities. Other works focus on using the available CSIT even though it is not perfect. For example, [16]

Work in this chapter has been published in IEEE Transactions on Communications, April 2016 [32], with a preliminary version presented at IEEE ICASSP 2015 [33] and IEEE ISIT 2015 [34].

and [36] show that even completely delayed CSIT provides a gain in achievable DoF. Scenarios with both delayed and imperfect current CSIT are considered in [17] and [37]. A combined setting where the CSIT alternates between perfect, delayed and unavailable is analysed in [38]. Literature mentioned so far assumes all transmitters have an identical view of the network; however, this is not always a pre-requisite. For example, in [19] transmitters only have perfect CSIT for a restricted subset of the global channel links, with this subset being specific to each transmitter. Additionally, situations where nodes have asymmetric local views of the global network structure are also considered in [39].

The TIM approach [20] (a general overview of which was provided earlier in Section 2.4.2) offers a new but complementary perspective to the interference management problem, starting from a limited CSIT availability perspective rather than an abundant one. The work in [40] considers such a setting for the case where transmitters cooperate via message sharing; results show that considerable DoF gains can be obtained for networks that are not fully connected. Additionally, [41] expands the DL hexagonal cellular network scenario from [20] to include multiple layers of interference and analyses how these affect the corresponding DoF gains. Throughout the studies in [20, 40, 41] it is assumed that network topology is fixed for the duration of communication. For this chapter, we move beyond such an assumption, and consider a scenario where inter-cell connectivity may vary in order to analyse the DoF gains that can be achieved. The overall setting is referred to as an alternating connectivity scenario and was considered for the two-user SISO IC and the X-channel in [42], and three-user SISO ICs with various restrictions in [43] and [44]. Here, we focus on the more complex IBC which has the additional challenge of intra-cell interference, and also introduce a mixed CSIT setting where global topological knowledge is combined with varying degrees of local CSIT.

The main contribution of our work is in the derivation of novel DoF outer bounds for the two-cell two-user-per-cell IBC with alternating connectivity. While our initial focus is on a SISO system, we also consider MISO and MIMO configurations as a means of resolving intra-cell interference. Global channel knowledge is restricted to topological information only; however, local CSIT availability varies depending on the system configuration itself, leading to a mixed CSIT setting for the MISO case. The achievability of the derived bounds is investigated for a variety of contexts. Additionally, we propose novel transmission schemes based

on joint coding across states that are applicable for arbitrary state probabilities and analyse their performance, both for the general case and for situations where all states are equiprobable. Results show that DoF higher than those conventionally obtained without global topological knowledge can be achieved, proving that even such a minimal level of global CSIT is still very useful.

The rest of this chapter is organised as follows. Section 3.2 provides some preliminaries by introducing the system model, the alternating connectivity scenario, and the local CSIT availability. Next, in Section 3.3 we present the DoF outer bound for the SISO system, while the MISO and MIMO counterpart is provided in Section 3.4. The achievability of the derived DoF outer bounds is investigated in Section 3.5 and Section 3.6 respectively. Section 3.7 shows how the wireless network DoF results can be translated into capacity results for the corresponding wired network instances. Finally, Section 3.8 provides some concluding remarks. Additionally, there are three appendices; the first two provide extra details required to complete the outer bound derivations, while the third contains a useful lemma relavant to this chapter.

3.2 Preliminaries

3.2.1 System model

We consider a two-cell two-user-per cell IBC. This consists of two adjacent cells in a wireless network, where the first cell includes BS A and receivers $a1$ and $a2$, while the second cell has BS B and receivers $b1$ and $b2$. The basic network structure is shown in Fig. 3.1, where inter-cell interference links are omitted and the solid arrows represent useful links over which the desired symbols are delivered. The general input-output relationship is given by,

$$Y^r[n] = H_{r,A}[n]X^A[n] + H_{r,B}[n]X^B[n] + Z^r[n] \tag{3.1}$$

where at channel use index $[n]$, $Y^r[n]$ is the signal observed at receiver r for $r \in \{a1, a2, b1, b2\}$, $X^T[n]$ is the signal sent from transmitter T for $T \in \{A, B\}$, $Z^r[n]$ represents unit variance AWGN at receiver r, and $H_{r,T}[n]$ is the channel link between transmitter T and receiver r, whose entries are i.i.d. and drawn from a continuous distribution. Additionally, $\mathbb{E}\{\|X^T[n]\|^2\} \leq P\mathbf{I}$, where P represents the transmit power constraint and is equal to SNR for unit power AWGN.

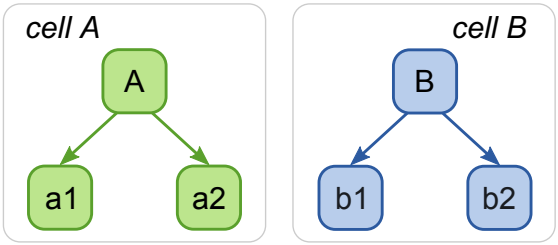


Figure 3.1: Two-cell two-user-per-cell network with omitted inter-cell interference links.

Note that throughout this chapter we use standard capital font for Y, X and H as a general notation to encompass their possible scalar, vector or matrix (only for H) forms depending on the SISO, MISO or MIMO configurations as highlighted later in Section 3.2.2. Moreover, for notational simplicity, the channel use index $[n]$ will be omitted from here onwards. Additionally, since all noise terms are drawn from the same distribution, they are all statistically equivalent, thus the general notation Z will be used throughout the rest of this chapter.

Within the scenario under analysis, inter-cell interference can occur between any of the users and the non-corresponding BS. We consider an alternating connectivity context, where inter-cell connectivity is not fixed throughout the duration of the whole communication process. Connectivity can easily vary in wireless networks, where some links may go into deep fade making them effectively non-existent. Additionally, in frequency selective environments, frequency hopping or multi-carrier transmission may also create a variety of inter-cell connectivity states. For the system in Fig. 3.1, a total of 16 different states may occur, as shown in Fig. 3.2. Each of these states is associated with a probability of occurrence λ_k for $k = 1, \ldots, 16$, where $\sum_{k=1}^{16} \lambda_k = 1$. Note that to ensure the problem is non-degenerate, desired links are considered to be always present and able to support the desired rate in the absence of interference.

3.2.2 Antenna configuration and CSIT availability

For each cell, we define M_B as the number of antennas at the BS and N_d as the number of antennas at each of the two receivers. With an appropriate choice of M_B and N_d, spatial multiplexing can be applied within the cells to resolve intra-cell interference, such that each BS can simultaneously deliver one symbol to its corresponding two users. This results in total achievable DoF of 2 per cell,

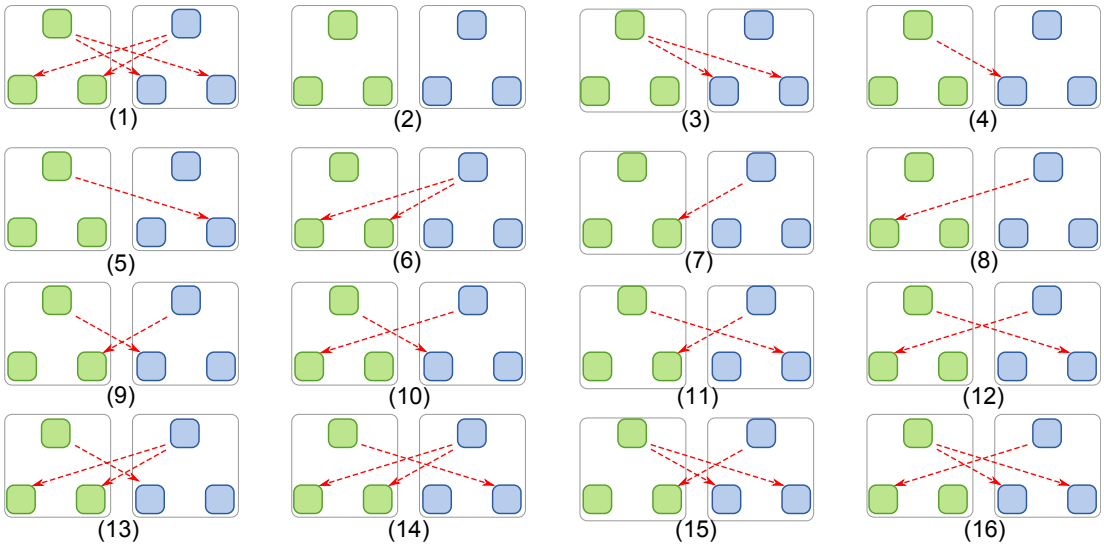


Figure 3.2: Set of all possible inter-cell connectivity states for the two-cell two-user-per-cell IBC. Cell A transmitters and receivers are on the left in green, while cell B elements are on the right in blue. The dashed red arrows represent interference links.

provided no inter-cell interference is present.

For a mixed CSIT[1] setting, where in addition to global topological CSIT, perfect current local CSIT is also available, the achievable DoF per cell are given by [45]

$$\min\{M_B, KN_d\}$$

where K represents the number of users in each cell. Here we consider a scenario where each cell has two users; thus, any $M_B \times 1$ system (where for MISO $M_B \geq 2$ by definition) achieves the required 2 DoF per cell. This is possible via zero-forcing (ZF) precoding. Consider a general cell C having users $c1$ and $c2$, where BS C transmits a combined symbol, X^C, consisting of s_{c1} intended for user $c1$ and s_{c2} intended for user $c2$. Given the availability of local CSIT, s_{c_1} can be precoded such that it is orthogonal to the channel from BS C to user $c2$, and s_{c2} can be precoded such that it is orthogonal to the channel from BS C to user $c1$. This allows users to extract their desired symbol from a single observation of X^C, thereby achieving 2 DOF within the cell if no inter-cell interference is present.

[1]Throughout this chapter the term *mixed CSIT* is used to refer to a mixture of global topological information and perfect current local CSIT. This is different to prior usage of the term in [17] and [18], where it is used to refer to a mixture of perfect delayed CSIT and imperfect current CSIT.

On the other hand, if local CSIT is not available, the achievable DoF per cell are equal to [46]

$$\min\{M_B, N_d\}\,.$$

Therefore, any $M_B \times 2$ or $2 \times N_d$ MIMO system (where by definition for MIMO $M_B \geq 2$ and $N_d \geq 2$), can achieve the required 2 DoF per cell. Consider a general cell C having users $c1$ and $c2$, where the BS transmits a combined symbol, X^C, consisting of s_{c1} and s_{c2}, and the antenna configuration is either $M_B \times 2$ or $2 \times N_d$. Due to the multiple antenna setting, each user can obtain at least two independent equations for the two unknown symbols s_{c1} and s_{c2}, and can therefore decode for the desired one. This results in achievable DoF of 2 per cell in the absence of inter-cell interference.

Note that for the SISO scenario spatial multiplexing is not an option, since by definition $M_B = N_d = 1$; thereby, even in the abscence of inter-cell interference, only 1 DoF per cell can be achieved.

Regardless of the system configuration, if no feedback is available with respect to the alternating global network topology, both transmitters have to assume full inter-cell connectivity throughout, i.e. State 1 in Fig. 3.2. This only allows for one possible transmission strategy, where BS A and BS B are provided with non-overlapping transmission opportunities and leads to a sum DoF across the two cells of: 1 for the SISO configuration, and 2 for the MISO system with local CSIT or the MIMO one without local CSIT. Considering all the states in Fig. 3.2, it is clear that assuming full connectivity throughout is wasteful in terms of network resource use - states 2 to 16 have a smaller amount of inter-cell interference links, and can potentially achieve higher sum DoF than the fully connected scenario in state 1.

Our interest lies in exploiting this opportunity, whilst keeping the global CSIT requirement to a minimum. Therefore, while varying degrees of local CSIT are considered, global knowledge of the inter-cell interference channels is restricted to topological information only. This requires just one bit of CSIT per inter-cell interference link, used to indicate whether interference may be experienced over that link or not. Receivers can compare the nominal received power from an undesired link to a pre-established threshold value and assign a '0' to those links for which the received power is below the threshold and a '1' to those links for which the received power exceeds it; this information is then fed back to the transmitters making them aware of the network's topological structure [20].

3.2.3 Reformulation for SISO scenario

While the setting described so far is sufficient to analyse the MISO and MIMO IBC scenarios, we also consider a SISO scenario which requires further reformulation. For SISO systems, maximum achievable DoF per cell equal to 1 can be achieved simply by avoiding intra-cell interference and serving only one user at a time. Hence from a DoF outer bound perspective, we can consider the case where for every instant each BS selects one user to be its designated user to serve, according to what is most advantageous in terms of achievable sum rate. Let us define U as the cell A designated user, $U \in \{a1, a2\}$, and V as the cell B designated user, $V \in \{b1, b2\}$. For any given U and V, the original network in Fig. 3.1 can be represented with the equivalent one in Fig. 3.3, where only four (U, V) combinations may occur, i.e.

$$(U, V) \in \{\ (a1, b1),\ (a2, b1),\ (a1, b2),\ (a2, b2)\ \}\ . \tag{3.2}$$

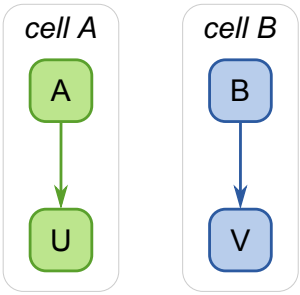


Figure 3.3: Equivalent network for SISO scenario, where U represents the cell A designated user and V represents the cell B designated user.

Having defined an equivalent network for the SISO scenario, the set of 16 alternating states from Fig. 3.2 can be mapped to a reduced set of only 4 possible states, as in Fig. 3.4. For any given (U, V) combination it only matters whether inter-cell interference affects the designated users. For example, when $(U, V) = (a1, b1)$ in state 9 from Fig. 3.2, $a1$ is free from inter-cell interference while b_1 receives interference from BS A; thus, from the perspective of this particular (U, V) combination, state 9 corresponds to state S in Fig. 3.4. Next, consider state 9 from the perspective of $(U, V) = (a1, b2)$; in this case both $a1$ and $b2$ are free from inter-cell interference, hence state 9 is mapped to the no interference state T in Fig. 3.4. Similar arguments can be made for all (U, V) combinations listed in (3.2) and all the states depicted in Fig. 3.2. The mapping of the

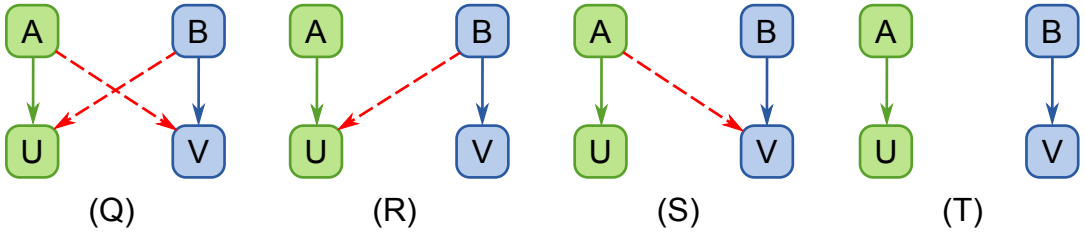


Figure 3.4: Reduced set of states used to replace original ones from Fig. 3.2 when considering the equivalent network for the SISO scenario in Fig. 3.3.

Table 3.1: Mapping of original states from Fig. 3.2 to the equivalent reduced set in Fig. 3.4 for each possible (U, V) combination.

$(\mathbf{U}, \mathbf{V})$	**Q**	**R**	**S**	**T**
$(a1, b1)$	$1, 10, 13, 16$	$6, 8, 12, 14$	$3, 4, 9, 15$	$2, 5, 7, 11$
$(a2, b1)$	$1, 9, 13, 15$	$6, 7, 11, 14$	$3, 4, 10, 16$	$2, 5, 8, 12$
$(a1, b2)$	$1, 12, 14, 16$	$6, 8, 10, 13$	$3, 5, 11, 15$	$2, 4, 7, 9$
$(a2, b2)$	$1, 11, 14, 15$	$6, 7, 9, 13$	$3, 5, 12, 16$	$2, 4, 8, 10$

original set of states from Fig. 3.2 to the reduced one in Fig. 3.4 for each (U, V) combination is provided in Table 3.1.

3.3 DOF outer bound for SISO IBC

In this section we present a DoF outer bound for the two-cell two-user-per-cell SISO IBC with alternating connectivity.

Theorem 3.1. *For the two-cell two-user-per-cell SISO IBC with alternating connectivity, the sum DoF, $d_{\Sigma,S}$, can be characterised as*

$$d_{\Sigma,S} \leq 2 - \Lambda$$

where

$$\Lambda = \max \begin{cases} \lambda_3 + \lambda_6 \\ \lambda_1 + \lambda_6 + \lambda_{13} + \lambda_{14} \\ \lambda_1 + \lambda_3 + \lambda_{15} + \lambda_{16}\,. \end{cases} \tag{3.3}$$

Proof. The overall outer bound consists of merging together bounds originating from different sources; one comes from the summation of the achievable rates per

cell, and an additional pair arises from genie aided bounds for each cell. Here, we present a compact version of the proof focusing on how the cell A expressions are obtained, details for their cell B counterparts are provided in Appendix 3.A.

3.3.1 Sum bound

For the sum bound, first we obtain separate expressions for the achievable rate within each cell, these are then combined to give an overall outer bound for the sum DoF across the two cells. Starting with the cell A achievable rate, we have

$$nR_{A,S} \leq I(W^A; Y_1^U, \ldots, Y_{16}^U) + n\epsilon$$

where W^A is the message set from BS A, and Y_k^U is the signal received by the cell A designated user U during state k. This can be further expressed as

$$\begin{aligned} nR_{A,S} &\leq h(Y_1^U, \ldots, Y_{16}^U) - h(Y_1^U, \ldots, Y_{16}^U \mid W^A) + n\epsilon \\ &\overset{(a)}{=} h(Y_1^U, \ldots, Y_{16}^U) - h(Y_Q^U, Y_R^U, Y_S^U, Y_T^U \mid W^A) + n\epsilon \\ &= h(Y_1^U, \ldots, Y_{16}^U) - h(Y_R^U, Y_S^U, Y_T^U \mid W^A) - \underbrace{h(Y_Q^U \mid W^A, Y_R^U, Y_S^U, Y_T^U)}_{\substack{\geq h(Y_Q^U \mid W^A, Y_R^U, Y_S^U, Y_T^U, W^B) \\ = no(\log_2 P)}} + n\epsilon \\ &\overset{(b)}{\leq} h(Y_1^U, \ldots, Y_{16}^U) - h(Y_R^U, Y_S^U, Y_T^U \mid W^A) + no(\log_2 P) + n\epsilon \end{aligned} \tag{3.4}$$

where (a) follows since the original set of 16 states are all contained within states Q, R, S and T for the equivalent SISO scenario, and (b) follows since conditioning reduces entropy and the effect of noise disappears at high SNR. Note that $o(\cdot)$ comes from the standard Landau notation, where $f(x) = o(g(x))$ implies $\lim_{x\to\infty} f(x)/g(x) = 0$.

Considering (3.4) and the state configurations in Fig. 3.4, it can be noticed that the received signal for the cell A designated user U in states S and T consists only of an X^A component and noise. The X^A component has no effect on entropy since it is solely a function of W^A, while the effect of noise disappears as $P \to \infty$ and can be integrated into the $no(\log_2 P)$ term, resulting in

$$nR_{A,S} \leq h(Y_1^U, \ldots, Y_{16}^U) - h(Y_R^U \mid W^A) + no(\log_2 P) + n\epsilon\,. \tag{3.5}$$

For all states corresponding to R, the cell A received signal is combination of X^A,

X^B and noise, i.e. it takes the form of $Y_R^U = H_{U,A}X^A + H_{U,B}X^B + Z$. The X^A component is negligible with respect to entropy. The X^B and noise components are independent of W^A. Additionally, since $H_{U,B}$ and $H_{V,B}$ are independently drawn from the same distribution, they are statistically equivalent and interchangeable [46]. Therefore, the X^B and noise terms can be represented by the signal received at the cell B designated user V, provided that V itself has no inter-cell interference. Comparing the list of all R states from Table 3.1, it can be noticed that this substitution is guaranteed as being always possible regardless of the current (U, V) combination only for state 6. Using this information, the cell A rate outer bound from (3.5) can be expressed as

$$\begin{aligned} nR_{A,S} &\leq h(Y_1^U, \ldots, Y_{16}^U) - h(H_{U,A}X_6^A + H_{U,B}X_6^B + Z \mid W^A) + no(\log_2 P) + n\epsilon \\ &= h(Y_1^U, \ldots, Y_{16}^U) - h(H_{U,B}X_6^B + Z) + no(\log_2 P) + n\epsilon \\ &= h(Y_1^U, \ldots, Y_{16}^U) - h(H_{V,B}X_6^B + Z) + no(\log_2 P) + n\epsilon \\ &= h(Y_1^U, \ldots, Y_{16}^U) - h(Y_6^V) + no(\log_2 P) + n\epsilon \\ &\leq h(Y_1^U) + \cdots + h(Y_{16}^U) - h(Y_6^V) + no(\log_2 P) + n\epsilon\,. \end{aligned} \tag{3.6}$$

Following a similar process from the perspective of cell B, we obtain the cell B rate outer bound as (3.7) below; additional details on how to derive this expression are provided in Appendix 3.A.1.

$$nR_{B,S} \leq h(Y_1^V) + \cdots + h(Y_{16}^V) - h(Y_3^U) + no(\log_2 P) + n\epsilon \tag{3.7}$$

The separate expressions from (3.6) and (3.7) are combined together to obtain an outer bound for the achievable rate across the whole network as

$$\begin{aligned} nR_{\Sigma(SB),S} &= nR_{A,S} + nR_{B,S} \\ &\leq h(Y_1^U) + h(Y_2^U) + h(Y_4^U) + \cdots + h(Y_{16}^U) + h(Y_1^V) + \cdots + h(Y_5^V) \\ &\quad + h(Y_7^V) + \cdots + h(Y_{16}^V) + no(\log_2 P) + n\epsilon \\ &\overset{(a)}{\leq} n(\lambda_1 + \lambda_2 + \lambda_4 + \cdots + \lambda_{16} + \lambda_1 + \cdots + \lambda_5 + \lambda_7 + \cdots + \lambda_{16})(\log_2 P) \\ &\quad + no(\log_2 P) + n\epsilon \end{aligned} \tag{3.8}$$

where λ_k represents the probability of occurrence of the corresponding state k and reflects the effect of alternating connectivity, and (a) follows from the fact that Gaussian distribution maximises differential entropy. Applying $\sum_{k=1}^{16} \lambda_k = 1$

to (3.8), we obtain

$$nR_{\Sigma(SB),S} \leq n(2-\lambda_3-\lambda_6)(\log_2 P) + no(\log_2 P) + n\epsilon\,.$$

Normalising by $n(\log_2 P)$ and letting $P \to \infty$, results in

$$d_{\Sigma(SB),S} \leq 2-\lambda_3-\lambda_6\,. \tag{3.9}$$

3.3.2 Genie aided bounds

Genie aided bounds are obtained by finding an outer bound on the rate achievable at a single cell after providing it with enough extra information, i.e. 'genies', such that the data required across the two cells can be decoded within that cell.

Starting with the genie aided bound for cell A, we have

$$nR_{\Sigma(GA),S} \leq I(W^A, W^B; Y_1^U, \ldots, Y_{16}^U, G_S^A) + n\epsilon \tag{3.10}$$

where G_S^A represents the genie set required by cell A. Genies are necessary in cases where no cell B data is received at cell A, thus G_S^A consists of all the original states from Fig. 3.2 that correspond to states S and T in Fig. 3.4. Considering the corresponding entries from Table 3.1, we obtain

$$G_S^A = \left\{Y_2^V, \ldots, Y_5^V, Y_7^V, \ldots, Y_{12}^V, Y_{15}^V, Y_{16}^V\right\}\,.$$

Having defined G_S^A, the initial expression in (3.10) can be represented as

$$\begin{aligned}
nR_{\Sigma(GA),S} &\leq h(Y_1^U, \ldots, Y_{16}^U, G_S^A) - \underbrace{h(Y_1^U, \ldots, Y_{16}^U, G_S^A \mid W^A, W^B)}_{=no(\log_2 P)} + n\epsilon \\
&\leq h(Y_1^U) + \cdots + h(Y_{16}^U) + h(Y_2^V) + \cdots + h(Y_5^V) + h(Y_7^V) + \cdots + h(Y_{12}^V) \\
&\quad + h(Y_{15}^V) + h(Y_{16}^V) + no(\log_2 P) + n\epsilon \\
&\overset{(a)}{\leq} n(1+\lambda_2+\cdots+\lambda_5+\lambda_7+\cdots+\lambda_{12}+\lambda_{15}+\lambda_{16})(\log_2 P) \\
&\quad + no(\log_2 P) + n\epsilon \\
&\overset{(b)}{=} n(2-\lambda_1-\lambda_6-\lambda_{13}-\lambda_{14})(\log_2 P) + no(\log_2 P) + n\epsilon
\end{aligned} \tag{3.11}$$

where (a) follows from the fact that Gaussian distribution maximises differential entropy, and (b) follows from the fact that $\sum_{k=1}^{16}\lambda_k = 1$. Normalising by $n(\log_2 P)$

and letting $P \to \infty$, we have

$$d_{\Sigma(GA),S} \leq 2 - \lambda_1 - \lambda_6 - \lambda_{13} - \lambda_{14}\,. \tag{3.12}$$

Following a similar process for cell B, details for which are provided in Appendix 3.A.2, we obtain the cell B genie aided outer bound as in (3.13) below.

$$d_{\Sigma(GB),S} \leq\ 2 - \lambda_1 - \lambda_3 - \lambda_{15} - \lambda_{16} \tag{3.13}$$

Finally the expression for $d_{\Sigma,S}$ in Theorem 3.1 results by combining the separate bounds from (3.9), (3.12) and (3.13). □

3.4 DOF outer bound for MISO/MIMO IBC

As outlined in Section 3.2, a MISO system with local CSIT and $M \geq 2$ transmit antennas achieves 2 DoF per cell provided there is no inter-cell interference. Similarly, a MIMO system with no local CSIT having either $M = 2$ and $N \geq 2$ or $M \geq 2$ and $N = 2$ can also achieve 2 DoF per cell. This makes the two settings equivalent from an achievable DoF perspective, since both apply spatial multiplexing to resolve intra-cell interference. Based on this equivalence, it follows that the same outer bound applies to both cases. Therefore, in this section we present a DoF outer bound for two-cell two-user-per-cell MISO/MIMO IBC systems which handle intra-cell interference via spatial multiplexing.

Theorem 3.2. *For the two-cell two-user-per-cell MISO/MIMO IBC with alternating connectivity, where intra-cell interference is handled via spatial multiplexing, the sum DoF, $d_{\Sigma,M}$, can be characterised as*

$$d_{\Sigma,M} \leq 2 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8 + \Gamma$$

where

$$\Gamma = \min \begin{cases} 2\lambda_1 \\ 2\lambda_3 + \lambda_4 + \lambda_5 + \lambda_9 + \cdots + \lambda_{12} + \lambda_{15} + \lambda_{16} \\ 2\lambda_6 + \lambda_7 + \cdots + \lambda_{14}\,. \end{cases} \tag{3.14}$$

Proof. To obtain the overall outer bound, bounds originating from different sources are merged together; one comes from the summation of outer bounds for the

achievable rate at each user, while another two arise from genie aided bounds obtained on a per cell basis. Due to the length of the proof itself, we present an abbreviated version in this section; additional details are provided in Appendix 3.B.

3.4.1 Sum bound

To obtain the sum DoF outer bound, we require separate expressions for the achievable rate at each user, which are then combined together. Starting with the achievable rate at user $a1$, we have

$$nR_{a1,M} \leq I(W^A; Y_1^{a1}, \ldots, Y_{16}^{a1}) + n\epsilon$$

where W^A is the message set from BS A and Y_k^{a1} is the signal received by user $a1$ during state k. This can be further represented as

$$\begin{aligned} &nR_{a1,M} \\ &\leq h(Y_1^{a1}, \ldots, Y_{16}^{a1}) - h(Y_1^{a1}, \ldots, Y_{16}^{a1} \mid W^A) + n\epsilon \\ &= h(Y_1^{a1}, \ldots, Y_{16}^{a1}) - h(Y_2^{a1}, \ldots, Y_{15}^{a1} \mid W^A) - \underbrace{h(Y_1^{a1}, Y_{16}^{a1} \mid W^A, Y_2^{a1}, \ldots, Y_{15}^{a1})}_{E_{a1}} + n\epsilon. \end{aligned} \tag{3.15}$$

Next, it can be observed that $Y_2^{a1}, \ldots, Y_{15}^{a1}$ can be divided into two sets as follows

$$L_1' = \{2, 3, 4, 5, 7, 9, 11, 15\} \text{ and } L_1 = \{6, 8, 10, 12, 13, 14\}$$

where for the L_1' set signals received at $a1$ consist only of an X^A component, which has no effect on entropy, and noise, whose contribution can be represented as $no(\log_2 P)$. For the L_1 set, data received at $a1$ is a combination of X^A, X^B and noise. Using this information (3.15), can be represented as

$$\begin{aligned} &nR_{a1,M} \\ &\leq h(Y_1^{a1}, \ldots, Y_{16}^{a1}) - h(Y_{L_1}^{a1} \mid W^A) - E_{a1} + no(\log_2 P) + n\epsilon \\ &\overset{(a)}{=} h(Y_1^{a1}, \ldots, Y_{16}^{a1}) - h(H_{a1,A}X_6^A + H_{a1,B}X_6^B + Z, H_{a1,A}X_8^A + H_{a1,B}X_8^B + Z, \\ &\quad H_{a1,A}X_{10}^A + H_{a1,B}X_{10}^B + Z, H_{a1,A}X_{12}^A + H_{a1,B}X_{12}^B + Z, H_{a1,A}X_{13}^A + H_{a1,B}X_{13}^B \\ &\quad + Z, H_{a1,A}X_{14}^A + H_{a1,B}X_{14}^B + Z \mid W^A) - E_{a1} + no(\log_2 P) + n\epsilon \end{aligned}$$

$$
\begin{aligned}
&\overset{(b)}{=} h(Y_1^{a1},\ldots,Y_{16}^{a1}) - h(H_{a1,B}X_6^B + Z, H_{a1,B}X_8^B + Z, H_{a1,B}X_{10}^B + Z, H_{a1,B}X_{12}^B \\
&\quad + Z, H_{a1,B}X_{13}^B + Z, H_{a1,B}X_{14}^B + Z) - E_{a1} + no(\log_2 P) + n\epsilon \\
&\overset{(c)}{=} h(Y_1^{a1},\ldots,Y_{16}^{a1}) - h(H_{b2,B}X_6^B + Z, H_{b1,B}X_8^B + Z, H_{b2,B}X_{10}^B + Z, H_{b1,B}X_{12}^B \\
&\quad + Z, H_{b2,B}X_{13}^B + Z, H_{b1,B}X_{14}^B + Z) - E_{a1} + no(\log_2 P) + n\epsilon \\
&\overset{(d)}{=} h(Y_1^{a1},\ldots,Y_{16}^{a1}) - h(Y_6^{b2}, Y_8^{b1}, Y_{10}^{b2}, Y_{12}^{b1}, Y_{13}^{b2}, Y_{14}^{b1}) - E_{a1} + no(\log_2 P) + n\epsilon
\end{aligned} \tag{3.16}
$$

where (a) follows by expressing the received signals for the L_1 set in terms of their original components; (b) follows by removing the X^A parts since they have no effect on entropy and also removing the conditioning since X^B and Z are independent of W^A; (c) is obtained by replacing channel coefficients from BS B to user $a1$ with ones to cell B users, due to their statistical equivalence, and lastly (d) is obtained by representing the X^B and noise components in terms of the signals received at the corresponding inter-cell interference free cell B users. Finally, considering all elements of the first negative entropy term in (3.16) to be independent of each other results in

$$
\begin{aligned}
nR_{a1,M} \leq\ & h(Y_1^{a1}) + \cdots + h(Y_{16}^{a1}) - h(Y_6^{b2}) - h(Y_8^{b1}) - h(Y_{10}^{b2}) - h(Y_{12}^{b1}) - h(Y_{13}^{b2}) \\
& - h(Y_{14}^{b1}) - E_{a1} + no(\log_2 P) + n\epsilon\ .
\end{aligned} \tag{3.17}
$$

Following a similar process for the remaining users $a2$, $b1$ and $b2$ separately, we obtain the inequalities in (3.18), (3.19) and (3.20). Further details on how to derive these expressions are provided in Appendices 3.B.1, 3.B.2 and 3.B.3 respectively.

$$
\begin{aligned}
nR_{a2,M} \leq\ & h(Y_1^{a2}) + \cdots + h(Y_{16}^{a2}) - h(Y_6^{b1}) - h(Y_7^{b2}) - h(Y_9^{b2}) - h(Y_{11}^{b1}) - h(Y_{13}^{b2}) \\
& - h(Y_{14}^{b1} - E_{a2} + no(\log_2 P) + n\epsilon
\end{aligned} \tag{3.18}
$$

$$
\begin{aligned}
nR_{b1,M} \leq\ & h(Y_1^{b1}) + \cdots + h(Y_{16}^{b1}) - h(Y_3^{a1}) - h(Y_4^{a1}) - h(Y_9^{a1}) - h(Y_{10}^{a2}) - h(Y_{15}^{a1}) \\
& - h(Y_{16}^{a2}) - E_{b1} + no(\log_2 P) + n\epsilon
\end{aligned} \tag{3.19}
$$

$$
\begin{aligned}
nR_{b2,M} \leq\ & h(Y_1^{b2}) + \cdots + h(Y_{16}^{b2}) - h(Y_3^{a2}) - h(Y_5^{a2}) - h(Y_{11}^{a1}) - h(Y_{12}^{a2}) - h(Y_{15}^{a1}) \\
& - h(Y_{16}^{a2}) - E_{b2} + no(\log_2 P) + n\epsilon
\end{aligned} \tag{3.20}
$$

Combining (3.17) to (3.20), the achievable sum rate across the whole network is bounded as

$$
\begin{aligned}
&nR_{\Sigma(SB),M} \\
&= nR_{a1,M} + nR_{a2,M} + nR_{b1,M} + nR_{b2,M} \\
&\leq h(Y_1^{a1}) + h(Y_2^{a1}) + h(Y_5^{a1}) + \cdots + h(Y_8^{a1}) + h(Y_{10}^{a1}) + h(Y_{12}^{a1}) + h(Y_{13}^{a1}) \\
&\quad + h(Y_{14}^{a1}) + h(Y_{16}^{a1}) + h(Y_1^{a2}) + h(Y_2^{a2}) + h(Y_4^{a2}) + h(Y_6^{a2}) + \cdots + h(Y_9^{a2}) \\
&\quad + h(Y_{11}^{a2}) + h(Y_{13}^{a2}) + h(Y_{14}^{a2}) + h(Y_{15}^{a2}) + h(Y_1^{b1}) + \cdots + h(Y_5^{b1}) + h(Y_7^{b1}) \\
&\quad + h(Y_9^{b1}) + h(Y_{10}^{b1}) + h(Y_{13}^{b1}) + h(Y_{15}^{b1}) + h(Y_{16}^{b1}) + h(Y_1^{b2}) + \cdots + h(Y_5^{b2}) \\
&\quad + h(Y_8^{b2}) + h(Y_{11}^{b2}) + h(Y_{12}^{b2}) + h(Y_{14}^{b2}) + h(Y_{15}^{b2}) + h(Y_{16}^{b2}) - h(Y_{15}^{a1}) - h(Y_{16}^{a2}) \\
&\quad - h(Y_{14}^{b1}) - h(Y_{13}^{b2}) - E_{a1} - E_{a2} - E_{b1} - E_{b2} + no(\log_2 P) + n\epsilon \,. \qquad (3.21)
\end{aligned}
$$

Next, we consider the remaining negative terms in (3.21) and pair one of $\{E_{a1}, E_{a2}, E_{b1}, E_{b2}\}$ with one of $\{h(Y_{15}^{a1}), h(Y_{16}^{a2}), h(Y_{14}^{b1}), h(Y_{13}^{b2})\}$ to find a joint lower bound. Starting with E_{a1}, we can express it as

$$
\begin{aligned}
E_{a1} &= h(Y_1^{a1}, Y_{16}^{a1} \mid W^A, Y_2^{a1}, \ldots, Y_{15}^{a1}) \\
&= h(Y_{16}^{a1} \mid W^A, Y_2^{a1}, \ldots, Y_{15}^{a1}) + h(Y_1^{a1} \mid W^A, Y_2^{a1}, \ldots, Y_{16}^{a1}) \\
&\overset{(a)}{=} h(H_{a1,A} X_{16}^A + H_{a1,B} X_{16}^B + Z \mid W^A) + \underbrace{h(Y_1^{a1} \mid W^A, Y_2^{a1}, \ldots, Y_{16}^{a1})}_{\substack{\geq h(Y_1^{a1} \mid W^A, Y_2^{a1}, \ldots, Y_{16}^{a1}, W^B) \\ = no(\log_2 P)}} \\
&\overset{(b)}{\geq} h(H_{a1,B} X_{16}^B + Z) + no(\log_2 P)
\end{aligned}
$$

where (a) follows by expressing Y_{16}^{a1} in terms of its original components and considering it to be independent of $Y_2^{a1}, \ldots, Y_{15}^{a1}$, and (b) follows by neglecting the X^A component since its effect is negligible with respect to entropy and removing the conditioning since the remaining terms are independent of W^A. Pairing E_{a1} with $h(Y_{16}^{a2})$, we obtain

$$
\begin{aligned}
h(Y_{16}^{a2}) + E_{a1} &\overset{(a)}{\geq} h(H_{a2,A} X_{16}^A + Z) + h(H_{a1,B} X_{16}^B + Z) + no(\log_2 P) \\
&\overset{(b)}{=} h(H_{a1,A} X_{16}^A) + h(H_{a1,B} X_{16}^B) + no(\log_2 P) \qquad (3.22)
\end{aligned}
$$

where (a) follows by representing Y_{16}^{a2} in terms of its original components, and (b) follows by applying the fact that $H_{a2,A}$ and $H_{a1,A}$ are statistically equivalent, and removing the noise components since their effect disappears with high SNR and can therefore be integrated into the $no(\log_2 P)$ term. Additionally, considering $h(Y_{16}^{a1})$ and the fact that $Y_{16}^{a1} = H_{a1,A} X_{16}^A + H_{a1,B} X_{16}^B + Z$, we can apply Lemma 3.1 from Appendix 3.C to obtain

$$h(Y_{16}^{a1}) \leq h(H_{a1,A}X_{16}^{A}) + h(H_{a1,B}X_{16}^{B}) + no(\log_2 P). \tag{3.23}$$

Subtracting (3.22) from (3.23), we obtain

$$h(Y_{16}^{a1}) - h(Y_{16}^{a2}) - E_{a1} \leq no(\log_2 P)\ . \tag{3.24}$$

Applying a similar process to different pairings we can also establish the following inequalities

$$h(Y_{15}^{a2}) - h(Y_{15}^{a1}) - E_{a2} \leq no(\log_2 P)\,, \tag{3.25}$$
$$h(Y_{13}^{b1}) - h(Y_{13}^{b2}) - E_{b1} \leq no(\log_2 P)\,, \tag{3.26}$$
$$h(Y_{14}^{b2}) - h(Y_{14}^{b1}) - E_{b2} \leq no(\log_2 P)\,. \tag{3.27}$$

Using (3.24) to (3.27) in the total rate expression from (3.21), we obtain

$$\begin{aligned}
&nR_{\Sigma(SB),M} \\
&\leq h(Y_1^{a1}) + h(Y_2^{a1}) + h(Y_5^{a1}) + \cdots + h(Y_8^{a1}) + h(Y_{10}^{a1}) + h(Y_{12}^{a1}) + h(Y_{13}^{a1}) \\
&\quad + h(Y_{14}^{a1}) + h(Y_1^{a2}) + h(Y_2^{a2}) + h(Y_4^{a2}) + h(Y_6^{a2}) + \cdots + h(Y_9^{a2}) + h(Y_{11}^{a2}) \\
&\quad + h(Y_{13}^{a2}) + h(Y_{14}^{a2}) + h(Y_1^{b1}) + \cdots + h(Y_5^{b1}) + h(Y_7^{b1}) + h(Y_9^{b1}) + h(Y_{10}^{b1}) \\
&\quad + h(Y_{15}^{b1}) + h(Y_{16}^{b1}) + h(Y_1^{b2}) + \cdots + h(Y_5^{b2}) + h(Y_8^{b2}) + h(Y_{11}^{b2}) + h(Y_{12}^{b2}) \\
&\quad + h(Y_{15}^{b2}) + h(Y_{16}^{b2}) + no(\log_2 P) + n\epsilon \\
&\overset{(a)}{\leq} n(2 + 2\lambda_1 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8)(\log_2 P) + no(\log_2 P) + n\epsilon
\end{aligned} \tag{3.28}$$

where (a) follows by using the fact that Gaussian distribution maximises differential entropy and applying $\sum_{k=1}^{16} \lambda_k = 1$. Finally, normalising by $n(\log_2 P)$ and letting $P \to \infty$, we obtain the desired DoF sum bound as

$$d_{\Sigma(SB),M} \leq 2 + 2\lambda_1 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8\,. \tag{3.29}$$

3.4.2 Genie aided bounds

The genie aided bounds for the MISO/MIMO scenario are obtained in a similar way to the SISO ones from Section 3.3.2. However, in this case the number of genies provided must ensure that 2 symbols from the other cell can be retrieved.

Starting with the cell A genie aided DoF bound, we have

$$nR_{\Sigma(GA),M} \leq I(W^A, W^B; Y_1^{a1}, \ldots, Y_{16}^{a1}, Y_1^{a2}, \ldots, Y_{16}^{a2}, G_M^A) + n\epsilon \tag{3.30}$$

where G_M^A represents the additional genie set required such that cell B data may be reconstructed within cell A. The amount of genies required is either one or two, depending on the number of signals containing cell B information reaching cell A. Looking at all the possible topologies in Fig. 3.2, this corresponds to

$$G_M^A = \left\{2 \times [Y_2^B, Y_3^B, Y_4^B, Y_5^B], Y_7^B, \ldots, Y_{12}^B, Y_{15}^B, Y_{16}^B\right\}$$

where B is used to represent either $b1$ or $b2$. Having defined G_M^A, this can be integrated into (3.30) as

$$\begin{aligned}
nR_{\Sigma(GA),M} &\leq h(Y_1^{a1}, \ldots, Y_{16}^{a1}, Y_1^{a2}, \ldots, Y_{16}^{a2}, G_M^A) + n\epsilon \\
&\quad - \underbrace{h(Y_1^{a1}, \ldots, Y_{16}^{a1}, Y_1^{a2}, \ldots, Y_{16}^{a2}, G_M^A \mid W^A, W^B)}_{=no(\log_2 P)} \\
&\leq h(Y_1^{a1}) + \cdots + h(Y_{16}^{a1}) + h(Y_1^{a2}) + \cdots + h(Y_{16}^{a2}) + 2h(Y_2^B) + 2h(Y_3^B) \\
&\quad + 2h(Y_4^B) + 2h(Y_5^B) + h(Y_7^B) + \cdots + h(Y_{12}^B) + h(Y_{15}^B) + h(Y_{16}^B) \\
&\quad + no(\log_2 P) + n\epsilon \\
&\overset{(a)}{\leq} n(2 + 2\lambda_2 + 2\lambda_3 + 2\lambda_4 + 2\lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \lambda_{15} + \ \lambda_{16})(\log_2 P) \\
&\quad + no(\log_2 P) + \ n\epsilon
\end{aligned} \tag{3.31}$$

where (a) follows by using the fact that Gaussian distribution maximises differential entropy and applying $\sum_{k=1}^{16} \lambda_k = 1$. Normalising by $n(\log_2 P)$ and letting $P \to \infty$, results in

$$d_{\Sigma(GA),M} \leq 2 + 2\lambda_2 + 2\lambda_3 + 2\lambda_4 + 2\lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \lambda_{15} + \lambda_{16}\,. \tag{3.32}$$

Following a similar process for cell B, details for which are provided in Appendix 3.B.4, we have

$$d_{\Sigma(GB),M} \leq 2 + 2\lambda_2 + \lambda_4 + \lambda_5 + 2\lambda_6 + 2\lambda_7 + 2\lambda_8 + \lambda_9 + \cdots + \lambda_{14}\,. \tag{3.33}$$

Finally, the result for $d_{\Sigma,M}$ in Theorem 3.2 is obtained by combining the bounds from (3.29), (3.32) and (3.33) together. □

Remark 3.1. *Similarities can be observed between the IBC outer bounds in Theorem 3.1 and Theorem 3.2, and the one for the two-user SISO IC from [42]. This is expected since the IC is essentially a subset of the IBC having only one user per cell. Before drawing any similarities we need to express the outer bound from Theorem 3.1 in an alternative way. Using the fact that* $\sum_{k=1}^{16} \lambda_k = 1$ *this can be represented as*

$$d_{\Sigma,S} \leq 1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \Psi \tag{3.34}$$

where

$$\Psi = \min \begin{cases} \lambda_1 + \lambda_{13} + \cdots + \lambda_{16} \\ \lambda_3 + \lambda_{15} + \lambda_{16} \\ \lambda_6 + \lambda_{13} + \lambda_{14}\,. \end{cases}$$

The reformulated version of Theorem 3.1, alongside with the outer bound in Theorem 3.2, and the SISO IC result in [42] can collectively be summarised as

$$d_\Sigma \leq d_c + \tilde{\lambda}_d + \tilde{\lambda}_e$$

where d_c *is the achievable DoF per cell when no inter-cell interference is present. This is equal to* 1 *for the two-user SISO IC and the two-cell two-user-per-cell SISO IBC, and corresponds to* 2 *for the MISO/MIMO IBC counterpart. For all scenarios,* $\tilde{\lambda}_d$ *consists exclusively of the probability of occurrence of all the states that directly obtain higher DoF than the fully connected one; its fixed presence in the outer bound reflects the corresponding DoF gain. Finally,* $\tilde{\lambda}_e$ *depends on which bound is the most restrictive, but is always a function of the probability of occurrence of the states which inherently obtain less DoF than the inter-cell interference free one.*

3.5 Achievable DoF for SISO IBC

Without knowledge of the network's topological structure, a fully connected scenario has to be assumed at all times, achieving a sum DoF of 1 across all states for the SISO system. However, if global topological CSIT is provided, the BSs can adapt their transmission strategies to exploit the partially connected states and obtain a DoF gain.

3.5.1 Single state has a probability of occurrence of one

This is an extreme case for the scenario considered, with $\lambda_i = 1$ and $\lambda_j = 0$ for $j = 1, \ldots, 16, j \neq i$. It essentially implies connectivity is fixed in state i throughout the whole transmission process.

For $i \in \{2, 4, 5, 7, \ldots, 12\}$ there is at least one user per cell that is free from inter-cell interference. These states represent the best case scenario from an achievable DoF perspective, with the outer bound in Theorem 3.1 corresponding to $d_{\Sigma,S} \leq 2$. Having knowledge of the network's structure, both BSs can operate simultaneously and serve one inter-cell interference free user per cell, achieving 2 DoF across the whole network. This is equal to the outer bound itself, and corresponds to a twofold increase over the no global topological CSIT case.

For the remaining states $i \in \{1, 3, 6, 13, \ldots, 16\}$, at least one of the two cells has both users experiencing inter-cell interference and the outer bound from Theorem 3.1 corresponds to $d_{\Sigma,S} \leq 1$. Sum DoF of 1 can be achieved simply by operating one BS at a time and serving one user within the corresponding cell.

3.5.2 Arbitrary state probabilities

Without global topological CSIT, only 1 DoF can be achieved regardless of the current connectivity state. However, if this information is available, BSs can use it to adapt their transmission strategy accordingly. Both BSs operate simultaneously for states where there is at least one inter-cell interference free user in each cell, delivering a symbol each to two users from different cells. For the remaining states, only one BS needs to be operated, delivering one symbol across the whole network. Therefore, considering all the states in Fig. 3.2 it is possible to obtain

$$\begin{aligned} \text{DoF} &= \begin{cases} 1 \text{ for states } 1, 3, 6, 13, 14, 15, 16 \\ 2 \text{ for states } 2, 4, 5, 7, 8, 9, 10, 11, 12 \end{cases} \\ &= 1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} \,. \end{aligned} \tag{3.35}$$

Higher DoF can be achieved via *joint coding across states*. Within the SISO IBC setting, joint coding can be used across a variety of state combinations to deliver a total of 4 symbols over 3 states. Considering the alternating connectivity states in Fig. 3.2, it can be noticed that the same interference links appear twice over states $\{3, 13, 14\}$. Thus, the three states in this set can be combined together

to resolve inter-cell interference. Defining s_r as the symbol intended for user r, then for scheme $\mathcal{S}_1$, which performs joint coding across states $\{3, 13, 14\}$, BSs transmit according to Table 3.2. User $b2$ obtains s_{b2} directly from the signal received in state 13, while the combination of received signals at the remaining users allows for interference cancellation decoding. For users $a1$ and $a2$, received signals are functions of s_{a1}, s_{a2} and s_{b2}. Having three independent equations in terms of three different symbols, then the desired data can be obtained at the respective users. For user $b1$, all received signals are functions of $(s_{a1} + s_{a2})$, s_{b1} and s_{b2}. Considering $(s_{a1} + s_{a2})$ as a single symbol, we have three independent equations for three unknowns and can solve for s_{b1}. Therefore 1 symbol each is transmitted to all 4 users in 3 channel uses, leading to an average of $\frac{4}{3}$ DoF per state.

Table 3.2: Transmission strategy for scheme $\mathcal{S}_1$.

Transmitted symbols	State 3	State 13	State 14
X^A	$(s_{a1} + s_{a2})$	$(s_{a1} + s_{a2})$	s_{a1}
X^B	s_{b1}	s_{b2}	s_{b2}

Joint coding can also be applied across other sets of states. In particular, states $\{6, 15, 16\}$ can be combined together using scheme $\mathcal{S}_2$ in Table 3.3 and states $\{1, 3, 6\}$ can be combined via scheme $\mathcal{S}_3$ in Table 3.4. In each case $\frac{4}{3}$ DoF per state are achieved. Additionally for quasi-static fading channels, where the value of the channel links does not change across the states involved in the scheme, it is also possible to code across states $\{13, 15, 16\}$ using scheme $\mathcal{S}_4$ in Table 3.5 or across states $\{14, 15, 16\}$ via scheme $\mathcal{S}_5$ in Table 3.6.

Table 3.3: Transmission strategy for scheme $\mathcal{S}_2$.

Transmitted symbols	State 6	State 15	State 16
X^A	s_{a2}	s_{a1}	s_{a1}
X^B	$(s_{b1} + s_{b2})$	$(s_{b1} + s_{b2})$	s_{b1}

Table 3.4: Transmission strategy for scheme $\mathcal{S}_3$.

Transmitted symbols	State 1	State 3	State 6
X^A	$(s_{a1} + s_{a2})$	$(s_{a1} + s_{a2})$	s_{a1}
X^B	$(s_{b1} + s_{b2})$	s_{b1}	$(s_{b1} + s_{b2})$

Table 3.5: Transmission strategy for scheme $\mathcal{S}_4$.

Transmitted symbols	State 13	State 15	State 16
X^A	$(s_{a1} + s_{a2})$	s_{a1}	s_{a2}
X^B	s_{b2}	s_{b1}	s_{b2}

Table 3.6: Transmission strategy for scheme $\mathcal{S}_5$.

Transmitted symbols	State 14	State 15	State 16
X^A	$(s_{a1} + s_{a2})$	s_{a1}	s_{a2}
X^B	s_{b1}	s_{b1}	s_{b2}

Due to the repetition of the states involved in schemes $\mathcal{S}_1$ to $\mathcal{S}_5$, no more than two can be combined together. The possible combinations are: $\mathcal{S}_1$ and $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$, or $\mathcal{S}_3$ and $\mathcal{S}_5$. With arbitrary state probabilities, achievable DoF for each combination can be characterised as follows.

(i) *Schemes $\mathcal{S}_1$ and $\mathcal{S}_2$:*

$$\begin{aligned} d_{\Sigma,\mathcal{S}-\mathcal{S}_1,\mathcal{S}_2} &= \lambda_1 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + 4\alpha_1 + 4\beta_1 + (\lambda_3 - \alpha_1) \\ &\quad + (\lambda_{13} - \alpha_1) + (\lambda_{14} - \alpha_1) + (\lambda_6 - \beta_1) + (\lambda_{15} - \beta_1) + (\lambda_{16} - \beta_1) \\ &= 1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \theta_1 + \psi_1 \end{aligned}$$

where $\theta_1 = \min\{\lambda_3, \lambda_{13}, \lambda_{14}\}$ and $\psi_1 = \min\{\lambda_6, \lambda_{15}, \lambda_{16}\}$.

(ii) *Schemes $\mathcal{S}_3$ and $\mathcal{S}_4$:*

$$\begin{aligned} d_{\Sigma,\mathcal{S}-\mathcal{S}_3,\mathcal{S}_4} &= 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \lambda_{14} + 4\alpha_2 + 4\beta_2 + (\lambda_1 - \alpha_2) \\ &\quad + (\lambda_3 - \alpha_2) + (\lambda_6 - \alpha_2) + (\lambda_{13} - \beta_2) + (\lambda_{15} - \beta_2) + (\lambda_{16} - \beta_2) \\ &= 1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \theta_2 + \psi_2 \end{aligned}$$

where $\theta_2 = \min\{\lambda_1, \lambda_3, \lambda_6\}$ and $\psi_2 = \min\{\lambda_{13}, \lambda_{15}, \lambda_{16}\}$.

(iii) *Schemes $\mathcal{S}_3$ and $\mathcal{S}_5$:*

$$\begin{aligned} d_{\Sigma,\mathcal{S}-\mathcal{S}_3,\mathcal{S}_4} &= 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{13} + 4\alpha_2 + 4\beta_3 + (\lambda_1 - \alpha_2) \\ &\quad + (\lambda_3 - \alpha_2) + (\lambda_6 - \alpha_2) + (\lambda_{14} - \beta_3) + (\lambda_{15} - \beta_3) + (\lambda_{16} - \beta_3) \\ &= 1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \theta_2 + \psi 3 \end{aligned}$$

where θ_2 is as defined in (ii) and $\psi_3 = \min\{\lambda_{14}, \lambda_{15}, \lambda_{16}\}$.

Combining $d_{\Sigma,S-\mathcal{S}_1,\mathcal{S}_2}$, $d_{\Sigma,S-\mathcal{S}_3,\mathcal{S}_4}$ and $d_{\Sigma,S-\mathcal{S}_3,\mathcal{S}_5}$, into a single expression for the maximum achievable DoF we obtain

$$d_{\Sigma,S-Ach} = 1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \cdots + \lambda_{12} + \tilde{\Psi} \tag{3.36}$$

where $\tilde{\Psi} = \max\{\theta_1 + \psi_1, \theta_2 + \psi_2, \theta_2 + \psi_3\}$ for quasi-static fading channels, and $\tilde{\Psi} = \theta_1 + \psi_1$ for the fast-fading scenario.

Note that the only difference between the outer bound in Theorem 3.1 (see the reformulated expression in (3.34)) and the achievable DoF expression in (3.36) is the final term. In fact for any state probabilities such that $\tilde{\Psi} = \Psi$, the two are equal, resulting in an outer bound which is achievable. For example, this happens for $\lambda_1 = \lambda_3 = \lambda_6 = \lambda_{13} = \cdots = \lambda_{16} = 0$ and general values of λ_i where $i \in \mathcal{I} = \{2, 4, 5, 7, \ldots, 12\}$ and $\sum_{i\in\mathcal{I}} \lambda_i = 1$.

3.5.3 Equal state probabilities

To understand the advantages obtained from this type of scenario on average, we now consider the case where all states are equiprobable, i.e. $\lambda_1 = \cdots = \lambda_{16} = \frac{1}{16}$, which can occur in a fast fading context. Using the result of Theorem 3.1 we can establish the following corollary.

Corollary 3.1. *For the two-cell two-user-per-cell SISO IBC with alternating connectivity and equiprobable states,* $d_{\Sigma,S} \leq 1\frac{3}{4}$.

Without global topological CSIT, only sum DoF of 1 can be achieved. However if this information is available the DoF in (3.35) can be obtained; with equiprobable states this implies 25 symbols are transmitted in 16 channel uses on average, equivalent to $1\frac{9}{16}$ DoF. While this is an improvement of $\frac{9}{16}$ over the no global topological CSIT case, it is still $\frac{3}{16}$ DoF away from the outer bound in Corollary 3.1. Applying joint coding across states the DoF in (3.36) can be achieved. With equiprobable states this is equal to $1\frac{11}{16}$ and corresponds to a gain of $\frac{11}{16}$ DoF over the no global topological CSIT setting. While this is not equivalent to the outer bound in Corollary 3.1, there is only a difference of $\frac{1}{16}$ DoF between the two, i.e. 96.4% of the outer bound is achieved.

3.6 Achievable DoF for MISO/MIMO IBC

In this section we investigate the achievability of the outer bound in Theorem 3.2. Without global topological CSIT, a fully connected network has to be assumed at all times. This allows only for one possible transmission strategy where the BSs are given unique transmission opportunities, thereby achieving 2 DoF across all states. However, if global topological information is provided, the BSs can adapt their strategies in order to exploit the partially connected states and achieve higher DoF.

3.6.1 Single state has a probability of occurrence of one

This represents an extreme case for the scenario considered in this work, where connectivity is fixed in a single state throughout the whole transmission process, i.e. $\lambda_i = 1$ and $\lambda_j = 0$ for $j = 1, \ldots, 16, j \neq i$.

For $i = 2$, Theorem 3.2 can be represented as $d_{\Sigma,M} \leq 4$. From an achievable DoF perspective, this situation corresponds to the best case scenario, since all users are inter-cell interference free. Having knowledge of the network's topology, both BSs can operate simultaneously and deliver a symbol each to their respective users, thereby achieving 4 DoF across the whole network. This is equal to the outer bound itself, and corresponds to a twofold increase in achievable DoF over the no global topological CSIT case.

For $i \in \{4, 5, 7, 8\}$, three out of the four users are free from inter-cell interference and the outer bound in Theorem 3.2 corresponds to $d_{\Sigma,M} \leq 3$. Since network topology is known, both BSs can operate simultaneously to serve the three inter-cell interference free users, while the fourth user is not served due to inter-cell interference. This achieves 3 DoF over the whole network, which is equal to the derived outer bound and provides a gain of 1 DoF over the case where global topological CSIT is not provided.

For the remaining states $i \in \{1, 3, 6, 9, \ldots, 16\}$, the outer bound from Theorem 3.1 is $d_{\Sigma,M} \leq 2$. Sum DoF of 2 can be achieved simply by operating one BS at a time and delivering one symbol each to the two users in the corresponding cell.

3.6.2 Arbitrary state probabilities

Without global topological CSIT, only 2 DoF can be achieved regardless of the current connectivity state; however, if this information is provided, the BSs can

adapt their transmission strategies to exploit the partially connected states. Considering the set of states in Fig. 3.2, it is possible to achieve

$$\begin{aligned}\text{DoF} &= \begin{cases} 2 \text{ for states } 1,3,6,9,10,11,12,13,14,15,16 \\ 3 \text{ for states } 4,5,7,8 \\ 4 \text{ for state } 2 \end{cases} \\ &= 2 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8 \,. \end{aligned} \tag{3.37}$$

Higher DoF can be obtained via scheme $\mathcal{S}_6$ which applies joint coding across states. Looking at the states in Fig. 3.2, it can be noticed that the interference links present in states 3 and 6 are contained within state 1; therefore, state 1 can be used to resolve them. The transmission strategy for scheme $\mathcal{S}_6$ is outlined in Table 3.7.

Table 3.7: Transmission strategy for scheme $\mathcal{S}_6$.

Transmitted symbols	State 3	State 6	State 1
X^A	$\mathbf{s}_A$	$\underline{\mathbf{s}}_A$	$\mathbf{s}_A$
X^B	$\mathbf{s}_B$	$\underline{\mathbf{s}}_B$	$\underline{\mathbf{s}}_B$

For the MISO case, we define the signals transmitted from BS A as $\mathbf{s}_A = (\mathbf{v}_{a1}s_{a1} + \mathbf{v}_{a2}s_{a2})$ and $\underline{\mathbf{s}}_A = (\underline{\mathbf{v}}_{a1}\underline{s}_{a1} + \underline{\mathbf{v}}_{a2}\underline{s}_{a2})$, where $\mathbf{v}_i$ and $\underline{\mathbf{v}}_i$ are $M \times 1$ ZF precoders. These are constructed using local CSIT knowledge according to the orthogonality principles outlined in Section 3.2, and ensure that each user can extract the desired symbol from the combined signal transmitted by the corresponding BS. The symbols transmitted by BS B are defined in a similar manner. By following the transmission strategy in Table 3.7, signals received over the three states at users $a1$ and $a2$ consist only of $\mathbf{s}_A$, $\underline{\mathbf{s}}_A$ and $\underline{\mathbf{s}}_B$, thus both users can decode for $\mathbf{s}_A$ and $\underline{\mathbf{s}}_A$. Additionally, due to the ZF precoding, users only see the desired part of the combined signal, thus a_1 obtains $\{s_{a1}, \underline{s}_{a1}\}$, while $a2$ obtains $\{s_{a2}, \underline{s}_{a2}\}$. A similar decoding process is carried out at users $b1$ and $b2$ to obtain $\{s_{b1}, \underline{s}_{b1}\}$ and $\{s_{b2}, \underline{s}_{b2}\}$ respectively. Therefore, across the 3 states a total of 8 new symbols are delivered, 2 for each user. Note that since the precoders depend on the corresponding channels, for the MISO case this scheme can only be applied in a slow-fading scenario, where the channel value remains constant across states $\{1, 3, 6\}$.

For the MIMO case, we define the signals transmitted from BS A as $\mathbf{s}_A = (\mathbf{v}_{a1}s_{a1} + \mathbf{v}_{a2}s_{a2})$ and $\underline{\mathbf{s}}_A = (\underline{\mathbf{v}}_{a1}\underline{s}_{a1} + \underline{\mathbf{v}}_{a2}\underline{s}_{a2})$, where $\mathbf{v}_i$ and $\underline{\mathbf{v}}_i$ are $M \times 1$ are

pseudo-random precoders which change for every state involved in scheme $\mathcal{S}_6$. The symbols transmitted by BS B are similarly defined. Considering users $a1$ and $a2$, it can be seen that across the whole set of states, received signals consist only of s_{a1}, s_{a2}, $\underline{s}_{a1}$, $\underline{s}_{a2}$, $\underline{s}_{b1}$ and $\underline{s}_{b2}$. Due to the multiple antenna configuration at the transmitters and receivers, both users $a1$ and $a2$ are in possession of six independent observations (two from each state) and can thus decode for their desired symbols, $\{s_{a1}, \underline{s}_{a1}\}$ and $\{s_{a2}, \underline{s}_{a2}\}$ respectively. A similar decoding process is carried out at users $b1$ and $b2$ to obtain $\{s_{b1}, \underline{s}_{b1}\}$ and $\{s_{b2}, \underline{s}_{b2}\}$ respectively. Therefore, by applying this scheme, each user obtains 2 new symbols across 3 states. For the MIMO case this scheme is applicable to both fast-fading or slow-fading channel scenarios.

Using scheme $\mathcal{S}_6$, for arbitrary state probabilities, achievable DoF can be characterised as

$$\begin{aligned} d_{\Sigma,M-Ach} &= 4\lambda_2 + 3(\lambda_4 + \lambda_5 + \lambda_7 + \lambda_8) + 2(\lambda_9 + \cdots + \lambda_{16}) + 8\gamma \\ &\quad + 2(\lambda_1 - \gamma) + 2(\lambda_3 - \gamma) + 2(\lambda_6 - \gamma) \\ &= 2 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8 + 2\tilde{\Gamma} \end{aligned} \tag{3.38}$$

where $2\tilde{\Gamma} = \min\{\lambda_1, \lambda_3, \lambda_6\}$.

Note the similarity between the achievable DoF in (3.38) and the outer bound in Theorem 3.2. The only difference is in the final term, such that for any state probabilities that result in $2\tilde{\Gamma} = \Gamma$ the two are equal, leading to a tight outer bound. For example, this occurs when $\lambda_1 \leq \min\{\lambda_3, \lambda_6\}$ for arbitrary values of $\lambda_i \ \forall i = 1, \ldots, 16$.

3.6.3 Equal state probabilities

We now consider the case where all states are equiprobable, i.e. $\lambda_1 = \cdots = \lambda_{16} = \frac{1}{16}$. This can occur for a fast fading context, and provides us with an understanding of the advantages that can be obtained from this type of scenario on average. For equiprobable states, the result in Theorem 3.2 can be used to establish the following corollary.

Corollary 3.2. *For the two-cell two-user-per-cell MISO/MIMO IBC with alternating connectivity and equiprobable states, where intra-cell interference is handled via spatial multiplexing,* $d_{\Sigma,M} \leq 2\frac{1}{2}$.

Without global topological CSIT, only a sum DoF of 2 can be achieved. However if topological CSIT is available, the DoF in (3.37) can be obtained; with equiprobable states this is equivalent to $2\frac{3}{8}$ DoF, since 38 symbols are transmitted in 16 channel uses on average. While this is an improvement of $\frac{3}{8}$ DoF over the no global topological CSIT case, it is still $\frac{1}{8}$ DoF away from the outer bound value established in Corollary 3.2. Applying joint coding across states, the DoF in (3.38) can be achieved, with equiprobable states this results in $2\frac{1}{2}$ DoF. This corresponds to a gain of $\frac{1}{2}$ DoF over the no global topological CSIT setting, and is equal to the outer bound value from Corollary 3.2.

3.7 Applicability to wired network equivalents

In [20] it was established that under the TIM framework, the capacity of a wireless network and the corresponding wired instance are equivalent in their normalised forms. The term 'corresponding' implies that the two networks have the same underlying noiseless linear network structure. For wireless networks, normalised capacity represents the achievable rate normalised by $\log_2\text{SNR}$ as $\text{SNR} \to \infty$, i.e. DoF. For wired networks, normalised capacity refers to the capacity of the network divided by the capacity of a single link, i.e. divided by $\log_2|\mathbb{GF}|$, where $\mathbb{GF}$ represents the finite Galois field.

This equivalence essentially implies that all networks (either wired or wireless) with the same logical end-to-end topology have the same normalised capacity, and requires wired networks to be SISO ones where each source has only one outgoing edge and each destination has only one incoming edge. Both wireless scenarios considered in this work can be mapped to such wired equivalent networks, examples of which are provided in Fig. 3.5 and Fig. 3.6. Here sources are in black, destinations are in white and intermediate nodes are in grey. Both figures represent the fully connected state, i.e. state 1 in Fig. 3.2, however the presence of the dotted red lines is variable to reflect the alternating connectivity. For wired networks connectivity can change due to variations in the linear network coding coefficients. Therefore, from the results of Theorems 3.1 and 3.2 we can establish the corresponding wired network results as stated in Corollaries 3.3 and 3.4.

Corollary 3.3. *The normalised sum capacity of a wired network with the same end-to-end topology as the wireless SISO network considered in this chapter is upper bounded by* $2 - \Lambda$*, where* Λ *is defined as in (3.3).*

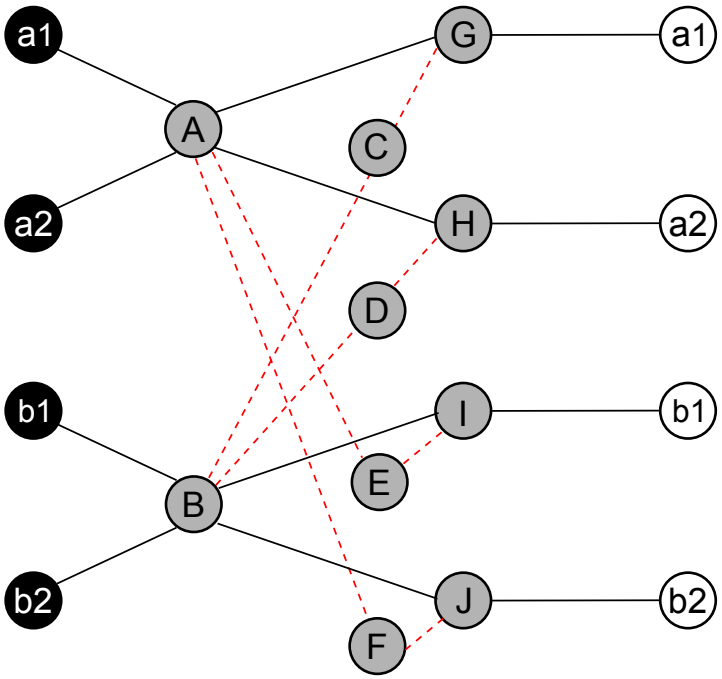


Figure 3.5: Wired network equivalent for SISO scenario.

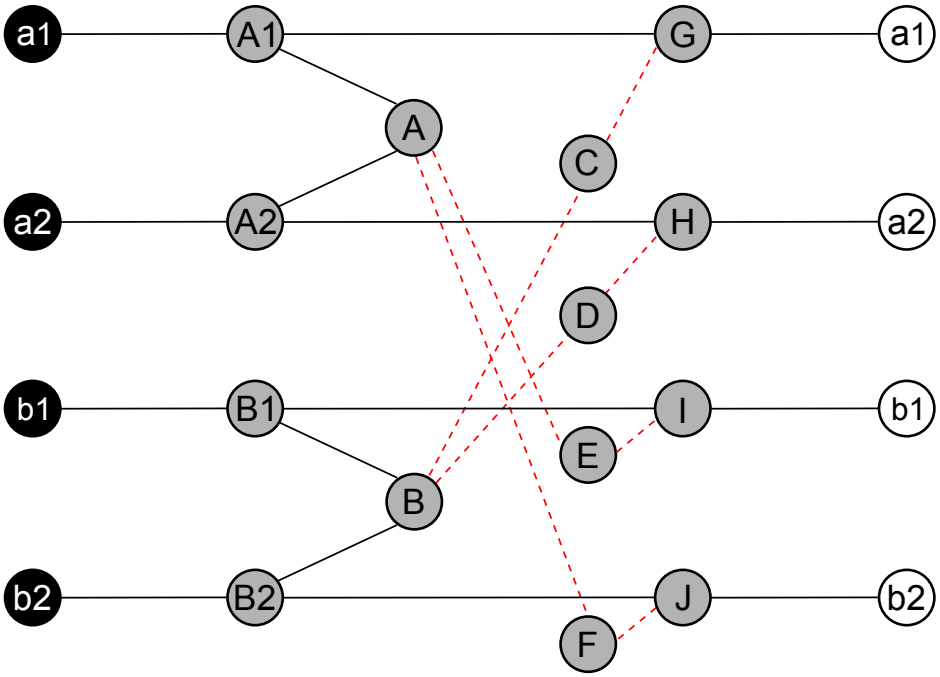


Figure 3.6: Wired network equivalent for MISO/MIMO scenario.

Corollary 3.4. *The normalised sum capacity of a wired network with the same end-to-end topology as the wireless MISO/MIMO network considered in this chapter is upper bounded by* $2 + 2\lambda_2 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8 + \Gamma$*, where* Γ *is defined as in (3.14).*

Note that the bounds in these corollaries have also been confirmed by deriving the outer bounds for the wired scenarios, i.e. using discrete rather than differential entropy and omitting noise considerations. Details are not provided, since our main focus is on wireless networks.

3.8 Conclusion

In this chapter we studied the DoF of a two-cell two-user-per-cell IBC with alternating connectivity and global topological interference management. Our analysis was first carried out for SISO systems, and later extended to MISO and MIMO ones. For each setting, we derived novel DoF outer bounds and investigated their achievability. We also proposed new transmission schemes based on joint coding across states and demonstrated under what conditions the derived outer bounds are achievable. In particular, when a single state has a probability of occurrence equal to one, the bounds are tight and for the best case scenario there is a twofold increase in achievable DoF over the no global topological CSIT case. Additionally, when all states are equiprobable, the SISO system obtains a gain of $\frac{11}{16}$ DoF and achieves 96.4% of the derived outer bound. For the corresponding MISO/MIMO scenario, there is a gain of $\frac{1}{2}$ DoF and the outer bound itself can be achieved. Our results clearly show that significant DoF gains can be obtained when transmitters are provided with global topological information, indicating that even such a minimal level of global CSIT is still highly useful.

Appendix 3.A
Additional details for proof of Theorem 3.1

3.A.1 Derivation of cell B rate outer bound in (3.7)

Considering the cell B achievable rate, we have

$$nR_{B,S} \leq I(W^B; Y_1^V, \ldots, Y_{16}^V) + n\epsilon$$

$$\begin{aligned}
&\overset{(a)}{=} h(Y_1^V,\ldots,Y_{16}^V) - h(Y_Q^V,Y_R^V,Y_S^V,Y_T^V \mid W^B) + n\epsilon \\
&= h(Y_1^V,\ldots,Y_{16}^V) - h(Y_R^V,Y_S^V,Y_T^V \mid W^B) - \underbrace{h(Y_Q^V, \mid W^B,Y_R^V,Y_S^V,Y_T^V)}_{\substack{\geq H(Y_Q^V,\mid W^B,Y_R^V,Y_S^V,Y_T^V,W^A) \\ = no(\log_2 P)}} + n\epsilon \\
&\overset{(b)}{\leq} h(Y_1^V,\ldots,Y_{16}^V) - h(Y_R^V,Y_S^V,Y_T^V \mid W^B) + no(\log_2 P) + n\epsilon \qquad (3.39)
\end{aligned}$$

where (a) follows since the original set of 16 states is contained in states Q, R, S and T, and (b) follows since conditioning reduces entropy and the effect of noise disappears at high SNR.

Considering (3.39) and the state configurations in Fig. 3.4, it can be noticed that the cell B received signal in states R and T consists only of an X^B component and noise. The X^B component has no effect on entropy, and the effect of noise can be represented as $no(\log_2 P)$. For all states corresponding to S, the cell B received signals are a combination of X^A, X^B and noise. The X^B component is negligible with respect to entropy. The X^A and noise components are independent of W^B, and due to the statistical equivalence of $H_{V,A}$ and $H_{U,A}$ can be represented in terms of the signal received at user U, provided that U is free from inter-cell interference. Comparing the list of S states from Table 3.1, this is guaranteed as being always possible regardless of the current (U,V) combination only for state 3. Using this information we can express (3.39) as

$$\begin{aligned}
nR_{B,S} &\leq h(Y_1^V,\ldots,Y_{16}^V) - h(H_{U,A}X_3^A + H_{U,B}X_3^B + Z) + no(\log_2 P) + n\epsilon \\
&= h(Y_1^V,\ldots,Y_{16}^V) - h(H_{U,A}X_3^A + Z) + no(\log_2 P) + n\epsilon \\
&\leq h(Y_1^V) + \cdots + h(Y_{16}^V) - h(Y_3^U) + no(\log_2 P) + n\epsilon
\end{aligned}$$

which is equivalent to (3.7).

3.A.2 Derivation of cell B genie aided DoF bound in (3.13)

The genie aided bound for the cell B achievable rate is given by

$$nR_{\Sigma(GB),S} \leq I(W^A,W^B;Y_1^V,\ldots,Y_{16}^V,G_S^B) + n\epsilon \qquad (3.40)$$

where G_S^B represents the genie set required at cell B such that the data required across the two cells can be decoded within cell B. Genies are necessary when no cell A data is received in cell B, which corresponds to states R and T in Table 3.1, resulting in

$$G_S^B = \{Y_2^U, Y_4^U, \ldots, Y_{14}^U\}\,. \tag{3.41}$$

Integrating G_S^B into (3.40), we obtain

$$\begin{aligned}
&nR_{\Sigma(GB),S}\\
&\leq h(Y_1^V,\ldots,Y_{16}^V,G_S^B) - \underbrace{h(Y_1^V,\ldots,Y_{16}^V,G_S^B \mid W^A,W^B)}_{=no(\log_2 P)} + n\epsilon\\
&\leq h(Y_1^V)+\cdots+h(Y_{16}^V)+h(Y_2^U)+h(Y_4^U)+\cdots+h(Y_{14}^U)+no(\log_2 P)+n\epsilon\\
&\overset{(a)}{\leq} n(2-\lambda_1-\lambda_3-\lambda_{15}-\lambda_{16})+no(\log_2 P)+n\epsilon
\end{aligned}$$

where (a) follows from the fact that Gaussian distribution maximises differential entropy and using $\sum_{k=1}^{16}\lambda_k = 1$. Normalising by $n(\log_2 P)$ and letting $P\to\infty$, we obtain the cell B genie aided DoF bound in (3.13).

Appendix 3.B

Additional details for proof of Theorem 3.2

3.B.1 Derivation of $a2$'s achievable rate bound in (3.18)

For user $a2$, we have

$$\begin{aligned}
&nR_{a2,M}\\
&\leq I(W^A;Y_1^{a2},\ldots,Y_{16}^{a2})+n\epsilon\\
&= h(Y_1^{a2},\ldots,Y_{16}^{a2}) - h(Y_1^{a2},\ldots,Y_{16}^{a2}\mid W^A)+n\epsilon\\
&= h(Y_1^{a2},\ldots,Y_{16}^{a2}) - h(Y_{L_2'}^{a2},Y_{L_2}^{a2}\mid W^A) - \underbrace{h(Y_1^{a2},Y_{15}^{a2}\mid W^A,Y_2^{a2},\ldots,Y_{14}^{a2},Y_{16}^{a2})}_{E_{a2}} + n\epsilon
\end{aligned} \tag{3.42}$$

where $L_2' = \{2,3,4,5,8,10,12,16\}$ and $L_2 = \{6,7,9,11,13,14\}$.

For the L_2' set, the $a2$ received signal consists only of an X^A component, which has no effect on entropy and can therefore be ignored, and noise whose contribution is $no(\log_2 P)$. For the L_2 set, data received at $a2$ is a combination of X^A, X^B and noise. Using this information, (3.42) can be expressed as

$$
\begin{aligned}
&nR_{a2,M} \\
&\leq h(Y_1^{a2},\ldots,Y_{16}^{a2}) - h(Y_{L_2}^{a2} \mid W^A) - E_{a2} + no(\log_2 P) + n\epsilon \\
&\overset{(a)}{=} h(Y_1^{a2},\ldots,Y_{16}^{a2}) - h(H_{a2,A}X_6^A + H_{a2,B}X_6^B + Z, H_{a2,A}X_7^A + H_{a2,B}X_7^B + Z, \\
&\quad H_{a2,A}X_9^A + H_{a2,B}X_9^B + Z, H_{a2,A}X_{11}^A + H_{a2,B}X_{11}^B + Z, H_{a2,A}X_{13}^A + H_{a2,B}X_{13}^B \\
&\quad + Z, H_{a2,A}X_{14}^A + H_{a2,B}X_{14}^B + Z \mid W^A) - E_{a2} + no(\log_2 P) + n\epsilon \\
&\overset{(b)}{=} h(Y_1^{a2},\ldots,Y_{16}^{a2}) - h(H_{b1,B}X_6^B + Z, H_{b2,B}X_7^B + Z, H_{b2,B}X_9^B + Z, H_{b1,B}X_{11}^B \\
&\quad + Z, H_{b2,B}X_{13}^B + Z, H_{b1,B}X_{14}^B + Z) - E_{a2} + no(\log_2 P) + n\epsilon \\
&\overset{(c)}{=} h(Y_1^{a2},\ldots,Y_{16}^{a2}) - h(Y_6^{b1}, Y_7^{b2}, Y_9^{b2}, Y_{11}^{b1}, Y_{13}^{b2}, Y_{14}^{b1}) - E_{a2} + no(\log_2 P) + n\epsilon
\end{aligned}
\tag{3.43}
$$

where (a) follows by expressing the signal in the L_2 set in terms of their separate components; (b) follows by ignoring the X^A parts since they don't have any effect on entropy and replacing channel coefficients from BS B to user $a2$ with ones to cell B users, due to their statistical equivalence, and (c) is obtained by replacing the X^B and noise components with the equivalent received signals at cell B inter-cell interference free users. Finally, by considering all the components of the first negative term in (3.43) to be independent of each other, we obtain (3.18).

3.B.2 Derivation of $b1$'s achievable rate bound in (3.19)

For user $b1$, we have

$$
\begin{aligned}
nR_{b1,M} &\leq I(W^B; Y_1^{b1},\ldots,Y_{16}^{b1}) + n\epsilon \\
&= h(Y_1^{b1},\ldots,Y_{16}^{b1}) - h(Y_1^{b1},\ldots,Y_{16}^{b1} \mid W^B) + n\epsilon \\
&= h(Y_1^{b1},\ldots,Y_{16}^{b1}) - h(Y_{L_3'}^{b1}, Y_{L_3}^{b1} \mid W^B) \\
&\quad - \underbrace{h(Y_1^{b1}, Y_{13}^{b1} \mid W^B, Y_2^{b1},\ldots,Y_{12}^{b1}, Y_{14}^{b1}, Y_{15}^{b1}, Y_{16}^{b1})}_{E_{b1}} + n\epsilon
\end{aligned}
\tag{3.44}
$$

where $L_3' = \{2,5,6,7,8,11,12,14\}$ and $L_3 = \{3,4,9,10,15,16\}$.

For the L_3' set, the signal received at $b1$ consists only of an X^B component which is a function of W^B and is therefore negligible with respect to entropy, and noise whose contribution is $no(\log_2 P)$. Taking this into consideration, and expressing the received signals for the L_3 set in terms of their original components, allows us to express

$$
\begin{aligned}
&nR_{b1,M} \\
&\leq h(Y_1^{b1},\ldots,Y_{16}^{b1}) - h(H_{b1,B}X_3^A + H_{b1,B}X_3^B + Z, H_{b1,B}X_4^A + H_{b1,B}X_4^B + Z, \\
&\quad H_{b1,B}X_9^A + H_{b1,B}X_9^B + Z, H_{b1,B}X_{10}^A + H_{b1,B}X_{10}^B + Z, H_{b1,B}X_{15}^A + H_{b1,B}X_{15}^B \\
&\quad + Z, H_{b1,B}X_{16}^A + H_{b1,B}X_{16}^B + Z, \mid W^B) - E_{b1} + no(\log_2 P) + n\epsilon \\
&\overset{(a)}{=} h(Y_1^{b1},\ldots,Y_{16}^{b1}) - E_{b1} - h(H_{a1,A}X_3^A + Z, H_{a1,A}X_4^A + Z, H_{a1,A}X_9^A + Z, \\
&\quad H_{a2,A}X_{10}^A + Z, H_{a1,A}X_{15}^A + Z, H_{a2,A}X_{16}^A + Z) + no(\log_2 P) + n\epsilon \\
&\overset{(b)}{=} h(Y_1^{b1},\ldots,Y_{16}^{b1}) - h(Y_3^{a1}, Y_4^{a1}, Y_9^{a1}, Y_{10}^{a2}, Y_{15}^{a1}, Y_{16}^{a2}) - E_{b1} + no(\log_2 P) + n\epsilon
\end{aligned}
\tag{3.45}
$$

where (a) follows by removing the X^B components, since they have no effect on entropy, and replacing the channel coefficients from BS B to user $b1$ with ones to cell A users, since they are statistically equivalent, and (b) follows by representing the X^A and noise components using the signals received at inter-cell interference free cell A users. Finally, (3.19) is obtained by considering all the components of the first negative term in (3.45) to be independent of each other.

3.B.3 Derivation of $b2$'s achievable rate bound in (3.20)

For user $b2$, we have

$$
\begin{aligned}
nR_{b2,M} &\leq I(W^B; Y_1^{b2},\ldots,Y_{16}^{b2}) + n\epsilon \\
&= h(Y_1^{b2},\ldots,Y_{16}^{b2}) - h(Y_1^{b2},\ldots,Y_{16}^{b2} \mid W^B) + n\epsilon \\
&= h(Y_1^{b2},\ldots,Y_{16}^{b2}) - h(Y_{L_4'}^{b2}, Y_{L_4}^{b2} \mid W^B) \\
&\quad - \underbrace{h(Y_1^{b2}, Y_{14}^{b2} \mid W^B, Y_2^{b2},\ldots,Y_{13}^{b2}, Y_{15}^{b2}, Y_{16}^{b2})}_{E_{b2}} + n\epsilon
\end{aligned}
\tag{3.46}
$$

where $L_4' = \{2,4,6,7,8,9,10,13\}$ and $L_4 = \{3,5,11,12,15,16\}$.

For the L_4' set, the signal received at $b2$ consists only of an X^B component which has no effect on entropy, and noise whose contribution is $no(\log_2 P)$. For the L_4 set, the data received at $b2$ consists of X^A, X^B and noise. Therefore we can express (3.46) as

$$
\begin{aligned}
&nR_{b2,M} \\
&\leq h(Y_1^{b1},\ldots,Y_{16}^{b1}) - h(H_{b2,B}X_3^A + H_{b2,B}X_3^B + Z, H_{b2,B}X_5^A + H_{b2,B}X_5^B + Z,
\end{aligned}
$$

$$
\begin{aligned}
&H_{b2,B}X_{11}^{A}+H_{b2,B}X_{11}^{B}+Z,H_{b2,B}X_{12}^{A}+H_{b2,B}X_{12}^{B}+Z,H_{b2,B}X_{15}^{A}+H_{b2,B}X_{15}^{B}\\
&+Z,H_{b2,B}X_{16}^{A}+H_{b2,B}X_{16}^{B}+Z,\mid W^{B})-E_{b2}+no(\log_2 P)+n\epsilon\\
&\overset{(a)}{=}h(Y_1^{b2},\ldots,Y_{16}^{b2})-h(H_{a2,A}X_3^{A}+Z,H_{a2,A}X_5^{A}+Z,H_{a1,B}X_{11}^{A}+Z,H_{a2,B}X_{12}^{A}\\
&+Z,H_{a1,B}X_{15}^{A}+Z,H_{a2,B}X_{16}^{A}+Z)-E_{b2}+no(\log_2 P)+n\epsilon\\
&\overset{(b)}{=}h(Y_1^{b2},\ldots,Y_{16}^{b2})-h(Y_3^{a2},Y_5^{a2},Y_{11}^{a1},Y_{12}^{a2},Y_{15}^{a1},Y_{16}^{a2})-E_{b2}+no(\log_2 P)+n\epsilon
\end{aligned}
\tag{3.47}
$$

where (a) follows by removing the X^B components, since they are functions of W^B and thus negligible with respect to entropy, and also replacing the channel coefficients from BS B to user $b2$ with ones to cell A users, since they are statistically equivalent, and (b) is obtained by representing the remaining X^A and noise components in terms of the signals received at inter-cell interference free cell A users. Finally, by considering all the components of the first negative entropy term in (3.47) to be independent of each other we obtain (3.20).

3.B.4 Derivation of cell B genie aided DoF bound in (3.33)

The genie aided bound for the cell B achievable rate is given by

$$
nR_{\Sigma(GB),M}\leq I(W^A,W^B;Y_1^{b1},\ldots,Y_{16}^{b1},Y_1^{b2},\ldots,Y_{16}^{b2},G_M^B)+n\epsilon \tag{3.48}
$$

where G_M^B represents the genie set required by cell B, given by

$$
G_M^B=\left\{2\times[Y_2^A,Y_6^A,Y_7^A,Y_8^A],Y_4^A,Y_5^A,Y_9^A,\ldots,Y_{14}^A\right\}.
$$

Having defined G_M^B , this can be integrated into (3.48) to obtain

$$
\begin{aligned}
nR_{\Sigma(GB),M}&\leq h(Y_{16}^{b1},\ldots,Y_{16}^{b1},Y_1^{b2},\ldots,Y_{16}^{b2},G_M^B)+n\epsilon\\
&\quad-\underbrace{h(Y_{16}^{b1},Y_1^{b2},\ldots,Y_{16}^{b2},G_M^B\mid W^A,W^B)}_{=no(\log_2 P)}\\
&\leq h(Y_1^{b1})+\cdots+h(Y_{16}^{b1})+h(Y_1^{b2})+\cdots+h(Y_{16}^{b2})+2h(Y_2^A)+h(Y_4^A)\\
&\quad+h(Y_5^A)+2h(Y_6^A)+2h(Y_7^A)+2h(Y_8^A)+h(Y_9^A)+\cdots+h(Y_{14}^A)\\
&\quad+no(\log_2 P)+n\epsilon\\
&\overset{(a)}{\leq} n(2+2\lambda_2+\lambda_4+\lambda_5+2\lambda_6+2\lambda_7+2\lambda_8+\lambda_9+\cdots+\lambda_{14})(\log_2 P)\\
&\quad+no(\log_2 P)+n\epsilon
\end{aligned}
\tag{3.49}
$$

where (a) follows by using the fact that Gaussian distribution maximises differential entropy and applying $\sum_{k=1}^{16} \lambda_k = 1$. Normalising by $n(\log_2 P)$ and letting $P \to \infty$, we obtain the cell B genie aided DoF bound in (3.33).

Appendix 3.C

Useful Lemma

Lemma 3.1. *For independent $H_{i,A}X_k^A$, $H_{i,B}X_k^B$ and Z, $h(H_{i,A}X_k^A + H_{i,B}X_k^B + Z) \leq h(H_{i,A}X_k^A) + h(H_{i,B}X_k^B) + no(\log_2 P)$.*

Proof. Starting with the following equality [47], for D and E independent of each other

$$h(D) + h(E) = h(D, E) = h(D, D + E) = h(D + E) + h(E|D + E) \, .$$

Letting $D = H_{i,A}X_k^A$ and $E = H_{i,B}X_k^B + Z$, we have

$$\begin{aligned}
&h(H_{i,A}X_k^A + H_{i,B}X_k^B + Z) \\
&= h(H_{i,A}X_k^A) + h(H_{i,B}X_k^B + Z) - \underbrace{h(H_{i,B}X_k^B + Z|H_{i,A}X_k^A + H_{i,B}X_k^B + Z)}_{\substack{\geq h(H_{i,B}X_k^B + Z|H_{i,A}X_k^A + H_{i,B}X_k^B + Z, W^B) \\ = h(Z|H_{i,A}X_k^A + H_{i,B}X_k^B + Z, W^B) \\ = no(\log_2 P)}} \\
&\overset{(a)}{\leq} h(H_{i,A}X_k^A) + h(H_{i,B}X_k^B) + no(\log_2 P)
\end{aligned}$$

where (a) follows since the effect of noise disappears as $P \to \infty$ and can thus be represented as $no(\log_2 P)$. □

Chapter 4

Interference Alignment for MIMO Interference Broadcast Channels with Imperfect CSI

4.1 Introduction

In Chapter 3, we considered a two-cell two-user-per-cell scenario where only topological information is available with respect to the inter-cell links. The topological CSIT assumption is on the worst case end of the CSIT availability range depicted in Fig. 2.5. In the rest of this thesis, we will move further along the channel information availability spectrum, and consider the case where global CSI is available but imperfect. Moreover, in this chapter we shift our attention to the general G-cell K-user-per-cell MIMO IBC setting, and consider the use of linear IA to manage the resultant interference under an imperfect CSI assumption.

While IA has the potential to achieve full DoF, this generally requires the highly idealistic assumption of the availability of perfect CSI at both the transmitter and the receiver. Therefore, it is important to fully understand to what extent imperfect CSI knowledge degrades IA performance. A substantial amount of work in literature focuses on imperfection due to quantisation. For example, it has been shown that with quantised CSI feedback, IA can still achieve optimal

Work in this chapter has been published in IEEE Transactions on Communications, April 2015 [48].

DoF, as long as the feedback bit rate scales sufficiently fast with SNR for both SISO [49] and MIMO ICs [50]. Aside from imperfection due to quantisation, performance analysis of IA under generalised CSI mismatch is of great interest; however, due to the complex nature of the issue, different works deal with various CSI uncertainty aspects separately. For example, the DoF achievable by IA over correlated channels with imperfect CSI has been investigated in [51], while [52] deals with the performance analysis of IA in systems with analogue channel state feedback. Also, [53] derives upper and lower bounds on the sum mutual information where the variance of the CSI error is considered as a constant. The literature highlighted so far deals with either multiple point-to-point interfering links or a single transmitter communicating with multiple users. Within the context of multi-cell systems with more than one user per cell, [54] develops a scheme for the IMAC that approaches full DoF as the number of users in each cell increases, while [55] studies the feasibility of IA in the symmetric MIMO IMAC. Moving on to the MIMO IBC, the achievable DoF under perfect CSI were initially studied for two-cell systems in [56, 57] and later investigated for systems with a varying number of cells in [58–61].

In this chapter, we focus on the performance of linear IA in the presence of imperfect CSI in a MIMO IBC setting. The CSI mismatch model used is highly versatile and allows us to the treat error variance either as a function of SNR or as independent of it. Given this error model, we derive a bound on the sum rate loss, and quantify the achievable DoF for the MIMO IBC. Results show that when the error variance is inversely proportional to SNR, full DoF can be achieved and the asymptotic sum rate loss is bounded by a derived value dependent on both the system parameters and the CSI error parameters. When the error variance depends on SNR to the power of a proper fraction, we quantify the ensuing DoF loss and also show that the asymptotic sum rate loss is unbounded.

Next, we consider two linear IA schemes for the MIMO IBC, namely the Max-SINR and the Min-LI algorithms. Both techniques were initially proposed for the MIMO IC in [10]. Subsequent works have shown that a straightforward extension to the MIMO IBC does not provide optimal results [62–64]. Here, we present our adaptations to the multi-cell multi-stream MIMO IBC setting, and use them to verify the derived bounds. Additionally, we consider performance improvement for the Max-SINR algorithm under CSI mismatch. This algorithm is given high relevance in literature since it has been found to outperform other techniques.

For example, [65] establishes its optimality within the class of linear beamforming algorithms at high SNR, while [66] shows that it achieves better throughput than sum rate gradient algorithms at low-to-intermediate SNRs. Therefore, inspired by the imperfect CSI model used to derive the bounds, we propose a novel version of the Max-SINR algorithm for the MIMO IBC that exploits knowledge of the CSI error variance in order to counter its negative impact. Results show that the proposed method, which we refer to as Max-SINR with statistical knowledge of the CSI error (Max-SINR-SKCE), does indeed improve performance over the standard version, without any additional computational costs.

This chapter is organised as follows. Section 4.2 provides some preliminaries, including, the system model, the CSI error model and the signal recovery process at the receivers. Section 4.3 gives an overview of the performance of IA under perfect CSI. In Section 4.4 we deal with the performance analysis of IA under imperfect CSI conditions, presenting two theorems that separately define the asymptotic sum rate loss and quantify the achievable DoF. Next, Section 4.5 presents IA schemes for the MIMO IBC; the first part focuses on adaptations of standard schemes used to verify the derived theorems, while the second part introduces the Max-SINR-SKCE algorithm. Section 4.7 provides simulation results, and finally Section 4.8 presents some concluding remarks. Additionally, there is an appendix which contains a number of lemmas used throughout this chapter.

4.2 Preliminaries

4.2.1 System model

We consider the symmetric[1] G-cell MIMO IBC network depicted in Fig. 4.1, where every cell has K users, each equipped with N_d antennas and requiring b_d streams. There is one BS having M_B antennas per cell, and it is assumed that the choice of system parameters is such that IA is feasible. The signal received by user k in cell g is given by

$$\mathbf{y}_{k_g} = \underbrace{\mathbf{U}_{k_g}^H \mathbf{H}_{k_g,g} \mathbf{V}_{k_g} \mathbf{s}_{k_g}}_{\text{desired signal}} + \underbrace{\sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{U}_{k_g}^H \mathbf{H}_{k_g,g} \mathbf{V}_{l_g} \mathbf{s}_{l_g}}_{\text{intra-cell interference}} + \underbrace{\sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{l=1}^{K} \mathbf{U}_{k_g}^H \mathbf{H}_{k_g,j} \mathbf{V}_{l_j} \mathbf{s}_{l_j}}_{\text{inter-cell interference}} + \underbrace{\mathbf{U}_{k_g}^H \mathbf{z}_{k_g}}_{\text{noise}} \quad (4.1)$$

[1] We consider a symmetric system for notational and presentational simplicity. This analysis can also be carried for non-symmetric systems.

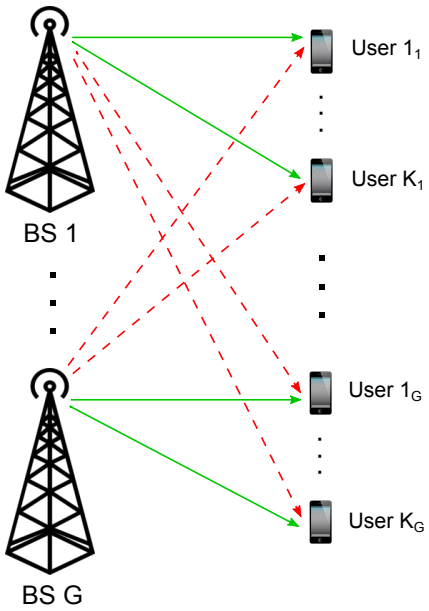


Figure 4.1: G-cell K-user-per-cell MIMO IBC with green solid arrows representing direct links and red dashed arrows representing inter-cell interference links.

where $\mathbf{s}_{k_g} \in \mathbb{C}^{b\times 1}$ is the transmitted symbol vector intended for k_g, satisfying $\mathbb{E}\{\mathbf{s}_{k_g}\mathbf{s}_{k_g}^H\} \leq P\mathbf{I}$; $\mathbf{V}_{k_g} \in \mathbb{C}^{M\times b}$ is the transmit beamforming matrix for data transmitted to k_g; $\mathbf{U}_{k_g} \in \mathbb{C}^{N\times b}$ is the receive beamforming matrix applied at k_g; $\mathbf{H}_{k_g,j} \in \mathbb{C}^{N\times M}$ is the channel link from BS j to user k_g with each element being drawn from a complex normal distribution with zero mean and variance one, and $\mathbf{z}_{k_g} \in \mathbb{C}^{N\times 1}$ represents AWGN with variance σ^2.

4.2.2 Imperfect CSI considerations

We are concerned with the effect of imperfect CSI on IA performance, thus we define the following model for the CSI mismatch

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{E} \tag{4.2}$$

where $\widehat{\mathbf{H}}$ represents the observed mismatched CSI, $\text{vec}(\mathbf{H}) \sim \mathcal{CN}(0, \mathbf{I})$ is the actual channel matrix, and $\mathbf{E}$ is the error matrix representing the degree of inaccuracy in the observed CSI. The error matrix, $\mathbf{E}$, is assumed to be independent of $\mathbf{H}$. Defining nominal SNR as $\rho \triangleq \frac{P}{\sigma^2}$, then $\mathbf{E}$ is modelled as [67]

$$\text{vec}(\mathbf{E}) \sim \mathcal{CN}(0, \eta\mathbf{I}) \;\text{ with }\; \eta \triangleq \beta\rho^{-\alpha}\,. \tag{4.3}$$

With this model, the error variance, η, can be used to capture a variety of CSI acquisition scenarios for any constants $\alpha \geq 0$ and $\beta > 0$. Of particular interest are the following instances.

- *Perfect CSI:* As $\alpha \to \infty$, perfect CSI is obtained for $\rho \geq 1$.

- *Reciprocal channels:* In reciprocal systems, like TDD, users transmit pilots over the UL based on which channel information is obtained at the BS. This CSI knowledge is applicable for both UL and DL channels, due to the reciprocity. Therefore, the CSI error is dependent on the ratio of pilot power to noise level at the BS, i.e. it is inversely proportional to SNR, and the error can be modelled by setting $\alpha = 1$ in (4.3).

- *Mismatched reciprocal channels:* There may be instances of reciprocal channels where the BS and the user equipment have power levels that vary significantly from each other (for example they are not in the same order of magnitude). Such scenarios are referred to as mismatched reciprocal channels. These can be represented by having $0 < \alpha < 1$, depending on the level of the power mismatch.

- *CSI feedback:* In non-reciprocal systems, like FDD, UL and DL are independent. Given the lack of reciprocity, pilots are transmitted by the BS to the users, allowing the receivers to obtain the DL channel information. This information is supplied to the BS via a dedicated feedback link. Data sent over this link is quantised before transmission and received back at the BS with a certain feedback delay, therefore the major contributors to the CSI mismatch are the quantisation and feedback processes. Since the resulting channel error is independent of SNR, it can be modelled by setting $\alpha = 0$ in (4.3).

Alternatively the error variance, η, as a whole can be interpreted as a single parameter that encapsulates the quality of the CSI. Its value may be assumed to be known *a priori*, and can be determined depending on the channel dynamics and the channel estimation schemes applied. Some examples are highlighted below, for additional details see [68] and references within.

- For a block Rayleigh fading channel with coherence time T, using orthonormal training signals, $\eta = 1/(1 + \frac{T_\tau}{t} P_\tau)$ where T_τ is the training interval

length, P_τ is the transmit power of the training symbols and t is the size of the channel input [69].

- For a continuously time varying Rayleigh fading channel with a band-limited low-pass spectrum with cutoff frequency, F, using pilot signals with sampling rate $1/L \geq 2F$, $\eta = 1/(1 + \frac{1}{2FL}P_\tau)$ [70].
- For a CSI feedback scenario where each user feeds back its channel index, $\eta = 2^{-B/M}$ where B is the number of feedback bits that represent the index of the quantised complex channel vector of length M [71].

For our performance analysis we require the statistical properties of the actual channel $\mathbf{H}$ conditioned on $\widehat{\mathbf{H}}$. Since $\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{E}$, with $\mathbf{H}$ and $\mathbf{E}$ being statistically independent Gaussian variables, then $\widehat{\mathbf{H}}$ and $\mathbf{H}$ are jointly Gaussian. Therefore conditioned on $\widehat{\mathbf{H}}$, $\mathbf{H}$ is Gaussian distributed with mean $\widehat{\mathbf{H}}/(1+\eta)$ and statistically independent elements of variance $\eta/(1+\eta)$ [72]. Thus, the actual channel can be expressed as

$$\mathbf{H} = \frac{1}{1+\eta}\widehat{\mathbf{H}} + \boldsymbol{\Upsilon} \tag{4.4}$$

where $\boldsymbol{\Upsilon}$ is independent of $\widehat{\mathbf{H}}$, and distributed as

$$\mathrm{vec}(\boldsymbol{\Upsilon}) \sim \mathcal{CN}(0, \frac{\eta}{1+\eta}\mathbf{I})\,. \tag{4.5}$$

4.2.3 Signal recovery at the receivers

Here we focus on the data recovery process at the user nodes using linear receivers. In particular, without loss of generality, we consider a ZF equaliser. Note that more sophisticated non-linear receiver architectures can also be applied, in which case the considerations outlined in this section would not be directly applicable.

For perfect CSI at both the receiver and the transmitter, defining the effective channel as $\bar{\mathbf{H}}_{k_g,l_j} = \mathbf{U}_{k_g}^H \mathbf{H}_{k_g,j} \mathbf{V}_{l_j}$ and the ZF equaliser as $\mathbf{G}_{k_g}$, the recovered signal at user k_g can be expressed as

$$\mathbf{G}_{k_g}\mathbf{y}_{k_g} = \underbrace{\mathbf{G}_{k_g}\bar{\mathbf{H}}_{k_g,k_g}\mathbf{s}_{k_g}}_{\text{desired signal}} + \underbrace{\left(\mathbf{G}_{k_g}\sum_{\substack{l=1\\l\neq k}}^{K}\bar{\mathbf{H}}_{k_g,l_g}\mathbf{s}_{l_g} + \sum_{\substack{j=1\\j\neq g}}^{G}\sum_{l=1}^{K}\bar{\mathbf{H}}_{k_g,l_j}\mathbf{s}_{l_j} + \mathbf{U}_{k_g}^H\mathbf{z}_{k_g}\right)}_{\text{interference-plus-noise}} \tag{4.6}$$

where the interference-plus-noise term disappears at high SNR since we are applying IA, and the ZF equaliser $\mathbf{G}_{k_g}$ can be expressed as $\mathbf{G}_{k_g} = (\bar{\mathbf{H}}_{k_g,k_g})^{\dagger}$.

For imperfect CSI at the transmitter, the recovered signal at user k_g can be expressed as

$$\begin{aligned}\widehat{\mathbf{G}}_{k_g}\mathbf{y}_{k_g} = &\underbrace{\widehat{\mathbf{G}}_{k_g}\widehat{\mathbf{U}}_{k_g}^{H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,g}+\boldsymbol{\Upsilon}_{k_g,g}\right)\widehat{\mathbf{V}}_{k_g}^{H}\mathbf{s}_{k_g}}_{\text{desired signal}} \\ &+\underbrace{\left(\widehat{\mathbf{G}}_{k_g}\sum_{\substack{l=1\\l\neq k}}^{K}\tilde{\mathbf{H}}_{k_g,l_g}\mathbf{s}_{l_g}+\sum_{\substack{j=1\\j\neq g}}^{G}\sum_{l=1}^{K}\tilde{\mathbf{H}}_{k_g,l_j}\mathbf{s}_{l_j}+\widehat{\mathbf{U}}_{k_g}^{H}\mathbf{z}_{k_g}\right)}_{\text{interference-plus-noise}}\end{aligned} \tag{4.7}$$

where $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$ are the receive and transmit beamformers calculated with imperfect CSI, $\tilde{\mathbf{H}}_{k_g,l_j}$ is the effective channel with imperfect CSI given by $\tilde{\mathbf{H}}_{k_g,l_j} = \widehat{\mathbf{U}}_{k_g}^{H}(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,j}+\boldsymbol{\Upsilon}_{k_g,j})\widehat{\mathbf{V}}_{l_j} = \widehat{\mathbf{U}}_{k_g}^{H}\mathbf{H}_{k_g,j}\widehat{\mathbf{V}}_{l_j}$, and $\widehat{\mathbf{G}}_{k_g}$ is the ZF equaliser for imperfect CSI. The value of $\widehat{\mathbf{G}}_{k_g}$ depends on the quality of the CSI available at the receiver and is specified as follows.

- *For perfect CSI at the receiver:* $\widehat{\mathbf{G}}_{k_g} = (\widehat{\mathbf{U}}_{k_g}^{H}\mathbf{H}_{k_g,g}\widehat{\mathbf{V}}_{k_g}^{H})^{\dagger}$.
- *For imperfect CSI at the receiver:* $\widehat{\mathbf{G}}_{k_g} = (1+\eta)(\widehat{\mathbf{U}}_{k_g}^{H}\widehat{\mathbf{H}}_{k_g,g}\widehat{\mathbf{V}}_{k_g}^{H})^{\dagger}$, i.e. the signal needs to be scaled by $(1+\eta)$ for signal recovery.

4.3 Performance with perfect CSI

Our analysis is concerned with performance degradation due to imperfect CSI. Since our focus is on the sum rate and DoF loss, it is necessary to first define these metrics under perfect CSI for comparison purposes. For the system specified in (4.7) with perfect CSI, the following conditions need to be satisfied to achieve IA

$$\begin{aligned}&|\mathbf{u}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^{d}| > 0 \;\forall\; d,k,g \\ &\mathbf{u}_{k_g}^{d\,H}\mathbf{H}_{k_g,j}\mathbf{v}_{l_j}^{m} = 0 \;\forall\; d,k,g,(d,k,g)\neq(m,l,j)\end{aligned} \tag{4.8}$$

where $\mathbf{u}_{k_g}^{d}$ and $\mathbf{v}_{k_g}^{d}$ refer to the dth column of $\mathbf{U}_{k_g}$ and $\mathbf{V}_{k_g}$ respectively.

Considering i.i.d. Gaussian inputs, and the fact there is no interference leakage with perfect CSI, then the achievable sum rate is given by

$$R = \sum_{g=1}^{G} \sum_{k=1}^{K} \sum_{d=1}^{b_d} \log_2 \left(1 + \frac{P|\mathbf{u}_{k_g}^{d\,H} \mathbf{H}_{k_g,g} \mathbf{v}_{k_g}^{d}|^2}{\sigma^2} \right) . \tag{4.9}$$

Sum rate is related to DoF as

$$\text{DoF} = \lim_{P \to \infty} \frac{\text{Rate}}{\log_2 P} . \tag{4.10}$$

Using this relationship, provided that the IA feasibility conditions are met, we obtain the DoF achievable with perfect CSI as

$$D = \lim_{P \to \infty} \frac{R}{\log_2 P} = GKb_d . \tag{4.11}$$

4.4 Performance analysis under imperfect CSI

When it comes to the imperfect CSI scenario, the information available for the calculation of the precoders and the receivers is only an imperfect observation of the actual channel state; thus, all beamformers are calculated using $\widehat{\mathbf{H}}$ rather than $\mathbf{H}$. This implies that instead of the original IA conditions in (4.8), the alignment conditions observed are

$$\begin{aligned} &|\widehat{\mathbf{u}}_{k_g}^{d\,H} \widehat{\mathbf{H}}_{k_g,g} \widehat{\mathbf{v}}_{k_g}^{d}| > 0 \;\forall\; d,k,g \\ &\widehat{\mathbf{u}}_{k_g}^{d\,H} \widehat{\mathbf{H}}_{k_g,j} \widehat{\mathbf{v}}_{l_j}^{m} = 0 \;\forall\; d,k,g,(d,k,g) \neq (m,l,j) \end{aligned} \tag{4.12}$$

where $\widehat{\mathbf{u}}_{k_g}^{d}$ and $\widehat{\mathbf{v}}_{k_g}^{d}$ refer to the dth column of the beamformers calculated with imperfect CSI, $\widehat{\mathbf{U}}_{k_g}$ and $\widehat{\mathbf{V}}_{k_g}$ respectively. Satisfying the modified IA conditions in (4.12) leads to an amount of residual leakage interference equivalent to

$$\begin{aligned} \widehat{J}_{k_g}^{d} = &\sum_{\substack{m=1 \\ m \neq d}}^{b_d} P \left| \widehat{\mathbf{u}}_{k_g}^{d\,H} \mathbf{H}_{k_g,g} \widehat{\mathbf{v}}_{k_g}^{m} \right|^2 + \sum_{\substack{l=1 \\ l \neq k}}^{K} \sum_{m=1}^{b_d} P \left| \widehat{\mathbf{u}}_{k_g}^{d\,H} \mathbf{H}_{k_g,g} \widehat{\mathbf{v}}_{l_g}^{m} \right|^2 \\ &+ \sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{l=1}^{K} \sum_{m=1}^{b_d} P \left| \widehat{\mathbf{u}}_{k_g}^{d\,H} \mathbf{H}_{k_g,j} \widehat{\mathbf{v}}_{l_j}^{m} \right|^2 . \end{aligned} \tag{4.13}$$

Residual leakage has a negative impact on the sum rate and DoF achievable by the system. Understanding the extent of this loss is fundamental, because it gives

a more realistic characterisation of system performance. Here, we present two theorems that separately quantify the asymptotic sum rate loss and the decrease in achievable DoF.

4.4.1 Sum rate loss with imperfect CSI

The mean sum rate loss, ΔR, is a measure of the difference between the expected value for the sum rate achievable with perfect CSI, R, and the expected value for the sum rate achievable with imperfect CSI, $\widehat{R}$, i.e.

$$\Delta R = \mathbb{E}_{\mathbf{H}}\{R\} - \mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}\}\} \tag{4.14}$$

where R is defined in (4.9) and $\widehat{R}$ is defined as

$$\widehat{R} = \sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d} \log_2\left(1 + \frac{P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2}{\widehat{J}_{k_g}^{d} + \sigma^2}\right). \tag{4.15}$$

Given this definition, we can now refer to the following theorem.

Theorem 4.1. *Consider a G-cell MIMO IBC with K users per cell, each requiring b streams. For such a system under imperfect CSI with error variance, $\eta = \beta\rho^{-\alpha}$, at asymptotically high SNR: ΔR tends to zero for $\alpha > 1$, tends to infinity for $0 \leq \alpha < 1$, and for $\alpha = 1$ is finite and upper bounded by $GKb_d \log_2(1 + \beta(GKb_d - 1))$, i.e.*

$$\lim_{SNR\to\infty} \Delta R \begin{cases} = 0 & \alpha > 1 \\ \leq GKb_d \log_2(1 + \beta(GKb_d - 1)) & \alpha = 1 \\ \to \infty & 0 \leq \alpha < 1\,. \end{cases} \tag{4.16}$$

Proof. Starting with the expression from (4.14) and replacing R with (4.9) and $\widehat{R}$ with (4.15), we obtain

$$\begin{aligned}\Delta R = \mathbb{E}_{\mathbf{H}}&\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d} \log_2\left(1 + \frac{P|\mathbf{u}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^{d}|^2}{\sigma^2}\right)\right\} \\ &- \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d} \log_2\left(1 + \frac{P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2}{\widehat{J}_{k_g}^{d} + \sigma^2}\right)\right\}\right\}. \end{aligned} \tag{4.17}$$

After some algebraic manipulations (4.17) can be further represented as (4.18).

$$\begin{aligned}\Delta R = \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{P|\mathbf{u}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^{d}|^2}{\sigma^2}\right)\right\}\\ -\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g}^{d}+P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2}{\sigma^2}\right)\right\}\right\}\\ +\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g}^{d}}{\sigma^2}\right)\right\}\right\}\end{aligned} \tag{4.18}$$

Since for unitary beamformers, analogous to [73, Lemma 2] it can be shown that $|\mathbf{u}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^{d}|^2$ and $|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2$ are exponentially distributed with both mean and variance one, we can establish the following inequality

$$\begin{aligned}\mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{P|\mathbf{u}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^{d}|^2}{\sigma^2}\right)\right\}\leq\\ \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g}^{d}+P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2}{\sigma^2}\right)\right\}\right\}.\end{aligned} \tag{4.19}$$

Next, taking into account (4.19), we can express (4.18) as

$$\begin{aligned}\Delta R &\leq \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g}^{d}}{\sigma^2}\right)\right\}\right\}\\ &\overset{(a)}{\leq}\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}}{\sigma^2}\right)\end{aligned} \tag{4.20}$$

where (a) follows from Jensen's inequality.

Therefore to quantify ΔR, we need to find an expression for the expected value of the leakage interference, $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}\}$. Combining the expression for $\widehat{J}_{k_g}^{d}$ from (4.13), with the channel model from (4.4) we obtain

$$\begin{aligned}\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}\} = &\sum_{\substack{m=1\\m\neq d}}^{b_d}P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g}^{d\,H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,g}+\boldsymbol{\Upsilon}_{k_g,g}\right)\widehat{\mathbf{v}}_{k_g}^{m}\right|^2\right\}\\ &+\sum_{\substack{l=1\\l\neq k}}^{K}\sum_{m=1}^{b_d}P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g}^{d\,H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,g}+\boldsymbol{\Upsilon}_{k_g,g}\right)\widehat{\mathbf{v}}_{l_g}^{m}\right|^2\right\}\end{aligned}$$

$$+\sum_{\substack{j=1\\j\neq g}}^{G}\sum_{l=1}^{K}\sum_{m=1}^{b_d} P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g}^{d\,H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,j}+\boldsymbol{\Upsilon}_{k_g,j}\right)\widehat{\mathbf{v}}_{l_j}^{m}\right|^2\right\}. \tag{4.21}$$

This can be further simplified by considering the IA conditions for imperfect CSI in (4.12). Applying the fact that $\left[\widehat{\mathbf{u}}_{k_g}^{d\,H}\widehat{\mathbf{H}}_{k_g,j}\widehat{\mathbf{v}}_{j_l}^{m}=0\;\forall\; d,k,g,(d,k,g)\neq(m,l,j)\right]$, results in

$$\begin{aligned}\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}\} &= \sum_{\substack{m=1\\m\neq d}}^{b_d} P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g}^{d\,H}\boldsymbol{\Upsilon}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{m}\right|^2\right\}+\sum_{\substack{l=1\\l\neq k}}^{K}\sum_{m=1}^{b_d} P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g}^{d\,H}\boldsymbol{\Upsilon}_{k_g,g}\widehat{\mathbf{v}}_{l_g}^{m}\right|^2\right\}\\ &\quad+\sum_{\substack{j=1\\j\neq g}}^{G}\sum_{l=1}^{K}\sum_{m=1}^{b_d} P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g}^{d\,H}\boldsymbol{\Upsilon}_{k_g,j}\widehat{\mathbf{v}}_{l_j}^{m}\right|^2\right\}\\ &\overset{(a)}{=} P\frac{\eta}{(\eta+1)}(GKb_d-1)\end{aligned} \tag{4.22}$$

where (a) follows by using Lemma 4.1 from Appendix 4.A.

This expression for $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}\}$ is next applied into the inequality from (4.20) to obtain

$$\Delta R \leq \sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{(\eta+1)}(GKb_d-1)\right) \tag{4.23}$$

which after evaluating the summation and replacing η with $\beta\rho^{-\alpha}$, becomes

$$\Delta R \leq GKb_d\log_2\left(1+(GKb_d-1)\frac{\beta\rho^{(1-\alpha)}}{1+\beta\rho^{-\alpha}}\right). \tag{4.24}$$

Taking a high SNR approximation, the asymptotic sum rate loss can be defined as in (4.16), proving Theorem 4.1 as originally stated. □

4.4.2 DoF loss with imperfect CSI

The DoF loss, ΔD, is a measure of the difference between the DoF achievable with perfect CSI, D, and the DoF achievable under imperfect CSI, $\widehat{D}$, i.e. $\Delta D = D-\widehat{D}$, where

$$\widehat{D}=\lim_{P\to\infty}\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}\}\}}{\log_2 P} \tag{4.25}$$

Given this definition, we can now refer to the following theorem.

Theorem 4.2. *Consider a G-cell MIMO IBC with K users per cell, each requiring b streams. For such a system under imperfect CSI with error variance, $\eta = \beta\rho^{-\alpha}$, full DoF can be achieved for values of $\alpha \geq 1$, while in the range of $0 \leq \alpha < 1$ the DoF loss is equivalent to a fraction of $(1-\alpha)$ of the full DoF, i.e.*

$$\Delta D = \begin{cases} 0 & \alpha \geq 1 \\ (1-\alpha)GKb_d & 0 \leq \alpha < 1\,. \end{cases} \tag{4.26}$$

Proof. We have already established that the DoF achievable under perfect CSI is given by $D = GKb$ from (4.11). Therefore, we only need to evaluate the achievable DoF under imperfect CSI, $\widehat{D}$.

Starting with (4.25) and including the imperfect CSI sum rate expression from (4.15), we obtain

$$\begin{aligned} \widehat{D} &= \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(1+\frac{P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2}{\widehat{J}_{k_g}^{d}+\sigma^2}\right)\right\}\right\}}{\log_2 P} \\ &= \underbrace{\lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2+\widehat{J}_{k_g}^{d}+\sigma^2\right)\right\}\right\}}{\log_2 P}}_{\text{A}} \\ &\quad - \underbrace{\lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(\widehat{J}_{k_g}^{d}+\sigma^2\right)\right\}\right\}}{\log_2 P}}_{\text{B}} \end{aligned} \tag{4.27}$$

which after discarding the interference-plus-noise term in part A, and applying Jensen's inequality to part B, can be expressed as

$$\begin{aligned} \widehat{D} &\geq \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(P|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2\right)\right\}\right\}}{\log_2 P} \\ &\quad - \lim_{P\to\infty} \frac{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d}\log_2\left(\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}\}+\sigma^2\right)}{\log_2 P}\,. \end{aligned} \tag{4.28}$$

Since for unitary beamformers, analogous to [73, Lemma 2] it can be shown that $|\widehat{\mathbf{u}}_{k_g}^{d\,H}\mathbf{H}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^{d}|^2$ is exponentially distributed with both mean and variance one. Using this result and replacing $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g}^{d}\}\}$ by the actual expression from (4.22), the DoF expression in (4.27) becomes

$$\widehat{D} = GKb_d - \lim_{P\to\infty} \frac{\sum_{g=1}^{G}\sum_{k=1}^{K}\sum_{d=1}^{b_d} \log_2\left(P\frac{\eta}{(\eta+1)}(GKb_d-1)+\sigma^2\right)}{\log_2 P}\,. \tag{4.29}$$

After replacing η with $\beta\rho^{-\alpha} = \beta P^{-\alpha}\sigma^{2\alpha}$ and letting $P\to\infty$, the achievable DoF can be characterised as

$$\widehat{D} = \begin{cases} GKb_d & \alpha \geq 1 \\ \alpha GKb_d & 0 \leq \alpha < 1\,. \end{cases} \tag{4.30}$$

Noting that $\Delta D = D - \widehat{D}$ and also that $D = GKb_d$ from (4.11), we obtain the result in (4.26), thereby proving the DoF loss is as originally stated in Theorem 4.2. □

Remark 4.1. *Note that the implications of the two theorems presented in this section are intrinsically related. For example, in the range of $\alpha \geq 1$ Theorem 4.1 indicates the sum rate loss is either zero or finite, which is directly reflected in Theorem 4.2 where no DoF loss is expected within the same α range. On the other hand for the case where $0 \leq \alpha < 1$, Theorem 4.2 indicates that a DoF loss is inevitable. This is also reflected in Theorem 4.1, which states that the sum rate loss increases unboundedly with SNR for the same range of α values.*

4.5 IA schemes adapted to the MIMO IBC

In order to test the bounds presented in Section 4.4, we require linear IA schemes for the MIMO IBC. Two iterative IA solutions are the Max-SINR and Min-LI algorithms originally proposed for the MIMO IC in [10]. Having originally been developed for the IC, the algorithms from [10] are unable to cater for intra-cell interference. Various works propose different alternatives on how to handle this additional interference component, here we outline our adaptations for the multi-stream MIMO IBC.

4.5.1 Max-SINR for the MIMO IBC

This algorithm focuses on maximising the signal-to-interference-plus-noise ratio (SINR) on a per stream basis, in order to create a desired signal subspace that contains the required number of interference free dimensions. SINR is defined as the ratio of the power of the signal of interest to the sum power of the interference and noise, i.e.

$$\text{SINR} = \frac{\text{signal of interest power}}{\text{interference-plus-noise power}}\,. \tag{4.31}$$

A direct extension of the algorithm from [10] would involve calculating both the transmit and receive filters based on the total interference-plus-noise covariance matrix, which for the MIMO IBC also includes intra-cell interference. However, [62, 63] show that a direct extension does not always achieve the desired alignment results over the whole SNR range, particularly in the high SNR region where saturation may occur. Solutions proposed in [62] and [63] separately ignore intra-cell interference in the transmit subspace. The adapted Max-SINR algorithm outlined in Algorithm 4.1 applies a similar principle, while still retaining an underlying structure that mirrors the original algorithm from [10]. Thus, the receive filters are concerned only with inter-cell interference, while the transmit filters deal with both inter-cell and intra-cell interference when calculating the forward and backward interference-plus-noise covariance matrices, given by $\mathbf{Q}_{k_g}^d$ in (4.32) and $\overleftarrow{\mathbf{Q}}_{k_g}^d$ in (4.33) respectively.

$$\mathbf{Q}_{k_g}^d = \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{l=1}^{K}\sum_{m=1}^{b_d} P\mathbf{H}_{k_g,j}\mathbf{v}_{l_j}^m \mathbf{v}_{l_j}^{m\,H}\mathbf{H}_{k_g,j}^H + \sigma^2\mathbf{I} \tag{4.32}$$

$$\begin{aligned}\overleftarrow{\mathbf{Q}}_{k_g}^d = &\sum_{\substack{m=1\\ m\neq d}}^{b_d} P\overleftarrow{\mathbf{H}}_{g,k_g}\mathbf{u}_{k_g}^m \mathbf{u}_{k_g}^{m\,H}\overleftarrow{\mathbf{H}}_{g,k_g}^H + \sum_{\substack{l=1\\ l\neq k}}^{K}\sum_{m=1}^{b_d} P\overleftarrow{\mathbf{H}}_{g,l_g}\mathbf{u}_{l_g}^m \mathbf{u}_{l_g}^{m\,H}\overleftarrow{\mathbf{H}}_{g,l_g}^H \\ &+ \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{l=1}^{K}\sum_{m=1}^{b_d} P\overleftarrow{\mathbf{H}}_{g,l_j}\mathbf{u}_{l_j}^m \mathbf{u}_{l_j}^{m\,H}\overleftarrow{\mathbf{H}}_{g,l_j}^H + \sigma^2\mathbf{I}\end{aligned} \tag{4.33}$$

Note that to obtain the backward interference-plus-noise covariance matrix, $\overleftarrow{\mathbf{Q}}_{k_g}^d$, in (4.33) we operate the channel in the reverse direction with users transmitting data to the corresponding BSs. Channel notations of the form $\overleftarrow{\mathbf{H}}_{x,y} = \mathbf{H}_{y,x}^H$ are used to represent the channel going from y to x in the reciprocal network.

Algorithm 4.1: Max-SINR algorithm for the MIMO IBC

1 Initialise $\mathbf{v}_{k_g}^d$ as random unit-norm vectors $\forall\, d, k, g$.

2 Compute the inter-cell interference-plus-noise covariance matrix in the forward communication channel as $\mathbf{Q}_{kg}^d$ from (4.32) $\forall\, d, k, g$.

3 Calculate receive filters $\forall\, d, k, g$ using

$$\mathbf{u}_{k_g}^d = \frac{(\mathbf{Q}_{k_g}^d)^{-1}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^d}{\parallel (\mathbf{Q}_{k_g}^d)^{-1}\mathbf{H}_{k_g,g}\mathbf{v}_{k_g}^d \parallel}\,.$$

4 Reverse the direction of communication and compute the total interference-plus-noise covariance matrix as $\mathbf{B}_{kg}^d$ from (4.33) $\forall\, d, k, g$.

5 Calculate transmit beamformers $\forall\, d, k, g$ using

$$\mathbf{v}_{k_g}^d = \frac{(\overleftarrow{\mathbf{Q}}_{k_g}^d)^{-1}\overleftarrow{\mathbf{H}}_{g,k_g}\mathbf{u}_{k_g}^d}{\parallel (\overleftarrow{\mathbf{Q}}_{k_g}^d)^{-1}\overleftarrow{\mathbf{H}}_{g,k_g}\mathbf{u}_{k_g}^d \parallel}\,.$$

6 Repeat the process from Step 2 until convergence or for a fixed number of iterates.

4.5.2 Min-LI for the MIMO IBC

The principle behind this algorithm is to design beamformers that limit the interference experienced from all other users within the same system. The original algorithm was proposed for the IC in [10] and does not cater for intra-cell interference. The key aspect in adapting it to the IBC is to treat intra-cell and inter-cell interference separately. This can be achieved by applying iterative leakage minimisation only with respect to inter-cell interference, and then using an additional cascaded precoder to handle intra-cell interference on its own. Leakage minimisation with cascaded filters was proposed in [64] for a single-stream setting, here we apply its multi-stream counterpart as outlined in Algorithm 4.2.

4.6 Max-SINR algorithm with statistical knowledge of the CSI error

In this section, inspired by the CSI mismatch model used for the performance analysis, we focus on the Max-SINR algorithm and propose a novel version, Max-SINR-SKCE, that exploits statistical knowledge of the CSI error in order to improve performance. The key difference between the naive Max-SINR technique

Algorithm 4.2: Min-LI algorithm for the MIMO IBC

1 Initialise $\widetilde{\mathbf{V}}_g$ as a random unitary matrix $\forall\ g$.
2 Calculate the inter-cell interference covariance matrix in the forward communication channel $\forall\ k, g$ using
$$\mathbf{B}_{k_g} = \sum\nolimits_{\substack{j=1 \\ j\neq g}}^{G} \mathbf{H}_{k_g,j}\widetilde{\mathbf{V}}_j\widetilde{\mathbf{V}}_j^H\mathbf{H}_{k_g,j}^H\,.$$
3 The receive filter at each user is given by $\mathbf{U}_{kg} = \mathcal{V}_b\left[\mathbf{B}_{k_g}\right]$.
4 Calculate the inter-cell interference covariance matrix for the BSs in the reciprocal network $\forall\ g$ using
$$\overleftarrow{\mathbf{B}}_g = \sum\nolimits_{\substack{j=1 \\ j\neq g}}^{G}\sum\nolimits_{k=1}^{K} \overleftarrow{\mathbf{H}}_{g,l_j}\mathbf{U}_{l_j}\mathbf{U}_{l_j}^H\overleftarrow{\mathbf{H}}_{g,l_j}^H\,.$$
5 The first part of the transmit beamformer at BS g is given by $\widetilde{\mathbf{V}}_g = \mathcal{V}_{Kb}\left[\mathbf{B}_g\right]$, where $\mathcal{V}_{Kb}\left[\mathbf{B}_g\right]$ represents the set of eigenvectors corresponding to the Kb smallest eigenvalues of $\mathbf{B}_g$.
6 Repeat the process from Step 2 until convergence or for a fixed number of iterates.
7 Calculate $\bar{\mathbf{V}}_g$, the additional ZF precoder that handles intra-cell interference using
$$\bar{\mathbf{V}}_g = \begin{bmatrix} \mathbf{U}_{1_g}^H\mathbf{H}_{1_g,g}\widetilde{\mathbf{V}}_g \\ \vdots \\ \mathbf{U}_{K_g}^H\mathbf{H}_{K_g,g}\widetilde{\mathbf{V}}_g \end{bmatrix}^{\dagger}\,. \tag{4.34}$$
8 Overall transmit beamformer at BS g is given by, $\mathbf{V}_g = \widetilde{\mathbf{V}}_g\bar{\mathbf{V}}_g$.
9 Take b consecutive columns of $\mathbf{V}_g$ separately for each user k and normalise to obtain $\mathbf{V}_{k_g}\ \forall\ k, g$.

in Algorithm 4.1 and Max-SINR-SKCE is in the way the interference-plus-noise covariance matrices are calculated when the available CSI is imperfect. The naive version uses imperfect CSI directly in place of the actual channels without any consideration for effects that the channel mismatch may have. Thus, the beamformers are calculated by replacing $\mathbf{H}$ with $\widehat{\mathbf{H}}$ directly in (4.32) and (4.33). However, in the design of the Max-SINR-SKCE algorithm we take advantage of statistical knowledge with respect to the CSI mismatch and replace $\mathbf{H}$ with the expression in (4.4). This leads to the calculation of more accurate interference-plus-noise covariance matrices in both forward and backward directions.

Starting with the forward channel inter-cell interference covariance matrix in (4.32) and replacing $\mathbf{H}$ with (4.4), we obtain

$$\mathbf{Q}_{k_g}^d = \sum_{\substack{j=1 \\ j\neq g}}^{G}\sum_{l=1}^{K}\sum_{m=1}^{b_d} P\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,j} + \boldsymbol{\Upsilon}_{k_g,j}\right)\mathbf{v}_{l_j}^m\mathbf{v}_{l_j}^{m\,H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g,j} + \boldsymbol{\Upsilon}_{k_g,j}\right)^H + \sigma^2\mathbf{I}$$

$$= \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{l=1}^{K} \sum_{m=1}^{b_d} P \Bigg[\frac{1}{(1+\eta)^2} \widehat{\mathbf{H}}_{k_g,j} \mathbf{v}_{l_j}^m \mathbf{v}_{l_j}^{m\,H} \widehat{\mathbf{H}}_{k_g,j}^H + \underbrace{\boldsymbol{\Upsilon}_{k_g,j} \mathbf{v}_{l_j}^m \mathbf{v}_{l_j}^{m\,H} \boldsymbol{\Upsilon}_{k_g,j}^H}_{\text{A}} + \frac{1}{(1+\eta)} \underbrace{\left(\widehat{\mathbf{H}}_{k_g,j} \mathbf{v}_{l_j}^m \mathbf{v}_{l_j}^{m\,H} \boldsymbol{\Upsilon}_{k_g,j}^H + \boldsymbol{\Upsilon}_{k_g,j} \mathbf{v}_{l_j}^m \mathbf{v}_{l_j}^{m\,H} \widehat{\mathbf{H}}_{k_g,j}^H \right)}_{\text{B}} \Bigg] + \sigma^2 \mathbf{I} \,. \tag{4.35}$$

Since the only information available with respect to the channel uncertainty is statistical, we replace all elements of (4.35) containing $\boldsymbol{\Upsilon}$ by their expected values. Using Lemma 4.2 from Appendix 4.A $\mathbb{E}_{\boldsymbol{\Upsilon}}\{\text{A}\} = \frac{\eta}{1+\eta}\mathbf{I}$, and from Lemma 4.3 $\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\text{B}\} = 0$. Therefore, the improved expression for the inter-cell interference-plus-noise covariance matrix in the forward direction is given by

$$\widehat{\mathbf{Q}}_{k_g}^d = \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{l=1}^{K} \sum_{m=1}^{b_d} \gamma \widehat{\mathbf{H}}_{k_g,j} \mathbf{v}_{l_j}^m \mathbf{v}_{l_j}^{m\,H} \widehat{\mathbf{H}}_{k_g,j}^H + \xi \mathbf{I} \tag{4.36}$$

where

$$\gamma = \frac{P}{(1+\eta)^2} \tag{4.37}$$

and

$$\xi = P \frac{\eta}{(1+\eta)} (G-1) K b_d + \sigma^2 \,. \tag{4.38}$$

Reversing the direction of communication, such that users are now transmitting data to the corresponding BSs, we can also calculate the interference-plus-noise covariance matrix for the backward channel. This is done using a method similar to the one applied in the forward communication channel, and results in

$$\begin{aligned} \overleftarrow{\widehat{\mathbf{Q}}}_{k_g}^d = & \sum_{\substack{m=1 \\ m\neq d}}^{b_d} \gamma \overleftarrow{\mathbf{H}}_{g,k_g} \mathbf{u}_{k_g}^m \mathbf{u}_{k_g}^{m\,H} \overleftarrow{\mathbf{H}}_{g,k_g}^H + \sum_{\substack{l=1 \\ l\neq k}}^{K} \sum_{m=1}^{b_d} \gamma \overleftarrow{\mathbf{H}}_{g,l_g} \mathbf{u}_{l_g}^m \mathbf{u}_{l_g}^{m\,H} \overleftarrow{\mathbf{H}}_{g,l_g}^H \\ & + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{l=1}^{K} \sum_{m=1}^{b_d} \gamma \overleftarrow{\mathbf{H}}_{g,l_j} \mathbf{u}_{l_j}^m \mathbf{u}_{l_j}^{m\,H} \overleftarrow{\mathbf{H}}_{g,l_j}^H + \overleftarrow{\xi} \mathbf{I} \end{aligned} \tag{4.39}$$

where γ is as defined in (4.37) and

$$\overleftarrow{\xi} = P \frac{\eta}{(1+\eta)} (GKb_d - 1) + \sigma^2 \,. \tag{4.40}$$

Having obtained improved expressions for the interference-plus-noise matrices in both directions, the novel Max-SINR-SKCE algorithm is as outlined in Algorithm 4.3 on the following page.

Remark 4.2. *The advantage of the Max-SINR-SKCE algorithm is its ability to calculate improved interference covariance matrices by proper specification of* γ, ξ *and* $\overleftarrow{\xi}$. *Setting* $\gamma = P$ *and* $\xi_f = \xi_b = \sigma^2$ *in the first step of Algorithm 4.3 would cause it to behave exactly in the same manner as the naive version in Algorithm 4.1. Therefore, any performance advantages obtained by the use of the novel version are obtained at no extra computational cost.*

Algorithm 4.3: Max-SINR-SKCE algorithm for the MIMO IBC

1 Set γ, ξ and $\overleftarrow{\xi}$ according to (4.37), (4.38) and (4.40).
2 Initialise $\widehat{\mathbf{v}}_{k_g}^d$ as random unit-norm vectors, $\forall\, d, k, g$.
3 Calculate $\widehat{\mathbf{Q}}_{kg}^d$ using the improved expression in (4.36) $\forall\, d, k, g$.
4 The receive filters $\forall\, d, k, g$ are given by

$$\widehat{\mathbf{u}}_{kg}^d = \frac{(\widehat{\mathbf{Q}}_{k_g}^d)^{-1}\widehat{\mathbf{H}}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^d}{\parallel (\widehat{\mathbf{Q}}_{k_g}^d)^{-1}\widehat{\mathbf{H}}_{k_g,g}\widehat{\mathbf{v}}_{k_g}^d \parallel}\,.$$

5 Compute $\overleftarrow{\widehat{\mathbf{Q}}}_{k_g}^d$ using the improved expression in (4.39) $\forall\, d, k, g$.
6 The precoders $\forall\, d, k, g$ are calculated as

$$\widehat{\mathbf{v}}_{kg}^d = \frac{(\overleftarrow{\widehat{\mathbf{Q}}}_{k_g}^d)^{-1}\overleftarrow{\widehat{\mathbf{H}}}_{g,k_g}\widehat{\mathbf{u}}_{k_g}^d}{\parallel (\overleftarrow{\widehat{\mathbf{Q}}}_{k_g}^d)^{-1}\overleftarrow{\widehat{\mathbf{H}}}_{g,k_g}\widehat{\mathbf{u}}_{k_g}^d \parallel}\,.$$

7 Repeat the process from Step 2 until convergence or for a fixed number of iterates.

4.7 Simulation results

This section provides numerical results to validate the analyses presented so far. It is divided into two main parts; first we confirm the validity of the bounds derived in Section 4.4, next we consider the performance of the Max-SINR-SKCE algorithm proposed in Section 4.6 compared to the standard one under imperfect CSI conditions.

Assuming all interference is treated as noise, throughout our simulations we calculate the achieved sum rate across all users as

$$\bar{R} = \sum_{g=1}^{G} \sum_{k=1}^{K} \log_2 \det \left(\mathbf{I} + \left(\mathbf{X}_{k_g} + \sigma^2 \mathbf{I} \right)^{-1} P \, \mathbf{U}_{k_g}^H \mathbf{H}_{k_g,g} \mathbf{V}_{k_g} \mathbf{V}_{k_g}^H \mathbf{H}_{k_g,g}^H \mathbf{U}_{k_g} \right)$$

where $\mathbf{X}_{k_g}$ represents the interference covariance matrix, given by

$$\mathbf{X}_{k_g} = \sum_{\substack{j=1 \\ (j,l) \neq (g,k)}}^{G} \sum_{l=1}^{K} P \mathbf{U}_{k_g}^H \mathbf{H}_{k_g,j} \mathbf{V}_{l_j} \mathbf{V}_{l_j}^H \mathbf{H}_{k_g,j}^H \mathbf{U}_{k_g} \, .$$

For imperfect CSI situations, the transmit and receive filters $\mathbf{U}$ and $\mathbf{V}$ are replaced by $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$, since they are calculated based on the available imperfect CSI. Note that calculating the rate in this manner results in a lower bound on the actual achievable rate. In truth higher rates can be obtained via the use of improved receivers, for example by considering the availability of perfect CSI at the receiver to obtain more accurate beamformers, or by applying more sophisticated receiver strategies such as for example maximum likelihood detectors.

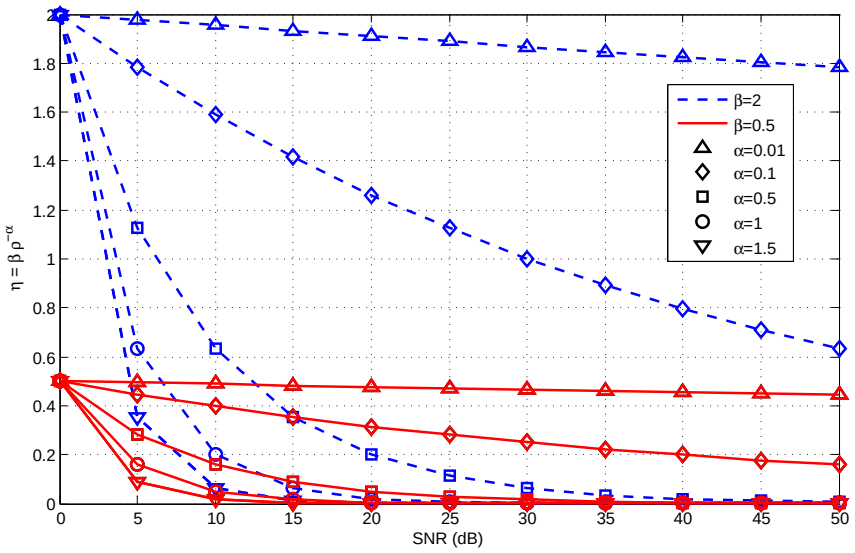


Figure 4.2: Error variance, η, against SNR for different α and β combinations.

Throughout all our simulations the noise variance, σ^2, is fixed at 1 making the transmit signal power equivalent to the network SNR. Additionally, all results provided are averaged over a number of channel realisations in a Monte Carlo fashion. Furthermore, β values are always chosen to be significantly larger than α. Fig. 4.2 plots the error variance, η, against SNR for various α and β combinations.

Considering for example $\beta = 0.5$, it can be noticed that for $\alpha \in \{1.5, 1, 0.5\}$ the resultant error variance is close to 0 for SNR $\geq$ 25 dB, however for the lower α values of 0.1 and 0.01 the error variance is significantly higher for the same SNR range. Considering $\alpha = 0.5$, increasing the β value from 0.5 to 2 shifts the quasi-zero error variance point from 25 dB up to 40 dB. Since our work focuses on IA, which is mainly concerned with DoF, and consequently the high SNR region, throughout our simulations β is always at least an order of magnitude larger than α. This ensures that the effect of the CSI mismatch is experienced across the whole SNR range, as confirmed by the η vs. SNR plots in Fig. 4.2.

4.7.1 Results for theoretically derived bounds

In this subsection we verify the validity of the bounds derived in Theorems 4.1 and 4.2. We simulate a system with $G = 3$, $K = 2$, $b_d = 1$ and $M_B = N_d = 4$ using the naive Max-SINR method in Algorithm 4.1 to obtain the results in Fig. 4.3, and a system with $G = K = b_d = 2$, $M_B = 4$ and $N_d = 6$ using the Min-LI technique from Algorithm 4.2 to obtain the results in Fig. 4.4.

Considering Fig. 4.3, the full DoF achievable with perfect CSI are equal to $GKb = 6$. Theorem 4.2 predicts no DoF loss for values of $\alpha \geq 1$, which can easily be verified by focusing on the $\alpha = 1.5$ and $\alpha = 1$ results in Fig. 4.3, since the slope for both is exactly equal to the one achieved with perfect CSI. One important difference between the $\alpha = 1.5$ and $\alpha = 1$ curves is the fact at high SNR the former is exactly in line with the perfect CSI result, while the latter runs parallel to it achieving lower sum rate values overall. This behaviour is expected from the bound in Theorem 4.1. For $\alpha > 1$, no sum rate loss is expected at high SNR, which is exactly what happens for $\alpha = 1.5$. However, at $\alpha = 1$ the same theorem indicates there should be a finite asymptotic sum rate loss upper bounded by $GKb_d \log_2 (1 + \beta(GKb_d - 1)) \approx 34.03$ bits per channel use for $\beta = 10$. Measuring the actual loss from Fig. 4.3 at SNR = 60 dB we obtain a value of 33.26 bits per channel use; this approaches the derived value closely, verifying that the upper bound it is not too loose.

When it comes to the $\alpha < 1$ range, from Theorem 4.1 we expect the sum rate loss to be unbounded. This can be confirmed via the $\alpha = 0.75$ and $\alpha = 0$ curves in Fig. 4.3. All three diverge from the perfect CSI result, indicating that the sum rate loss grows with increasing SNR. Within the same α range, a DoF loss is

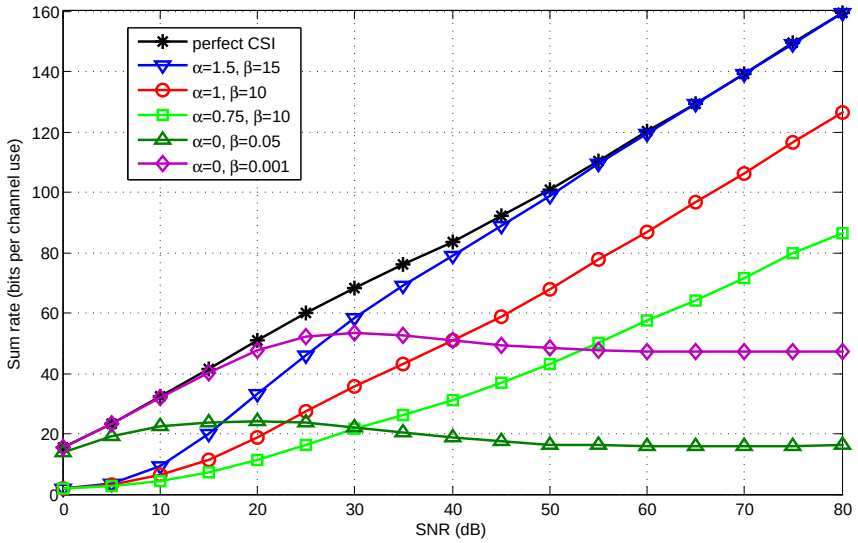


Figure 4.3: Average sum rates achieved by Max-SINR algorithm (Algorithm 4.1) under various imperfect CSI conditions for system with $G = 3$, $K = 2$, $b_d = 1$ and $M_B = N_d = 4$.

expected from Theorem 4.2. For example, at $\alpha = 0.75$ DoF equal to $\frac{3}{4}GKb_d = 4.5$ may be achieved, this can be easily verified from the slope of the curve itself at high SNR. Theorem 4.2 also indicates that 0 DoF are achievable at $\alpha = 0$, which is directly reflected in the flatness of the corresponding results in the high SNR region of Fig. 4.3.

Comparing the two $\alpha = 0$ results in Fig. 4.3 provides an insight into the impact of the β parameter. While for asymptotic analysis its effect is limited and does not determine the general sum rate performance trend, it can be noticed that at $\alpha = 0$, β has a heavier impact. In such situations the error variance η is no longer inversely proportional to SNR. Thus, for any fixed β, $\alpha = 0$ represents the worst case scenario, where the error variance is equal to β itself. In the lower SNR region (up to around 30 dB for $\beta = 0.001$ and 20 dB for $\beta = 0.05$) the power of the leakage is reasonably small, since the power levels we are dealing with are low, allowing for performance improvement in this range. However, once power levels increase, the performance starts to degrade until it eventually settles to a steady state value; this is due to the higher power levels of leakage interference. For the two $\alpha = 0$ results in Fig. 4.3, at $\beta = 0.05$ there is a much larger error variance than at $\beta = 0.001$, which is why the latter settles at a significantly higher rate value. The larger the level of the CSI mismatch, the more inaccurate are the derived IA beamformers with respect to the actual

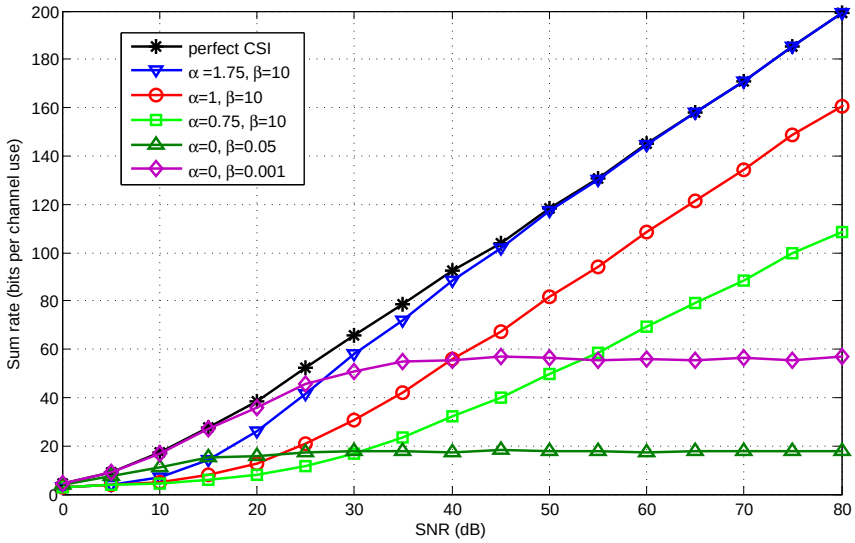


Figure 4.4: Average sum rates achieved by Min-LI algorithm (Algorithm 4.2) under various imperfect CSI conditions for system with $G = K = b_d = 2$, $M_B = 4$ and $N_d = 6$.

channel value, resulting in a higher level of residual leakage interference and sum rate loss. This behaviour does not improve with increasing SNR; since all users are allocated equal power, increasing the desired signal power inherently increases the power of the interfering signals and results in a more significant amount of interference leakage. Thus, the network becomes interference limited, causing sum rate saturation and leading to no advantage overall.

Similar rate and DoF behaviour can also be noticed in Fig. 4.4. Results for $\alpha \geq 1$ all have the same slope as the perfect CSI curve, indicating that full DoF equal to 8 are achieved as expected from Theorem 4.2. Additionally, at asymptotically high SNR for the $\alpha = 1.5$ result there is no asymptotic sum rate loss as expected from Theorem 4.1. For $\alpha = 1, \beta = 10$ the same theorem indicates that the asymptotic sum rate loss should be upper bounded by 49.19 bits per channel use. Measuring the actual gap from Fig. 4.4 a value of 37.38 bits per channel use is obtained. Finally Theorem 4.1 states that the asymptotic sum rate loss is unbounded for the range of $\alpha < 1$; this can be confirmed from the $\alpha = 0.75$ and $\alpha = 0$ curves in Fig. 4.4, both of which diverge from the perfect CSI result. Within the same α range we expect the achievable DoF to be equal to αGKb. Thus, at $\alpha = 0.75$ DoF equal to 6 are achievable, as verified from the slope of the curve in Fig. 4.4. Similarly, both $\alpha = 0$ results saturate at high SNR, denoting that $\alpha GKb = 0$ DoF are obtained.

4.7.2 Results for Max-SINR-SKCE algorithm

Here we compare the novel Max-SINR-SKCE algorithm proposed in Algorithm 4.3 to the standard one outlined in Algorithm 4.1. A system configuration with $G = 3$, $K = 3$, $b_d = 1$ and $M_B = N_d = 5$ is used to produce the sum rate and bit error rate (BER) results in Fig. 4.5 and Fig. 4.6 respectively. We focus on the range $\alpha \leq 1$, since both Theorem 4.1 and Theorem 4.2 indicate that the system becomes asymptotically equivalent to the perfect CSI case for $\alpha > 1$.

As can be seen from Fig. 4.5 and Fig. 4.6, the Max-SINR-SKCE algorithm outperforms the standard one, both in terms of sum rate and BER. In fact Max-SINR-SKCE achieves higher sum rates throughout, for example at $\alpha = 1, \beta = 10$ we obtain a 10.1 bits per channel use gain at an SNR of 30 dB, while for $\alpha = 0.75, \beta = 10$ the gain is equal to 11.83 bits per channel use at the same SNR. When it comes to the $\alpha = 0, \beta = 0.1$ case we observe that as SNR increases, the sum rate achievable by both versions of the algorithm settles at a steady value. This value is approximately 14 bits per channel use higher for the Max-SINR-SKCE algorithm in comparison to the standard one. As observed earlier in Section 4.7.1, at $\alpha = 0$ we have the highest level of channel uncertainty for any given β; for such a significant CSI mismatch the network becomes interference limited and increasing transmission power provides no advantage. This effect is also mirrored into the BER results in Fig. 4.6, where the results for $\alpha = 0$ both start to flatten out for increasing SNR. However, the result obtained using Max-SINR-SKCE settles at lower a BER value than that obtained using standard Max-SINR, indicating the superior performance of the former. For larger α values, the Max-SINR-SKCE algorithm still achieves a lower BER than the standard one. For example, at $\alpha = 1, \beta = 10$ the standard algorithm requires an SNR of approximately 28.36 dB to achieve a BER of 10^{-2}, while Max-SINR-SKCE achieves the same BER at around 20.7 dB. Similarly, for $\alpha = 0.75, \beta = 10$ Max-SINR-SKCE requires 13.47 dB less than standard Max-SINR to achieve a BER of 10^{-2}.

4.8 Conclusion

IA is a very promising technique, and while it has been shown to provide many benefits under a perfect CSI assumption, it is also important to consider the more

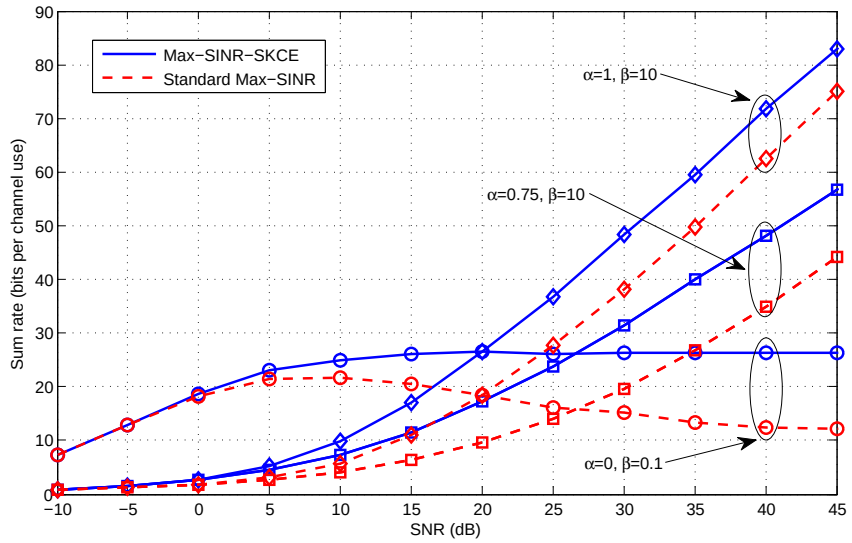


Figure 4.5: Average sum rates achieved for system with $G = 3$, $K = 3$, $b_d = 1$ and $M_B = N_d = 5$ under various imperfect CSI scenarios.

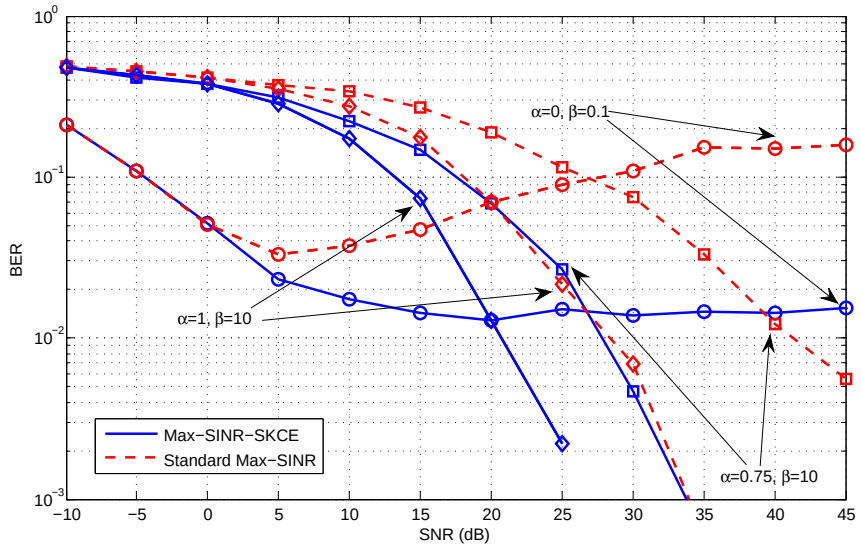


Figure 4.6: BER achieved for system with $G = 3$, $K = 3$, $b_d = 1$ and $M_B = N_d = 5$ under various imperfect CSI scenarios, using QPSK modulation.

realistic imperfect CSI scenario. In this chapter, we analysed the performance of linear IA under CSI mismatch for the MIMO IBC by deriving a bound on the asymptotic mean loss in sum rate compared to the perfect CSI case and characterising the achievable DoF under CSI mismatch. These properties are found to be dependent on the number of cells and the amount of users per cell, in addition to the CSI mismatch parameters themselves. Results show that the way

error variance scales with SNR is highly significant. When the two parameters are inversely proportional, full DoF can be achieved and the asymptotic sum rate loss is finite. However, in cases where the error variance depends inversely on SNR to the power of a proper fraction, full DoF cannot be achieved and the asymptotic sum rate loss is unbounded. Additionally, inspired by the CSI mismatch model used, we also designed a novel Max-SINR-SKCE algorithm which exploits statistical knowledge of the CSI error in order to improve performance over standard Max-SINR, without any additional computational costs.

Appendix 4.A

Useful Lemmas

Lemma 4.1. $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\widehat{\mathbf{u}}_{k_g}^{d\,H}\boldsymbol{\Upsilon}_{k_g,j}\widehat{\mathbf{v}}_{l_j}^{m}|^2\}$ *is equal to* $\eta/(1+\eta)\ \forall\ k,g,d,l,j,m.$

Proof. From the error model definition in Section 4.2.2 we know that $\widehat{\mathbf{H}}_{k_g,j}$ and $\boldsymbol{\Upsilon}_{k_g,j}$ are independent. Since $\widehat{\mathbf{u}}_{k_g}^{d}$ and $\widehat{\mathbf{v}}_{l_j}^{m}$ are calculated on $\widehat{\mathbf{H}}_{k_g,j}$, this makes the transmit and receive beamformers automatically independent of $\boldsymbol{\Upsilon}_{k_g,j}$. In addition, $\boldsymbol{\Upsilon}_{k_g,j}$ is Gaussian and bi-unitarily invariant [74], thereby for unitary beamformers the product $\widehat{\mathbf{u}}_{k_g}^{d\,H}\boldsymbol{\Upsilon}_{k_g,j}\widehat{\mathbf{v}}_{l_j}^{m}\quad\forall\ d,k,g,m,l,j$ is a Gaussian random variable with zero mean and variance $\eta/(1+\eta)$. Finally, using central absolute moments we can evaluate $\mathbb{E}\{|\widehat{\mathbf{u}}_{k_g}^{d\,H}\boldsymbol{\Upsilon}_{k_g,j}\widehat{\mathbf{v}}_{l_j}^{m}|^2\}$, which is equal to $\eta/(1+\eta)$. □

Lemma 4.2. *If* $\mathbf{A}\in\mathbb{C}^{M\times N}$ *is a Gaussian matrix whose elements are i.i.d. with zero mean and variance* ω*, and* $\mathbf{b}\in\mathbb{C}^{N\times 1}$ *is a unit-norm vector independent of* $\mathbf{A}$*, then* $\mathbb{E}_{\mathbf{A}}\{\mathbf{A}\mathbf{b}\mathbf{b}^{H}\mathbf{A}^{H}\}=\omega I.$

Proof. Consider a unitary matrix $\mathbf{B}\in\mathbb{C}^{N\times N}$ that is independent of $\mathbf{A}$. Since $\mathbf{A}$ is a Gaussian matrix, it is bi-unitarily invariant [74], hence the joint distribution of the entries of $\mathbf{AB}$ is equal to that of the entries of $\mathbf{A}$ itself. Additionally, since matrix $\mathbf{B}$ consists of N unit-norm vectors of size $N\times 1$, the vector $\mathbf{b}$ described in the definition of Lemma 4.2 can take the role of any column vector within $\mathbf{B}$. Thus, vector $\mathbf{Ab}$ can be considered as a column vector of matrix $\mathbf{AB}$, implying that the entries of $\mathbf{Ab}$ have zero mean and covariance ω. □

Lemma 4.3. $\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\widehat{\mathbf{H}}_{k_g,j}\mathbf{v}_{l_j}^{m}\mathbf{v}_{l_j}^{m\,H}\boldsymbol{\Upsilon}_{k_g,j}^{H}\}=\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\boldsymbol{\Upsilon}_{k_g,j}\mathbf{v}_{l_j}^{m}\mathbf{v}_{l_j}^{m\,H}\widehat{\mathbf{H}}_{k_g,j}^{H}\}=0\,\forall\,l,m,j.$

Proof. Beamforming elements are calculated using $\widehat{\mathbf{H}}_{k_g,j}$, thus they are automatically independent of $\boldsymbol{\Upsilon}_{k_g,j}$ from the definition of the imperfect CSI model in Section 4.2. □

Chapter 5

Weighted Sum Rate Maximisation in Full-Duplex Multi-User Multi-Cell MIMO Networks

5.1 Introduction

As highlighted earlier in Chapter 2, the application of FD nodes instead of standard HD ones has significant potential to increase spectral efficiency. Therefore, while in the prior Chapters 3 and 4 we considered completely HD systems, in the upcoming Chapters 5 and 6 we will focus on systems with FD BSs. The network considered is still a multi-user multi-cell network; the use of FD BSs now allows us to serve both legacy HD DL and UL users simultaneously. For this chapter our analysis will focus on proposing WSR maximisation algorithms to manage the resulting complex interference scenario under both perfect and imperfect CSI.

The promise of increased spectral efficiency, alongside with the newfound ability to mitigate SI (see Section 2.5), has motivated a wide range of research into FD communication and its possible applications. For example, the use of FD operation in relays [76, 77] and cognitive radio systems [78, 79] has proven to be

Work in this chapter has been published in IEEE Transactions on Communications, April 2017 [75].

highly effective. Additionally FD operation, either at the BS only [80] or at both the users and the BS [81], has been found to be particularly suited for small cell scenarios due to the low transmit powers and small transmission distances involved. Different to [76–81], which consider single-cell systems, here we focus on a more practical multi-cell scenario with FD BSs and HD users; the multi-cell aspect introduces the additional challenge of managing CCI from nodes in other cells. A similar network was considered in [82] where the authors focus on user selection and power allocation methods. A stochastic geometry approach for system performance characterisation of FD multi-cell systems has been considered in [83–85]. In contrast to [82–85], which assume all nodes are equipped with a single antenna, we consider a MIMO system and focus on beamformer design for WSR maximisation.

As was hinted earlier, in this chapter we focus on a multi-cell scenario where each BS serves multiple HD users; however unlike traditional systems, the BSs operate in FD mode serving all of their corresponding DL and UL users simultaneously. The FD capability at the BSs and the inherent structure of the network lead to a large amount of interference at the different receivers. Fig. 5.1 provides a simple illustration of the network under consideration, having G cells and one DL and one UL user per cell. It can be seen that, apart from the usual standard HD network interference components, for UL communication BSs have additional SI and BS-to-BS interference, while DL users have additional CCI from UL users both from the same cell and from other cells. Therefore, finding ways to manage this complex interference scenario, while still delivering good service to all users, is a fundamental challenge that needs to be addressed for the practical realisation of FD enabled wireless networks.

Since our main focus is on small cell networks where coverage distances are short, and BSs and users have similar transmission powers [86], we consider the case where none of the interference components may be ignored. This is in contrast to prior studies which assume that CCI can be avoided via scheduling [87], allocating different sub-carriers [88] or assuming channels between UL and DL users to be sufficiently weak [89]. Additionally, we also take into consideration the effect of transmitter and receiver distortion. These hardware impairments are a natural consequence of non-ideal amplifiers, oscillators, ADCs and digital-to-analogue converters (DACs), and cannot be avoided in practice [77, 79, 90].

Within this context, our aim is to investigate under what conditions replacing

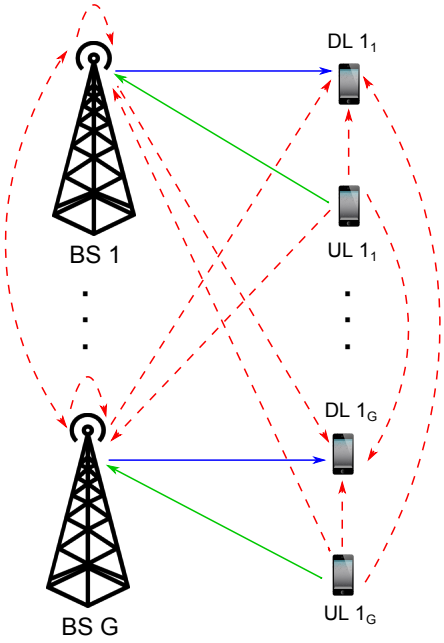


Figure 5.1: G-cell network with an FD BS, and one DL and one UL user per cell. Solid arrows represent desired links, while dashed ones represent interference links.

HD BSs with FD ones is beneficial. Such a multi-cell scenario has not yet been investigated from a beamformer design perspective, here we develop filter design algorithms for WSR maximisation. Since WSR problems are non-convex, we map each of them to a WMMSE problem; this technique is less computationally complex than gradient-based alternatives, is guaranteed to converge, and has been proven to work for various types of HD networks [91–93]. The rate to MSE relationship was also used for transceiver design in MIMO ICs with FD nodes under a perfect CSI assumption in [90]. Unlike [90], in this chapter we consider HD users and cater for multiple users per cell, additionally we consider imperfect CSI under two different models. The first one is a norm-bounded error model, particularly suited for situations where the CSI error is mainly due to quantisation. The second one is a stochastic CSI error model, more suited to errors occurring during the channel estimation process itself. Results show that FD communication can indeed achieve higher rates than the baseline HD schemes for intermediate to low distortion levels, and confirm the robust performance of the imperfect CSI designs. Furthermore, we also extend our original design to one which maximises the total DL rate subject to each UL user achieving a pre-established target rate; this can be used to ensure each UL user is served in every time slot, which is not guaranteed with the joint design.

The rest of the chapter is organised as follows. Section 5.2 provides some preliminaries, including the system model, the CSI error models and the relationship between sum rate and MSE. In Section 5.3 we present the WSR problem under perfect CSI. Sections 5.4 and 5.5 tackle the norm-bounded error and the stochastic error problems respectively. Next, in Section 5.6 we consider the extension to a weighted DL rate maximisation problem subject to a minimum per UL user target rate. Simulation results are presented in Section 5.7. Section 5.8 provides an insight on the implementation and complexity of the proposed algorithms. Finally, Section 5.9 presents some concluding remarks. Additionally, there are two appendices, the first includes a lemma used within this chapter, while the second provides details for the proof of the theorem presented in Section 5.4.

5.2 Preliminaries

5.2.1 System model

We consider a G-cell system where each cell g has one FD BS, K_g^d DL users requiring b_d streams each and K_g^u UL users requiring b_u streams each. A simplified version of this network with one DL and one UL user per cell is depicted in Fig. 5.1. BSs are equipped with M_B FD antennas, DL users are equipped with N_d HD antennas and UL users are equipped with N_u HD antennas. The maximum transmit power is given by P_B at the BSs and P_U at the UL users.

The signal received at user k_g^d, the kth DL user in cell g, and at BS g are given by $\mathbf{y}_{k_g^d}$ and $\mathbf{y}_g$, respectively defined as

$$\mathbf{y}_{k_g^d} = \sum_{j=1}^{G} \mathbf{H}_{k_g^d,j} \sum_{i=1}^{K_j^d} (\mathbf{V}_{i_j^d} \mathbf{s}_{i_j^d} + \mathbf{c}_{i_j^d}) + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathbf{H}_{k_g^d,i_j^u} (\mathbf{V}_{i_j^u} \mathbf{s}_{i_j^u} + \mathbf{c}_{i_j^u}) + \mathbf{n}_{k_g^d} + \mathbf{e}_{k_g^d} \quad (5.1)$$

$$\mathbf{y}_g = \sum_{j=1}^{G} \mathbf{H}_{g,j} \sum_{i=1}^{K_j^d} (\mathbf{V}_{i_j^d} \mathbf{s}_{i_j^d} + \mathbf{c}_{i_j^d}) + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathbf{H}_{g,i_j^u} (\mathbf{V}_{i_j^u} \mathbf{s}_{i_j^u} + \mathbf{c}_{i_j^u}) + \mathbf{n}_g + \mathbf{e}_g \,. \quad (5.2)$$

Here, $\mathbf{H}_{k_g^d,j} \in \mathbb{C}^{N_d \times M_B}$ represents the channel from BS j to DL user k_g^d, $\mathbf{H}_{k_g^d,i_j^u} \in \mathbb{C}^{N_d \times N_u}$ is the channel from UL user i_j^u to DL user k_g^d, $\mathbf{H}_{g,j} \in \mathbb{C}^{M_B \times M_B}$ is the channel from BS j to BS g, and $\mathbf{H}_{g,i_j^u} \in \mathbb{C}^{M_B \times N_u}$ is the channel from UL user i_j^u to BS g. $\mathbf{V}_{i_j^d} \in \mathbb{C}^{M_B \times b_d}$ is the precoder for $\mathbf{s}_{i_j^d}$, with $\mathbf{s}_{i_j^d} \in \mathbb{C}^{b_d \times 1}$ being the data intended for the ith DL user in cell j, where $\mathbb{E}\{\mathbf{s}_{i_j^d} \mathbf{s}_{i_j^d}^H\} = \mathbf{I}$. $\mathbf{V}_{i_j^u} \in \mathbb{C}^{N_u \times b_u}$ is

the precoder for $\mathbf{s}_{i_j^u}$, with $\mathbf{s}_{i_j^u} \in \mathbb{C}^{b_u \times 1}$ being the data transmitted by the ith UL user in cell j, where $\mathbb{E}\{\mathbf{s}_{i_j^u}\mathbf{s}_{i_j^u}^H\} = \mathbf{I}$. Moreover, $\mathbf{n}_{k_g^d}$ and $\mathbf{n}_g$ represent AWGN with zero mean and variance σ_U^2 and σ_B^2 respectively. Finally, $\mathbf{c}_{i_j^u}$ and $\mathbf{c}_{i_j^d}$ represent transmitter distortion at the UL users and the BSs respectively, while $\mathbf{e}_{k_g^d}$ and $\mathbf{e}_g$ represent receiver distortion at the DL users and the BSs respectively.

Transmitter distortion models the effect of limited transmitter dynamic range by approximating the combined effects of additive power-amplifier noise, oscillator phase noise, and non-linearities in the DAC and the power amplifier. This distortion is statistically independent from the transmitted signal and can be modelled as [77]

$$\mathbf{c}_{i_j^d} \sim \mathcal{CN}\left(\mathbf{0}, \kappa_B \text{diag}(\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H)\right)$$
$$\mathbf{c}_{i_j^u} \sim \mathcal{CN}\left(\mathbf{0}, \kappa_U \text{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\right)$$

where $\kappa_U, \kappa_B \ll 1$.

Receiver distortion models the effect of limited receiver dynamic range by capturing the combined effects of oscillator phase noise, additive gain control noise, and non-linearities in the ADC and gain-control. It is statistically independent from the received signal itself and can be modelled as [77]

$$\mathbf{e}_{k_g^d} \sim \mathcal{CN}\Big(\mathbf{0}, \iota_U \text{diag}\big(\text{cov}(\mathbf{y}_{k_g^d} - \mathbf{e}_{k_g^d})\big)\Big)$$
$$\mathbf{e}_g \sim \mathcal{CN}\Big(\mathbf{0}, \iota_B \text{diag}\big(\text{cov}(\mathbf{y}_g - \mathbf{e}_g)\big)\Big)$$

where $\iota_U, \iota_B \ll 1$.

Finally, since with perfect CSI $\mathbf{H}_{g,g}\sum_{i=1}^{K_g^d}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d}$ is known at BS g, this can be subtracted from $\mathbf{y}_g$ resulting in (5.3) with $\Theta = 0$ [77].[1] The effect of residual SI is then captured in the term $\mathbf{H}_{g,g}\sum_{i=1}^{K_g^d}\mathbf{c}_{i_g^d} + \mathbf{e}_g$. The parameter Θ is a binary term used to differentiate between the perfect and imperfect CSI scenarios. For the perfect CSI case $\Theta = 0$, whilst for the imperfect CSI case $\Theta = 1$ leading to an extra residual SI term; further details for the imperfect CSI case are provided in Section 5.2.2.

[1]The SI channel can be estimated using pilot signals. For SI channel estimation, the FD node transmitting the pilot signal is also the one receiving it, this implies that the signal is received with high power. Having a strong signal allows for accurate estimation of $\mathbf{H}_{g,g}$ [76], which implies that $\mathbf{H}_{g,g}\sum_{i=1}^{K_g^d}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d}$ can be considered as available at BS g under the perfect CSI assumption.

$$\tilde{\mathbf{y}}_g = \sum_{\substack{j=1 \\ j\neq g}}^{G} \mathbf{H}_{g,j} \sum_{i=1}^{K_j^d} (\mathbf{V}_{i_j^d} \mathbf{s}_{i_j^d} + \mathbf{c}_{i_j^d}) + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathbf{H}_{g,i_j^u} (\mathbf{V}_{i_j^u} \mathbf{s}_{i_j^u} + \mathbf{c}_{i_j^u}) + \mathbf{H}_{g,g} \sum_{i=1}^{K_g^d} \mathbf{c}_{i_g^d}$$
$$+ \underbrace{\Theta \sum_{i=1}^{K_g^d} \mathbf{\Delta}_{g,g} \mathbf{V}_{i_j^d} \mathbf{s}_{i_j^d}}_{\substack{\text{extra residual SI for} \\ \text{imperfect CSI scenarios}}} + \mathbf{n}_g + \mathbf{e}_g \tag{5.3}$$

Similar to the majority of literature dealing with beamforming and interference management, our proposed algorithms require CSI knowledge in order to be implemented. While going into exact details is beyond the scope of this thesis, it is important to highlight the fact that all relevant channels can indeed be learned. Channels from users to BSs, from BSs to users, and between different BSs can be estimated using standard 3GPP LTE channel estimation protocols for HD systems. Channels between the users themselves can be learned via neighbour discovery methods applicable to device-to-device (D2D) communication, such as sounding reference signals (SRS) in 3GPP LTE. See for example [76, 82, 94] and references therein for further details on channel estimation.

5.2.2 Imperfect CSI considerations

Whilst perfect CSI formulations provide a useful baseline to highlight the advantages of FD over HD, it is important to recognise that the perfect CSI assumption is idealistic - in practice only an imperfect estimate is available. Therefore, moving on beyond the initial perfect CSI assumption, we also consider the design of robust beamformers. The channels are modelled as

$$\begin{aligned} \mathbf{H}_{k_g^d, i_j^u} &= \widehat{\mathbf{H}}_{k_g^d, i_j^u} + \Delta_{k_g^d, i_j^u} \\ \mathbf{H}_{k_g^d, j} &= \widehat{\mathbf{H}}_{k_g^d, j} + \Delta_{k_g^d, j} \\ \mathbf{H}_{g, i_j^u} &= \widehat{\mathbf{H}}_{g, i_j^u} + \Delta_{g, i_j^u} \\ \mathbf{H}_{g,j} &= \widehat{\mathbf{H}}_{g,j} + \Delta_{g,j} \end{aligned} \tag{5.4}$$

where $\mathbf{H}$ indicates the perfect channel, $\widehat{\mathbf{H}}$ is the imperfect channel and Δ is the CSI error. For the imperfect CSI case, only $\widehat{\mathbf{H}}_{g,g} \sum_{i=1}^{K_g^d} \mathbf{V}_{i_g^d} \mathbf{s}_{i_g^d}$ is known at BS g. This can be subtracted from $\mathbf{y}_g$ resulting in (5.3) with $\Theta = 1$, where there is an extra residual SI component compared to the perfect CSI case.

In this chapter the CSI error will be modelled in two different ways. First, we consider a norm-bounded error model, outlined in Section 5.2.2.1, which is more suited to cases where the CSI imperfection is due to quantisation errors; here we adopt an approach that can handle all errors as long their norm does not exceed a pre-established threshold, resulting in a more complex semi-definite programming (SDP) / determinant maximisation (Max-Det) based approach (see Section 5.4). Second, we consider a stochastic error model, outlined in Section 5.2.2.2, which is more suited to situations where the CSI error is due to estimation issues; this results in a simpler to implement closed form solution approach (see Section 5.5).

5.2.2.1 Norm-bounded error model

For the deterministic norm-bounded error model, the Frobenius norm of the CSI errors cannot exceed a pre-established upper bound, and the CSI error is expressed as

$$\begin{aligned}
\{\Delta_{k_g^d,i_j^u} : ||\Delta_{k_g^d,i_j^u}||_F \leq \varepsilon_{k_g^d,i_j^u}\} \ \ \forall\, k,g,i,j \\
\{\Delta_{k_g^d,j} : ||\Delta_{k_g^d,j}||_F \leq \varepsilon_{k_g^d,j}\} \ \ \forall\, k,g,j \\
\{\Delta_{g,i_j^u} : ||\Delta_{g,i_j^u}||_F \leq \varepsilon_{g,i_j^u}\} \ \ \forall\, g,i,j \\
\{\Delta_{g,j} : ||\Delta_{g,j}||_F \leq \varepsilon_{g,j}\} \ \ \forall\, g,j
\end{aligned} \tag{5.5}$$

where ε represents the upper limit on the Frobenius norm of the error. This model considers the case where the imperfect CSI is allowed to fall anywhere within an uncertainty region around the perfect CSI value, and is particularly suited to situations where quantisation errors dominate the imperfection in the available CSI. It is well established in literature and has been considered for beamformer design in a variety of systems, for example MIMO relay networks [95], MIMO IBCs [96], DL multi-user MIMO systems [97] and point-to-point MIMO communication [98, 99].

5.2.2.2 Stochastic error model

For the stochastic error model, the CSI errors are assumed to be independent of the perfect channel, $\mathbf{H}$, and distributed as

$$\begin{aligned}
\text{vec}(\Delta_{k_g^d,i_j^u}) &\sim \mathcal{CN}(0, \eta_{UU}\mathbf{I}) \\
\text{vec}(\Delta_{k_g^d,j}) &\sim \mathcal{CN}(0, \eta_{UB}\mathbf{I})
\end{aligned}$$

$$\begin{aligned}\text{vec}(\Delta_{g,i_j^u}) &\sim \mathcal{CN}(0, \eta_{BU}\mathbf{I}) \\ \text{vec}(\Delta_{g,j}) &\sim \mathcal{CN}(0, \eta_{BB}\mathbf{I})\end{aligned} \tag{5.6}$$

where η represents the variance of the CSI error, and the subscripts B and U indicate the BS and user respectively. This type of error model is suitable for cases where the channel error is mainly due to estimation inaccuracies. The parameter η can be assumed to be known *a priori* depending on the channel dynamics and the channel estimation scheme applied. It may be viewed either as a whole [68], or modelled as [67]

$$\eta_{rt} = \beta\rho_{rt}^{-\alpha}$$

where $r, t \in \{B, U\}$, ρ represents the SNR and parameters α and β capture a variety of CSI acquisition scenarios. More details are available in Section 4.2.2 for the corresponding parameters in (4.3).

5.2.3 Relationship between achievable rate and MSE

While we are ultimately interested in WSR maximisation, our approach is based on minimising the MSE, therefore we start by establishing a link between the achievable rate and the MSE covariance matrix. Starting with the DL users, the general DL MSE matrix is given by

$$\begin{aligned}\mathbf{E}_{k_g^d} &= \mathbb{E}\{(\mathbf{U}_{k_g^d}\mathbf{y}_{k_g^d} - \mathbf{s}_{k_g^d})(\mathbf{U}_{k_g^d}\mathbf{y}_{k_g^d} - \mathbf{s}_{k_g^d})^H\} \\ &= (\mathbf{I} - \mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d})(\mathbf{I} - \mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d})^H + \mathbf{U}_{k_g^d}\mathbf{\Phi}_{k_g^d}\mathbf{U}_{k_g^d}^H\end{aligned} \tag{5.7}$$

where the expectation is taken with respect to $\mathbf{s}$ and $\mathbf{n}$ under an independence assumption, and $\mathbf{U}_{k_g^d} \in \mathbb{C}^{b_d \times M_d}$ is the receiver applied by user k_g^d. Here, $\mathbf{\Phi}_{k_g^d}$ represents the DL interference-plus-noise covariance matrix, expressed as

$$\mathbf{\Phi}_{k_g^d} = \sum_{\substack{j=1 \\ (i,j\neq k,g)}}^{G}\sum_{i=1}^{K_j^d} \mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{k_g^d,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H + \mathbf{F}_{k_g^d} + \sigma_U^2\mathbf{I}$$

with $\mathbf{F}_{k_g^d}$, defined in (5.8), representing the combined contribution of the transmitter and receiver distortion. The approximation is obtained by omitting terms involving the multiplication of κ_B, κ_U and ι_U with each other since their product is negligibly small.

$$\mathbf{F}_{k_g^d} \approx \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \kappa_B \mathbf{H}_{k_g^d,j}\mathrm{diag}(\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H)\mathbf{H}_{k_g^d,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \kappa_U \mathbf{H}_{k_g^d,i_j^u}\mathrm{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\mathbf{H}_{k_g^d,i_j^u}^H$$
$$+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \iota_U \mathrm{diag}(\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{k_g^d,j}^H) + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \iota_U \mathrm{diag}(\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H) + \iota_U\sigma_U^2\mathbf{I} \tag{5.8}$$

Fixing all the precoders, the optimal DL receiver is an MMSE one, defined as

$$\bar{\mathbf{U}}_{k_g^d} = \arg\min_{\mathbf{U}} \mathrm{Tr}(\mathbf{E}_{k_g^d})$$
$$= \mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H\left[\sum_{j=1}^{G}\sum_{i=1}^{K_j^d}\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{k_g^d,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H + \mathbf{F}_{k_g^d} + \sigma_U^2\mathbf{I}\right]^{-1}. \tag{5.9}$$

Using (5.9) in (5.7) allows us to express the DL MSE matrix when using an MMSE receiver as

$$\begin{aligned}\bar{\mathbf{E}}_{k_g^d} &= \mathbf{I} - \mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H(\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H + \mathbf{\Phi}_{k_g^d})^{-1}\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d} \\ &= (\mathbf{I} + \mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H\mathbf{\Phi}_{k_g^d}^{-1}\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d})^{-1}. \end{aligned}\tag{5.10}$$

Under Gaussian signalling, the rate achieved by user k_g^d is given by

$$\begin{aligned}R_{k_g^d} &= \log_2\det\left(\mathbf{I} + \mathbf{\Phi}_{k_g^d}^{-1}\mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d}\right) \\ &\overset{(a)}{=} \log_2\det\left(\mathbf{I} + \mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H\mathbf{\Phi}_{k_g^d}^{-1}\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d}\right) \\ &\overset{(b)}{=} \log_2\det\left(\bar{\mathbf{E}}_{k_g^d}^{-1}\right). \end{aligned}\tag{5.11}$$

where (a) follows from Sylvester's determinant theorem and (b) comes from (5.10).

The rate achieved by user k_g^u is given by

$$R_{k_g^u} = \log_2\det\left(\mathbf{I} + \mathbf{\Phi}_{k_g^u}^{-1}\mathbf{V}_{k_g^u}^H\mathbf{H}_{g,k_g^u}^H\mathbf{H}_{g,k_g^u}\mathbf{V}_{k_g^u}\right)$$

where $\mathbf{\Phi}_{k_g^u}$ is the UL interference-plus-noise covariance matrix, expressed as

$$\mathbf{\Phi}_{k_g^u} = \sum_{\substack{j=1\\(j\neq g)}}^{G}\sum_{i=1}^{K_j^d}\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H + \sum_{\substack{j=1\\(i,j\neq k,g)}}^{G}\sum_{i=1}^{K_j^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H + \mathbf{F}_g + \sigma_B^2\mathbf{I}$$

with $\mathbf{F}_g$, defined in (5.12), representing the effect of transmitter and receiver distortion.

$$\mathbf{F}_g \approx \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \kappa_B \mathbf{H}_{g,j}\mathrm{diag}(\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H)\mathbf{H}_{g,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \kappa_U \mathbf{H}_{g,i_j^u}\mathrm{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\mathbf{H}_{g,i_j^u}^H$$
$$+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \iota_B \mathrm{diag}(\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H) + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \iota_B \mathrm{diag}(\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H) + \iota_B\sigma_B^2\mathbf{I} \tag{5.12}$$

Using a process similar to the one outlined for the DL, we can relate the UL rate to the UL MSE covariance matrix as

$$R_{k_g^u} = \log_2 \det\left(\bar{\mathbf{E}}_{k_g^u}^{-1}\right) \tag{5.13}$$

where $\bar{\mathbf{E}}_{k_g^u} = \mathbb{E}\{(\bar{\mathbf{U}}_{k_g^u}\tilde{\mathbf{y}}_{k_g^u} - \mathbf{s}_{k_g^u})(\bar{\mathbf{U}}_{k_g^u}\tilde{\mathbf{y}}_{k_g^u} - \mathbf{s}_{k_g^u})^H\}$ is the UL MSE matrix when using the optimal MMSE receiver given by

$$\bar{\mathbf{U}}_{k_g^u} = \arg\min_{\mathbf{U}} \mathrm{Tr}(\mathbf{E}_{k_g^u})$$
$$= \mathbf{V}_{k_g^u}^H\mathbf{H}_{g,k_g^u}^H\left[\sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_j^d}\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H + \mathbf{F}_g + \sigma_B^2\mathbf{I}\right]^{-1}. \tag{5.14}$$

5.3 Weighted sum rate maximisation

Starting with the perfect CSI case, we want to find the optimal precoders that maximise the WSR subject to transmit power constraints, i.e.

$$\begin{aligned}\max_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\mu_{k_g^d}R_{k_g^d} + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\mu_{k_g^u}R_{k_g^u} \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_U \;\; \forall\, k,g \\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_B \;\; \forall\, g\,. \end{aligned} \tag{5.15}$$

where $\mu_{k_g^d}$ and $\mu_{k_g^u}$ $\forall\, k,g$ denote pre-defined weights.

Theorem 5.1. *The WSR problem in (5.15) is equivalent to the WMMSE problem in (5.16), such that the global optimal solution for the precoders of the two problems are identical.*

$$\begin{aligned}\min_{\mathbf{U},\mathbf{W},\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\left[\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) - \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) - \frac{\mu_{k_g^d}}{\ln 2}b_d\right] \\ & + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\left[\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) - \mu_{k_g^u}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^u}}\mathbf{W}_{k_g^u}\right) - \frac{\mu_{k_g^u}}{\ln 2}b_u\right] \\ s.t. \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \ \ \forall\, k,g \\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \ \ \forall\, g \end{aligned} \tag{5.16}$$

Proof. First we define the metrics $\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d})$ and $\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u})$ used in (5.16) as in (5.17) and (5.18). Here, $\mathbf{B}_{k_g^d}$ comes from the decomposition of $\mathbf{W}_{k_g^d}$ as $\mathbf{W}_{k_g^d} = \mathbf{B}_{k_g^d}^H\mathbf{B}_{k_g^d}$, and $\mathbf{B}_{k_g^u}$ comes from the decomposition of $\mathbf{W}_{k_g^u}$ as $\mathbf{W}_{k_g^u} = \mathbf{B}_{k_g^u}^H\mathbf{B}_{k_g^u}$. (Note that while the weight decomposition is not exploited in this theorem, we introduce the notation here for conformity, since it is necessary at a number of points later in this chapter.)

$$\begin{aligned}\mathrm{Tr}\left(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}\right) = \mathrm{Tr}\Big(&\mathbf{B}_{k_g^d}(\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d} - \mathbf{I})(\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d} - \mathbf{I})^H\mathbf{B}_{k_g^d}^H\Big) \\ & + \sum_{j=1}^{G}\sum_{i=1}^{K_j^d}\mathrm{Tr}\left(\kappa_B\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,j}\mathrm{diag}(\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H)\,\mathbf{H}_{k_g^d,j}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\ & + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\mathrm{Tr}\left(\kappa_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u}\,\mathrm{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\ & + \sum_{j=1}^{G}\sum_{i=1}^{K_j^d}\mathrm{Tr}\left(\iota_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathrm{diag}(\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{k_g^d,j}^H)\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\ & + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\mathrm{Tr}\left(\iota_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathrm{diag}(\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H)\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\ & + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u}\mathrm{Tr}\left(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\ & + \sum_{\substack{j=1\\(i,j\neq k,g)}}^{G}\sum_{i=1}^{K_j^d}\mathrm{Tr}\left(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{k_g^d,j}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\ & + (\sigma_U^2 + \iota_U\sigma_U^2)\mathrm{Tr}\left(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \end{aligned} \tag{5.17}$$

$$
\begin{aligned}
\mathrm{Tr}\left(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}\right) = \mathrm{Tr}\Big(&\mathbf{B}_{k_g^u}(\mathbf{U}_{k_g^u}\mathbf{H}_{g,k_g^u}\mathbf{V}_{k_g^u} - \mathbf{I})(\mathbf{U}_{k_g^u}\mathbf{H}_{g,k_g^u}\mathbf{V}_{k_g^u} - \mathbf{I})\mathbf{B}_{k_g^u}^H\Big) \\
&+ \sum_{\substack{j=1 \\ (i,j\neq k,g)}}^{G}\sum_{i=1}^{K_j^u} \mathrm{Tr}\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ \sum_{\substack{j=1 \\ (j\neq g)}}^{G}\sum_{i=1}^{K_j^d} \mathrm{Tr}\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \mathrm{Tr}\left(\kappa_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j}\mathrm{diag}(\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H)\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \mathrm{Tr}\left(\kappa_U\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathrm{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\mathbf{H}_{g,i_j^u}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \mathrm{Tr}\left(\iota_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathrm{diag}(\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H)\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \mathrm{Tr}\left(\iota_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathrm{diag}(\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H)\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ (\sigma_B^2 + \iota_B\sigma_B^2)\mathrm{Tr}\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
&+ \Theta\sum_{i=1}^{K_j^d} \mathrm{Tr}\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\Delta_{g,g}\mathbf{V}_{i_g^d}\mathbf{V}_{i_g^d}^H\Delta_{g,g}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \qquad (5.18)
\end{aligned}
$$

Considering (5.16), it can be seen that the optimal $\mathbf{U}$ are the standard MMSE receivers $\bar{\mathbf{U}}_{k_g^d}$ and $\bar{\mathbf{U}}_{k_g^u}$ in (5.9) and (5.14) respectively. Next, fixing $\mathbf{U}$ and $\mathbf{V}$ we can derive expressions for the weights. Ignoring all parts that do not contain $\mathbf{W}$, since these have no effect on the resultant weights, the Lagrangian of (5.16) can be written as

$$
\begin{aligned}
\mathcal{L}_{\mathcal{W}} = &\sum_{g=1}^{G}\sum_{k=1}^{K_g^d} \mathrm{Tr}\left(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}\right) - \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) \\
&+ \sum_{g=1}^{G}\sum_{k=1}^{K_g^u} \mathrm{Tr}\left(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}\right) - \mu_{k_g^u}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^u}}\mathbf{W}_{k_g^u}\right).
\end{aligned}
$$

Taking partial derivatives of $\mathcal{L}_{\mathcal{W}}$ and setting them to zero, we obtain the optimal weights as

$$
\bar{\mathbf{W}}_{k_g^d} = \frac{\mu_{k_g^d}}{\ln 2}\bar{\mathbf{E}}_{k_g^d}^{-1} \quad \text{and} \quad \bar{\mathbf{W}}_{k_g^u} = \frac{\mu_{k_g^u}}{\ln 2}\bar{\mathbf{E}}_{k_g^u}^{-1}. \qquad (5.19)
$$

Finally, substituting for optimal $\mathbf{U}$ and $\mathbf{W}$ in (5.16), results in

$$\begin{aligned}\min_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} -\mu_{k_g^d}\log_2\det\left(\bar{\mathbf{E}}_{k_g^d}^{-1}\right) + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u} -\mu_{k_g^u}\log_2\det\left(\bar{\mathbf{E}}_{k_g^u}^{-1}\right) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \;\; \forall\, k, g \\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \;\; \forall\, g\end{aligned}$$

which considering (5.11) and (5.13) is the same as the original problem in (5.15). □

Since (5.16) is only separately convex in $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$, we apply an alternating minimisation approach to solve the problem as outlined in Algorithm 5.1.

Algorithm 5.1: Alternating optimisation process for WMMSE problems

1. Initialise $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$ $\forall\, k, g$.
2. Calculate $\mathbf{U}_{k_g^d}$ and $\mathbf{U}_{k_g^u}$ $\forall\, k, g$.
3. Calculate $\mathbf{W}_{k_g^d}$ and $\mathbf{W}_{k_g^u}$ $\forall\, k, g$.
4. Compute $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$ $\forall\, k, g$.
5. Repeat from Step 2 until convergence or for a fixed number of iterates.

Having closed form expressions for $\mathbf{U}$ and $\mathbf{W}$, we need to focus on obtaining $\mathbf{V}$. Fixing $\mathbf{U}$ and $\mathbf{W}$, (5.16) can be expressed as

$$\begin{aligned}\min_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \;\; \forall\, k, g \\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \;\; \forall\, g\,. \end{aligned} \tag{5.20}$$

With respect to $\mathbf{V}$, the Lagrange dual objective function of (5.20) is given by

$$\begin{aligned}\mathcal{L}_{\mathcal{V}} = & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) \\ & + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\varpi_{k_g^u}\Big[\mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) - P_U\Big] + \sum_{g=1}^{G}\varrho_g\Big[\sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) - P_B\Big]\end{aligned}$$

where $\varpi_{k_g^u}$ and ϱ_g are the Lagrange multipliers associated with the transmit power constraints. Taking the partial derivatives of $\mathcal{L}_{\mathcal{V}}$ and setting them to zero, we obtain closed form solutions for the optimal precoders as

$$\begin{aligned}\bar{\mathbf{V}}_{k_g^d} &= [\mathbf{X}_g + \varrho_g \mathbf{I}]^{-1} \mathbf{H}_{k_g^d,g}^H \mathbf{U}_{k_g^d}^H \mathbf{W}_{k_g^d} \\ \bar{\mathbf{V}}_{k_g^u} &= [\mathbf{X}_{k_g^u} + \varpi_{k_g^u} \mathbf{I}]^{-1} \mathbf{H}_{g,k_g^u}^H \mathbf{U}_{k_g^u}^H \mathbf{W}_{k_g^u}\end{aligned} \tag{5.21}$$

where $\mathbf{X}_g$ and $\mathbf{X}_{k_g^u}$ are defined in (5.22) and (5.23) respectively.

$$\begin{aligned}\mathbf{X}_g = &\sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \mathbf{H}_{i_j^d,g}^H \mathbf{U}_{i_j^d}^H \mathbf{W}_{i_j^d} \mathbf{U}_{i_j^d} \mathbf{H}_{i_j^d,g} + \sum_{\substack{j=1 \\ j\neq g}}^{G}\sum_{i=1}^{K_j^u} \mathbf{H}_{j,g}^H \mathbf{U}_{i_j^u}^H \mathbf{W}_{i_j^u} \mathbf{U}_{i_j^u} \mathbf{H}_{j,g} \\ &+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \kappa_{BS} \mathrm{diag}(\mathbf{H}_{i_j^d,g}^H \mathbf{U}_{i_j^d}^H \mathbf{W}_{i_j^d} \mathbf{U}_{i_j^d} \mathbf{H}_{i_j^d,g}) + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \kappa_{BS} \mathrm{diag}(\mathbf{H}_{j,g}^H \mathbf{U}_{i_j^u}^H \mathbf{W}_{i_j^u} \mathbf{U}_{i_j^u} \mathbf{H}_{j,g}) \\ &+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \iota_{us} \mathbf{H}_{i_j^d,g}^H \mathrm{diag}(\mathbf{U}_{i_j^d}^H \mathbf{W}_{i_j^d} \mathbf{U}_{i_j^d}) \mathbf{H}_{i_j^d,g} + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \iota_{BS} \mathbf{H}_{j,g}^H \mathrm{diag}(\mathbf{U}_{i_j^u}^H \mathbf{W}_{i_j^u} \mathbf{U}_{i_j^u}) \mathbf{H}_{j,g}\end{aligned} \tag{5.22}$$

$$\begin{aligned}\mathbf{X}_{k_g^u} = &\sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \mathbf{H}_{i_j^d,k_g^u}^H \mathbf{U}_{i_j^d}^H \mathbf{W}_{i_j^d} \mathbf{U}_{i_j^d} \mathbf{H}_{i_j^d,k_g^u} + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \mathbf{H}_{j,k_g^u}^H \mathbf{U}_{i_j^u}^H \mathbf{W}_{i_j^u} \mathbf{U}_{i_j^u} \mathbf{H}_{j,k_g^u} \\ &+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \kappa_{us} \mathrm{diag}(\mathbf{H}_{i_j^d,k_g^u}^H \mathbf{U}_{i_j^d}^H \mathbf{W}_{i_j^d} \mathbf{U}_{i_j^d} \mathbf{H}_{i_j^d,k_g^u}) + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \kappa_{us} \mathrm{diag}(\mathbf{H}_{j,k_g^u}^H \mathbf{U}_{i_j^u}^H \mathbf{W}_{i_j^u} \mathbf{U}_{i_j^u} \mathbf{H}_{j,k_g^u}) \\ &+ \sum_{j=1}^{G}\sum_{i=1}^{K_j^d} \iota_{us} \mathbf{H}_{i_j^d,k_g^u}^H \mathrm{diag}(\mathbf{U}_{i_j^d}^H \mathbf{W}_{i_j^d} \mathbf{U}_{i_j^d}) \mathbf{H}_{i_j^d,k_g^u} + \sum_{j=1}^{G}\sum_{i=1}^{K_j^u} \iota_{BS} \mathbf{H}_{j,k_g^u}^H \mathrm{diag}(\mathbf{U}_{i_j^u}^H \mathbf{W}_{i_j^u} \mathbf{U}_{i_j^u}) \mathbf{H}_{j,k_g^u}\end{aligned} \tag{5.23}$$

The Lagrange multipliers $\varpi_{k_g^u}$ and ϱ_g should be either zero, or positive numbers that satisfy the following equations

$$\begin{aligned}&\mathrm{Tr}\left(\mathbf{V}_{k_g^u}(\varpi_{k_g^u}) \mathbf{V}_{k_g^u}^H(\varpi_{k_g^u})\right) = P_U \quad \forall\, k, g \\ &\sum_{k=1}^{K_g^d} \mathrm{Tr}\left(\mathbf{V}_{k_g^d}(\varrho_g) \mathbf{V}_{k_g^d}^H(\varrho_g)\right) = P_B \quad \forall\, g\,.\end{aligned} \tag{5.24}$$

The equalities in (5.24) can alternatively be expressed as

$$\sum_{m=1}^{M_u} \frac{[\mathbf{G}_{k_g^u}]_m}{([\mathbf{D}_{k_g^u}]_m + \varpi_{k_g^u})^2} = P_U \ \ \forall\, k, g$$
$$\sum_{m=1}^{M_B} \frac{[\mathbf{G}_g]_m}{([\mathbf{D}_g]_m + \varrho_g)^2} = P_B \ \ \forall\, g$$

where $\mathbf{D}_{k_g^u}$ comes from the decomposition $\mathbf{X}_{k_g^u} = \mathbf{C}_{k_g^u}\mathbf{D}_{k_g^u}\mathbf{C}_{k_g^u}^H$, $\mathbf{G}_{k_g^u} = \mathbf{C}_{k_g^u}^H\mathbf{H}_{g,k_g^u}^H\mathbf{U}_{k_g^u}^H \mathbf{W}_{k_g^u}\mathbf{W}_{k_g^u}^H\ \mathbf{U}_{k_g^u}\mathbf{H}_{g,k_g^u}\mathbf{C}_{k_g^u}$, $\mathbf{D}_g$ comes from the decomposition $\mathbf{X}_g = \mathbf{C}_g\mathbf{D}_g\mathbf{C}_g^H$ and $\mathbf{G}_g = \sum_{k=1}^{K_g^d} \mathbf{C}_g^H\mathbf{H}_{k_g^d,g}^H\mathbf{U}_{k_g^d}^H\ \mathbf{W}_{k_g^d}\mathbf{W}_{k_g^d}^H\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,g}\mathbf{C}_g$. These can be respectively solved for $\nu_{k_g^u}$ and μ_g using linear search techniques, such as the bisection method [92].

Therefore to solve (5.16), we can follow the process in Algorithm 5.1, where in Step 2 we use (5.9) and (5.14) to calculate the receivers as $\mathbf{U}_{k_g^d} = \bar{\mathbf{U}}_{k_g^d}$ and $\mathbf{U}_{k_g^u} = \bar{\mathbf{U}}_{k_g^u}$. The weights in Step 3 are calculated as $\mathbf{W}_{k_g^d} = \bar{\mathbf{W}}_{k_g^d}$ and $\mathbf{W}_{k_g^u} = \bar{\mathbf{W}}_{k_g^u}$ using (5.19). Finally, (5.21) is used to calculate $\mathbf{V}_{k_g^d} = \bar{\mathbf{V}}_{k_g^d}$ and $\mathbf{V}_{k_g^u} = \bar{\mathbf{V}}_{k_g^u}$ in Step 4.

Remark 5.1. *The alternating minimisation process used to solve the WMMSE problem decreases the cost function monotonically with each step. Since the cost function is lower bounded, then the algorithm is guaranteed to converge. Additionally, using an argument parallel to the one in [92, Appendix C], convergence to a stationary point of the original WSR problem can also be proven.*

5.4 Robust design with norm-bounded error model

Next, we want to solve the WSR problem from the prior section with additional considerations for norm-bounded CSI errors, i.e.

$$\begin{aligned}
\max_{\mathbf{V}} \min_{\Delta} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} \mu_{k_g^d} R_{k_g^d} + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u} \mu_{k_g^u} R_{k_g^u} \\
\text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \ \ \forall\, k, g \\
& \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \ \ \forall\, g \\
& \{\Delta_{k_g^d,i_j^u} : \|\Delta_{k_g^d,i_j^u}\|_F \le \varepsilon_{k_g^d,i_j^u}\} \ \ \forall\, k, g, i, j \\
& \{\Delta_{k_g^d,j} : \|\Delta_{k_g^d,j}\|_F \le \varepsilon_{k_g^d,j}\} \ \ \forall\, k, g, j
\end{aligned}$$

$$\{\Delta_{g,i_j^u} : \|\Delta_{g,i_j^u}\|_F \leq \varepsilon_{g,i_j^u}\} \ \forall\, g,i,j$$
$$\{\Delta_{g,j} : \|\Delta_{g,j}\|_F \leq \varepsilon_{g,j}\} \ \forall\, g,j\,. \tag{5.25}$$

We apply an iterative approach to solve our non-convex optimisation problem, this involves solving a convex sub-problem at each iteration step and has been proven to converge [96, 97]. Having already established an equivalence between (5.15) and (5.16) for the perfect CSI case, it directly follows that the cost function of (5.25) can be mapped to

$$\max_{\mathbf{V}} \min_{\Delta} \max_{\mathbf{U},\mathbf{W}} \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\left[-\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) + \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) + \frac{\mu_{k_g^d}}{\ln 2}b_d\right] + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\left[-\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) + \mu_{k_g^u}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^u}}\mathbf{W}_{k_g^u}\right) + \frac{\mu_{k_g^u}}{\ln 2}b_u\right]. \tag{5.26}$$

By applying the max-min inequality, which states that for any function $f(w,z)$ then $\min_w \max_z f(w,z) \geq \max_z \min_w f(w,z)$, rather than using the cost function in (5.26), we can instead focus on solving the following problem

$$\begin{aligned}
\max_{\mathbf{V},\mathbf{U},\mathbf{W}} \min_{\Delta} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\left[-\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) + \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) + \frac{\mu_{k_g^d}}{\ln 2}b_d\right] \\
& + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\left[-\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) + \mu_{k_g^u}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^u}}\mathbf{W}_{k_g^u}\right) + \frac{\mu_{k_g^u}}{\ln 2}b_u\right] \\
\text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_U \ \forall\, k,g \\
& \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_B \ \forall\, g \\
& \{\Delta_{k_g^d,i_j^u} : \|\Delta_{k_g^d,i_j^u}\|_F \leq \varepsilon_{k_g^d,i_j^u}\} \ \forall\, k,g,i,j \\
& \{\Delta_{k_g^d,j} : \|\Delta_{k_g^d,j}\|_F \leq \varepsilon_{k_g^d,j}\} \ \forall\, k,g,j \\
& \{\Delta_{g,i_j^u} : \|\Delta_{g,i_j^u}\|_F \leq \varepsilon_{g,i_j^u}\} \ \forall\, g,i,j \\
& \{\Delta_{g,j} : \|\Delta_{g,j}\|_F \leq \varepsilon_{g,j}\} \ \forall\, g,j\,.
\end{aligned} \tag{5.27}$$

The cost function of (5.27) is not equivalent to the original one in (5.26), however the ensuing formulation is still a valid one. Firstly, the new cost function

is a lower bound on (5.26), implying that the resultant rate is surely achievable. Secondly, the formulation in (5.27) ensures that none of the optimisation variables depend on perfect CSI, which is the ultimate aim of a robust beamforming approach.

Theorem 5.2. *The optimisation problem in (5.27) is equivalent to the reformulation in (5.28) such that the optimal* $\mathbf{U}$*,* $\mathbf{V}$ *and* $\mathbf{W} = \mathbf{B}^H\mathbf{B}$ *for the two problems are identical.*

$$
\begin{aligned}
\max_{\mathbf{V},\mathbf{U},\mathbf{B},m,\nu} \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} \Bigg[& -\sum_{j=1}^{G}\sum_{i=1}^{K_j^u} m_{1,k_g i_j} - \sum_{j=1}^{G} m_{2,k_g j} - (\sigma_U^2 + \iota_U\sigma_U^2)||\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}||_F^2 \\
& + \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{B}_{k_g^d}^H\mathbf{B}_{k_g^d}\right) + \frac{\mu_{k_g^d}}{\ln 2}b_d\Bigg] \\
+ \sum_{g=1}^{G}\Bigg[& -\sum_{j=1}^{G}\sum_{i=1}^{K_j^u} m_{3,gi_j} - \sum_{j=1}^{G} m_{4,gj} + \sum_{k=1}^{K_g^u}\Bigg(-(\sigma_B^2 + \iota_B\sigma_B^2)||\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}||_F^2 \\
& + \mu_{k_g^u}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^u}}\mathbf{B}_{k_g^u}^H\mathbf{B}_{k_g^u}\right) + \frac{\mu_{k_g^u}}{\ln 2}b_u\Bigg)\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
s.t.\quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \;\forall\; k,g \\
& \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \;\forall\; g \\
& \begin{bmatrix} m_{1,k_g i_j} - \nu_{1,k_g i_j} & \boldsymbol{\omega}_{1,k_g i_j}^H & \mathbf{0} \\ \boldsymbol{\omega}_{1,k_g i_j} & \mathbf{I} & -\varepsilon_{k_g^d,i_j^u}\boldsymbol{\Omega}_{1,k_g i_j} \\ \mathbf{0} & -\varepsilon_{k_g^d,i_j^u}\boldsymbol{\Omega}_{1,k_g i_j}^H & \nu_{1,k_g i_j}\mathbf{I} \end{bmatrix} \succeq \mathbf{0} \;\forall\; k,g,i,j \\
& \begin{bmatrix} m_{2,k_g j} - \nu_{2,k_g j} & \boldsymbol{\omega}_{2,k_g j}^H & \mathbf{0} \\ \boldsymbol{\omega}_{2,k_g j} & \mathbf{I} & -\varepsilon_{k_g^d,j}\boldsymbol{\Omega}_{2,k_g j} \\ \mathbf{0} & -\varepsilon_{k_g^d,j}\boldsymbol{\Omega}_{2,k_g j}^H & \nu_{2,k_g j}\mathbf{I} \end{bmatrix} \succeq \mathbf{0} \;\forall\; k,g,j \\
& \begin{bmatrix} m_{3,gi_j} - \nu_{3,gi_j} & \boldsymbol{\omega}_{3,gi_j}^H & \mathbf{0} \\ \boldsymbol{\omega}_{3,gi_j} & \mathbf{I} & -\varepsilon_{g,i_j^u}\boldsymbol{\Omega}_{3,gi_j} \\ \mathbf{0} & -\varepsilon_{g,i_j^u}\boldsymbol{\Omega}_{3,gi_j}^H & \nu_{3,gi_j}\mathbf{I} \end{bmatrix} \succeq \mathbf{0} \;\forall\; g,i,j \\
& \begin{bmatrix} m_{4,gj} - \nu_{4,gj} & \boldsymbol{\omega}_{4,gj}^H & \mathbf{0} \\ \boldsymbol{\omega}_{4,gj} & \mathbf{I} & -\varepsilon_{g,j}\boldsymbol{\Omega}_{4,gj} \\ \mathbf{0} & -\varepsilon_{g,j}\boldsymbol{\Omega}_{4,gj}^H & \nu_{4,gj}\mathbf{I} \end{bmatrix} \succeq \mathbf{0} \;\forall\; g,j \\
& \nu_{1,k_g i_j} \ge 0, \quad \nu_{2,k_g j} \ge 0, \quad \nu_{3,gi_j} \ge 0, \quad \nu_{4,gj} \ge 0 \quad \forall\; k,g,i,j
\end{aligned} \tag{5.28}
$$

In (5.28), m and ν represent additional scalar variables introduced during the reformulation, and the $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ terms are defined as

$$\boldsymbol{\omega}_{1k_g,i_j} = \begin{bmatrix} \mathrm{vec}(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\hat{\mathbf{H}}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}) \\ (\kappa_U)^{\frac{1}{2}}\left\lfloor \left((\mathbf{S}_n\mathbf{V}_{i_j^u})^T \otimes (\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d})\right)\mathrm{vec}(\hat{\mathbf{H}}_{k_g^d,i_j^u})\right\rfloor_{n=1\ldots M_u} \\ (\iota_U)^{\frac{1}{2}}\left\lfloor \left(\mathbf{V}_{i_j^u}^T \otimes \left((\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H)^T\mathbf{S}_n\right)\right)\mathrm{vec}(\hat{\mathbf{H}}_{k_g^d,i_j^u})\right\rfloor_{n=1\ldots M_d} \end{bmatrix}$$

$$\boldsymbol{\omega}_{2,k_gj} = \begin{bmatrix} \left\lfloor \mathrm{vec}(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\hat{\mathbf{H}}_{k_g^d,j}\mathbf{V}_{i_j^d} - \delta_{i,j}^{k,g}\mathbf{B}_{k_g^d})\right\rfloor_{i=1\ldots K_j^d} \\ \left\lfloor (\kappa_B)^{\frac{1}{2}}\left\lfloor \left((\mathbf{S}_n\mathbf{V}_{i_j^d})^T \otimes \mathrm{vec}(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d})\right)\mathrm{vec}(\hat{\mathbf{H}}_{k_g^d,j})\right\rfloor_{n=1\ldots M_B}\right\rfloor_{i=1\ldots K_j^d} \\ \left\lfloor (\iota_U)^{\frac{1}{2}}\left\lfloor \left(\mathbf{V}_{i_j^d}^T \otimes \left((\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H)^T\mathbf{S}_n\right)\right)\mathrm{vec}(\hat{\mathbf{H}}_{k_g^d,j})\right\rfloor_{n=1\ldots M_d}\right\rfloor_{i=1\ldots K_j^d} \end{bmatrix}$$

$$\boldsymbol{\omega}_{3,gi_j} = \begin{bmatrix} \left\lfloor \mathrm{vec}(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\hat{\mathbf{H}}_{g,i_j^u}\mathbf{V}_{i_j^u} - \delta_{i,j}^{k,g}\mathbf{B}_{k_g^d})\right\rfloor_{k=1\ldots K_g^u} \\ \left\lfloor (\kappa_U)^{\frac{1}{2}}\left\lfloor \left((\mathbf{S}_n\mathbf{V}_{i_j^u})^T \otimes (\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u})\right)\mathrm{vec}(\hat{\mathbf{H}}_{g,i_j^u})\right\rfloor_{n=1\ldots M_u}\right\rfloor_{k=1\ldots K_g^u} \\ \left\lfloor (\iota_B)^{\frac{1}{2}}\left\lfloor \left(\mathbf{V}_{i_j^u}^T \otimes \left((\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H)^T\mathbf{S}_n\right)\right)\mathrm{vec}(\hat{\mathbf{H}}_{g,i_j^u})\right\rfloor_{n=1\ldots M_B}\right\rfloor_{k=1\ldots K_g^u} \end{bmatrix}$$

$$\boldsymbol{\omega}_{4,gj} = \begin{bmatrix} \left\lfloor \vartheta_j^g\mathrm{vec}(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\hat{\mathbf{H}}_{g,j}\mathbf{V}_{i_j^d})\right\rfloor_{\substack{k=1\ldots K_g^u \\ i=1\ldots K_j^d}} \\ \left\lfloor (\kappa_B)^{\frac{1}{2}}\left\lfloor \left((\mathbf{S}_n\mathbf{V}_{i_j^d})^T \otimes (\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u})\right)\mathrm{vec}(\hat{\mathbf{H}}_{g,j})\right\rfloor_{n=1\ldots M_B}\right\rfloor_{\substack{k=1\ldots K_g^u \\ i=1\ldots K_j^d}} \\ \left\lfloor (\iota_B)^{\frac{1}{2}}\left\lfloor \left(\mathbf{V}_{i_j^d}^T \otimes \left((\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H)^T\mathbf{S}_n\right)\right)\mathrm{vec}(\hat{\mathbf{H}}_{g,j})\right\rfloor_{n=1\ldots M_B}\right\rfloor_{\substack{k=1\ldots K_g^u \\ i=1\ldots K_j^d}} \end{bmatrix}$$

$$\boldsymbol{\Omega}_{1k_g,i_j} = \begin{bmatrix} \left(\mathbf{V}_{i_j^u}^T \otimes \mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\right) \\ (\kappa_U)^{\frac{1}{2}}\left\lfloor (\mathbf{S}_n\mathbf{V}_{i_j^u})^T \otimes (\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d})\right\rfloor_{n=1\ldots M_u} \\ (\iota_U)^{\frac{1}{2}}\left\lfloor \mathbf{V}_{i_j^u}^T \otimes \left((\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H)^T\mathbf{S}_n\right)\right\rfloor_{n=1\ldots M_d} \end{bmatrix}$$

$$\mathbf{\Omega}_{2,k_g j} = \begin{bmatrix} \Big\lfloor \left(\mathbf{V}_{i_j^d}^T \otimes \mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\right) \Big\rfloor_{i=1\ldots K_j^d} \\ \left\lfloor (\kappa_B)^{\frac{1}{2}} \left\lfloor (\mathbf{S}_n \mathbf{V}_{i_j^d})^T \otimes \mathrm{vec}(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}) \right\rfloor_{n=1\ldots M_B} \right\rfloor_{i=1\ldots K_j^d} \\ \left\lfloor (\iota_U)^{\frac{1}{2}} \left\lfloor \mathbf{V}_{i_j^d}^T \otimes \left((\mathbf{U}_{k_g^d}^H \mathbf{B}_{k_g^d}^H)^T \mathbf{S}_n\right) \right\rfloor_{n=1\ldots M_d} \right\rfloor_{i=1\ldots K_j^d} \end{bmatrix}$$

$$\mathbf{\Omega}_{3,g i_j} = \begin{bmatrix} \lfloor \left(\mathbf{V}_{i_j^u}^T \otimes \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\right) \rfloor_{k=1\ldots K_g^u} \\ \left\lfloor (\kappa_U)^{\frac{1}{2}} \lfloor (\mathbf{S}_n \mathbf{V}_{i_j^u})^T \otimes (\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}) \rfloor_{n=1\ldots M_u} \right\rfloor_{k=1\ldots K_g^u} \\ \left\lfloor (\iota_B)^{\frac{1}{2}} \left\lfloor \mathbf{V}_{i_j^u}^T \otimes \left((\mathbf{U}_{k_g^u}^H \mathbf{B}_{k_g^u}^H)^T \mathbf{S}_n\right) \right\rfloor_{n=1\ldots M_B} \right\rfloor_{k=1\ldots K_g^u} \end{bmatrix}$$

$$\mathbf{\Omega}_{4,gj} = \begin{bmatrix} \Big\lfloor \left(\mathbf{V}_{i_j^d}^T \otimes \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\right) \Big\rfloor_{\substack{k=1\ldots K_g^u \\ i=1\ldots K_j^d}} \\ \left\lfloor (\kappa_B)^{\frac{1}{2}} \left\lfloor (\mathbf{S}_n \mathbf{V}_{i_j^d})^T \otimes (\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}) \right\rfloor_{n=1\ldots M_B} \right\rfloor_{\substack{k=1\ldots K_g^u \\ i=1\ldots K_j^d}} \\ \left\lfloor (\iota_B)^{\frac{1}{2}} \left\lfloor \mathbf{V}_{i_j^d}^T \otimes \left((\mathbf{U}_{k_g^u}^H \mathbf{B}_{k_g^u}^H)^T \mathbf{S}_n\right) \right\rfloor_{n=1\ldots M_B} \right\rfloor_{\substack{k=1\ldots K_g^u \\ i=1\ldots K_j^d}} \end{bmatrix}$$

where $\mathbf{S}_n$ *is a selection matrix consisting of all zeros except for the* n*th element along the diagonal which is equal to* 1*,*

$$\delta_{i,j}^{k,g} = \begin{cases} 1 & \text{if } (k,g) = (i,j) \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \vartheta_j^g = \begin{cases} 0 & \text{if } g = j \\ 1 & \text{otherwise} \end{cases}.$$

Proof. The problem formulation in (5.28) is based on finding an equivalent form for the inner maximisation of (5.27). Note that $\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d})$ and $\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u})$ are given by (5.17) and (5.18) where $\Theta = 1$ since we are dealing with imperfect CSI. Also the CSI error, Δ, appears in these terms when we replace $\mathbf{H}$ with $\widehat{\mathbf{H}} + \Delta$ from (5.4).

Next, it can be noticed that the problem is separable over each occurrence of the different types of CSI error [96]. Therefore, we can separate the problem over $\Delta_{k_g^d,i_j^u}$, $\Delta_{k_g^d,j}$, Δ_{g,i_j^u} and $\Delta_{g,j}$, and focus on one of them at a time to obtain a more useful formulation. Starting with $\Delta_{k_g^d,i_j^u}$, this only appears in terms containing $\mathbf{H}_{k_g^d,i_j^u}$, since $\mathbf{H}_{k_g^d,i_j^u} = \widehat{\mathbf{H}}_{k_g^d,i_j^u} + \Delta_{k_g^d,i_j^u}$. Thus, from the overall cost function of (5.27), from the perspective of each $\Delta_{k_g^d,i_j^u}$, we are only concerned with

$$
\begin{aligned}
T_{1,k_g i_j} &= \mathrm{Tr}\left(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\
&\quad + \mathrm{Tr}\left(\kappa_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u}\,\mathrm{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\
&\quad + \mathrm{Tr}\left(\iota_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathrm{diag}(\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H)\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\
&= \mathrm{Tr}\left(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\
&\quad + \sum_{n=1}^{M_u}\mathrm{Tr}\left(\kappa_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u}\mathbf{S}_n\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{S}_n^H\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right) \\
&\quad + \sum_{n=1}^{M_d}\mathrm{Tr}\left(\iota_U\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{S}_n\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{S}_n^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\right).
\end{aligned}
$$

Using, $\mathrm{Tr}(\mathbf{X}\mathbf{X}^H) = \|\mathrm{vec}(\mathbf{X})\|^2$ and $\mathrm{vec}(\mathbf{X}\mathbf{Y}\mathbf{Z}) = (\mathbf{Z}^T \otimes \mathbf{X})\mathrm{vec}(\mathbf{Y})$, and introducing slack variable $m_{1,k_g i_j}$, this can be expressed as

$$
T_{1,k_g i_j} = \|\boldsymbol{\omega}_{1,k_g i_j} + \boldsymbol{\Omega}_{1,k_g i_j}\mathrm{vec}(\Delta_{k_g^d,i_j^u})\|^2 \leq m_{1,k_g i_j}. \tag{5.29}
$$

Thus, the inner minimisation in (5.27) from the perspective of each occurrence of $\Delta_{k_g^d,i_j^u}$ is given by

$$
\begin{aligned}
&\max_{m} -m_{1,k_g i_j} \\
\text{s.t. } &\|\boldsymbol{\omega}_{1,k_g i_j} + \boldsymbol{\Omega}_{1,k_g i_j}\mathrm{vec}(\Delta_{k_g^d,i_j^u})\|^2 \leq m_{1,k_g i_j} \\
&\forall\{\Delta_{k_g^d,i_j^u} : ||\mathrm{vec}(\Delta_{k_g^d,i_j^u})|| \leq \varepsilon_{k_g^d,i_j^u}\}\,.
\end{aligned} \tag{5.30}
$$

Next, representing the inequality in (5.29) as

$$
m_{1,k_g i_j} - \big(\boldsymbol{\omega}_{1,k_g i_j} + \boldsymbol{\Omega}_{1,k_g i_j}\mathrm{vec}(\Delta_{k_g^d,i_j^u})\big)^H\mathbf{I}\big(\boldsymbol{\omega}_{1,k_g i_j} + \boldsymbol{\Omega}_{1,k_g i_j}\mathrm{vec}(\Delta_{k_g^d,i_j^u})\big) \geq 0
$$

we can apply the Schur Complement Lemma, to formulate the constraints of (5.30) as

$$
\begin{bmatrix} m_{1,k_g i_j} & \boldsymbol{\omega}_{1,k_g i_j}^H \\ \boldsymbol{\omega}_{1,k_g i_j} & \mathbf{I} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathrm{vec}(\Delta_{k_g^d,i_j^u})^H\boldsymbol{\Omega}_{1,k_g i_j}^H \\ \boldsymbol{\Omega}_{1,k_g i_j}\mathrm{vec}(\Delta_{k_g^d,i_j^u}) & \mathbf{0} \end{bmatrix} \succeq \mathbf{0}\,.
$$

Additionally, applying Lemma 5.1 from Appendix 5.A with $\xi = \varepsilon_{k_g^d,i_j^u}$, $\mathbf{B} = [\mathbf{0}\ \boldsymbol{\Omega}_{1,k_g i_j}^H]$, $\mathbf{C} = [-1\ \mathbf{0}]$, $\mathbf{D} = \mathrm{vec}(\Delta_{k_g^d,i_j^u})$ and

$$
\mathbf{A} = \begin{bmatrix} m_{1,k_g i_j} & \boldsymbol{\omega}_{1,k_g i_j}^H \\ \boldsymbol{\omega}_{1,k_g i_j} & \mathbf{I} \end{bmatrix}
$$

this can be further represented as

$$\nu_{1,k_g i_j} \geq 0, \quad \begin{bmatrix} m_{1,k_g i_j} - \nu_{1,k_g i_j} & \boldsymbol{\omega}_{1,k_g i_j}^H & \mathbf{0} \\ \boldsymbol{\omega}_{1,k_g i_j} & \mathbf{I} & -\varepsilon_{k_g^d, i_j^u} \boldsymbol{\Omega}_{1,k_g i_j} \\ \mathbf{0} & -\varepsilon_{k_g^d, i_j^u} \boldsymbol{\Omega}_{1,k_g i_j}^H & \nu_{1,k_g i_j} \mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \tag{5.31}$$

Using the same separation of variables principle we can also treat the remaining norm-bounded errors $\Delta_{k_g^d,j}$, Δ_{g,i_j^u} and $\Delta_{g,j}$ in an analogous manner, and reformulate the problem for each one in terms of the corresponding m, ν, $\boldsymbol{\omega}$ and $\boldsymbol{\Omega}$ parameters. Details on these processes are provided in Appendices 5.B.1, 5.B.2 and 5.B.3 respectively. After going through this procedure, we can express the original cost function from (5.27) as a summation of the slack variables and some additional terms in order to obtain the final problem formulation in (5.28). □

Since problem (5.28) is not jointly convex in $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{B}$ we apply the alternating optimisation approach in Algorithm 5.1 to solve it [96][2]. In Step 2, to compute $\mathbf{U}$, we fix $\mathbf{V}$ and $\mathbf{B}$ and solve the resulting SDP problem. In Step 3, instead of finding $\mathbf{W}$, we now want to find $\mathbf{B}$ where $\mathbf{W} = \mathbf{B}^H\mathbf{B}$. Therefore, after replacing terms of the form $\mu \log_2|\,(\ln 2/\mu)\,\mathbf{B}^H\mathbf{B}|$ with $2\,\mu \log_2|\,(\ln 2/\mu)^{\frac{1}{2}}\,\mathbf{B}|$, we fix $\mathbf{V}$ and $\mathbf{U}$, and solve the resulting Max-Det problem [100]. Finally, in Step 4, to compute $\mathbf{V}$ we fix $\mathbf{U}$ and $\mathbf{B}$ and solve the resulting SDP problem. All problems may be solved using standard convex optimisation solvers.

Note that the alternating maximisation approach applied here to solve (5.27) converges. This follows because each step of the iterations leads to a monotonic increase of the objective function, since the objective function is upper bounded, convergence is guaranteed.

5.5 Robust design with stochastic error model

For the stochastic CSI error model, all nodes have access to $\widehat{\mathbf{H}}$ instead of $\mathbf{H}$. Therefore, instead of focusing on the actual achievable DL and UL rates, we consider their lower bounds $R_{k_g^d}^S$ and $R_{k_g^u}^S$ (defined later in (5.37)), where channel estimation errors are treated as noise [69].

[2]Note that some additional minor reformulations are required when solving for $\mathbf{U}$ and $\mathbf{B}$. In particular, we introduce slack variables to handle terms of the form $\|\mathbf{BU}\|_F^2$, similar to the process applied to (5.29).

Starting with the DL rate, under Gaussian signalling, $R^S_{k_g^d} = \log_2 \det\big(\mathbf{I} + \widehat{\mathbf{\Phi}}^{-1}_{k_g^d} \widehat{\mathbf{H}}_{k_g^d,g}\ \mathbf{V}_{k_g^d} \mathbf{V}^H_{k_g^d} \widehat{\mathbf{H}}^H_{k_g^d,g}\big)$, where $\widehat{\mathbf{\Phi}}_{k_g^d}$ is given by

$$\widehat{\mathbf{\Phi}}_{k_g^d} = \sum_{\substack{j=1 \\ (i,j\neq k,g)}}^{G} \sum_{i=1}^{K_j^d} \widehat{\mathbf{H}}_{k_g^d,j} \mathbf{V}_{i_j^d} \mathbf{V}^H_{i_j^d} \widehat{\mathbf{H}}^H_{k_g^d,j} + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \widehat{\mathbf{H}}_{k_g^d,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}^H_{i_j^u} \widehat{\mathbf{H}}^H_{k_g^d,i_j^u} + \widehat{\mathbf{F}}_{k_g^d} + (\sigma_U^2 + f_{k_g^d})\mathbf{I}.$$

Here, $\widehat{\mathbf{F}}_{k_g^d}$ is defined similarly to (5.8) but has all instances of $\mathbf{H}$ replaced by $\widehat{\mathbf{H}}$. Additionally, $f_{k_g^d}$ reflects the effect of the imperfect CSI and is given by

$$f_{k_g^d} \approx \eta_{UB}(1 + \kappa_B + \iota_U) \sum_{j=1}^{G} \sum_{i=1}^{K_j^d} \mathrm{Tr}\left(\mathbf{V}_{i_j^d} \mathbf{V}^H_{i_j^d}\right) + \eta_{UU}(1 + \kappa_U + \iota_U) \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathrm{Tr}\left(\mathbf{V}_{i_j^u} \mathbf{V}^H_{i_j^u}\right).$$

For the UL rate, assuming Gaussian signalling, we have $R^S_{k_g^u} = \log_2 \det\big(\mathbf{I} + \widehat{\mathbf{\Phi}}^{-1}_{k_g^u} \widehat{\mathbf{H}}_{g,k_g^u} \mathbf{V}_{k_g^u}\ \mathbf{V}^H_{k_g^u} \widehat{\mathbf{H}}^H_{g,k_g^u}\big)$, where $\widehat{\mathbf{\Phi}}_{k_g^u}$ is defined as

$$\widehat{\mathbf{\Phi}}_{k_g^u} = \sum_{\substack{j=1 \\ (j\neq g)}}^{G} \sum_{i=1}^{K_j^d} \widehat{\mathbf{H}}_{g,j} \mathbf{V}_{i_j^d} \mathbf{V}^H_{i_j^d} \widehat{\mathbf{H}}^H_{g,j} + \sum_{\substack{j=1 \\ (i,j\neq k,g)}}^{G} \sum_{i=1}^{K_j^u} \widehat{\mathbf{H}}_{g,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}^H_{i_j^u} \widehat{\mathbf{H}}^H_{g,i_j^u} + \widehat{\mathbf{F}}_g + (\sigma_B^2 + f_g)\mathbf{I}.$$

Here, $\widehat{\mathbf{F}}_g$ is defined parallel to (5.12) with $\mathbf{H}$ replaced by $\widehat{\mathbf{H}}$, and f_g is given by

$$f_g \approx \eta_{BB}(1 + \kappa_B + \iota_B) \sum_{j=1}^{G} \sum_{i=1}^{K_j^d} \mathrm{Tr}\left(\mathbf{V}_{i_j^d} \mathbf{V}^H_{i_j^d}\right) + \eta_{BU}(1 + \kappa_U + \iota_B) \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \mathrm{Tr}\left(\mathbf{V}_{i_j^u} \mathbf{V}^H_{i_j^u}\right).$$

Therefore, for the stochastic CSI error model, our problem is given by

$$\begin{aligned} \max_{\mathbf{V}} \quad & \sum_{g=1}^{G} \sum_{k=1}^{K_g^d} \mu_{k_g^d} R^S_{k_g^d} + \sum_{g=1}^{G} \sum_{k=1}^{K_g^u} \mu_{k_g^u} R^S_{k_g^u} \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u} \mathbf{V}^H_{k_g^u}) \leq P_U \ \ \forall\, k, g \\ & \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d} \mathbf{V}^H_{k_g^d}) \leq P_B \ \ \forall\, g\,. \end{aligned} \tag{5.32}$$

Similar to the perfect CSI case, we solve this problem by transforming it into a WMMSE one. To obtain the MSE matrices, we start with $\mathbf{E}_{k_g^d} = \mathbb{E}\{(\mathbf{U}_{k_g^d}\mathbf{y}_{k_g^d} - \mathbf{s}_{k_g^d})(\mathbf{U}_{k_g^d}\mathbf{y}_{k_g^d} - \mathbf{s}_{k_g^d})^H\}$ and $\mathbf{E}_{k_g^u} = \mathbb{E}\{(\mathbf{U}_{k_g^u}\mathbf{y}_g - \mathbf{s}_{k_g^u})(\mathbf{U}_{k_g^u}\mathbf{y}_g - \mathbf{s}_{k_g^u})^H\}$ and replace $\mathbf{H}$ with $\widehat{\mathbf{H}} + \Delta$ from (5.4). Taking the expectation over $\mathbf{s}$, $\mathbf{n}$ and $\mathbf{\Delta}$ under an independence assumption, we obtain $\mathbf{E}^S_{k_g^d}$ in (5.33) for the DL and $\mathbf{E}^S_{k_g^u}$ in (5.34) for the UL.

$$\mathbf{E}_{k_g^d}^S = \mathbf{U}_{k_g^d} \sum_{j=1}^{G} \sum_{i=1}^{K_j^d} \widehat{\mathbf{H}}_{k_g^d,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \widehat{\mathbf{H}}_{k_g^d,j}^H \mathbf{U}_{k_g^d}^H + \mathbf{U}_{k_g^d} \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \widehat{\mathbf{H}}_{k_g^d,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \widehat{\mathbf{H}}_{k_g^d,i_j^u}^H \mathbf{U}_{k_g^d}^H - \mathbf{U}_{k_g^d} \widehat{\mathbf{H}}_{k_g^d,g} \mathbf{V}_{k_g^d} - \mathbf{V}_{k_g^d}^H \widehat{\mathbf{H}}_{k_g^d,g}^H \mathbf{U}_{k_g^d}^H + (\sigma_U^2 + f_{k_g^d}) \mathbf{U}_{k_g^d} \mathbf{U}_{k_g^d}^H + \mathbf{U}_{k_g^d} \widehat{\mathbf{F}}_{k_g^d} \mathbf{U}_{k_g^d}^H + \mathbf{I} \tag{5.33}$$

$$\mathbf{E}_{k_g^u}^S = \mathbf{U}_{k_g^u} \sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{i=1}^{K_j^d} \widehat{\mathbf{H}}_{g,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \widehat{\mathbf{H}}_{g,j}^H \mathbf{U}_{k_g^u}^H + \mathbf{U}_{k_g^u} \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \widehat{\mathbf{H}}_{g,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \widehat{\mathbf{H}}_{g,i_j^u}^H \mathbf{U}_{k_g^u}^H - \mathbf{U}_{k_g^u} \widehat{\mathbf{H}}_{g,k_g^u} \mathbf{V}_{k_g^u} - \mathbf{V}_{k_g^u}^H \widehat{\mathbf{H}}_{g,k_g^u}^H \mathbf{U}_{k_g^u}^H + (\sigma_B^2 + f_g) \mathbf{U}_{k_g^u} \mathbf{U}_{k_g^u}^H + \mathbf{U}_{k_g^u} \widehat{\mathbf{F}}_g \mathbf{U}_{k_g^u}^H + \mathbf{I} \tag{5.34}$$

The MMSE receivers can be obtained by solving $\bar{\mathbf{U}}_{k_g^d}^S = \arg\min_{\mathbf{U}} \mathrm{Tr}(\mathbf{E}_{k_g^d}^S)$ and $\bar{\mathbf{U}}_{k_g^u}^S = \arg\min_{\mathbf{U}} \mathrm{Tr}(\mathbf{E}_{k_g^u}^S)$, resulting in (5.35) and (5.36) respectively.

$$\bar{\mathbf{U}}_{k_g^d}^S = \mathbf{V}_{k_g^d}^H \widehat{\mathbf{H}}_{k_g^d,g}^H \left[\sum_{j=1}^{G} \sum_{i=1}^{K_j^d} \widehat{\mathbf{H}}_{k_g^d,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \widehat{\mathbf{H}}_{k_g^d,j}^H + \widehat{\mathbf{F}}_{k_g^d} + (\sigma_U^2 + f_{k_g^d}) \mathbf{I} + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \widehat{\mathbf{H}}_{k_g^d,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \widehat{\mathbf{H}}_{k_g^d,i_j^u}^H \right]^{-1} \tag{5.35}$$

$$\bar{\mathbf{U}}_{k_g^u}^S = \mathbf{V}_{k_g^u}^H \widehat{\mathbf{H}}_{g,k_g^u}^H \left[\sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{i=1}^{K_j^d} \widehat{\mathbf{H}}_{g,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \widehat{\mathbf{H}}_{g,j}^H + \widehat{\mathbf{F}}_g + (\sigma_B^2 + f_g) \mathbf{I} + \sum_{j=1}^{G} \sum_{i=1}^{K_j^u} \widehat{\mathbf{H}}_{g,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \widehat{\mathbf{H}}_{g,i_j^u}^H \right]^{-1} \tag{5.36}$$

Applying the MMSE receivers from (5.35) and (5.36), the MSE matrices in (5.33) and (5.34) can respectively be expressed as $\bar{\mathbf{E}}_{k_g^d}^S = (\mathbf{I} + \mathbf{V}_{k_g^d}^H \widehat{\mathbf{H}}_{k_g^d,g}^H \widehat{\mathbf{\Phi}}_{k_g^d}^{-1} \widehat{\mathbf{H}}_{k_g^d,g} \mathbf{V}_{k_g^d})^{-1}$ and $\bar{\mathbf{E}}_{k_g^u}^S = (\mathbf{I} + \widehat{\mathbf{H}}_{g,k_g^u} \mathbf{V}_{k_g^u} \widehat{\mathbf{\Phi}}_{k_g^u}^{-1} \mathbf{V}_{k_g^u}^H \widehat{\mathbf{H}}_{g,k_g^u}^H)^{-1}$. Finally, using an argument parallel to the one applied in the perfect CSI case, it can easily be shown that

$$R_{k_g^d}^S = \log_2 \det \left(\left(\bar{\mathbf{E}}_{k_g^d}^S \right)^{-1} \right) \quad \text{and} \quad R_{k_g^u}^S = \log_2 \det \left(\left(\bar{\mathbf{E}}_{k_g^u}^S \right)^{-1} \right). \tag{5.37}$$

This rate to MSE relationship allows us to establish the following theorem.

Theorem 5.3. *The stochastic CSI error WSR problem in (5.32) is equivalent to the WMMSE problem in (5.38), such that the global optimal solution for the precoders of the two problems are identical.*

$$\begin{aligned}\min_{\mathbf{U},\mathbf{W},\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\left[\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}^S) - \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) - \frac{\mu_{k_g^d}}{\ln 2}b_d\right] \\ & + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u}\left[\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}^S) - \mu_{k_g^u}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^u}}\mathbf{W}_{k_g^u}\right) - \frac{\mu_{k_g^u}}{\ln 2}b_u\right] \\ s.t. \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \ \ \forall\, k, g \\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \ \ \forall\, g\,. \end{aligned} \tag{5.38}$$

Proof. It can be noticed that the optimal $\mathbf{U}$ for (5.38) are the MMSE receivers $\bar{\mathbf{U}}_{k_g^d}^S$ and $\bar{\mathbf{U}}_{k_g^u}^S$ in (5.35) and (5.36) respectively. Secondly, for fixed $\mathbf{U}$ and $\mathbf{V}$, checking the first order optimality conditions for the weights we obtain

$$\bar{\mathbf{W}}_{k_g^d}^S = \frac{\mu_{k_g^d}}{\ln 2}\left(\bar{\mathbf{E}}_{k_g^d}^S\right)^{-1} \ \text{ and } \ \bar{\mathbf{W}}_{k_g^u}^S = \frac{\mu_{k_g^u}}{\ln 2}\left(\bar{\mathbf{E}}_{k_g^u}^S\right)^{-1}. \tag{5.39}$$

Substituting for optimal $\mathbf{U}$ and $\mathbf{W}$ in (5.38) results in

$$\begin{aligned}\min_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} -\mu_{k_g^d}\log_2\det\left(\left(\bar{\mathbf{E}}_{k_g^d}^S\right)^{-1}\right) + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u} -\mu_{k_g^u}\log_2\det\left(\left(\bar{\mathbf{E}}_{k_g^u}^S\right)^{-1}\right) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \ \ \forall\, k, g \\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \ \ \forall\, g \end{aligned}$$

which considering (5.37) is the same as (5.32). □

Since (5.38) is not jointly convex in $\mathbf{V}$, $\mathbf{U}$ and $\mathbf{W}$ the alternating optimisation process from Algorithm 5.1 is applied to solve it. For Step 2, we use (5.35) and (5.36) to calculate the optimal receivers as $\mathbf{U}_{k_g^d} = \bar{\mathbf{U}}_{k_g^d}^S$ and $\mathbf{U}_{k_g^u} = \bar{\mathbf{U}}_{k_g^u}^S$. In Step 3, the weights are calculated as $\mathbf{W}_{k_g^d} = \bar{\mathbf{W}}_{k_g^d}^S$ and $\mathbf{W}_{k_g^u} = \bar{\mathbf{W}}_{k_g^u}^S$ using (5.39). The optimal precoders can be obtained similar to the perfect CSI case using the Lagrangian method. Therefore, in Step 4, we calculate the precoders $\mathbf{V}_{k_g^d} = \bar{\mathbf{V}}_{k_g^d}^S$ and $\mathbf{V}_{k_g^u} = \bar{\mathbf{V}}_{k_g^u}^S$ using

$$\bar{\mathbf{V}}^S_{k^d_g} = [\mathbf{X}^S_g + \varpi^S_g \mathbf{I}]^{-1} \widehat{\mathbf{H}}^H_{k^d_g,g} \mathbf{U}^H_{k^d_g} \mathbf{W}_{k^d_g}$$
$$\bar{\mathbf{V}}^S_{k^u_g} = [\mathbf{X}^S_{k^u_g} + \varrho^S_{k^u_g} \mathbf{I}]^{-1} \widehat{\mathbf{H}}^H_{g,k^u_g} \mathbf{U}^H_{k^u_g} \mathbf{W}_{k^u_g} \tag{5.40}$$

where $\mathbf{X}^S_{g^d}$ and $\mathbf{X}^S_{k^u_g}$ are defined in (5.41) and (5.43). Here, ϖ^S_g and $\varrho^S_{k^u_g}$ are the Lagrange multipliers, and $\widehat{\mathbf{X}}_g$ and $\widehat{\mathbf{X}}_{k^u_g}$ are defined similar to $\mathbf{X}_g$ and $\mathbf{X}_{k^u_g}$ from (5.22) and (5.23) respectively but with $\mathbf{H}$ replaced by $\widehat{\mathbf{H}}$.

$$\mathbf{X}^S_g = \widehat{\mathbf{X}}_g + \eta_{UB}(1 + \kappa_B + \iota_U) \sum_{j=1}^{G} \sum_{i=1}^{K^d_j} \mathrm{Tr}\left(\mathbf{U}_{i^d_j} \mathbf{U}^H_{i^d_j} \mathbf{W}_{i^d_j}\right) \tag{5.41}$$

$$+ \eta_{BB}(1 + \kappa_B + \iota_B) \sum_{j=1}^{G} \sum_{i=1}^{K^u_j} \mathrm{Tr}\left(\mathbf{U}_{i^d_j} \mathbf{U}^H_{i^u_j} \mathbf{W}_{i^u_j}\right) \tag{5.42}$$

$$\mathbf{X}^S_{k^u_g} = \widehat{\mathbf{X}}_{k^u_g} + \eta_{UU}(1 + \kappa_U + \iota_U) \sum_{j=1}^{G} \sum_{i=1}^{K^d_j} \mathrm{Tr}\left(\mathbf{U}_{i^d_j} \mathbf{U}^H_{i^d_j} \mathbf{W}_{i^d_j}\right)$$
$$+ \eta_{BU}(1 + \kappa_U + \iota_B) \sum_{j=1}^{G} \sum_{i=1}^{K^u_j} \mathrm{Tr}\left(\mathbf{U}_{i^d_j} \mathbf{U}^H_{i^u_j} \mathbf{W}_{i^u_j}\right) \tag{5.43}$$

Note that the convergence considerations in Remark 5.1 are also applicable to the alternating minimisation approach applied to solve the stochastic CSI error problem in (5.38).

5.6 Weighted DL rate maximisation subject to a per UL user target rate

In addition to the total rate maximisation design we also consider sum DL rate maximisation subject to each UL user achieving a target rate of R_{UL}. The motivation behind this design is due to the fact that even if FD outperforms HD, this does not guarantee that all UL users are served evenly in every time slot. In some instances a UL user may achieve a lower rate in order to reduce the amount of interference present in the system. Therefore, we consider the following problem

$$\max_{\mathbf{V}} \quad \sum_{g=1}^{G} \sum_{k=1}^{K^d_g} \mu_{k^d_g} R_{k^d_g}$$
$$\text{s.t.} \quad R_{k^u_g} \geq R_{UL} \ \ \forall\, k, g$$

$$\begin{aligned}&\mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_U \;\; \forall\, k,g\\ &\sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_B \;\; \forall\, g\,.\end{aligned} \tag{5.44}$$

Similar to the total rate problems from the previous sections, we exploit the rate to MSE relationship to obtain an equivalent WMMSE problem as in (5.45) for the WSR one in (5.44). To our knowledge the use of this property in the constraint of a problem (rather than in the cost function) has not been applied in prior literature.

Theorem 5.4. *The WSR problem in (5.44) is equivalent to the WMMSE problem in (5.45), such that the global optimal solutions for the precoders of the two problems are identical.*

$$\begin{aligned}\min_{\mathbf{U},\mathbf{W},\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d}\left[\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) - \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) - \frac{\mu_{k_g^d}}{\ln 2}b_d\right]\\ s.t. \quad & \left[\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) - \log_2\det\left(\ln 2\,\mathbf{W}_{k_g^u}\right) - \frac{1}{\ln 2}b_u\right] \leq -R_{UL} \;\; \forall\, k,g\\ & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_U \;\; \forall\, k,g\\ & \sum_{k=1}^{K_g^d}\mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_B \;\; \forall\, g\,.\end{aligned} \tag{5.45}$$

Proof. Firstly, it can be seen that the optimal $\mathbf{U}$ for (5.45) are the standard MMSE receivers $\bar{\mathbf{U}}_{k_g^d}$ and $\bar{\mathbf{U}}_{k_g^u}$ in (5.9) and (5.14) respectively. Secondly, fixing $\mathbf{U}$ and $\mathbf{V}$ and checking the first order optimality conditions for the weights we obtain their optimal values as

$$\bar{\mathbf{W}}_{k_g^d}^c = \frac{\mu_{k_g^d}}{\ln 2}\bar{\mathbf{E}}_{k_g^d}^{-1} \quad \text{and} \quad \bar{\mathbf{W}}_{k_g^u}^c = \frac{1}{\ln 2}\bar{\mathbf{E}}_{k_g^u}^{-1}\,. \tag{5.46}$$

Substituting for optimal $\mathbf{U}$ and $\mathbf{W}$ in (5.45) results in

$$\begin{aligned}\min_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} -\mu_{k_g^d}\log_2\det\left(\bar{\mathbf{E}}_{k_g^d}^{-1}\right)\\ \text{s.t.} \quad & -\log_2\det\left(\bar{\mathbf{E}}_{k_g^u}^{-1}\right) \leq -R_{UL} \;\; \forall\, k,g\\ & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_U \;\; \forall\, k,g\end{aligned}$$

$$\sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_B \ \ \forall\, g \tag{5.47}$$

which considering (5.11) and (5.13) is the same as (5.44). □

Since (5.45) is not jointly convex in $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$, but is separately convex in each variable, it can be solved via alternating maximisation. Having already obtained closed form expressions for optimal $\mathbf{U}$ and $\mathbf{W}$, we focus on obtaining $\mathbf{V}$. For fixed $\mathbf{U}$ and $\mathbf{W}$, we can express (5.45) as

$$\begin{aligned}
\min_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) \\
\text{s.t.} \quad & \mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) \leq \Psi_{k_g^u} \ \ \forall\, k, g \\
& \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_u \ \ \forall\, k, g \\
& \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_{BS} \ \ \forall\, g\ .
\end{aligned} \tag{5.48}$$

where $\Psi_{k_g^u} = -R_{UL} + \log_2 \det\left(\ln 2\, \mathbf{W}_{k_g^u}\right) + b_u/\ln 2$.

Next, using $\mathrm{Tr}(\mathbf{A}\mathbf{A}^H) = \|\mathrm{vec}(\mathbf{A})\|^2$ and $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$, we can rewrite $\mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d})$ and $\mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u})$ as $\|\boldsymbol{\phi}_{k_g^d}\|^2$ in (5.50) and $\|\boldsymbol{\phi}_{k_g^u}\|^2$ in (5.51) respectively (see definitions on next page). This reformulation allows us to introduce slack variable t, such that $\|\boldsymbol{\phi}_{k_g^d}\|^2 \leq t_{k_g^d}$, and cast (5.48) as the following problem

$$\begin{aligned}
\min_{\mathbf{V},t} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} t_{k_g^d} \\
\text{s.t.} \quad & \|\boldsymbol{\phi}_{k_g^d}\|^2 \leq t_{k_g^d} \ \ \forall\, k, g \\
& \|\boldsymbol{\phi}_{k_g^u}\|^2 \leq \Psi_{k_g^u} \ \ \forall\, k, g \\
& \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq P_u \ \ \forall\, k, g \\
& \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq P_{BS} \ \ \forall\, g
\end{aligned} \tag{5.49}$$

which after additional minor reformulations can be transformed into a second-order cone programming (SOCP) problem, and then solved using standard convex optimisation solvers.

$$\|\boldsymbol{\phi}_{k_g^d}\|^2 = \mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) =$$

$$\left\| \begin{array}{c} (\mathbf{I}\otimes \mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,g})\mathrm{vec}(\mathbf{V}_{k_g^d}) - \mathrm{vec}(\mathbf{B}_{k_g^d}) \\ \lfloor (\mathbf{I}\otimes \mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,j})\mathrm{vec}(\mathbf{V}_{i_j^d})\rfloor_{\forall\, j=1\ldots G, i=1\ldots K_j^d,(i,j\neq k,g)} \\ \lfloor (\mathbf{I}\otimes \mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u})\mathrm{vec}(\mathbf{V}_{i_j^u})\rfloor_{\forall\, j=1\ldots G, i=1\ldots K_j^u} \\ \kappa_B^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left(\mathrm{diag}(\mathbf{H}_{k_g^d,j}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,j})\right)^{\frac{1}{2}}\right)\mathrm{vec}(\mathbf{V}_{i_j^d})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^d}} \\ \kappa_U^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left(\mathrm{diag}(\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{H}_{k_g^d,i_j^u})\right)^{\frac{1}{2}}\right)\mathrm{vec}(\mathbf{V}_{i_j^u})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^u}} \\ \iota_U^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left[\left(\mathrm{diag}(\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d})\right)^{\frac{1}{2}}\mathbf{H}_{k_g^d,j}\right]\right)\mathrm{vec}(\mathbf{V}_{i_j^d})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^d}} \\ \iota_U^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left[\left(\mathrm{diag}(\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d})\right)^{\frac{1}{2}}\mathbf{H}_{k_g^d,i_j^u}\right]\right)\mathrm{vec}(\mathbf{V}_{i_j^u})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^u}} \\ (\sigma_U^2+\iota_U\sigma_U^2)^{\frac{1}{2}}\mathrm{Tr}(\mathbf{B}_{k_g^d}\mathbf{U}_{k_g^d}\mathbf{U}_{k_g^d}^H\mathbf{B}_{k_g^d}^H)^{\frac{1}{2}} \end{array} \right\|^2 \tag{5.50}$$

$$\|\boldsymbol{\phi}_{k_g^u}\|^2 = \mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) =$$

$$\left\| \begin{array}{c} (\mathbf{I}\otimes \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,k_g^u})\mathrm{vec}(\mathbf{V}_{k_g^u}) - \mathrm{vec}(\mathbf{B}_{k_g^u}) \\ \lfloor (\mathbf{I}\otimes \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j})\mathrm{vec}(\mathbf{V}_{i_j^d})\rfloor_{\forall\, j=1\ldots G, i=1\ldots K_g^d,(j\neq g)} \\ \lfloor (\mathbf{I}\otimes \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u})\mathrm{vec}(\mathbf{V}_{i_j^u})\rfloor_{\forall\, j=1\ldots G, i=1\ldots K_g^u,(i,j\neq k,g)} \\ \kappa_B^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left(\mathrm{diag}(\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j})\right)^{\frac{1}{2}}\right)\mathrm{vec}(\mathbf{V}_{i_j^d})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^d}} \\ \kappa_U^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left(\mathrm{diag}(\mathbf{H}_{g,i_j^u}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u})\right)^{\frac{1}{2}}\right)\mathrm{vec}(\mathbf{V}_{i_j^u})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^u}} \\ \iota_B^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left[\left(\mathrm{diag}(\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u})\right)^{\frac{1}{2}}\mathbf{H}_{g,j}\right]\right)\mathrm{vec}(\mathbf{V}_{i_j^d})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^d}} \\ \iota_B^{\frac{1}{2}}\lfloor \left(\mathbf{I}\otimes \left[\left(\mathrm{diag}(\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u})\right)^{\frac{1}{2}}\mathbf{H}_{g,i_j^u}\right]\right)\mathrm{vec}(\mathbf{V}_{i_j^u})\rfloor_{\substack{\forall\, j=1\ldots G \\ i=1\ldots K_j^u}} \\ (\sigma_B^2+\iota_B\sigma_B^2)^{\frac{1}{2}}\mathrm{Tr}(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H)^{\frac{1}{2}} \end{array} \right\|^2 \tag{5.51}$$

Therefore to solve (5.45) we apply the alternating optimisation process from Algorithm 5.1. The optimal weights in Step 2 are calculated as $\mathbf{U}_{k_g^d} = \bar{\mathbf{U}}_{k_g^d}$ and $\mathbf{U}_{k_g^u} = \bar{\mathbf{U}}_{k_g^u}$ using (5.9) and (5.14). In Step 3, the optimal weights $\mathbf{W}_{k_g^d} = \bar{\mathbf{W}}_{k_g^d}^c$ and $\mathbf{W}_{k_g^u} = \bar{\mathbf{W}}_{k_g^u}^c$ are found using (5.46). In Step 4, the optimal precoders $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$ are found by solving (5.49).

Proposition 5.1. *The alternating optimisation process applied to solve (5.45) produces a convergent monotonically decreasing objective value sequence.*

Proof. Defining the following parameters

$$C_{k_g^d}(\mathbf{U},\mathbf{W},\mathbf{V}) = \mathrm{Tr}(\mathbf{W}_{k_g^d}\mathbf{E}_{k_g^d}) - \mu_{k_g^d}\log_2\det\left(\frac{\ln 2}{\mu_{k_g^d}}\mathbf{W}_{k_g^d}\right) - \frac{\mu_{k_g^d}}{\ln 2}b_d$$

$$C_{k_g^u}(\mathbf{U},\mathbf{W},\mathbf{V}) = \mathrm{Tr}(\mathbf{W}_{k_g^u}\mathbf{E}_{k_g^u}) - \log_2\det\left(\ln 2\,\mathbf{W}_{k_g^u}\right) - \frac{b_u}{\ln 2}$$

we can express (5.45) as

$$\begin{aligned}\min_{\mathbf{U},\mathbf{W},\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} C_{k_g^d}(\mathbf{U},\mathbf{W},\mathbf{V}) \\ \text{s.t.} \quad & C_{k_g^u}(\mathbf{U},\mathbf{W},\mathbf{V}) \le -R_{UL} \;\; \forall\, k,g \\ & \mathrm{Tr}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \le P_U \;\; \forall\, k,g \\ & \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \le P_B \;\; \forall\, g\,. \end{aligned} \tag{5.52}$$

Assume that for (5.52) we have feasible solution $\{\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\}$ at the end of the (i)th iterate, and feasible solution $\{\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i+1)\}$ at the end of the $(i+1)$th iterate. At the beginning of the $(i+1)$th iterate, to perform Step 2 of Algorithm 5.1, we fix the weights and precoders to $\mathbf{W}(i)$ and $\mathbf{V}(i)$ in order to obtain the updated receivers $\mathbf{U}(i+1)$. Since these receivers are MMSE ones, they are unique optimisers, therefore

$$C_{k_g^d}\big(\mathbf{U}(i+1),\mathbf{W}(i),\mathbf{V}(i)\big) \le C_{k_g^d}\big(\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\big) \quad \forall\, k,g \tag{5.53}$$

$$C_{k_g^u}\big(\mathbf{U}(i+1),\mathbf{W}(i),\mathbf{V}(i)\big) \le C_{k_g^u}\big(\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\big) \overset{(a)}{\le} -R_{UL} \quad \forall\, k,g \tag{5.54}$$

where (a) follows since $\{\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\}$ is feasible.

Next, in Step 3, we fix the receivers and precoders to $\mathbf{U}(i+1)$ and $\mathbf{V}(i)$ in order to obtain the new weights $\mathbf{W}(i+1)$. The weights are updated using (5.46), which are unique optimisers, therefore

$$\begin{aligned} C_{k_g^d}\big(\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i)\big) &\le C_{k_g^d}\big(\mathbf{U}(i+1),\mathbf{W}(i),\mathbf{V}(i)\big)\big) \\ &\overset{(a)}{\le} C_{k_g^d}\big(\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\big) \quad \forall\, k,g \end{aligned} \tag{5.55}$$

$$\begin{aligned} C_{k_g^u}\big(\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i)\big) &\le C_{k_g^u}\big(\mathbf{U}(i+1),\mathbf{W}(i),\mathbf{V}(i)\big) \\ &\overset{(b)}{\le} C_{k_g^u}\big(\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\big) \le -R_{UL} \quad \forall\, k,g \end{aligned}$$

where (a) follows from (5.53) and (b) follows from (5.54).

At this stage we have intermediate solution $\{\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i)\}$ which is a feasible point, and the value of the cost function is given by

$$\sum_{g=1}^{G}\sum_{k=1}^{K_g^d} C_{k_g^d}\big(\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i)\big) \overset{(a)}{\leq} \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} C_{k_g^d}\big(\mathbf{U}(i),\mathbf{W}(i),\mathbf{V}(i)\big)$$

where (a) follows from (5.55). Next, in Step 4 we fix the receivers and weights to $\mathbf{U}(i+1)$ and $\mathbf{W}(i+1)$ and solve (5.52) to obtain the new precoders $\mathbf{V}(i+1)$. Since with $\{\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i)\}$ the problem is known to be feasible, it follows that

$$\sum_{g=1}^{G}\sum_{k=1}^{K_g^d} C_{k_g^d}\big(\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i+1)\big) \leq \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} C_{k_g^d}\big(\mathbf{U}(i+1),\mathbf{W}(i+1),\mathbf{V}(i)\big).$$

As can be seen from the above process, the alternating optimisation method applied to solve (5.45) produces a convergent monotonically decreasing objective value sequence. □

5.7 Simulation results

Our simulations follow the 3GPP LTE [86] specifications for multi-cell pico scenarios outlined in Table 5.1, with all channel gains assumed to be i.i.d. Channel gains between BSs and users, and between the BSs themselves, are modelled as $\mathbf{H}_{r,t} = \sqrt{\varrho}\tilde{\mathbf{H}}_{r,t}$, where r represents the receiver, t represents the transmitter, $\tilde{\mathbf{H}}_{r,t}$ has elements distributed as $\mathcal{CN}(0,1)$ and $\varrho = 10^{-PL/10}$ with PL being the pathloss calculated according to Table 5.1, depending on r and t. The SI channel, $\mathbf{H}_{g,g}$, is modelled as $\mathcal{CN}\big(\sqrt{K_H/(1+K_H)}\bar{\mathbf{H}}_{g,g}, (1/(1+K_H))\mathbf{I}_{M_B}\otimes\mathbf{I}_{M_B}\big)$ [22], where K_H is the Rician factor and $\bar{\mathbf{H}}_{g,g}$ is a deterministic matrix[3].

Throughout all simulations we fix $\mu_{k_g^d} = \mu_{k_g^u} = 1\,\forall\, k, g$. We also set $\kappa_B = \kappa_U = \kappa$ and $\iota_B = \iota_U = \iota$. Parameters κ and ι jointly reflect the amount of transmitter and receiver distortion, and more importantly they reflect the amount of residual SI at the FD BS as can be seen from (5.3). The larger their value, the larger both distortion and residual SI. Additionally, for all algorithms we consider random precoder initialisation and average the rate results in a Monte Carlo fashion over a number of randomly generated scenario realisations.

[3]Without loss of generality, we set $K_H = 1$ and $\bar{\mathbf{H}}_{g,g}$ to be a matrix of all ones similar to [79, 90].

Table 5.1: Parameter settings for simulations [86].

Parameter	**Setting**
Cell radius	40 m
Bandwidth	10 MHz
Thermal noise density	174 dBm/Hz
Noise figure	BS: 13 dB, user: 9 dB
Maximum transmit power	$P_B = 24$ dBm, $P_U = 23$ dBm
Minimum distance	$r_{BS,BS-min} = 40$ m $r_{BS,user-min} = 10$ m
BS to BS pathloss (in dB, r in km)	LOS if $r < 2/3$: $98.4 + 20\log_{10}(r)$ LOS if $r \geq 2/3$: $101.9 + 40\log_{10}(r)$ NLOS: $101.9 + 40\log_{10}(r)$
BS to user pathloss (in dB, r in km)	LOS: $103.8 + 20.9\log_{10}(r)$ NLOS: $145.4 + 37.5\log_{10}(r)$
User to user pathloss (in dB, r in km)	if $r \leq 50$ m: $98.45 + 20\log_{10}(r)$ if $r > 50$ m: $175.78 + 40\log_{10}(r)$
Shadowing standard deviation (in dB)	between BS & users, LOS: 3, NLOS: 4 between cells: 6
LOS probability (r in km)	$0.5 - \min(0.5, 5\exp(-0.156/r))$ $+ \min(0.5, 5\exp(-r/0.003))$

5.7.1 Perfect CSI results

The aim of this section is to compare results obtained by FD and HD beamformer design under perfect CSI, in order to understand under what conditions FD operation provides performance advantages. We set $G = 2$, $K_g^d = K_g^u = 1$ $\forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ to obtain the results in Fig. 5.2 and Fig. 5.3. Fig. 5.2 compares the sum rates achieved by the FD beamformer design from Section 5.3 with HD operation. For HD we consider the case where the BSs serve their corresponding DL and UL users separately in alternate channel uses, with the aim in each case being to maximise either the DL rate or the UL rate accordingly. As can be seen from Fig. 5.2 for $\kappa = \iota = -50$ dB both HD and FD systems obtain similar rates, however FD outperforms HD for values of $\kappa = \iota < -50$ dB. The amount of gain achieved varies with the $\kappa = \iota$ value. This is mainly due to the fact that the higher the distortion, $\kappa = \iota$, the more residual SI there is. The residual SI is a limiting factor for the UL rate, which contributes a smaller portion of the total rate for larger $\kappa = \iota$.

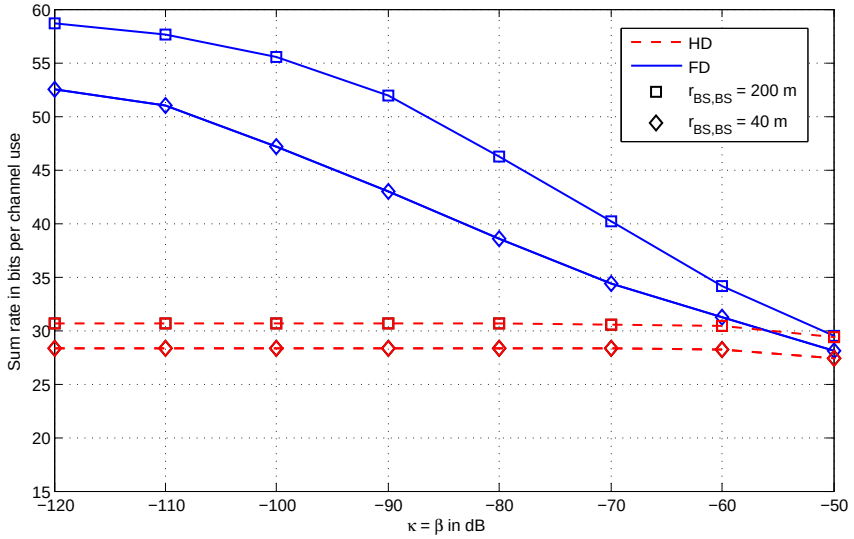


Figure 5.2: Total sum rates achieved for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$ and $N_d = N_u = 2$.

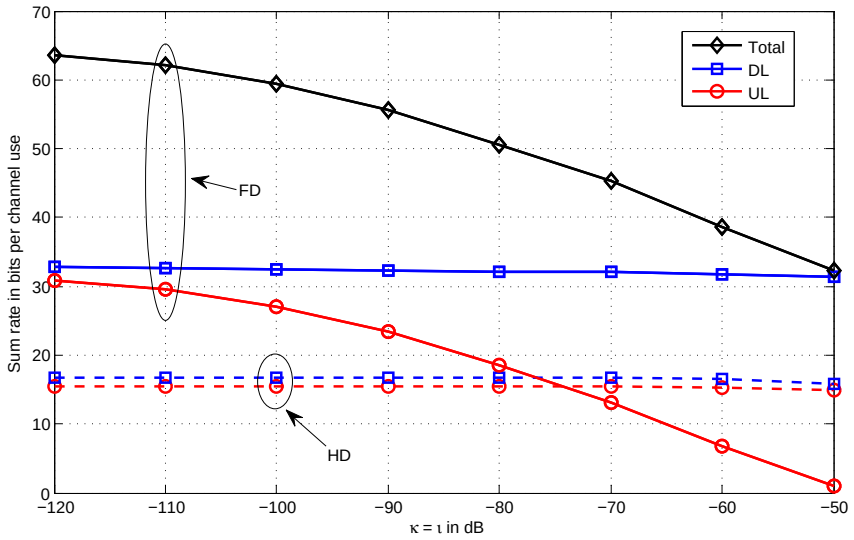


Figure 5.3: Total sum rates achieved for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 4$ and $r_{BS,BS} = 100$ m.

The effect of κ and ι separately can be understood from Fig. 5.4 and 5.5, where we plot the sum rate results for fixed κ and varying ι, and vice-versa. Considering first the results in Fig. 5.4, it can be noticed that the FD total rate at $\iota = -120$ dB are relatively close to each other in value $\forall\ \kappa \in \{-60, -90, -120\}$ dB, even though the initial starting points at $\iota = -50$ dB vary significantly. Similar behaviour can also be noticed when varying κ from -50 dB to -120 dB

$\forall\ \iota \in \{-60, -90, -120\}$ dB in Fig. 5.5. This indicates that it is not necessary to have both κ and ι smaller than -50 dB for FD to obtain significant gains over HD; having either one or the other low is sufficient. For example, from Fig. 5.4 at $\iota = -50$ dB we have a gain of 1.35 for $\kappa = -90$ dB and a gain of 1.54 for $\kappa = -120$ dB, also from Fig. 5.5 at $\kappa = -50$ dB we have a gain of 1.44 at $\iota = -90$ dB and a gain of 1.69 at $\iota = -120$ dB.

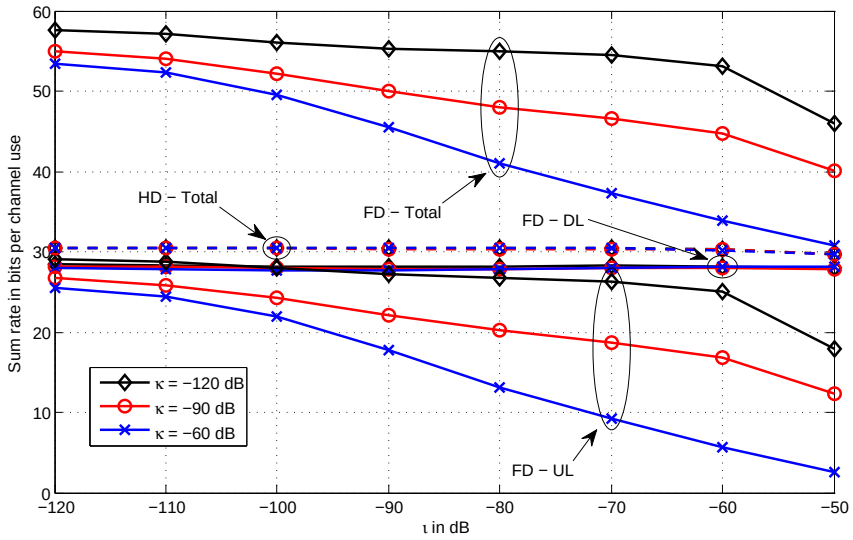


Figure 5.4: Total sum rates achieved for varying ι for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ and $r_{BS,BS} = 100$ m.

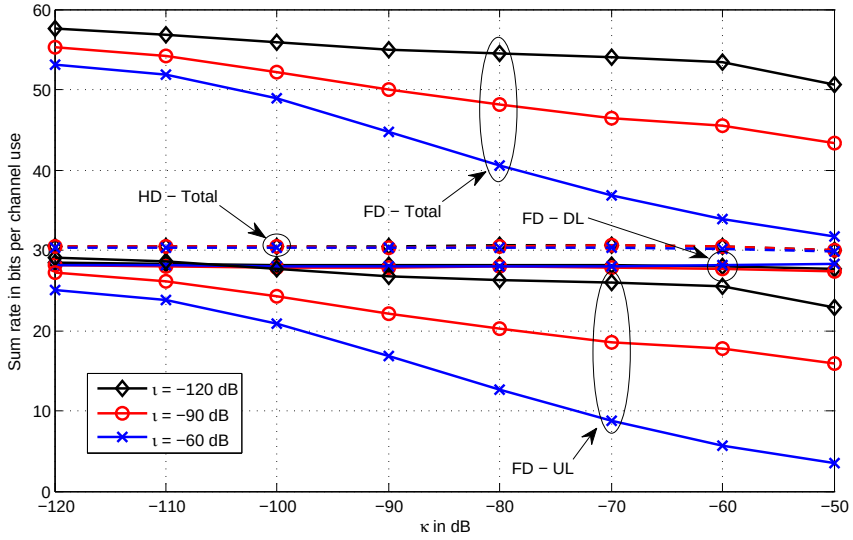


Figure 5.5: Total sum rates achieved for varying κ for for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ and $r_{BS,BS} = 100$ m.

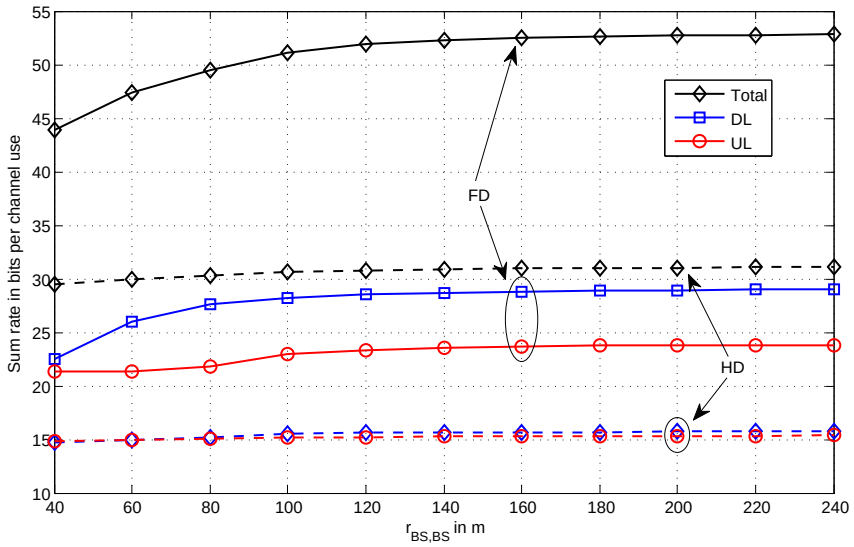
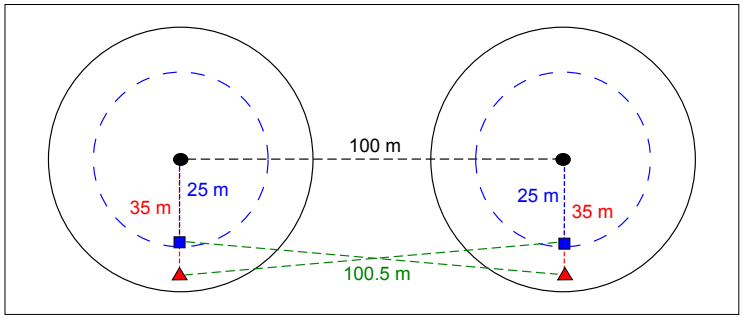


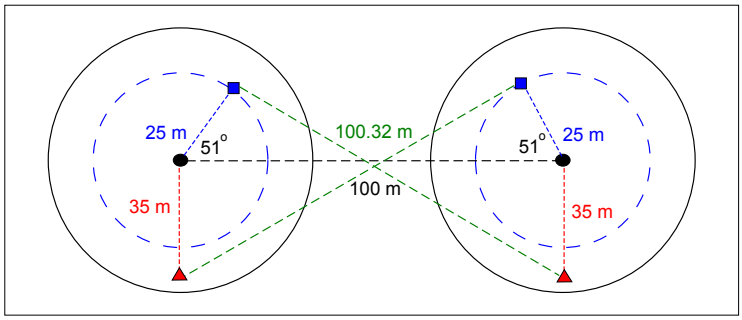
Figure 5.6: Total sum rates achieved for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ and $\kappa = \iota = -90$ dB.

Considering Fig. 5.2 to Fig.5.5 it can be noticed that as the value of κ and/or ι decreases the gain of FD over HD starts to increase significantly. In particular, for Fig. 5.2 at $\kappa = \iota = -120$ dB there is a gain of 1.92 for $r_{BS,BS} = 200$ m and a gain of 1.85 for $r_{BS,BS} = 40$ m. For FD the rate drop between achievable rates at $r_{BS,BS} = 200$ m and at $r_{BS,BS} = 40$ m is larger than the rate drop experienced by HD. This is due to the fact that when the BSs operate in FD there are more interference links than for HD, thus the negative impact of closer proximity between the cells affects FD more than HD.

The impact of inter-cell CCI on FD operation can be understood more clearly from Fig. 5.6, where we set $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ and $\kappa = \iota = -90$, and plot sum rate against the distance between BSs, $r_{BS,BS}$. As $r_{BS,BS}$ increases, inter-cell CCI decreases, thus the total achievable rate increases. An interesting effect can be noticed by looking at the separate FD DL and UL rate results in the range of 40 m to 80 m. The UL rate within this range remains approximately the same, however the DL rate has a significant increase. DL users experience inter-cell CCI from both BSs and UL users in other cells, thus a small increase in the distance between BSs contributes to a significant decrease in inter-cell CCI, allowing DL users to achieve higher rates. For this $r_{BS,BS}$ range, the BS to BS channel is very strong, implying that it is not advantageous in terms of the overall achievable rate to promote UL rate gain, hence the resulting small change in UL rate between 40 m and 80 m. A



Scenario A.



Scenario B.

Figure 5.7: Scenarios with same inter-cell CCI. Black circles represent the BSs, blue squares are UL users and red triangles are DL users.

$r_{BS,BS}$ of around 100 m or more is sufficient to overcome this issue, leading to a marked increase in UL rate at 100 m.

Having seen the effect of inter-cell CCI, next we investigate the effect of intra-cell CCI. In order to do so, we have devised two scenarios with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ and $r_{BS,BS} = 100$ m where we fix the location of the BSs and the users, as shown in Fig. 5.7. For both scenarios A and B, the BSs are 100 m apart and the distance between different cell DL and UL users is approximately 100 m (100.5 m for Scenario A and 100.32 m for Scenario B), implying that the effect of inter-cell CCI is the same. However, the distance between same cell DL and UL users is only 10 m for Scenario A and a much larger 56.49 m for Scenario B. Fig. 5.8 provides some simulation results. As can be seen, scenario B achieves higher rates throughout; this is expected since Scenario B represents the lower interference case. Considering Scenario A and looking at the separate DL and UL rates, it can be noticed that for example at $\kappa = \iota = -50$ dB the DL rate is around 27.5 bits per channel use and the UL rate is nearly zero. At this $\kappa = \iota$ value the SI component is very high, making UL communication very difficult, thus DL communication is given priority. However

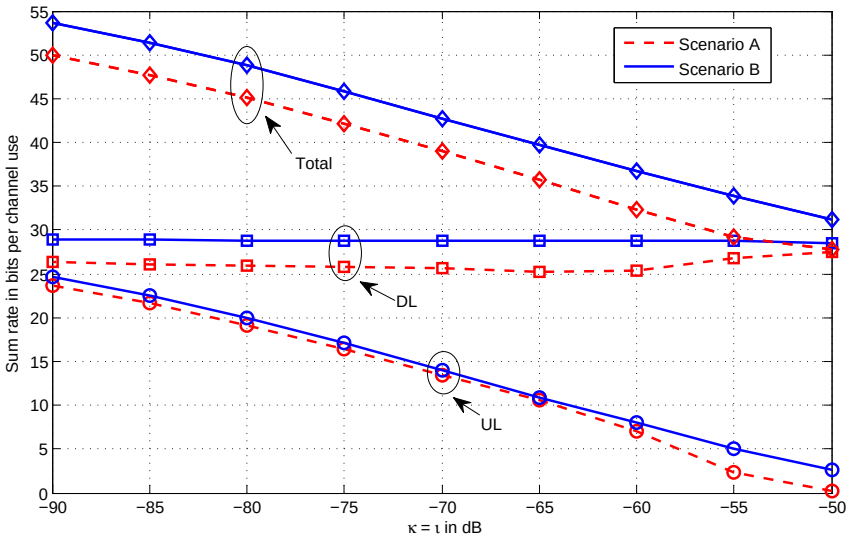


Figure 5.8: Total sum rates achieved for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$ and $r_{BS,BS} = 100$ m.

as SI decreases, the UL rate starts to increase. This increase in UL rate in the lower SI region comes at the expense of a slight decrease in the DL rate, due to the higher intra-cell CCI component. For scenario B, same cell UL and DL users are much further apart, thus the effect of intra-cell CCI is considerably reduced and this UL/DL rate trade-off does not occur.

5.7.2 Imperfect CSI results

After establishing the gains of FD systems over HD ones, our next goal is to show how the FD imperfect CSI designs fare. Starting with the norm-bounded error design from Section 5.4 we set $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 2$, $r_{BS,BS} = 100$ m and $\varepsilon_{k_g^d,i_j^u} = \varepsilon_{k_g^d,j} = \varepsilon_{g,i_j^u} = \varepsilon_{g,j} = \varepsilon\ \forall k, g, i, j$ to obtain the results in Fig. 5.9. Note that channel strengths generated using the 3GPP LTE model from [86] are in the order of -30 dB or lower, which is why for $\varepsilon = -30$ dB achievable rates are close to zero. This also highlights why in the range of $\varepsilon = -30$ dB to $\varepsilon = -35$ dB, there is only a small difference in the rates achieved for different $\kappa = \iota$ values. Within this region the CSI error is considerably large, varying from being of the same order of magnitude as the strongest channels at -30 dB to a third at -35 dB; with CSI errors being so large, the error is more of a limiting factor on rate performance than transmitter and receiver distortion. The converse is true for lower ε regions. As the norm

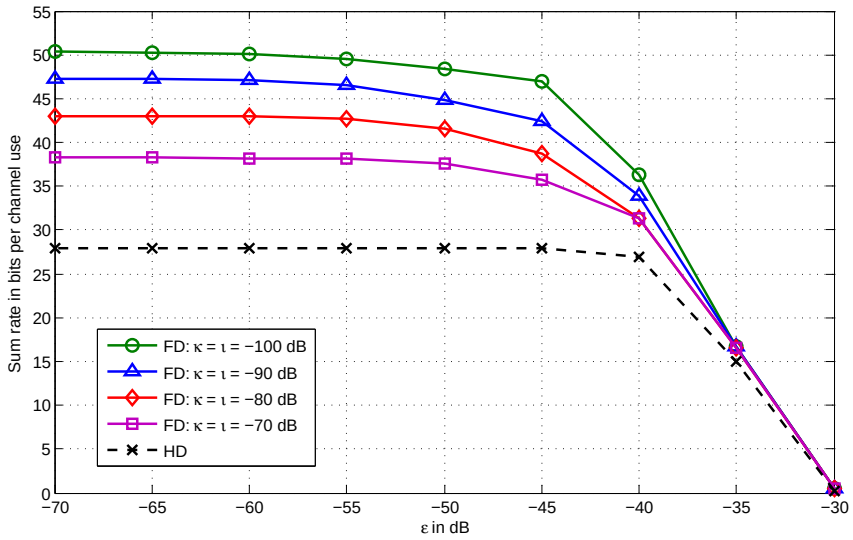


Figure 5.9: Total sum rates achieved for different norm-bounded errors for scenario with $G=2$, $K_g^d=K_g^u=1$ $\forall g$, $b_d=b_u=1$, $M_B=4$, $N_d=N_u=2$ and $r_{BS,BS}=100$ m.

of the CSI error starts to decrease, the curves achieved for different $\kappa = \iota$ values become more distinct, indicating that distortion effects are more of a rate limiting factor than the CSI error. Naturally, the curve for the lowest $\kappa = \iota$ settles at the highest rate value, which is expected since this corresponds to the least amount of distortion and residual SI.

For the stochastic CSI error model, in Fig. 5.10 we set $G = 2$, $K_g^d = K_g^u = 1$ $\forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 4$, $r_{BS,BS} = 100$ m and $\kappa = \iota = -90$ dB, and plot the achievable rate for varying values of β and α, where $\beta = 0$ corresponds to perfect CSI for any α. The robust design from Section 5.5 is compared with a naive version obtained by using the available imperfect CSI as if it were perfect, i.e. without any robustness considerations. For fixed α, rate decreases as β increases; this is expected since larger β values correspond to larger CSI errors. Additionally, it can be noticed that the lower the α the sharper is the rate decrease for varying β, and the larger is the gain between the rate achieved by the robust beamformer versus the naive one. For $\alpha = 1$, there is only a small difference between the performance of the robust and the naive designs, and the rate decrease for varying β is also small. This behaviour is a reflection of the fact that previous studies with a similar error model show that $\alpha = 1$ corresponds to perfect CSI from a DoF perspective. (Note that this has already been proven for HD systems in Chapter 4, and will also be shown for this specific type of system with FD BSs and HD users later on in Chapter 6).

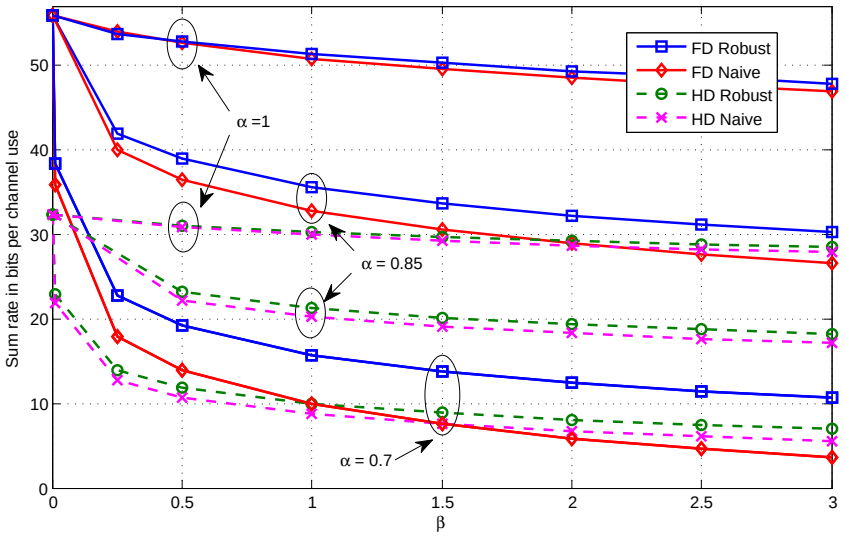


Figure 5.10: Total sum rates achieved for different stochastic errors for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_d = b_u = 1$, $M_B = 4$, $N_d = N_u = 4$, $r_{BS,BS} = 100$ m and $\kappa = \iota = -90$ dB.

Looking at both Fig. 5.9 and Fig. 5.10, it can be noticed that for both types of error the FD results deteriorate more than the HD ones for the same decrease in CSI quality. When using FD BSs there are more channel links between the various nodes than for the corresponding HD system. Having an increased amount of links with imperfect knowledge results in a sharper rate decrease, thereby stressing the added importance of channel estimation and robust beamformer design for FD systems.

5.7.3 Results for target UL rate problem

For the problem from Section 5.6, which considers weighted DL rate maximisation subject to a per UL user target rate, we set $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_u = b_d = 1$, $M_B = 4$ and $N_d = N_u = 4$ to obtain the results in Table 5.2. This table provides a comparison between the DL rates achieved by the constraint design and the corresponding HD system which maximises the total DL rate. Values written in the form of $(a)^{*b\%}$ indicate that the problem is not always feasible for the considered target rate R_{UL}. Here, $b\%$ represents the percentage of scenarios for which the problem was found to be feasible, and a represents the average rate achieved over these feasible scenarios. UL rate results are not included, since provided that the chosen target is feasible, a total UL rate of GKR_{UL} is achievable.

The gains of FD DL rates over HD DL rates range from 1.89 to 1.98 in Table 5.2. On the other hand for the joint problem in Fig. 5.3, which considers the same system with $r_{BS,BS} = 100$ m, there is a gain of 1.83 at $\kappa = \iota = -100$ dB and 1.40 at $\kappa = \iota = -70$ dB. Such a difference is mainly due to the fact that for FD κ and ι are not only related to distortion, but also to residual SI, which makes UL communication more difficult. Constricting both FD and HD to achieve the same target UL rate removes the latter factor, thereby leading to higher gains over HD for the target UL rate problem as opposed to the joint UL and DL maximisation one.

With respect to the feasibility of the chosen target rate, it can be noticed that for a fixed R_{UL}, the lower the distortion the more likely is the problem always feasible. For example at $r_{BS,BS} = 100$ m and $R_{UL} = 2.5$, feasibility goes from 12% at $\kappa = \iota = -70$ dB to 100% at $\kappa = \iota = -90$ dB. Such behaviour is expected because the higher the distortion, the stronger the SI and the more difficult it is to communicate in the UL. For the lowest distortion value of $\kappa = \iota = -100$ dB, R_{UL} of up to around 8 is generally always feasible for $r_{BS,BS} = 100$ m, this decreases to R_{UL} of up to around 5.5 for $r_{BS,BS} = 40$ m. Naturally, for $r_{BS,BS} = 100$ m higher R_{UL} can be achieved than for $r_{BS,BS} = 40$ m, this is due to the stronger interference present in the latter scenario. This trend can be confirmed by comparing all $r_{BS,BS} = 100$ m and $r_{BS,BS} = 40$ m results across Table 5.2.

Table 5.2: Sum DL rates achieved in bits per channel use for scenario with $G = 2$, $K_g^d = K_g^u = 1\ \forall g$, $b_u = b_d = 1$, $M_B = 4$ and $N_d = N_u = 4$.

$r_{BS,BS} = 40$ m				
$\kappa = \iota$	FD			HD
(in dB)	$R_{UL} = 0.5$	$R_{UL} = 1.5$	$R_{UL} = 2.5$	
-100	32.01	31.96	31.87	16.67
-90	31.81	31.66	31.60	16.66
-80	31.60	31.50	$(31.44)^{*92\%}$	16.66
-70	$(31.38)^{*99\%}$	$(31.02)^{*48\%}$	$(30.48)^{*7\%}$	16.64

$r_{BS,BS} = 100$ m				
$\kappa = \iota$	FD			HD
(in dB)	$R_{UL} = 0.5$	$R_{UL} = 1.5$	$R_{UL} = 2.5$	
-100	33.68	33.62	33.57	17.00
-90	33.50	33.42	33.40	16.98
-80	33.38	33.34	$(33.29)^{*98\%}$	16.98
-70	33.23	$(32.94)^{*64\%}$	$(33.28)^{*12\%}$	16.96

5.7.4 Convergence results

Fig. 6.9 illustrates the convergence behaviour of the proposed algorithms. For each algorithm we plot a randomly selected instance. In each case we set $\kappa = \iota = -90$ dB and run for 30 iterations. For the perfect CSI problem we consider the system setup from Fig. 5.3. For the norm-bounded error problem we simulate the system from Fig. 5.9 with $\varepsilon = -45$ dB. For the stochastic CSI error problem we consider the system from Fig. 5.10 with $\alpha = 0.85$ and $\beta = 0.5$. For the constraint problem from Section 5.6 we simulate the system in Table 5.2 at $r_{BS,BS} = 100$ m with $R_{UL} = 1.5$. As can be seen all algorithms converge monotonically within a few steps.

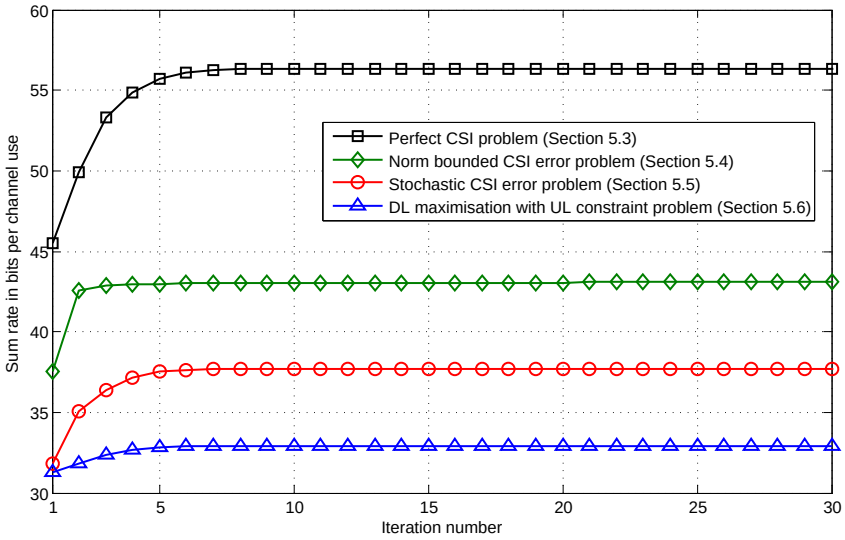


Figure 5.11: Convergence behaviour of the proposed algorithms.

5.8 Implementation and complexity analysis

In order to simplify the notation in our analysis, throughout this section we fix $M_B = N_d = N_u = M$, $K_d = K_u = \bar{K}$ and $b_d = b_u = b$.

5.8.1 Implementation

All proposed algorithms can be implemented in a centralised manner, where a central processing site (CPS) collects all the required CSI, computes the required variables, and then distributes them to the respective nodes. For this implementation a total of $M^2G^2(\bar{K}^2+2\bar{K}+1)$ CSI elements need to be made available at the

CPS to implement the algorithm. The CPS must then distribute the calculated precoders, resulting in $2G\bar{K}Mb$ matrix elements for all of $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$.

Additionally, the closed-form solution algorithms from Sections 5.3 and 5.5 may also be applied in a distributed manner. Similar to the implementations in [90, 92] and references therein, this requires all nodes to have knowledge of the channels directly linked to them, i.e. local CSI, and also assumes all receiving nodes can provide additional feedback information to transmitting nodes. Each receiving node locally estimates its interference-plus-noise covariance matrix, $\mathbf{\Phi}$. This metric is related to the MSE matrix which, when using an MMSE receiver, is given by $\bar{\mathbf{E}}_{k_g^d} = (\mathbf{I} + \mathbf{V}_{k_g^d}^H \mathbf{H}_{k_g^d,g}^H \mathbf{\Phi}_{k_g^d}^{-1} \mathbf{H}_{k_g^d,g} \mathbf{V}_{k_g^d})^{-1}$ for the DL and $\bar{\mathbf{E}}_{k_g^u} = (\mathbf{I} + \mathbf{V}_{k_g^u}^H \mathbf{H}_{g,k_g^u}^H \mathbf{\Phi}_{k_g^u}^{-1} \mathbf{H}_{g,k_g^u} \mathbf{V}_{k_g^u})^{-1}$ for the UL. Therefore, $\mathbf{\Phi}$ can be used to calculate $\mathbf{U}$ and $\mathbf{W}$, which can then be made available to the transmitting nodes to calculate $\mathbf{V}$. Thus, for a distributed implementation each node requires local CSI knowledge, resulting in a total of $2G\bar{K}M^2$ elements across all users. Additionally, $2G\bar{K}(Mb+b^2)$ elements per iteration need to be fedback to the transmitting nodes to account for all of $\mathbf{U}$ and $\mathbf{W}$.

5.8.2 Complexity analysis

Starting with the closed-form solutions, we evaluate the order of the number of flops required to calculate the optimisation variables using [101] which provides the number of flops required to perform standard mathematical operations. Taking (5.9) as an example, computing all receivers, $\mathbf{U}$, requires $\mathcal{O}\big(4G^2\bar{K}^2(6M^3 + 2Mb^2)\big)$ flops for multiplications inside the inverse if $\mathbf{\Phi}$ is unavailable or $\mathcal{O}\big(2G\bar{K}(4M^3 + 2Mb^2)\big)$ flops if $\mathbf{\Phi}$ is available, $\mathcal{O}(2G\bar{K}M^3)$ flops for the inverse and $\mathcal{O}(8G\bar{K}M^2b)$ flops for multiplying the inverse with the rest of the outside terms. To compute the weights, $\mathbf{W}$, we need to calculate the MSE. The interference-plus-noise covariance matrix, $\mathbf{\Phi}$, is already available since it was previously used in the calculation of $\mathbf{U}$, therefore we only need $\mathcal{O}(2G\bar{K}M^3)$ flops to calculate its inverse, and $\mathcal{O}\big(2G\bar{K}(4M^3 + 2M^2b + 2Mb^2)\big)$ flops for multiplication. Taking (5.21) as an example, the calculation of each precoder, $\mathbf{V}$, requires $\mathcal{O}\big(4G^2\bar{K}^2(6M^3 + 2M^2b + 2Mb^2)\big)$ flops for multiplications inside the inverse to compute $\mathbf{X}$, $\mathcal{O}(2G\bar{K}M^3)$ flops for the inverse, and $\mathcal{O}\big(2G\bar{K}(4M^2b + 2Mb^2)\big)$ flops for multiplying the inverse with the rest of the outside terms.

For the norm-bounded error model we solve a number of SDP problems, the

complexity of which is given by $\mathcal{O}(n^2 \sum_i^I m_i^2)$ [102]. Here, n represents the total size of the variables being solved for and I is the total number of constraints, with each constraint i being of dimension m_i. In our case the complexity can be expressed as $\mathcal{O}\big((x_1 + x_2)^2(z_1 + z_2)\big)$, where $x_2 = 8G^2\bar{K}^2$ and $z_2 = 4G^2\bar{K}^2(1 + M^2 + b^2 + 2Mb)^2$. When solving for $\mathbf{V}$, $x_1 = 2G\bar{K}Mb$ and $z_1 = (G\bar{K} + \bar{K})(Mb)^2$. When solving for $\mathbf{U}$, $x_1 = 2G\bar{K}Mb$ and $z_1 = 0$. When solving for $\mathbf{B}$ we have a Max-Det problem. This is of higher complexity than SDP, however using the SDP complexity as a lower bound we have $x_1 = 2G\bar{K}b^2$ and $z_1 = 0$.

In Section 5.6 we solve an SOCP problem to obtain $\mathbf{V}$. The complexity of solving a general SOCP problem is given by $\mathcal{O}(n^2 \sum_i^I m_i)$ [102], where the significance of the terms is the same as for the SDP complexity expression. Applying this to our problem we have $\mathcal{O}\big((2G\bar{K}Mb + G\bar{K})^2(G\bar{K}Mb + \bar{K}Mb + 8G^2\bar{K}^2Mb^3 + 4G^2\bar{K}^2M^2b^2 + 2)\big)$.

5.9 Conclusion

In this chapter we have addressed filter design for WSR maximisation in multi-user multi-cell MIMO networks with FD BSs and HD users, taking into consideration CCI, and transmitter and receiver distortion. Since WSR problems are non-convex, we transformed them into WMMSE problems and proposed alternating optimisation algorithms that are guaranteed to converge. Using the perfect CSI design as a starting point, we also considered robust beamformer design under two types of CSI error, namely norm-bounded error and stochastic CSI error. Simulation results for small cell scenarios show that replacing standard HD BSs with FD ones within this context can indeed increase achievable sum rate for low to intermediate distortion levels, and also confirm the robustness of the imperfect CSI designs. Additionally, we also proposed a DL rate maximisation problem subject to each UL user achieving a desired target rate, which can be used in cases where it is important for each UL user to be equally served in every time slot.

Appendix 5.A

Useful Lemma

Lemma 5.1. *[103] Let* $\mathbf{A}$*,* $\mathbf{B}$ *and* $\mathbf{C}$ *be given matrices, with* $\mathbf{A} = \mathbf{A}^H$*. Then, the relation* $\mathbf{A} \succeq \mathbf{B}^H\mathbf{D}\mathbf{C} + \mathbf{C}^H\mathbf{D}^H\mathbf{B}\ \forall\ \mathbf{D} : \|\mathbf{D}\| \leq \xi$ *is valid if, and only if, there*

exists $\nu \geq 0$ *such that*

$$\begin{bmatrix} \mathbf{A} - \nu \mathbf{C}^H \mathbf{C} & -\xi \mathbf{B}^H \\ -\xi \mathbf{B} & \nu \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \, .$$

Appendix 5.B
Additional details for proof of Theorem 5.2

5.B.1 Handling $\Delta_{k_g^d j}$ terms

Considering the cost function in (5.27) from the perspective of $\Delta_{k_g^d,j}$, we only need to focus on

$$\begin{aligned} T_{2,k_g j} = \sum_{i=1}^{K_j^d} \Bigg[& \mathrm{Tr}\left(\mathbf{B}_{k_g^d} (\mathbf{U}_{k_g^d} \mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} - \delta_{i,j}^{k,g} \mathbf{I})(\mathbf{U}_{k_g^d} \mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} - \delta_{i,j}^{k,g} \mathbf{I})^H \mathbf{B}_{k_g^d}^H \right) \\ & + \mathrm{Tr}\left(\kappa_B \mathbf{B}_{k_g^d} \mathbf{U}_{k_g^d} \mathbf{H}_{k_g^d,j} \mathrm{diag}(\mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H) \mathbf{H}_{k_g^d,j}^H \mathbf{U}_{k_g^d}^H \mathbf{B}_{k_g^d}^H \right) \\ & + \mathrm{Tr}\left(\iota_U \mathbf{B}_{k_g^d} \mathbf{U}_{k_g^d} \mathrm{diag}(\mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \mathbf{H}_{k_g^d,j}^H) \mathbf{U}_{k_g^d}^H \mathbf{B}_{k_g^d}^H \right) \Bigg] \\ = \sum_{i=1}^{K_j^d} \Bigg[& \mathrm{Tr}\left(\mathbf{B}_{k_g^d} (\mathbf{U}_{k_g^d} \mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} - \delta_{i,j}^{k,g} \mathbf{I})(\mathbf{U}_{k_g^d} \mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} - \delta_{i,j}^{k,g} \mathbf{I})^H \mathbf{B}_{k_g^d}^H \right) \\ & + \sum_{n=1}^{M_B} \mathrm{Tr}\left(\kappa_B \mathbf{B}_{k_g^d} \mathbf{U}_{k_g^d} \mathbf{H}_{k_g^d,j} \mathbf{S}_n \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \mathbf{S}_n^H \mathbf{H}_{k_g^d,j}^H \mathbf{U}_{k_g^d}^H \mathbf{B}_{k_g^d}^H \right) \\ & + \sum_{n=1}^{M_d} \mathrm{Tr}\left(\iota_U \mathbf{B}_{k_g^d} \mathbf{U}_{k_g^d} \mathbf{S}_n \mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \mathbf{H}_{k_g^d,j}^H \mathbf{S}_n^H \mathbf{U}_{k_g^d}^H \mathbf{B}_{k_g^d}^H \right) \Bigg] . \end{aligned}$$

This can be rewritten as

$$T_{2,k_g j} = \| \boldsymbol{\omega}_{2,k_g,j} + \boldsymbol{\Omega}_{2,k_g,j} \mathrm{vec}(\Delta_{k_g^d j}) \|^2 \leq m_{2,k_g j} \tag{5.56}$$

where $\boldsymbol{\omega}_{2,k_g,j}$ and $\boldsymbol{\Omega}_{2,k_g,j}$ are defined in Theorem 5.2, and $m_{2,k_g j}$ is a slack variable. The introduction of $m_{2,k_g j}$, allows to express the inner minimisation in (5.27) from the perspective of each $\Delta_{k_g^d,j}$ as

$$\begin{aligned} \max_{m} \quad & -m_{2,k_g j} \\ \text{s.t.} \quad & \|\boldsymbol{\omega}_{2,k_g,j} + \boldsymbol{\Omega}_{2,k_g,j}\text{vec}(\Delta_{k_g^d j})\|^2 \leq m_{2,k_g j} \\ & \forall\{\Delta_{k_g^d,j} : \|\text{vec}(\Delta_{k_g^d j})\| \leq \varepsilon_{k_g^d,j}\}\,. \end{aligned} \tag{5.57}$$

Representing (5.56) as

$$m_{2,k_g j} - \left(\boldsymbol{\omega}_{2,k_g,j} + \boldsymbol{\Omega}_{2,k_g,j}\text{vec}(\Delta_{k_g^d j}))\right)^H \mathbf{I}(\boldsymbol{\omega}_{2,k_g,j} + \boldsymbol{\Omega}_{2,k_g,j}\text{vec}(\Delta_{k_g^d j})) \geq 0$$

we can use the Schur Complement Lemma, to rewrite the constraints of (5.57) as

$$\begin{bmatrix} m_{2,k_g j} & \boldsymbol{\omega}_{2,k_g,j}^H \\ \boldsymbol{\omega}_{2,k_g,j} & \mathbf{I} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \text{vec}(\Delta_{k_g^d j})^H \boldsymbol{\Omega}_{2,k_g,j}^H \\ \boldsymbol{\Omega}_{2,k_g,j}\text{vec}(\Delta_{k_g^d j}) & \mathbf{0} \end{bmatrix} \succeq \mathbf{0}\,.$$

Additionally, applying Lemma 5.1 from Appendix 5.A with $\xi = \varepsilon_{k_g^d,j}$, $\mathbf{B} = [\mathbf{0}\ \boldsymbol{\Omega}_{2,k_g j}^H]$, $\mathbf{C} = [-1\ \mathbf{0}]$, $\mathbf{D} = \text{vec}(\Delta_{k_g^d,j})$ and

$$\mathbf{A} = \begin{bmatrix} m_{2,k_g j} & \boldsymbol{\omega}_{2,k_g j}^H \\ \boldsymbol{\omega}_{2,k_g j} & \mathbf{I} \end{bmatrix}$$

this can be further represented as

$$\nu_{2,k_g j} \geq 0, \quad \begin{bmatrix} m_{2,k_g j} - \nu_{2,k_g j} & \boldsymbol{\omega}_{2,k_g j}^H & \mathbf{0} \\ \boldsymbol{\omega}_{2,k_g j} & \mathbf{I} & -\varepsilon_{k_g^d,j}\boldsymbol{\Omega}_{2,k_g j} \\ \mathbf{0} & -\varepsilon_{k_g^d,j}\boldsymbol{\Omega}_{2,k_g j}^H & \nu_{2,k_g j}\mathbf{I} \end{bmatrix} \succeq \mathbf{0}\,.$$

5.B.2 Handling Δ_{g,i_j^u} terms

From the perspective of Δ_{g,i_j^u} we only need focus on

$$\begin{aligned} T_{3,gi_j} = \sum_{k=1}^{K_g^u} \Big[& \text{Tr}\left(\mathbf{B}_{k_g^u}(\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u} - \delta_{i,j}^{k,g}\mathbf{I})(\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u} - \delta_{i,j}^{k,g}\mathbf{I})\mathbf{B}_{k_g^u}^H\right) \\ & + \text{Tr}\left(\kappa_U \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\text{diag}(\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H)\mathbf{H}_{g,i_j^u}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\ & + \text{Tr}\left(\iota_B \mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\text{diag}(\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H)\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right)\Big] \\ = \sum_{k=1}^{K_g^u} \Big[& \text{Tr}\left(\mathbf{B}_{k_g^u}(\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u} - \delta_{i,j}^{k,g}\mathbf{I})(\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u} - \delta_{i,j}^{k,g}\mathbf{I})\mathbf{B}_{k_g^u}^H\right) \end{aligned}$$

$$+\sum_{n=1}^{M_u}\mathrm{Tr}\left(\kappa_U\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,i_j^u}\mathbf{S}_n\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{S}_n^H\mathbf{H}_{g,i_j^u}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right)$$
$$+\sum_{n=1}^{M_B}\mathrm{Tr}\left(\iota_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{S}_n\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H\mathbf{S}_n^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right)\Bigg]$$

from the the cost function of (5.27). This can be expressed as

$$T_{3,gi_j}=\|\boldsymbol{\omega}_{3,g,i_j}+\boldsymbol{\Omega}_{3,g,i_j}\mathrm{vec}(\Delta_{g,i_j^u})\|^2\leq m_{3,gi_j} \tag{5.58}$$

where $\boldsymbol{\omega}_{3,gi_j}$ and $\boldsymbol{\Omega}_{3,gi_j}$ are defined in Theorem 5.2, and m_{3,gi_j} is a slack variable. Thus, from the perspective of each Δ_{g,i_j^u}, we can represent the inner minimisation in (5.27) as

$$\begin{aligned}\max_{m}\quad & -m_{3,gi_j}\\ \text{s.t.}\quad & \|\boldsymbol{\omega}_{3,gi_j}+\boldsymbol{\Omega}_{3,gi_j}\mathrm{vec}(\Delta_{g,i_j^u})\|^2\leq m_{3,gi_j}\\ & \forall\{\Delta_{g,i_j^u}:\|\mathrm{vec}(\Delta_{g,i_j^u})\|\leq\varepsilon_{g,i_j^u}\}\,.\end{aligned} \tag{5.59}$$

Additionally, expressing (5.58) as

$$m_{3,gi_j}-\big(\boldsymbol{\omega}_{3,gi_j}+\boldsymbol{\Omega}_{3,gi_j}\mathrm{vec}(\Delta_{g,i_j^u}))\big)^H\mathbf{I}\big(\boldsymbol{\omega}_{3,gi_j}+\boldsymbol{\Omega}_{3,gi_j}\mathrm{vec}(\Delta_{g,i_j^u}))\big)\geq 0$$

we can use the Schur Complement Lemma to represent the constraints of (5.59) as

$$\begin{bmatrix} m_{3,gi_j} & \boldsymbol{\omega}_{3,gi_j}^H\\ \boldsymbol{\omega}_{3,gi_j} & \mathbf{I}\end{bmatrix}+\begin{bmatrix}\mathbf{0} & \mathrm{vec}(\Delta_{g,i_j^u})^H\boldsymbol{\Omega}_{3,gi_j}^H\\ \boldsymbol{\Omega}_{3,gi_j}\mathrm{vec}(\Delta_{g,i_j^u}) & \mathbf{0}\end{bmatrix}\succeq\mathbf{0}\,.$$

Applying Lemma 5.1 from Appendix 5.A with $\xi=\varepsilon_{g,i_j^u}$, $\mathbf{B}=[\mathbf{0}\ \boldsymbol{\Omega}_{3,gi_j}^H]$, $\mathbf{C}=[-1\ \mathbf{0}]$, $\mathbf{D}=\mathrm{vec}(\Delta_{g,i_j^u})$ and

$$\mathbf{A}=\begin{bmatrix} m_{3,gi_j} & \boldsymbol{\omega}_{3,gi_j}^H\\ \boldsymbol{\omega}_{3,gi_j} & \mathbf{I}\end{bmatrix}$$

this constraint can be represented as

$$\nu_{3,gi_j}\geq 0,\quad\begin{bmatrix} m_{3,gi_j}-\nu_{3,gi_j} & \boldsymbol{\omega}_{3,gi_j}^H & \mathbf{0}\\ \boldsymbol{\omega}_{3,gi_j} & \mathbf{I} & -\varepsilon_{g,i_j^u}\boldsymbol{\Omega}_{3,gi_j}\\ \mathbf{0} & -\varepsilon_{g,i_j^u}\boldsymbol{\Omega}_{3,gi_j}^H & \nu_{3,gi_j}\mathbf{I}\end{bmatrix}\succeq\mathbf{0}.$$

5.B.3 Handling $\Delta_{g,j}$ terms

Considering the cost function of (5.27), from the perspective of $\Delta_{g,j}$ we are only concerned with

$$\begin{aligned}
T_{4,gj} = \sum_{k=1}^{K_g^u}\sum_{i=1}^{K_j^d} \Big[& \vartheta_j^g \mathrm{Tr}\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
& + (1-\vartheta_j^g)\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\Delta_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\Delta_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
& + \mathrm{Tr}\left(\kappa_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j}\mathrm{diag}(\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H)\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
& + \mathrm{Tr}\left(\iota_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathrm{diag}(\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H)\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right)\Big] \\
= \sum_{k=1}^{K_g^u}\sum_{i=1}^{K_j^d} \Big[& \vartheta_j^g \mathrm{Tr}\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
& + (1-\vartheta_j^g)\left(\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\Delta_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\Delta_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
& + \sum_{n=1}^{M_B}\mathrm{Tr}\left(\kappa_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{H}_{g,j}\mathbf{S}_n\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{S}_n^H\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right) \\
& + \sum_{n=1}^{M_B}\mathrm{Tr}\left(\iota_B\mathbf{B}_{k_g^u}\mathbf{U}_{k_g^u}\mathbf{S}_n\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H\mathbf{S}_n^H\mathbf{U}_{k_g^u}^H\mathbf{B}_{k_g^u}^H\right)\Big].
\end{aligned}$$

This can be further expressed as

$$T_{4,gj} = \|\boldsymbol{\omega}_{4,gj} + \boldsymbol{\Omega}_{4,gj}\mathrm{vec}(\Delta_{g,j})\|^2 \leq m_{4,gj} \tag{5.60}$$

where $\boldsymbol{\omega}_{4,g,j}$ and $\boldsymbol{\Omega}_{4,g,j}$ are defined in Theorem 5.2, and $m_{4,gj}$ is a slack variable. The use of this slack variable allows to represent the inner minimisation in (5.27) from the perspective of each $\Delta_{g,j}$ as

$$\begin{aligned}
\max_{m} \quad & -m_{4,gj} \\
\text{s.t.} \quad & \|\boldsymbol{\omega}_{4,gj} + \boldsymbol{\Omega}_{4,gj}\mathrm{vec}(\Delta_{g,j})\|^2 \leq m_{4,gj} \\
& \forall\{\Delta_{g,j} : \|\mathrm{vec}(\Delta_{g,j})\| \leq \varepsilon_{g,j}\}.
\end{aligned} \tag{5.61}$$

Next, representing (5.60) as

$$m_{4,gj} - \big(\boldsymbol{\omega}_{4,gj} + \boldsymbol{\Omega}_{4,gj}\mathrm{vec}(\Delta_{g,j})\big)^H\mathbf{I}\big(\boldsymbol{\omega}_{4,gj} + \boldsymbol{\Omega}_{4,gj}\mathrm{vec}(\Delta_{g,j})\big) \geq 0$$

and applying the Schur Complement Lemma, the constraints of (5.61) can be expressed as

$$\begin{bmatrix} m_{4,gj} & \boldsymbol{\omega}_{4,gj}^H \\ \boldsymbol{\omega}_{4,gj} & \mathbf{I} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathrm{vec}(\Delta_{g,j})^H \boldsymbol{\Omega}_{4,gj}^H \\ \boldsymbol{\Omega}_{4,gj}\mathrm{vec}(\Delta_{g,j}) & \mathbf{0} \end{bmatrix} \succeq \mathbf{0}\,.$$

Finally, using Lemma 5.1 from Appendix 5.A with $\xi = \varepsilon_{g,j}$, $\mathbf{B} = [\mathbf{0}\ \boldsymbol{\Omega}_{4,gj}^H]$, $\mathbf{C} = [-1\ \mathbf{0}]$, $\mathbf{D} = \mathrm{vec}(\Delta_{g,j})$ and

$$\mathbf{A} = \begin{bmatrix} m_{4,gj} & \boldsymbol{\omega}_{4,gj}^H \\ \boldsymbol{\omega}_{4,gj} & \mathbf{I} \end{bmatrix}$$

this constraint can be represented as

$$\nu_{4,gj} \geq 0, \quad \begin{bmatrix} m_{4,gj} - \nu_{4,gj} & \boldsymbol{\omega}_{4,gj}^H & \mathbf{0} \\ \boldsymbol{\omega}_{4,gj} & \mathbf{I} & -\varepsilon_{g,j}\boldsymbol{\Omega}_{4,gj} \\ \mathbf{0} & -\varepsilon_{g,j}\boldsymbol{\Omega}_{4,gj}^H & \nu_{4,gj}\mathbf{I} \end{bmatrix} \succeq \mathbf{0}\,.$$

Chapter 6

Interference Alignment for Full-Duplex MIMO Networks with Imperfect CSI

6.1 Introduction

This chapter merges the use of IA and FD operation in order to provide a dual approach solution to address the spectrum demand problem in future generation wireless networks. The system considered is the multi-cell multi-user network with FD BSs and HD DL and UL users from Chapter 5, however, for the purpose of this chapter we consider the use of linear IA to manage the resultant interference under imperfect CSI.

There have been a number of information-theoretic studies with the aim of understanding the fundamental capacity limits of FD systems, particularly the characterisation of achievable DoF. These DoF studies exploit a variety of interference management solutions in order to maximise capacity. For example, [105] studies the DoF region for single-cell systems with one multi-antenna FD BS and K single antenna FD users, and proposes an achievable scheme based on ergodic IA. The DoF regions for an FD BS communicating with HD users [106, 107] and a point-to-point MIMO FD link [107] have also been studied, with the authors proposing achievable schemes based on asymptotic IA for each scenario. A simi-

Work from this chapter has been presented at IEEE ICC 2017 [104].

lar study was carried out in [108], where systems with an FD BS and either FD or HD users are considered; for each case the sum DoF are characterised, and achievable schemes based on a combination of interference nulling and asymptotic IA are proposed.

While the ergodic and asymptotic IA techniques exploited in the FD literature mentioned so far are beneficial from a theoretical standpoint, they are difficult to implement in practice as highlighted earlier in Section 2.4.1.1. Therefore, here we consider the use of linear IA which offers a more practical alternative [5]. The application of linear IA in FD enabled networks has previously been considered for single-cell MIMO systems with FD BSs and HD users in [8], where IA feasibility conditions are derived. Additionally, in [109] the authors derive a scaling law for the multiplexing gain of FD over HD in a multi-cell setting with network MIMO capability between the BSs, allowing for the BS-to-BS interference to be ignored. Different to [8,105–109] we consider an imperfect CSI scenario; moreover, moving beyond the no BS-to-BS interference assumption in the only multi-cell study available so far [109], we consider a fully connected multi-cell multi-user network, which leads to a more complex interference scenario.

Within this context, work in this chapter seeks to characterise the impact of imperfect CSI on system performance, and propose linear IA algorithms applicable to networks equipped with FD BSs serving HD users (both single-cell and multi-cell), since none are available in literature so far. Using the CSI error model previously applied in Chapter 4, in the first part of this chapter we derive bounds on the sum rate loss and achievable DoF. Results show that when the CSI error variance depends on SNR to the power of a proper fraction, full DoF cannot be achieved and the asymptotic sum rate loss is unbounded. Additionally, when the error is exactly inversely proportional to SNR, full DoF are achievable and the asymptotic rate loss is bounded by a fixed value dependent on the number of cells, UL users, DL users, required streams and the CSI error parameters.

In the second part of the chapter we propose two novel linear IA algorithms applicable to the FD network being considered. The proposed algorithms are inspired by techniques originally proposed for the HD IC, namely the MMSE algorithm from [11] and the Max-SINR algorithm from [10]. The resultant solutions are not straightforward extensions of the original HD ones; they are based on specific design criteria that separate the various interference components amongst the different available beamformers, and exploit statistical knowledge of the CSI

error leading to a more robust design. Moreover, they are designed to produce unitary beamformers. The use of unitary beamformers has gained significant attention in recent years due to its role in codebook design for limited feedback scenarios [110]. Additionally, it has been shown to lower complexity for MMSE based algorithms [111], and improve performance for Max-SINR based ones [66]. The algorithms are first derived for the single-cell case, since the feasibility of linear IA in such networks is already known [8], and later extended to the multi-cell case. Furthermore, we show that the two proposed algorithms produce equivalent beamformers under certain conditions. For the multi-cell case, we also derive the proper condition for IA feasibility. This condition, along with the multi-cell version of the algorithms, can serve as tools to help future efforts into the determination of a full set of linear IA feasibility conditions for multi-cell FD enabled systems.

The rest of this chapter is organised as follows. Section 6.2 presents some preliminaries, including the system model, the performance of IA with perfect CSI and the CSI error model. Section 6.3 deals with the performance of IA under imperfect CSI, presenting two theorems that characterise the sum rate and DoF loss. Next, in Section 6.4, we derive the MMSE and Max-SINR based algorithms for the single-cell context, and establish an equivalence between the two. Section 6.5 deals with multiple cell considerations, here we derive a proper condition for multi-cell systems and also present the multi-cell extensions for the IA algorithms. Section 6.6 presents simulation results, and finally conclusions are provided in Section 6.7. Additionally, there is an appendix which includes some lemmas used wtihin this chapter.

6.2 Preliminaries

6.2.1 System model

We consider a G-cell scenario, where each cell g has one FD BS, K_d DL users requiring b_d streams each and K_u UL users requiring b_u streams each. A simplified version of this network with one DL and one UL user per cell is depicted in Fig. 6.1. BSs are equipped with M_B FD antennas, DL users are equipped with N_d HD antennas and UL users are equipped with N_u HD antennas.

The signal received at user k_g^d, the kth DL user in cell g, and at BS g are

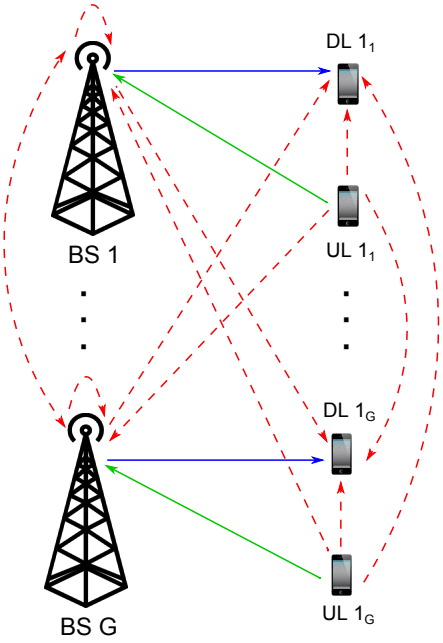


Figure 6.1: G-cell network with an FD BS, and one DL and one UL user per cell. Solid arrows represent desired links, while dashed ones represent interference links.

given by (6.1) and (6.2) respectively.

$$\mathbf{y}_{k_g^d} = \sum_{j=1}^{G}\sum_{i=1}^{K_d} \mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} + \sum_{j=1}^{G}\sum_{i=1}^{K_u} \mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u} + \mathbf{n}_{k_g^d} \tag{6.1}$$

$$\mathbf{y}_g = \sum_{j=1}^{G}\sum_{i=1}^{K_u} \mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u} + \sum_{j=1}^{G}\sum_{i=1}^{K_d} \mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} + \mathbf{n}_g \tag{6.2}$$

Here, $\mathbf{H}_{k_g^d,j} \in \mathbb{C}^{N_d \times M_B}$ represents the channel from BS j to DL user k_g^d, $\mathbf{H}_{k_g^d,i_j^u} \in \mathbb{C}^{N_d \times N_u}$ is the channel from UL user i_j^u to DL user k_g^d, $\mathbf{H}_{g,j} \in \mathbb{C}^{M_B \times M_B}$ is the channel from BS j to BS g, and $\mathbf{H}_{g,i_j^u} \in \mathbb{C}^{M_B \times N_u}$ is the channel from UL user i_j^u to BS g. All channel elements are drawn from a complex normal distribution with zero mean and variance one. $\mathbf{V}_{i_j^d} \in \mathbb{C}^{M_B \times b_d}$ is the precoder for $\mathbf{s}_{i_j^d}$, with $\mathbf{s}_{i_j^d} \in \mathbb{C}^{b_d \times 1}$ being the data intended for the ith DL user in cell j, such that $\mathbb{E}\{\mathbf{s}_{i_j^d}\mathbf{s}_{i_j^d}^H\} = P\mathbf{I}$. $\mathbf{V}_{i_j^u} \in \mathbb{C}^{N_u \times b_u}$ is the precoder for $\mathbf{s}_{i_j^u} \in \mathbb{C}^{b_u \times 1}$, with $\mathbf{s}_{i_j^u}$ being the data transmitted by the ith UL user in cell j, such that $\mathbb{E}\{\mathbf{s}_{i_j^u}\mathbf{s}_{i_j^u}^H\} = P\mathbf{I}$. Moreover, $\mathbf{n}_{k_g^d}$ and $\mathbf{n}_g$ represent AWGN with zero mean and variance σ^2.

The estimated DL and UL received signals are given by

$$\widehat{\mathbf{s}}_{k_g^d} = \sum_{j=1}^{G}\sum_{i=1}^{K_d} \mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} + \sum_{j=1}^{G}\sum_{i=1}^{K_u} \mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u} + \mathbf{U}_{k_g^d}^H\mathbf{n}_{k_g^d} \tag{6.3}$$

$$\widehat{\mathbf{s}}_{k_g^u} = \sum_{j=1}^{G}\sum_{i=1}^{K_u} \mathbf{U}_{k_g^u}^H \mathbf{H}_{g,i_j^u} \mathbf{V}_{i_j^u} \mathbf{s}_{i_j^u} + \sum_{\substack{j=1 \\ j\neq g}}^{G}\sum_{i=1}^{K_d} \mathbf{U}_{k_g^u}^H \mathbf{H}_{g,j} \mathbf{V}_{i_j^d} \mathbf{s}_{i_j^d} + \mathbf{U}_{k_g^u}^H \mathbf{n}_g + \underbrace{\Theta \sum_{i=1}^{K_d} \mathbf{U}_{k_g^u}^H \boldsymbol{\Upsilon}_{g,g} \mathbf{V}_{i_g^d} \mathbf{s}_{i_g^d}}_{\substack{\text{residual SI for imperfect} \\ \text{SI cancellation scenarios}}} \tag{6.4}$$

where $\mathbf{U}_{k_g^d} \in \mathbb{C}^{N_d \times b_d}$ is the receive beamformer applied at DL user k_g^d and $\mathbf{U}_{k_g^u} \in \mathbb{C}^{M_B \times b_u}$ is the receive beamformer applied at BS g to extract the data transmitted by UL user k_g^u. Here, (6.3) is obtained as $\mathbf{U}_{k_g^d}^H \mathbf{y}_{k_g^d}$. For the perfect CSI case $\mathbf{H}_{g,g} \sum_{i=1}^{K_d} \mathbf{V}_{i_g^d} \mathbf{s}_{i_g^d}$ is known at BS g, thus we obtain (6.4) with $\Theta = 0$ as $\mathbf{U}_{k_g^u}^H (\mathbf{y}_g - \mathbf{H}_{g,g} \sum_{i=1}^{K_d} \mathbf{V}_{i_g^d} \mathbf{s}_{i_g^d})$. The parameter Θ is a binary term used to differentiate between perfect SI cancellation and imperfect SI cancellation. For perfect CSI, similar to other FD DoF studies [8, 105–108, 112], we assume that SI is always perfectly cancelled, therefore $\Theta = 0$. For imperfect CSI, perfect SI cancellation is not guaranteed leading to a residual SI term, further details are provided in Section 6.2.3.

6.2.2 Achievable sum rate and DoF with perfect CSI

Whilst our analysis is concerned with imperfect CSI, since our aim is to characterise performance loss, it is useful to first define achievable sum rate and DoF with perfect CSI. The linear IA conditions for a multi-cell FD system with HD users are given by

$$\begin{aligned}
|\mathbf{u}_{k_g^d}^{n\,H} \mathbf{H}_{k_g^d,g} \mathbf{v}_{k_g^d}^n| &> 0 \quad \forall\, n,k,g \\
|\mathbf{u}_{k_g^u}^{n\,H} \mathbf{H}_{g,k_g^u} \mathbf{v}_{k_g^u}^n| &> 0 \quad \forall\, n,k,g \\
\mathbf{u}_{k_g^d}^{n\,H} \mathbf{H}_{k_g^d,j} \mathbf{v}_{i_j^d}^m &= 0 \quad \forall\, n,m,k,i,g,j,(n,k,g \neq m,i,j) \\
\mathbf{u}_{k_g^u}^{n\,H} \mathbf{H}_{g,j} \mathbf{v}_{i_j^d}^m &= 0 \quad \forall\, n,m,k,i,g,j,(g \neq j) \\
\mathbf{u}_{k_g^d}^{n\,H} \mathbf{H}_{k_g^d,i_j^u} \mathbf{v}_{i_j^u}^m &= 0 \quad \forall\, n,m,k,i,g,j \\
\mathbf{u}_{k_g^u}^{n\,H} \mathbf{H}_{g,i_j^u} \mathbf{v}_{i_j^u}^m &= 0 \quad \forall\, n,m,k,i,g,j,(n,k,g \neq m,i,j)
\end{aligned} \tag{6.5}$$

where $\mathbf{u}_{k_g^d}^n$, $\mathbf{u}_{k_g^u}^n$, $\mathbf{v}_{k_g^d}^n$ and $\mathbf{v}_{k_g^u}^n$ refer to the nth column of $\mathbf{U}_{k_g^d}$, $\mathbf{U}_{k_g^u}$, $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$ respectively.

Considering i.i.d. Gaussian inputs, and the fact that there is no interference leakage for perfect CSI, the achievable sum rate across the whole network is

$$
\begin{aligned}
R_{\text{FD TOT}} &= R_{\text{FD DL}} + R_{\text{FD UL}} \\
&= \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d} \log_2\left(1 + \frac{P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}|^2}{\sigma^2}\right) \\
&\quad + \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u} \log_2\left(1 + \frac{P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}|^2}{\sigma^2}\right). \qquad (6.6)
\end{aligned}
$$

Assuming that the system configuration is such that IA is feasible, we can use the rate to DoF relationship in (4.10) to calculate the total achievable DoF with perfect CSI as

$$D_{\text{FD TOT}} = G(K_d b_d + K_u b_u)\,. \qquad (6.7)$$

6.2.3 Imperfect CSI considerations

In this chapter we are concerned with the effect of imperfect CSI on IA performance, thus analogous to the CSI error model used for the HD MIMO IBC outlined in Section 4.2.2, we apply the following model for the CSI mismatch

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{E} \qquad (6.8)$$

where $\widehat{\mathbf{H}}$ represents the available imperfect CSI, $\text{vec}(\mathbf{H}) \sim \mathcal{CN}(0, \mathbf{I})$ is the perfect channel matrix and $\mathbf{E}$ is the error matrix representing the degree of inaccuracy in the available CSI. Matrix $\mathbf{E}$ is assumed to be independent of $\mathbf{H}$ and is modeled as $\text{vec}(\mathbf{E}) \sim \mathcal{CN}(0, \eta\mathbf{I})$, where $\eta = \beta\rho^{-\alpha}$ with $\rho = \frac{P}{\sigma^2}$ representing the nominal SNR, $\alpha \geq 0$ and $\beta > 0$.

Conditioned on $\widehat{\mathbf{H}}$, $\mathbf{H}$ is Gaussian distributed with mean $\widehat{\mathbf{H}}/(1+\eta)$ and statistically independent elements of variance $\eta/(1+\eta)$ [72], allowing us to express

$$\mathbf{H} = \frac{1}{1+\eta}\widehat{\mathbf{H}} + \boldsymbol{\Upsilon} \qquad (6.9)$$

where $\boldsymbol{\Upsilon}$ is independent of $\widehat{\mathbf{H}}$ and distributed as $\text{vec}(\boldsymbol{\Upsilon}) \sim \mathcal{CN}\left(0, \frac{\eta}{1+\eta}\mathbf{I}\right)$.

Generally for DoF and IA related studies perfect SI cancellation is assumed [8, 105–109, 112]. This can also be applied to our scenario, where it may be assumed that each BS has perfect knowledge of its SI channel and imperfect CSI for the remaining channels. However, if the SI channel is also known imper-

fectly, then only $\widehat{\mathbf{H}}_{g,g}$ is available at BS g instead of $\mathbf{H}_{g,g}$, implying that only $\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,g}\sum_{i=1}^{K_d}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d}$ is subtracted from (6.2). This results in a residual SI component, with the estimated UL data $\widehat{\mathbf{s}}_{k_g^u}$, being given by (6.4) with $\Theta = 1$.

6.3 Performance under imperfect CSI

For imperfect CSI scenarios, only $\widehat{\mathbf{H}}$ is available for beamformer calculation. Therefore, instead of the original IA conditions in (6.5) the alignment conditions observed are

$$\begin{aligned}
|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}| &> 0 \quad \forall\, n,k,g \\
|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\widehat{\mathbf{H}}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}| &> 0 \quad \forall\, n,k,g \\
\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m} &= 0 \quad \forall\, n,m,k,i,g,j,(n,k,g\neq m,i,j) \\
\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\widehat{\mathbf{H}}_{g,j}\widehat{\mathbf{v}}_{i_j^d}^{m} &= 0 \quad \forall\, n,m,k,i,g,j,(g\neq j) \\
\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m} &= 0 \quad \forall\, n,m,k,i,g,j \\
\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\widehat{\mathbf{H}}_{g,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m} &= 0 \quad \forall\, n,m,k,i,g,j,(n,k,g\neq m,i,j)
\end{aligned} \tag{6.10}$$

where $\widehat{\mathbf{u}}_{k_g^d}^{n}$, $\widehat{\mathbf{u}}_{k_g^u}^{n}$, $\widehat{\mathbf{v}}_{k_g^d}^{n}$ and $\widehat{\mathbf{v}}_{k_g^u}^{n}$ refer to the nth column of the beamformers calculated with imperfect CSI, namely $\widehat{\mathbf{U}}_{k_g^d}$, $\widehat{\mathbf{U}}_{k_g^u}$, $\widehat{\mathbf{V}}_{k_g^d}$ and $\widehat{\mathbf{V}}_{k_g^u}$ respectively.

Satisfying the IA conditions in (6.10) instead of those in (6.5) leads to residual leakage interference, given by (6.11) for the DL and (6.12) for the UL.

$$\widehat{J}_{k_g^d} = \sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}|^2 + \sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m}|^2 \tag{6.11}$$

$$\begin{aligned}
\widehat{J}_{k_g^u} = &\sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m}|^2 + \sum_{\substack{j=1\\j\neq g}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,j}\widehat{\mathbf{v}}_{i_j^d}^{m}|^2 \\
&+ \Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{v}}_{i_g^d}^{m}|^2
\end{aligned} \tag{6.12}$$

Residual leakage has an adverse effect on achievable sum rate and DoF. Understanding the extent of this negative impact is crucial in order to obtain a more

realistic characterisation of the system performance. Here, we present two theorems that quantify this effect in terms of asymptotic sum rate loss and decrease in achievable DoF.

6.3.1 Sum rate loss

The mean sum rate loss, ΔR_{FD}, is defined as the difference between the expected value for the sum rate achievable with perfect CSI from (6.6) and the expected value for the sum rate achievable with imperfect CSI defined as

$$\begin{aligned}\widehat{R}_{\text{FD TOT}} &= \widehat{R}_{\text{FD DL}} + \widehat{R}_{\text{FD UL}} \\ &= \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}|^2}{\widehat{J}_{k_g^d}+\sigma^2}\right) \\ &\quad+\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}|^2}{\widehat{J}_{k_g^u}+\sigma^2}\right). \end{aligned} \tag{6.13}$$

Given this definition we can now refer to the following theorem.

Theorem 6.1. *Consider a G-cell system where each cell has one FD BS, K_d DL users requiring b_d streams each and K_u UL users requiring b_u streams each. For this system, under imperfect CSI with error variance $\eta = \beta\rho^{-\alpha}$, at asymptotically high SNR: ΔR_{FD} tends to zero for $\alpha > 1$, tends to infinity for $0 \leq \alpha < 1$ and for $\alpha = 1$ is finite and upper bounded by Ω, defined as*

$$\begin{aligned}\Omega = {} & GK_db_d\Big(\log_2\big(1+\beta(GK_db_d+GK_ub_u-1)\big)\Big) \\ & + GK_ub_u\Big(\log_2\big(1+\beta(GK_db_d+GK_ub_u-1-\bar{\Theta}K_db_d)\big)\Big)\end{aligned}$$

where $\bar{\Theta} = \text{not}(\Theta)$, such that $\bar{\Theta} = 1$ for perfect SI cancellation and $\bar{\Theta} = 0$ for imperfect SI cancellation. Thus, the overall asymptotic sum rate loss can be summarised as

$$\lim_{SNR\to\infty}\Delta R_{FD}\begin{cases}=0 & \alpha>1\\ \leq\Omega & \alpha=1\\ \to\infty & 0\leq\alpha<1\,.\end{cases} \tag{6.14}$$

Proof. Considering that

$$\Delta R_{FD} = \mathbb{E}_{\mathbf{H}}\{R_{\text{FD TOT}}\} - \mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}_{\text{FD TOT}}\}\} \tag{6.15}$$

we can replace $R_{\text{FD TOT}}$ with (6.6) and $\widehat{R}_{\text{FD TOT}}$ with (6.13) to obtain

$$\begin{aligned}
\Delta R_{\text{FD}} = {} & \mathbb{E}_{\mathbf{H}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d} \log_2\left(1+\frac{P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}|^2}{\sigma^2}\right)\right\} \\
& - \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d} \log_2\left(1+\frac{P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}|^2}{\widehat{J}_{k_g^d}+\sigma^2}\right)\right\}\right\} \\
& + \mathbb{E}_{\mathbf{H}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u} \log_2\left(1+\frac{P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}|^2}{\sigma^2}\right)\right\} \\
& - \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u} \log_2\left(1+\frac{P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}|^2}{\widehat{J}_{k_g^u}+\sigma^2}\right)\right\}\right\}.
\end{aligned} \tag{6.16}$$

After some algebraic manipulations, this can be further represented as

$$\begin{aligned}
\Delta R_{\text{FD}} = {} & \mathbb{E}_{\mathbf{H}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d} \log_2\left(1+\frac{P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}|^2}{\sigma^2}\right)\right\} \\
& + \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d} \log_2\left(1+\frac{\widehat{J}_{k_g^d}}{\sigma^2}\right)\right\}\right\} \\
& - \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d} \log_2\left(1+\frac{\widehat{J}_{k_g^d}+P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}|^2}{\sigma^2}\right)\right\}\right\} \\
& + \mathbb{E}_{\mathbf{H}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u} \log_2\left(1+\frac{P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}|^2}{\sigma^2}\right)\right\} \\
& + \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u} \log_2\left(1+\frac{\widehat{J}_{k_g^u}}{\sigma^2}\right)\right\}\right\} \\
& - \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{ \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u} \log_2\left(1+\frac{\widehat{J}_{k_g^u}+P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}|^2}{\sigma^2}\right)\right\}\right\}.
\end{aligned} \tag{6.17}$$

Additionally, since for unitary beamformers, analogous to [73, Lemma 2] it can be shown that $|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}|^2$, $|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}|^2$, $|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}|^2$ and $|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}|^2$ are exponentially distributed with both mean and variance one, we can establish the following inequalities.

$$\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g^d}+P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}|^2}{\sigma^2}\right)\right\}\right\}$$
$$\geq \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}|^2}{\sigma^2}\right)\right\} \tag{6.18}$$

$$\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\widehat{J}_{k_g^u}+P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}|^2}{\sigma^2}\right)\right\}\right\}$$
$$\geq \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}|^2}{\sigma^2}\right)\right\} \tag{6.19}$$

Considering the expression from (6.17), taking into account (6.18) and (6.19), and applying Jensen's inequality, we can express the sum rate loss as

$$\Delta R_{\,\mathrm{FD}} \leq \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}}{\sigma^2}\right)$$
$$+\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}}{\sigma^2}\right). \tag{6.20}$$

Therefore to quantify $\Delta R_{\,\mathrm{FD}}$ we need to find expressions for $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}$ and $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}$. Starting with $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}$, having already defined $\widehat{J}_{k_g^d}$ in (6.12), we can combine it with the channel model from (6.9) to obtain

$$\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\} = \sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g^d,j}+\boldsymbol{\Upsilon}_{k_g^d,j}\right)\widehat{\mathbf{v}}_{i_j^d}^{m}\right|^2\right\}$$
$$+\sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g^d,i_j^u}+\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\right)\widehat{\mathbf{v}}_{i_j^u}^{m}\right|^2\right\}.$$

Using the IA conditions for imperfect CSI in (6.10), particularly the fact that $\left[\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}=0\ \forall\ n,m,k,i,g,j\ \ (n,k,g\neq m,i,j)\right]$ and $\left[\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m}=0\ \forall\ n,m,k,i,g,j\right]$, this can be further simplified as

$$\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\} = \sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}\right|^2\right\}$$

$$+\sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m}\right|^2\right\}$$

$$\overset{(a)}{=} P\frac{\eta}{1+\eta}\left(GK_db_d - 1 + GK_ub_u\right) \tag{6.21}$$

where (a) follows by integrating the result of Lemma 6.1 from Appendix 6.A.

Using a similar process, the expected value for the UL interference leakage can be expressed as

$$\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\} = P\frac{\eta}{1+\eta}\left(GK_ub_u - 1 + (G-\bar{\Theta})K_db_d\right)\,. \tag{6.22}$$

Applying the results from (6.21) and (6.22) into (6.20), we obtain

$$\Delta R_{\,\mathrm{FD}} \leq \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{1+\eta}\left(GK_db_d - 1 + GK_ub_u\right)\right)$$
$$+\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{1+\eta}\left(GK_ub_u - 1 + (G-\bar{\Theta})K_db_d\right)\right)$$

which after evaluating the summations and replacing η with $\beta\rho^{-\alpha}$, becomes

$$\Delta R_{\,\mathrm{FD}} \leq GK_db_d\left[\log_2\left(1+(GK_db_d+GK_ub_u-1)\frac{\beta\rho^{1-\alpha}}{1+\beta\rho^{-\alpha}}\right)\right]$$
$$+GK_ub_u\left[\log_2\left(1+(GK_db_d+GK_ub_u-1-\bar{\Theta}K_db_d)\frac{\beta\rho^{1-\alpha}}{1+\beta\rho^{-\alpha}}\right)\right]\,.$$

Finally, taking a high SNR approximation of this inequality, the asymptotic sum rate loss can be defined as in (6.14), proving Theorem 6.1 as originally stated. □

6.3.2 DoF loss

The DoF loss, $\Delta D_{\,\mathrm{FD}}$, is defined as the difference between the DoF achievable with perfect CSI and the DoF achievable under imperfect CSI, expressed as

$$\widehat{D}_{\,\mathrm{FD\ TOT}} = \lim_{P\to\infty}\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}_{\,\mathrm{FD\ DL}}\}\}}{\log_2 P} + \lim_{P\to\infty}\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}_{\,\mathrm{FD\ UL}}\}\}}{\log_2 P}\,. \tag{6.23}$$

Given this definition we can now refer to the following theorem.

Theorem 6.2. *Consider a G-cell system where each cell has one FD BS, K_d DL*

users requiring b_d streams each and K_u UL users requiring b_u streams each. For this system, under imperfect CSI with error variance $\eta = \beta\rho^{-\alpha}$, full DoF can be achieved for $\alpha \geq 1$, while for $0 \leq \alpha < 1$ achievable DoF are equal to a fraction α of the full DoF, i.e. the overall DoF loss is given by

$$\Delta D_{\,FD} = \begin{cases} 0 & \alpha \geq 1 \\ (1-\alpha)\, G\,(K_d b_d + K_u b_u) & 0 \leq \alpha < 1\,. \end{cases} \tag{6.24}$$

Proof. The imperfect CSI DoF expression (6.23) can be transformed into (6.25) by replacing $\widehat{R}_{\,\text{FD DL}}$ and $\widehat{R}_{\,\text{FD UL}}$ with the corresponding expressions from (6.13) and performing some additional algebraic manipulations.

$$\begin{aligned}
\widehat{D}_{\,\text{FD TOT}} = & \underbrace{\lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\widehat{J}_{k_g^d} + \sigma^2 + P\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}\right|^2\right)\right\}\right\}}{\log_2 P}}_{\text{A}} \\
& + \underbrace{\lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\widehat{J}_{k_g^u} + \sigma^2 + P\left|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}\right|^2\right)\right\}\right\}}{\log_2 P}}_{\text{B}} \\
& - \underbrace{\lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\widehat{J}_{k_g^d} + \sigma^2\right)\right\}\right\}}{\log_2 P}}_{\text{C}} \\
& - \underbrace{\lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\widehat{J}_{k_g^u} + \sigma^2\right)\right\}\right\}}{\log_2 P}}_{\text{D}}
\end{aligned} \tag{6.25}$$

Next, discarding the interference-plus-noise noise terms in parts A and B, and applying Jensen's inequality to parts C and D, results in

$$\begin{aligned}
\widehat{D}_{\,\text{FD TOT}} \geq & \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(P\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}\right|^2\right)\right\}\right\}}{\log_2 P} \\
& - \lim_{P\to\infty} \frac{\displaystyle\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\} + \sigma^2\right)}{\log_2 P}
\end{aligned}$$

$$+\lim_{P\to\infty}\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(P\left|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}\right|^2\right)\right\}\right\}}{\log_2 P}$$

$$-\lim_{P\to\infty}\frac{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}+\sigma^2\right)}{\log_2 P}. \tag{6.26}$$

Additionally, the DoF expression from (6.26) can be transformed into (6.27). This follows since for unitary beamformers, analogous to [73, Lemma 2] it can be shown that $|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}|^2$ and $|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}|^2$ are exponentially distributed with both mean and variance one, and also by replacing $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}$ and $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}$ with (6.21) and (6.22) respectively.

$$\widehat{D}_{\text{FD TOT}} = GK_db_d - \lim_{P\to\infty}\frac{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(P\frac{\eta}{1+\eta}\left(GK_db_d-1+GK_ub_u\right)+\sigma^2\right)}{\log_2 P}$$

$$+GK_ub_u - \lim_{P\to\infty}\frac{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(P\frac{\eta}{1+\eta}\left(GK_ub_u-1+(G-\bar{\Theta})K_db_d\right)+\sigma^2\right)}{\log_2 P} \tag{6.27}$$

Finally, replacing η with $\beta\rho^{-\alpha}=\beta P^{-\alpha}\sigma^{2\alpha}$ in (6.27) and taking $P\to\infty$, the achievable DoF with imperfect CSI can be characterised as

$$\widehat{D}_{\text{FD TOT}} = \begin{cases} G(K_db_d+K_ub_u) & \alpha\geq 1 \\ \alpha G(K_db_d+K_ub_u) & 0\leq\alpha<1. \end{cases} \tag{6.28}$$

Noting that $\Delta D_{\text{FD}} = D_{\text{FD TOT}} - \widehat{D}_{\text{FD TOT}}$, and having already established that $D_{\text{FD TOT}} = G(K_db_d+K_ub_u)$ in (6.7), we can use (6.28) to obtain (6.24), proving that the DoF loss is as originally stated in Theorem 6.2. □

6.4 Linear IA algorithms

While the bounds derived so far provide an understanding of the expected behaviour of linear IA within the system model considered, it is also necessary to have algorithms that work within this context. Such algorithms are not yet available in literature for systems with FD BSs and HD users, therefore here

we propose two different approaches: (a) an MMSE based solution, and (b) a Max-SINR based one.

The proposed algorithms are not straightforward extensions of the original HD ones from [11] and [10]; (a) they separate the various interference components amongst the different available beamformers rather than treating all interference equivalently, i.e. they are based on design principles that are specifically catered to the new system model, (b) they exploit statistical knowledge of the CSI error to provide added robustness, and (c) they result in unitary beamformers. The use of unitary beamformers has gained significant attention in recent years due to its role in codebook design for limited feedback scenarios. It has been selected for both single-user and multi-user mode operation for evolved universal terrestrial radio access, with advantages that include added simplicity of application and improved robustness to channel estimation errors [110]. Additionally, it has been shown to lower complexity for MMSE based algorithms [111] by avoiding the need for an extra linear search to enforce transmit power constraints when generating precoders, and also improve performance for Max-SINR based ones in multi-stream applications [66].

The interference that needs to be handled by our IA solutions can be classified into four main categories:

(i) *Intra-DL interference* - interference caused by undesired DL data for other users in the same cell;

(ii) *Intra-UL interference* - interference caused by undesired UL data for other users in the same cell;

(iii) *CCI-OC* - co-channel interference caused by nodes located in other cells (includes both DL and UL data);

(iv) *R-SI* - residual self-interference at the BSs due to imperfect CSI knowledge.

While it is possible to create beamformers that handle all the interference jointly, prior results for HD systems [62, 63] indicate that this approach is not suited to interference scenarios that are more complex than the initially studied HD IC. Similar behaviour has also been noted for our FD system, thus we base our IA algorithms on specific design principles.

Focus in this section will be on the derivation of the algorithms for a single-cell system, since the feasibility of such configurations has already been explored in

current literature [8], and also due to the relevant compactness of the corresponding expressions in comparison to multi-cell ones. The results for the multi-cell extensions are presented later in Section 6.5.2. Note that when considering the single-cell case, where by definition $G = 1$, we drop the use of index g to indicate which cell a user belongs to, i.e. we use k^u to indicate the kth UL user in the cell and k^d to represent the kth DL user. However, the notation g is still used in channel related indices to represent the BS. The single-cell versions of the algorithms follow Design Principle 1, outlined below.

Design Principle 1. *Intra-UL interference is only handled by the receivers. Intra-DL interference is only handled by the precoders. R-SI is handled by both the transmit and receive beamformers at the BS.*

6.4.1 MMSE based design for single-cell systems

This algorithm focuses on minimising the mean squared error, and designs beamformers which aim to find a balance between aligning the interference and ensuring that the signal level is suitably above noise. It was originally proposed for the IC with perfect CSI and a single-stream per user in [11], and later generalized to the multi-stream case in [113]. The designs in [11, 113] carry out a separate linear search (using techniques such as for example the bisection method) to enforce transmit power contraints for each of the precoders generated. The added computional cost incurred by the numerical search can be avoided by ensuring that the beamformers produced are unitary [111]. Our MMSE design incorporates this lower complexity feature, and produces unitary beamformers via the inclusion of QR decomposition stages (see Steps 4 and 6 in Algorithm 6.1).

Starting with UL communication in the intended direction, with fixed $\widehat{\mathbf{V}}$ and in accordance to Design Principle 1, the optimisation problem to find the BS receivers, $\widehat{\mathbf{U}}_{k^u}$, is given by

$$\min_{\widehat{\mathbf{U}}_{k^u}} \mathbb{E}\left\{\|\mathbf{f}_{k^u} - \mathbf{s}_{k^u}\|^2\right\} \ \forall\, k$$

where

$$\mathbf{f}_{k^u} = \widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_u} \mathbf{H}_{g,i^u} \widehat{\mathbf{V}}_{i^u} \mathbf{s}_{i^u} + \Theta \widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_d} \boldsymbol{\Upsilon}_{g,g} \widehat{\mathbf{V}}_{i^d} \mathbf{s}_{i^d} + \widehat{\mathbf{U}}_{k^u}^H \mathbf{n}_g \,.$$

The optimisation function can be defined as

$$
\begin{aligned}
F_{k^u} &= \mathbb{E}\left\{\|\mathbf{f}_{k^u} - \mathbf{s}_{k^u}\|^2\right\} \\
&= \mathbb{E}\left\{\mathrm{Tr}\left(\left(\widehat{\mathbf{U}}^H \sum_{i=1}^{K_u} \mathbf{H}_{g,i^u}\widehat{\mathbf{V}}_{i^u}\mathbf{s}_{i^u} + \Theta\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_d} \boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{V}}_{i^d}\mathbf{s}_{i^d} + \widehat{\mathbf{U}}_{k^u}^H\mathbf{n}_g - \mathbf{s}_{k^u}\right)\right.\right. \\
&\qquad \left.\left.\left(\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_u} \mathbf{H}_{g,i^u}\widehat{\mathbf{V}}_{i^u}\mathbf{s}_{i^u} + \Theta\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_d} \boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{V}}_{i^d}\mathbf{s}_{i^d} + \widehat{\mathbf{U}}_{k^u}^H\mathbf{n}_g - \mathbf{s}_{k^u}\right)^H\right)\right\} \\
&\overset{(a)}{=} \mathrm{Tr}\left(P\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_u} \mathbf{H}_{g,i^u}\widehat{\mathbf{V}}_{i^u}\widehat{\mathbf{V}}_{i^u}^H\mathbf{H}_{g,i^u}^H\widehat{\mathbf{U}}_{k^u} + \sigma^2\widehat{\mathbf{U}}_{k^u}^H\widehat{\mathbf{U}}_{k^u} - P\widehat{\mathbf{U}}_{k^u}^H\mathbf{H}_{g,k^u}\widehat{\mathbf{V}}_{k^u}\right. \\
&\qquad \left. - P\widehat{\mathbf{V}}_{k^u}^H\mathbf{H}_{g,k^u}^H\widehat{\mathbf{U}}_{k^u} + \Theta P\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_d} \boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{V}}_{i^d}\widehat{\mathbf{V}}_{i^d}^H\boldsymbol{\Upsilon}_{g,g}^H\widehat{\mathbf{U}}_{k^u}\right) + Pb_u
\end{aligned}
$$

where (a) follows since the transmitted data consists of i.i.d symbols, allowing us to use $\mathbb{E}\left\{\mathbf{s}_{k^u}\mathbf{s}_{i^u}^H\right\} = \mathbb{E}\left\{\mathbf{s}_{k^d}\mathbf{s}_{i^d}^H\right\} = \mathbf{0}\ \forall\ k, i, (k \neq i)$, $\mathbb{E}\left\{\mathbf{s}_{k^u}\mathbf{s}_{i^d}^H\right\} = \mathbb{E}\left\{\mathbf{s}_{k^d}\mathbf{s}_{i^u}^H\right\} = \mathbf{0}$ $\forall\ k, i$ and $\mathbb{E}\left\{\mathbf{s}_{k^u}\mathbf{s}_{k^u}^H\right\} = \mathbb{E}\left\{\mathbf{s}_{k^d}\mathbf{s}_{k^d}^H\right\} = P\mathbf{I}\ \forall\ k$. Differentiating with respect to $\widehat{\mathbf{U}}_{k^u}$ and replacing $\mathbf{H}$ by (6.9) results in

$$
\begin{aligned}
\frac{\partial F_{k^u}}{\partial \widehat{\mathbf{U}}_{k^u}} &= P\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_u}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,i^u} + \boldsymbol{\Upsilon}_{g,i^u}\right)\widehat{\mathbf{V}}_{i^u}\widehat{\mathbf{V}}_{i^u}^H\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,i^u} + \boldsymbol{\Upsilon}_{g,i^u}\right)^H \\
&\quad + \sigma^2\widehat{\mathbf{U}}_{k^u}^H - P\widehat{\mathbf{V}}_{k^u}^H\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,k^u} + \boldsymbol{\Upsilon}_{g,k^u}\right)^H + \Theta P\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_d}\boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{V}}_{i^d}\widehat{\mathbf{V}}_{i^d}^H\boldsymbol{\Upsilon}_{g,g}^H\,.
\end{aligned}
$$

This can be made dependent on the imperfect CSI, $\widehat{\mathbf{H}}$, only by using the statistical knowledge we have of the error. Thus, taking expectations with respect to $\boldsymbol{\Upsilon}$ and using Lemmas 4.2 and 6.2, we obtain

$$
\begin{aligned}
\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\frac{\partial F_{k^u}}{\partial \widehat{\mathbf{U}}_{k^u}}\right\} &= \frac{P}{(1+\eta)^2}\widehat{\mathbf{U}}_{k^u}^H \sum_{i=1}^{K_u}\widehat{\mathbf{H}}_{g,i^u}\widehat{\mathbf{V}}_{i^u}\widehat{\mathbf{V}}i^{uH}\widehat{\mathbf{H}}_{g,i^u}^H + \sigma^2\widehat{\mathbf{U}}_{k^u}^H - \frac{P}{(1+\eta)}\widehat{\mathbf{V}}_{k^u}^H\widehat{\mathbf{H}}_{g,k^u}^H \\
&\quad + \frac{P\eta}{(1+\eta)}\widehat{\mathbf{U}}_{k^u}^H(K_u b_u + \Theta K_d b_d)\,. \qquad (6.29)
\end{aligned}
$$

The receiver which minimises the UL mean squared error is obtained by setting (6.29) equal to zero, resulting in

$$
\widehat{\mathbf{U}}_{k^u} = \left(\sum_{i=1}^{K_u}\widehat{\mathbf{H}}_{g,i^u}\widehat{\mathbf{V}}_{i^u}\widehat{\mathbf{V}}_{i^u}^H\widehat{\mathbf{H}}_{g,i^u}^H + \gamma_{k^u}\mathbf{I}\right)^{-1}(1+\eta)\quad \widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{V}}_{k^u} \qquad (6.30)
$$

where

$$\gamma_u = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(K_u b_u + \Theta K_d b_d)\,. \tag{6.31}$$

Using a similar process for DL communication in the intended direction, with fixed $\widehat{\mathbf{V}}$ and in accordance to Design Principle 1, we solve

$$\min_{\widehat{\mathbf{U}}_{k^d}} \mathbb{E}\left\{\|\mathbf{f}_{k^d} - \mathbf{s}_{k^d}\|^2\right\} \;\forall\; k$$

where

$$\mathbf{f}_{k^d} = \widehat{\mathbf{U}}_{k^d}\mathbf{H}_{k^d,g}\widehat{\mathbf{V}}_{k^d}\mathbf{s}_{k^d} + \widehat{\mathbf{U}}_{k^d}\sum_{i=1}^{K_u}\mathbf{H}_{k^d,i^u}\widehat{\mathbf{V}}_{i^u}\mathbf{s}_{i^u} + \widehat{\mathbf{U}}_{k^d}\mathbf{n}_{k_d}\,.$$

This results in

$$\begin{aligned}\widehat{\mathbf{U}}_{k^d} = &\left(\widehat{\mathbf{H}}_{k^d,g}\widehat{\mathbf{V}}_{k^d}\widehat{\mathbf{V}}_{k^d}^H\widehat{\mathbf{H}}_{k^d,g}^H + \sum_{i=1}^{K_u}\widehat{\mathbf{H}}_{k^d,i^u}\widehat{\mathbf{V}}_{i^u}\widehat{\mathbf{V}}_{i^u}^H\widehat{\mathbf{H}}_{k^d,i^u}^H + \gamma_{k^d}\mathbf{I}\right)^{-1}\\ &\times(1+\eta)\quad\widehat{\mathbf{H}}_{k^d,g}\widehat{\mathbf{V}}_{k^d}\end{aligned} \tag{6.32}$$

with

$$\gamma_d = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(b_d + K_u b_u)\,. \tag{6.33}$$

Considering the reciprocal network we can also apply a similar method to solve for $\widehat{\mathbf{V}}$ with fixed $\widehat{\mathbf{U}}$. In the reciprocal network we assume that all directions of communication are reversed, i.e. UL users want to receive data from the BS, while DL users want to transmit data to the BS. Additionally, $\widehat{\mathbf{V}}$ now act as receive beamformers and $\widehat{\mathbf{U}}$ act as precoders. We use channel notations of the form $\overleftarrow{\mathbf{H}}_{x,y} = \mathbf{H}_{y,x}^H$ to represent the channel going from y to x in the reciprocal network.

For communication by UL users in the reciprocal channel, we solve

$$\min_{\widehat{\mathbf{V}}_{k^u}} \mathbb{E}\{\|\overleftarrow{\mathbf{f}}_{k^u} - \mathbf{s}_{k^u}\|^2\} \;\forall\; k$$

where

$$\overleftarrow{\mathbf{f}}_{k^u} = \widehat{\mathbf{V}}_{k^u}\overleftarrow{\mathbf{H}}_{k^u,g}\widehat{\mathbf{U}}_{k^u}\mathbf{s}_{k^u} + \widehat{\mathbf{V}}_{k^u}\sum_{i=1}^{K_d}\overleftarrow{\mathbf{H}}_{k^d,i^u}\widehat{\mathbf{U}}_{i^d}\mathbf{s}_{i^d} + \widehat{\mathbf{V}}_{k^u}\overleftarrow{\mathbf{n}}_{k_u}$$

to obtain

$$\widehat{\mathbf{V}}_{k^u} = \left(\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g} \widehat{\mathbf{U}}_{k^u} \widehat{\mathbf{U}}_{k^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{k^u,g}^H + \sum_{i=1}^{K_d} \overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d} \widehat{\mathbf{U}}_{i^d} \widehat{\mathbf{U}}_{i^d}^H \overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d}^H + \overleftarrow{\gamma}_{k^u} \mathbf{I} \right)^{-1} \times (1+\eta) \quad \overleftarrow{\widehat{\mathbf{H}}}_{k^u,g} \widehat{\mathbf{U}}_{k^u} \tag{6.34}$$

with

$$\overleftarrow{\gamma}_u = \frac{\sigma^2 (1+\eta)^2}{P} + \eta(1+\eta)(b_u + K_d b_d)\,. \tag{6.35}$$

For communication by DL users in the reciprocal network, we solve

$$\min_{\widehat{\mathbf{v}}_{k^d}} \mathbb{E}\{\| \overleftarrow{\mathbf{f}}_{k^d} - \mathbf{s}_{k^d} \|^2\} \; \forall \; k$$

where

$$\overleftarrow{\mathbf{f}}_{k^d} = \sum_{i=1}^{K_d} \widehat{\mathbf{V}}_{k^d} \overleftarrow{\mathbf{H}}_{g,i^d} \widehat{\mathbf{U}}_{i^d} \mathbf{s}_{i^d} + \Theta \widehat{\mathbf{V}}_{k^d} \sum_{i=1}^{K_u} \overleftarrow{\mathbf{\Upsilon}}_{g,g} \widehat{\mathbf{U}}_{i^u} \mathbf{s}_{i^u} + \widehat{\mathbf{V}}_{k^d} \overleftarrow{\mathbf{n}}_g$$

to obtain

$$\widehat{\mathbf{V}}_{k^d} = \left(\sum_{i=1}^{K_d} \overleftarrow{\widehat{\mathbf{H}}}_{g,i^d} \widehat{\mathbf{U}}_{i^d} \widehat{\mathbf{U}}_{i^d}^H \overleftarrow{\widehat{\mathbf{H}}}_{g,i^d}^H + \overleftarrow{\gamma}_{k^d} \mathbf{I} \right)^{-1} (1+\eta) \quad \overleftarrow{\widehat{\mathbf{H}}}_{g,k^d} \widehat{\mathbf{U}}_{k^d} \tag{6.36}$$

with

$$\overleftarrow{\gamma}_d = \frac{\sigma^2 (1+\eta)^2}{P} + \eta(1+\eta)(K_d b_d + \Theta K_u b_u)\,. \tag{6.37}$$

Having derived expressions for all beamformers, the resulting FD MMSE algorithm with statistical knowledge of the CSI error, which we refer to as FD-MMSE-SKCE, is as outlined in Algorithm 6.1.

6.4.2 Max-SINR based design for single-cell systems

In the design of our Max-SINR algorithm we take advantage of statistical knowledge with respect to the CSI mismatch. This results in the calculation of more accurate interference-plus-noise covariance matrices when compared to the naive approach, which would simply assume that the available CSI is perfect and take

Algorithm 6.1: FD-MMSE-SKCE algorithm for FD system

1 Set γ_u, γ_d, $\overleftarrow{\gamma}_u$ and $\overleftarrow{\gamma}_d$ according to (6.31), (6.33), (6.35) and (6.37).
2 Initialise $\widehat{\mathbf{V}}_{k^u}$ and $\widehat{\mathbf{V}}_{k^d}$ as random unitary matrices $\forall\ k$.
3 Obtain the receive filters $\widehat{\mathbf{U}}_{k^u}$ and $\widehat{\mathbf{U}}_{k^d}$ using (6.30) and (6.32) $\forall\ k$.
4 Set $\widehat{\mathbf{U}}_{k^u} = \mathbb{QR}(\widehat{\mathbf{U}}_{k^u})$ and $\widehat{\mathbf{U}}_{k^d} = \mathbb{QR}(\widehat{\mathbf{U}}_{k^d})$ $\forall\ k$.
5 Obtain the precoders $\widehat{\mathbf{V}}_{k_g^u}$ and $\widehat{\mathbf{V}}_{k_g^d}$ using (6.34) and (6.36) $\forall\ k$.
6 Set $\widehat{\mathbf{V}}_{k^u} = \mathbb{QR}(\widehat{\mathbf{V}}_{k^u})$ and $\widehat{\mathbf{V}}_{k^d} = \mathbb{QR}(\widehat{\mathbf{V}}_{k^d})$ $\forall\ k$.
7 Repeat the process from Step 2 until convergence or for a fixed number of iterates.

no additional measures in order to counter the effect of channel imperfections. Note that in addition to following Design Principle 1, our interference-plus-noise covariance matrices also take into account inter-stream interference for the data required at each node.

Starting with UL communication in the intended direction, the interference-plus-noise covariance matrix is given by (6.38), where (a) follows by replacing $\mathbf{H}$ with (6.9).

$$
\begin{aligned}
\mathbf{Q}_{k^u}^n &= \sum_{\substack{i=1\\(i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u} P\mathbf{H}_{g,i^u}\widehat{\mathbf{v}}_{i^u}^m\widehat{\mathbf{v}}_{i^u}^{m\,H}\mathbf{H}_{g,i^u}^H + \Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P\boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{v}}_{i^d}^m\widehat{\mathbf{v}}_{i^d}^{m\,H}\boldsymbol{\Upsilon}_{g,g}^H + \sigma^2\mathbf{I} \\
&\overset{(a)}{=} \sum_{\substack{i=1\\(i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u} P\Bigg[\frac{1}{(1+\eta)^2}\widehat{\mathbf{H}}_{g,i^u}\widehat{\mathbf{v}}_{i^u}^m\widehat{\mathbf{v}}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{g,i^u}^H + \underbrace{\sum_{\substack{i=1\\(i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u}\boldsymbol{\Upsilon}_{g,i^u}\widehat{\mathbf{v}}_{i^u}^m\widehat{\mathbf{v}}_{i^u}^{m\,H}\boldsymbol{\Upsilon}_{g,i^u}^H}_{\text{A}} \\
&\quad + \frac{1}{(1+\eta)}\underbrace{\left(\widehat{\mathbf{H}}_{g,i^u}\widehat{\mathbf{v}}_{i^u}^m\widehat{\mathbf{v}}_{i^u}^{m\,H}\boldsymbol{\Upsilon}_{g,i^u}^H + \boldsymbol{\Upsilon}_{g,i^u}^H\widehat{\mathbf{v}}_{i^u}^m\widehat{\mathbf{v}}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{g,i^u}\right)}_{\text{B}}\Bigg] \\
&\quad + \Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P\underbrace{\boldsymbol{\Upsilon}_{g,g}\widehat{\mathbf{v}}_{i^d}^m\widehat{\mathbf{v}}_{i^d}^{m\,H}\boldsymbol{\Upsilon}_{g,g}^H}_{\text{C}} + \sigma^2\mathbf{I}
\end{aligned}
\tag{6.38}
$$

Since statistical information is available with respect to the channel uncertainty, we can simplify (6.38) further by replacing all the elements that contain $\boldsymbol{\Upsilon}$ by their expected values. Using Lemma 6.2 from Appendix 6.A $\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\text{B}\} = \mathbf{0}$. Also from Lemma 4.2 $\mathbb{E}_{\boldsymbol{\Upsilon}}\{\text{A}\} = \mathbb{E}_{\boldsymbol{\Upsilon}}\{\text{C}\} = \eta/(1+\eta)\mathbf{I}$. Thus, instead of using (6.38), we can represent the UL interference-plus-noise covariance matrix in the intended direction as

$$\widehat{\mathbf{Q}}_{k^u}^{n} = \sum_{\substack{i=1 \\ (i,m)\neq(k,n)}}^{K_u} \sum_{m=1}^{b_u} \tau \widehat{\mathbf{H}}_{g,i^u} \widehat{\mathbf{v}}_{i^u}^{m} \widehat{\mathbf{v}}_{i^u}^{m\,H} \widehat{\mathbf{H}}_{g,i^u}^{H} + \xi_u \mathbf{I} \tag{6.39}$$

where

$$\tau = \frac{P}{(1+\eta)^2} \tag{6.40}$$

and

$$\xi_u = \sigma^2 + \frac{P\eta}{(1+\eta)}(K_u b_u - 1 + \Theta K_d b_d)\,. \tag{6.41}$$

Applying a method similar to the one used to derive (6.39), for DL communication in the intended direction we obtain the interference-plus-noise covariance matrix as

$$\widehat{\mathbf{Q}}_{k^d}^{n} = \sum_{i=1}^{K_u} \sum_{m=1}^{b_u} \tau \widehat{\mathbf{H}}_{k^d,i^u} \widehat{\mathbf{v}}_{i^u}^{m} \widehat{\mathbf{v}}_{i^u}^{m\,H} \widehat{\mathbf{H}}_{k^d,i^u}^{H} + \sum_{\substack{m=1 \\ m\neq n}}^{b_d} \tau \widehat{\mathbf{H}}_{k^d,g} \widehat{\mathbf{v}}_{k^d}^{m} \widehat{\mathbf{v}}_{k^d}^{m\,H} \widehat{\mathbf{H}}_{k^d,g}^{H} + \xi_d \mathbf{I} \tag{6.42}$$

where

$$\xi_d = \sigma^2 + \frac{P\eta}{(1+\eta)}(K_u b_u + b_d - 1)\,. \tag{6.43}$$

Next, reversing the direction of communication, we calculate the interference-plus-noise covariance matrices for the reciprocal network. Starting with communication by UL users in the reciprocal network, we obtain the following interference-plus-noise covariance matrix

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k^u}^{n} = \sum_{i=1}^{K_d} \sum_{m=1}^{b_d} \tau \overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d} \widehat{\mathbf{u}}_{i^d}^{m} \widehat{\mathbf{u}}_{i^d}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d}^{H} + \sum_{\substack{m=1 \\ m\neq n}}^{b_u} \tau \overleftarrow{\widehat{\mathbf{H}}}_{k^u,g} \widehat{\mathbf{u}}_{k^u}^{m} \widehat{\mathbf{u}}_{k^u}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{k^u,g} + \overleftarrow{\xi}_{u} \mathbf{I} \tag{6.44}$$

where

$$\overleftarrow{\xi}_{u} = \sigma^2 + \frac{P\eta}{(1+\eta)}(K_d b_d + b_u - 1)\,. \tag{6.45}$$

For communication by DL users in the reciprocal network, we obtain

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k^d}^{n} = \sum_{\substack{i=1 \\ (i,m)\neq(k,n)}}^{K_d} \sum_{m=1}^{b_d} \tau \overleftarrow{\widehat{\mathbf{H}}}_{g,i^d} \widehat{\mathbf{u}}_{i^d}^{m} \widehat{\mathbf{u}}_{i^d}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{g,i^d}^{H} + \overleftarrow{\xi}_{d} \mathbf{I} \tag{6.46}$$

where

$$\overleftarrow{\xi}_d = \sigma^2 + \frac{P\eta}{(1+\eta)}(K_d b_d - 1 + \Theta K_u b_u) \ . \tag{6.47}$$

The resulting Max-SINR algorithm which exploits statistical knowledge of the CSI error (Max-SINR-SKCE) is as outlined in Algorithm 6.2. Note that the original Max-SINR IA based algorithm from [10] does not contain a QR decomposition stage, but instead normalises the per-stream beamformers. Having unitary beamformers was later shown to improve performance for multi-stream applications [66]. By including a QR decomposition stage in Steps 5 and 8 of Algorithm 6.2 we produce unitary beamformers, thereby ensuring we obtain the multi-stream advantages, and also eliminating the need for separate normalisation steps since the resultant beamformers inherently consist of unit-norm vectors.

Algorithm 6.2: FD-Max-SINR-SKCE algorithm for FD system

1 Set τ, ξ_u, ξ_d, $\overleftarrow{\xi}_u$ and $\overleftarrow{\xi}_d$ according to (6.41), (6.40), (6.43), (6.45) and (6.47).
2 Initialise $\widehat{\mathbf{v}}_{k^u}^n$ and $\widehat{\mathbf{v}}_{k^d}^n$ as random unit-norm vectors $\forall\ n, k$.
3 Calculate $\widehat{\mathbf{Q}}_{k^u}^n$ and $\widehat{\mathbf{Q}}_{k^d}^n$ using (6.39) and (6.42) $\forall\ n, k$.
4 Obtain the receive filters as $\widehat{\mathbf{u}}_{k^u}^n = (\widehat{\mathbf{Q}}_{k^u}^n)^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n$ and $\widehat{\mathbf{u}}_{k^d}^n = (\widehat{\mathbf{Q}}_{k^d}^n)^{-1}\widehat{\mathbf{H}}_{k^d,g}\widehat{\mathbf{v}}_{k^d}^n$ $\forall\ n, k$.
5 Set $\widehat{\mathbf{U}}_{k^u} = \mathbb{QR}(\widehat{\mathbf{U}}_{k^u})$ and $\widehat{\mathbf{U}}_{k^d} = \mathbb{QR}(\widehat{\mathbf{U}}_{k^d})$ $\forall\ k$.
6 Compute $\overleftarrow{\widehat{\mathbf{Q}}}_{k^u}^n$ and $\overleftarrow{\widehat{\mathbf{Q}}}_{k^d}^n$ using (6.44) and (6.46) $\forall\ n, k$.
7 Obtain the precoders as $\widehat{\mathbf{v}}_{k^u}^n = (\overleftarrow{\widehat{\mathbf{Q}}}_{k^u}^n)^{-1}\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g}\widehat{\mathbf{u}}_{k^u}^n$ and $\widehat{\mathbf{v}}_{k^d}^n = (\overleftarrow{\widehat{\mathbf{Q}}}_{k^d}^n)^{-1}\overleftarrow{\widehat{\mathbf{H}}}_{g,k^d}\widehat{\mathbf{u}}_{k^d}^n$ $\forall\ n, k$.
8 Set $\widehat{\mathbf{V}}_{k^u} = \mathbb{QR}(\widehat{\mathbf{V}}_{k^u})$ and $\widehat{\mathbf{V}}_{k^d} = \mathbb{QR}(\widehat{\mathbf{V}}_{k^d})$ $\forall\ k$.
9 Repeat the process from Step 2 until convergence or for a fixed number of iterates.

Remark 6.1. *In case of perfect CSI or for imperfect CSI situations where statistical knowledge of the CSI error is unavailable, a naive version of Algorithms 6.1 and 6.2 can be implemented. For such situations we have $\eta = 0$ in the expressions for beamformer calculation. Thus, for FD-MMSE-Naive we set $\gamma_u = \gamma_d = \overleftarrow{\gamma}_u = \overleftarrow{\gamma}_d = \frac{\sigma^2}{P}$ in Step 1. While, for FD-Max-SINR-Naive we set $\tau = P$ and $\xi_u = \xi_d = \overleftarrow{\xi}_u = \overleftarrow{\xi}_d = \sigma^2$ in Step 1. Additionally, for the perfect CSI case $\mathbf{H}$ is used in place of $\widehat{\mathbf{H}}$ throughout, and the resulting beamformers are $\mathbf{U}$ and $\mathbf{V}$ instead of $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$. Note that the naive versions of the algorithms have the same computational complexity as those originally presented in Algorithms 6.1 and 6.2.*

6.4.3 Equivalence between MMSE and Max-SINR designs

Under certain conditions the beamformers obtained by the proposed MMSE and Max-SINR algorithms are equivalent, in the sense that at each iteration both result in identical precoders and receivers.

Starting with the expression for the UL receiver in the intended direction from Step 4 of Algorithm 6.2, and defining $\widehat{\mathbf{Q}}_{k^u}^n \triangleq \mathbf{A}_{k^u}^n \frac{P}{(1+\eta)^2}$, we can express

$$\widehat{\mathbf{u}}_{k^u}^n = (\mathbf{A}_{k^u}^n)^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n \frac{(1+\eta)^2}{P}\,.$$

Additionally, $\mathbf{A}_{k^u}^n$ may be represented as

$$\mathbf{A}_{k^u}^n = \mathbf{B}_{k^u} - \widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n \widehat{\mathbf{v}}_{k^u}^{n\,H}\widehat{\mathbf{H}}_{g,k^u}^H$$

where

$$\mathbf{B}_{k^u} = \sum_{i=1}^{K_u} \widehat{\mathbf{H}}_{g,i^u}\widehat{\mathbf{V}}_{i^u}\widehat{\mathbf{V}}_{i^u}^H\widehat{\mathbf{H}}_{g,i^u}^H + \xi_u \frac{(1+\eta)^2}{P}\mathbf{I}\,.$$

Applying Lemma 6.3 to $(\mathbf{A}_{k^u}^n)^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n$ we obtain

$$(\mathbf{A}_{k^u}^n)^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{g,k^u}^n = \frac{(\mathbf{B}_{k^u})^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n}{1 - \widehat{\mathbf{v}}_{k^u}^{n\,H}\widehat{\mathbf{H}}_{g,k^u}^H(\mathbf{B}_{k^u})^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n}\,.$$

Letting $\lambda_{k^u}^n = 1 - \widehat{\mathbf{v}}_{k^u}^{n\,H}\widehat{\mathbf{H}}_{g,k^u}^H(\mathbf{B}_{k^u})^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n$, we can represent the receiver as

$$\widehat{\mathbf{u}}_{k^u}^n = (\mathbf{B}_{k^u})^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{v}}_{k^u}^n \frac{(1+\eta)^2}{P}\frac{1}{\lambda_{k^u}^n}\,.$$

Next, $\widehat{\mathbf{u}}_{k^u}^n \;\forall\; n = 1 \dots b_u$ can be horizontally concatenated to obtain the receiver across all streams as

$$\widehat{\mathbf{U}}_{k^u} = (\mathbf{B}_{k^u})^{-1}\widehat{\mathbf{H}}_{g,k^u}\widehat{\mathbf{V}}_{k^u}\mathbf{\Lambda}_{k^u} \tag{6.48}$$

where

$$\mathbf{\Lambda}_{k^u} = \frac{(1+\eta)^2}{P}\begin{bmatrix} \frac{1}{\lambda_{k^u}^1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_{k^u}^{b_u}} \end{bmatrix}.$$

Comparing (6.48) with the MMSE derived expression for the same beamformer in (6.30), it can be noticed that they are very similar. For the naive and perfect CSI versions of the algorithms, where η is set to 0, the term inside the inverse for (6.48) and (6.30) is equivalent. The only difference is an additional post-multiplication by $\mathbf{\Lambda}_{k^u}$ in (6.48); this matrix essentially multiplies each column vector with a scalar and thus has no effect on the resultant unitary part after the QR decomposition, therefore both algorithms obtain the same $\widehat{\mathbf{U}}_{k^u}$. A similar argument can be made for each of $\widehat{\mathbf{U}}_{k^d}$, $\widehat{\mathbf{V}}_{k^u}$ and $\widehat{\mathbf{V}}_{k^d}$. Thus, for the case where η is equal to zero, or unknown and assumed to be zero, (i.e. FD-Max-SINR-Naive and FD-MMSE-Naive) the two algorithms are equivalent.

Remark 6.2. *Note that even in cases where the Max-SINR/MMSE equivalence holds, the MMSE algorithm is less computationally complex than the Max-SINR algorithm, since the former operates on a per-user basis whilst the latter operates on a per-stream basis. Consider for example the number of matrix inverses involved; the MMSE algorithm requires a total of* $2(K_u + K_d)$ *inverses per iteration to compute the beamformers in Steps 3 and 5, while the Max-SINR algorithm requires a total of* $2(K_u b_u + K_d b_d)$ *inverses in Steps 4 and 7.*

6.4.4 Convergence of the proposed algorithms

Firstly, it is important to note that the convergence of Max-SINR based algorithms to achieve IA cannot be proven analytically, not even for the simplest case of the HD IC [10]. Considering that the Max-SINR algorithm proposed in this paper is based in principle on the original one from [10], but with increased complexity in the resultant expressions (due to the more complex system model), it follows by extension that the convergence of our Max-SINR based approach cannot be analytically proven. However, the overall consensus in literature is that Max-SINR algorithms for IA generally seem to converge to a constant value, as shown numerically in [114], and proven for sufficiently high SNR in [115]. Finally, it is also important to note that these convergence remarks also apply to our MMSE based algorithm, due to the equivalence established in Section 6.4.3.

6.5 Multi-cell considerations

A significant body of literature related to linear IA focuses on the analytic derivation of feasibility conditions, for example, [6] studies this issue for the IC, [7, 62]

consider IBCs and [8] derives feasibility conditions for linear IA in single-cell systems with an FD BS communicating with both DL and UL users. However, no feasibility conditions are available in literature so far for multi-cell multi-user systems with FB BSs and HD users. Here we look into this issue by deriving the proper condition for this type of network, and also by extending the linear IA algorithms proposed in Sections 6.4.1 and 6.4.2 to the multi-cell case. The derived proper condition, along with the proposed algorithms, can aid future work in this direction by serving as starting point that provides insight into the theoretical feasibility of linear IA for different antenna configurations and DoF requirements.

6.5.1 Proper condition

The proper condition relates the feasibility of IA to the issue of determining the resolvability of a system represented by multivariate polynomial equations. A system of equations is classified as *proper* if the number of equations, N_e, does not exceed the number of variables, N_v, i.e. if $N_v \geq N_e$. Prior studies [6, 7] show that for systems classified as *improper*, IA is surely infeasible. However, classifying a system as *proper* is not a sufficient condition to prove IA feasibility, i.e. systems that are proper but for which IA is infeasible may exist.

We follow the method from [6] to derive expressions for N_v and N_e, and obtain the proper condition for the FD enabled multi-cell scenario considered in this work. Focusing on a symmetric system where $K_d = K_u = \bar{K}$, $b_d = b_u = b$ and $N_d = N_u = N$, we obtain

$$N_v = 2G\bar{K}b(M_B + N - 2b)$$

and

$$N_e = (\bar{K}b)^2(4G^2 - 2 - G)\,.$$

This allows us to express the proper condition as

$$\frac{2G(M_B + N)}{4G + \bar{K}(4G^2 - 2 - G)} \geq b\,. \tag{6.49}$$

6.5.2 Multi-cell algorithm extension

Here we extend our IA algorithms to the multi-cell case. The actual method applied to design these algorithms is analogous to the one followed for the single-

cell case - the only difference in the derivation process is that instead of following Design Principle 1, we follow Design Principle 2 which includes additional considerations for CCI-OC which is now present.

Design Principle 2. *Intra-UL interference is only handled by the receivers. Intra-DL interference is only handled by the precoders. R-SI is handled by both the transmit and receive beamformers at the BSs. CCI-OC is handled by all beamformers.*

6.5.2.1 Multi-cell version of MMSE algorithm

The multi-cell version of FD-MMSE-SKCE follows the general steps outlined for the single-cell version in Algorithm 6.1, with the following differences.

- In Step 1 set γ_d, γ_u, $\overleftarrow{\gamma}_d$ and $\overleftarrow{\gamma}_u$ as follows.

$$\begin{aligned}
\gamma_d &= \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_u b_u + (G-1)K_d b_d + b_d - 1) \\
\gamma_u &= \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_u b_u - 1 + (G-1+\Theta)K_d b_d) \\
\overleftarrow{\gamma}_d &= \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_d b_d - 1 + (G-1+\Theta)K_u b_u) \\
\overleftarrow{\gamma}_u &= \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_d b_d + (G-1)K_u b_u + b_u - 1)
\end{aligned}$$

- In Step 3 calculate $\widehat{\mathbf{U}}_{k_g^d}$ and $\widehat{\mathbf{U}}_{k_g^u}$ using (6.50) and (6.51).

$$\begin{aligned}
\widehat{\mathbf{U}}_{k_g^d} = \Bigg(&\sum_{j=1}^{G}\sum_{i=1}^{K_u} \widehat{\mathbf{H}}_{k_g^d,i_j^u}\widehat{\mathbf{V}}_{i_j^u}\widehat{\mathbf{V}}_{i_j^u}^H\widehat{\mathbf{H}}_{k_g^d,i_j^u}^H + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_d} \widehat{\mathbf{H}}_{k_g^d,j}\widehat{\mathbf{V}}_{i_j^d}\widehat{\mathbf{V}}_{i_j^d}^H\widehat{\mathbf{H}}_{k_g^d,j}^H \\
&+ \widehat{\mathbf{H}}_{k_g^d,g}\widehat{\mathbf{V}}_{k_g^d}\widehat{\mathbf{V}}_{k_g^d}^H\widehat{\mathbf{H}}_{k_g^d,g}^H + \gamma_d\mathbf{I}\Bigg)^{-1} (1+\eta) \quad \widehat{\mathbf{H}}_{k_g^d,g}\widehat{\mathbf{V}}_{k_g^d} \qquad (6.50)
\end{aligned}$$

$$\begin{aligned}
\widehat{\mathbf{U}}_{k_g^u} = \Bigg(&\sum_{j=1}^{G}\sum_{i=1}^{K_u} \widehat{\mathbf{H}}_{g,i_j^u}\widehat{\mathbf{V}}_{i_j^u}\widehat{\mathbf{V}}_{i_j^u}^H\widehat{\mathbf{H}}_{g,i_j^u}^H + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_d} \widehat{\mathbf{H}}_{g,j}\widehat{\mathbf{V}}_{i_j^d}\widehat{\mathbf{V}}_{i_j^d}^H\widehat{\mathbf{H}}_{g,j}^H + \gamma_u\mathbf{I}\Bigg)^{-1} \\
&\times (1+\eta) \quad \widehat{\mathbf{H}}_{g,k_g^u}\widehat{\mathbf{V}}_{k_g^u} \qquad (6.51)
\end{aligned}$$

- In Step 5 find $\widehat{\mathbf{V}}_{k_g^d}$ and $\widehat{\mathbf{V}}_{k_g^u}$ using (6.52) and (6.53).

$$\widehat{\mathbf{V}}_{k_g^u} = \left(\sum_{j=1}^{G} \sum_{i=1}^{K_d} \overleftarrow{\widehat{\mathbf{H}}}_{i_j^d,k_g^u} \widehat{\mathbf{U}}_{i_j^d} \widehat{\mathbf{U}}_{i_j^d}^H \overleftarrow{\widehat{\mathbf{H}}}_{i_j^d,k_g^u}^H + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_u} \overleftarrow{\widehat{\mathbf{H}}}_{j,k_g^u} \widehat{\mathbf{U}}_{i_j^u} \widehat{\mathbf{U}}_{i_j^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{j,k_g^u}^H \right.$$
$$\left. + \overleftarrow{\widehat{\mathbf{H}}}_{g,k_g^u} \widehat{\mathbf{U}}_{k_g^u} \widehat{\mathbf{U}}_{k_g^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{g,k_g^u}^H + \overleftarrow{\gamma}_u \mathbf{I} \right)^{-1} (1+\eta) \quad \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,g} \widehat{\mathbf{U}}_{k_g^u} \tag{6.52}$$

$$\widehat{\mathbf{V}}_{k_g^d} = \left(\sum_{j=1}^{G} \sum_{i=1}^{K_d} \overleftarrow{\widehat{\mathbf{H}}}_{g,i_j^d} \widehat{\mathbf{U}}_{i_j^d} \widehat{\mathbf{U}}_{i_j^d}^H \overleftarrow{\widehat{\mathbf{H}}}_{g,i_j^d}^H + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_u} \overleftarrow{\widehat{\mathbf{H}}}_{g,j} \widehat{\mathbf{U}}_{i_j^u} \widehat{\mathbf{U}}_{i_j^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{g,j}^H + \overleftarrow{\gamma}_d \mathbf{I} \right)^{-1}$$
$$\times (1+\eta) \quad \overleftarrow{\widehat{\mathbf{H}}}_{g,k_g^d} \widehat{\mathbf{U}}_{k_g^d} \tag{6.53}$$

6.5.2.2 Multi-cell version of Max-SINR algorithm

The multi-cell version of FD-Max-SINR-SKCE follows the general steps outlined for the single-cell version in Algorithm 6.2, with the following differences.

- In Step 1 set ξ_d, ξ_u, $\overleftarrow{\xi}_d$ and $\overleftarrow{\xi}_u$ as follows.

$$\begin{aligned}
\xi_d &= \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_u b_u + (G-1)K_d b_d + b_d) \\
\xi_u &= \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_u b_u + (G-1+\Theta)K_d b_d) \\
\overleftarrow{\xi}_d &= \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_d b_d + (G-1+\Theta)K_u b_u) \\
\overleftarrow{\xi}_u &= \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_d b_d + (G-1)K_u b_u + b_u)
\end{aligned}$$

- In Step 3 the forward interference-plus-noise covariances matrices $\widehat{\mathbf{Q}}_{k_g^d}^n$ and $\widehat{\mathbf{Q}}_{k_g^u}^n$ are calculated using (6.54) and (6.55).

$$\widehat{\mathbf{Q}}_{k_g^d}^n = \sum_{j=1}^{G} \sum_{i=1}^{K_u} \sum_{m=1}^{b_u} \tau \widehat{\mathbf{H}}_{k_g^d,i_j^u} \widehat{\mathbf{v}}_{i_j^u}^m \widehat{\mathbf{v}}_{i_j^u}^{m\,H} \widehat{\mathbf{H}}_{k_g^d,i_j^u}^H + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_d} \sum_{m=1}^{b_d} \tau \widehat{\mathbf{H}}_{k_g^d,j} \widehat{\mathbf{v}}_{i_j^d}^m \widehat{\mathbf{v}}_{i_j^d}^{m\,H} \widehat{\mathbf{H}}_{k_g^d,j}^H$$
$$+ \sum_{\substack{m=1 \\ m\neq n}}^{b_d} \tau \widehat{\mathbf{H}}_{k_g^d,g} \widehat{\mathbf{v}}_{k_g^d}^m \widehat{\mathbf{v}}_{k_g^d}^{m\,H} \widehat{\mathbf{H}}_{k_g^d,g}^H + \xi_d \mathbf{I} \tag{6.54}$$

$$\widehat{\mathbf{Q}}_{k_g^u}^n = \sum_{\substack{j=1 \\ (j,i,m)\neq(g,k,n)}}^{G} \sum_{i=1}^{K_u} \sum_{m=1}^{b_u} \tau \widehat{\mathbf{H}}_{g,i_j^u} \widehat{\mathbf{v}}_{i_j^u}^m \widehat{\mathbf{v}}_{i_j^u}^{m\,H} \widehat{\mathbf{H}}_{g,i_j^u}^H + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_d} \sum_{m=1}^{b_d} \tau \widehat{\mathbf{H}}_{g,j} \widehat{\mathbf{v}}_{i_j^d}^m \widehat{\mathbf{v}}_{i_j^d}^{m\,H} \widehat{\mathbf{H}}_{g,j}^H + \xi_u \mathbf{I} \tag{6.55}$$

- In Step 4 use $\widehat{\mathbf{u}}_{k_g^u}^n = (\widehat{\mathbf{Q}}_{k_g^u}^n)^{-1} \widehat{\mathbf{H}}_{g,k_g^u} \widehat{\mathbf{v}}_{k_g^u}^n$ and $\widehat{\mathbf{u}}_{k_g^d}^n = (\widehat{\mathbf{Q}}_{k_g^d}^n)^{-1} \widehat{\mathbf{H}}_{k_g^d,g} \widehat{\mathbf{v}}_{k_g^d}^n \; \forall \; n, k, g.$
- In Step 6 the backward interference-plus-noise covariance matrices $\overleftarrow{\mathbf{Q}}_{k_g^d}^n$ and $\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^u}^n$ are calculated using (6.56) and (6.57).

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^d}^n = \sum_{\substack{j=1 \\ (j,i,m)\neq(g,k,n)}}^{G} \sum_{i=1}^{K_d} \sum_{m=1}^{b_d} \tau \overleftarrow{\widehat{\mathbf{H}}}_{g,i_j^d} \widehat{\mathbf{u}}_{i_j^d}^m \widehat{\mathbf{u}}_{i_j^d}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{g,i_j^d}^H + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_u} \sum_{m=1}^{b_u} \tau \overleftarrow{\widehat{\mathbf{H}}}_{g,j} \widehat{\mathbf{u}}_{i_j^d}^m \widehat{\mathbf{u}}_{i_j^d}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{g,j}^H + \overleftarrow{\xi}_d \mathbf{I} \tag{6.56}$$

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^u}^n = \sum_{j=1}^{G} \sum_{i=1}^{K_d} \sum_{m=1}^{b_d} \tau \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,i_j^d} \widehat{\mathbf{u}}_{i_j^d}^m \widehat{\mathbf{u}}_{i_j^d}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,i_j^d}^H + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_u} \sum_{m=1}^{b_u} \tau \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,j} \widehat{\mathbf{u}}_{i_j^u}^m \widehat{\mathbf{u}}_{i_j^u}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,j}^H + \sum_{\substack{m=1 \\ m\neq n}}^{b_u} \tau \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,g} \widehat{\mathbf{u}}_{k_g^u}^m \widehat{\mathbf{u}}_{k_g^u}^{m\,H} \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,g}^H + \overleftarrow{\xi}_u \mathbf{I} \tag{6.57}$$

- In Step 7 use $\widehat{\mathbf{v}}_{k_g^u}^n = (\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^u}^n)^{-1} \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,g} \widehat{\mathbf{u}}_{k_g^u}^n$ and $\widehat{\mathbf{v}}_{k_g^d}^n = (\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^d}^n)^{-1} \overleftarrow{\widehat{\mathbf{H}}}_{g,k_g^d} \widehat{\mathbf{u}}_{k_g^d}^n \; \forall \; n, k, g.$

6.6 Simulation results

This section provides simulation results to validate the analyses presented so far. Throughout all our simulations the noise variance, σ^2, is fixed at 1 making the transmit signal power equivalent to the SNR, and we assume that for imperfect CSI scenarios SI cancellation is imperfect, i.e. $\bar{\Theta} = 0$. Additionally, all results are averaged in a Monte-Carlo fashion over a number of different channel realisations. Treating all interference as noise, we calculate the achieved sum rate as

$$\bar{R} = \sum_{g=1}^{G} \sum_{k=1}^{K_d} \log_2 \det\Big(\mathbf{I} + (\mathbf{X}_{k_g^d} + \sigma^2 \mathbf{I})^{-1} P \mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d,g} \mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \mathbf{H}_{k_g^d,g}^H \mathbf{U}_{k_g^d}\Big) + \sum_{g=1}^{G} \sum_{k=1}^{K_u} \log_2 \det\Big(\mathbf{I} + (\mathbf{X}_{k_g^u} + \sigma^2 \mathbf{I})^{-1} P \mathbf{U}_{k_g^u}^H \mathbf{H}_{g,k_g^u} \mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H \mathbf{H}_{g,k_g^u}^H \mathbf{U}_{k_g^u}\Big)$$

where

$$\mathbf{X}_{k_g^u} = \sum_{\substack{j=1 \\ (j,i)\neq(g,k)}}^{G} \sum_{i=1}^{K_u} P\mathbf{U}_{k_g^u}^H \mathbf{H}_{g,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \mathbf{H}_{g,i_j^u}^H \mathbf{U}_{k_g^u} + \sum_{\substack{j=1 \\ j\neq g}}^{G} \sum_{i=1}^{K_d} P\mathbf{U}_{k_g^u}^H \mathbf{H}_{g,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \mathbf{H}_{g,j}^H \mathbf{U}_{k_g^u}$$
$$+ \Theta \sum_{i=1}^{K_d} P\mathbf{U}_{k_g^u}^H \boldsymbol{\Upsilon}_{g,g} \mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \boldsymbol{\Upsilon}_{g,g}^H \mathbf{U}_{k_g^u}$$

and

$$\mathbf{X}_{k_g^d} = \sum_{\substack{j=1 \\ (j,i)\neq(g,k)}}^{G} \sum_{i=1}^{K_d} P\mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d,j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \mathbf{H}_{k_g^d,j}^H \mathbf{U}_{k_g^d} + \sum_{j=1}^{G} \sum_{i=1}^{K_u} P\mathbf{U}_{k_g^d}^H \mathbf{H}_{k_g^d,i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \mathbf{H}_{k_g^d,i_j^u}^H \mathbf{U}_{k_g^d} .$$

Here, $\mathbf{X}_{k_g^d}$ and $\mathbf{X}_{k_g^u}$ represent the DL and UL interference covariance matrices respectively. For imperfect CSI situations, in order to calculate R we replace $\mathbf{V}$ with $\widehat{\mathbf{V}}$ and $\mathbf{U}$ with $\widehat{\mathbf{U}}$, since the beamformers are calculated using the available imperfect CSI. Note that calculating the rate in this manner results in a lower bound on the actual achievable rate. In truth higher rates can be obtained via the use of improved receivers, for example by considering the availability of perfect CSI at the receiver to obtain more accurate beamformers, or by applying more sophisticated receiver strategies such as maximum likelihood detectors.

6.6.1 Results for theoretically derived bounds

Here we verify the validity of the bounds derived in Theorems 6.1 and 6.2 using the naive versions of the algorithms proposed in Section 6.4. We simulate a system having $G = 1$, $K_d = K_u = 2$, $b_d = b_u = 3$, $M_B = 6$ and $N_d = N_u = 7$ to obtain the results in Fig. 6.2 and Fig. 6.3. For this system IA is known to be feasible [8] and the total achievable DoF under perfect CSI are given by $G(K_d b_d + K_u b_u) = 12$. For the same antenna configuration with a HD BS serving all $K = K_d + K_u$ users, where the number of antennas at the users, N, is given by $N = N_d = N_u$, the total achievable DoF are given by $\min\{M_B, NK\} = 6$, which is exactly half the DoF achievable when using an FD BS. We also simulate another feasible system having $G = 1$, $K_d = K_u = 4$, $b_d = b_u = 1$, $M_B = 4$ and $N_d = N_u = 3$ to obtain the results in Fig. 6.4 and Fig. 6.5. For this scenario having an FD BS allows us to achieve 8 DoF, double those obtained by the HD counterpart.

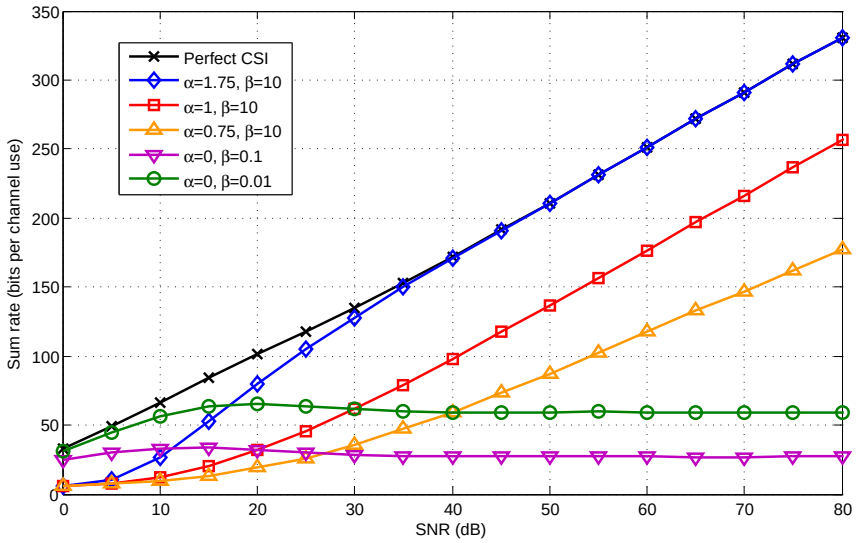


Figure 6.2: Average sum rates achieved by both FD-Max-SINR-Naive and FD-MMSE-Naive algorithms under various imperfect CSI conditions for system with $G = 1$, $K_d = K_u = 2$, $b_d = b_u = 3$, $M_B = 6$ and $N_d = N_u = 7$.

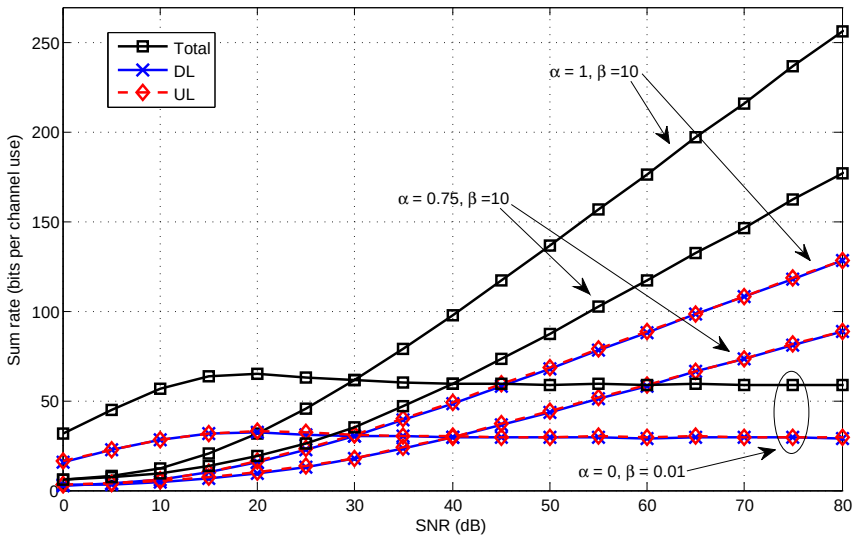


Figure 6.3: Average total, DL and UL rates achieved by both FD-Max-SINR-Naive and FD-MMSE-Naive algorithms under various imperfect CSI conditions for system with $G = 1$, $K_d = K_u = 2$, $b_d = b_u = 3$, $M_B = 6$ and $N_d = N_u = 7$.

From Theorem 6.2 we know that full DoF are achievable for values of $\alpha \geq 1$. This can be verified by focusing on the results for $\alpha = 1.75$ and $\alpha = 1$ in Fig. 6.2, both of which have the same slope as the perfect CSI curve. However, it can be noticed that while the $\alpha = 1.75$ result overlaps with the perfect CSI one at high SNR, the $\alpha = 1$ result runs parallel to it achieving lower sum rates overall.

Such behaviour is expected from Theorem 6.1; for $\alpha > 1$ there should be no sum rate loss at high SNR, while at $\alpha = 1$ the same theorem predicts a finite loss equal to Ω. For the system under consideration with $\beta = 10$, $\Omega \approx 81.5$ bits per channel use. From Fig. 6.2 it can be noticed that the perfect CSI result and the one for $\alpha = 1$ run parallel for SNR values of around 30 dB or higher, measuring the gap at 65 dB we obtain 74.6 bits per channel use. This confirms the validity of Theorem 6.1 and shows that the derived bound is not excessively loose.

Focusing on the results for $\alpha < 1$, from Theorem 6.1 we expect the sum rate loss to be unbounded. This is validated from the fact that the curves for $\alpha = 0.75$ and $\alpha = 0$ in Fig. 6.2 all diverge from the perfect CSI result, implying that the sum rate loss grows with SNR. From a DoF perspective, in the range of $\alpha < 1$ we expect a loss. For example for $\alpha = 0.75$, Theorem 6.2 predicts that only 75% of the full DoF are achievable. This can be confirmed by comparing the high SNR slopes for the perfect CSI curve, which achieves 12 DoF, and the one for $\alpha = 0.75$, which achieves 8 DoF. For $\alpha = 0$ the same theorem predicts 0 DoF achievable, and indeed both $\alpha = 0$ curves lie flat in the high SNR region. Additionally comparing the result for $\alpha = 0, \beta = 0.01$ and $\alpha = 0, \beta = 0.1$, it can be noticed that while the β value does not affect DoF behaviour, it still has a significant impact on achievable rate. In fact the curve for the smallest β settles at the highest rate, which is expected since this indicates the smallest error. Note that for any fixed β, $\alpha = 0$ represents the worst case scenario with the CSI error being equal to β itself; this causes a huge amount of interference leakage, making the network interference limited and eventually causing the sum rate to saturate.

For the $\alpha < 1$ range, Theorem 6.2 not only shows that a fraction equal to α of the full DoF are achievable, but it also indicates that this loss is distributed evenly between the DL and UL users, i.e. achievable DL DOF are equal to $\alpha G K_d b_d$ and achievable UL DOF are equal to $\alpha G K_u b_u$. This behaviour can be confirmed by considering Fig. 6.3, which plots DL and UL rates separately. As can be seen for the $\alpha = 1, \beta = 10$ curves in total 12 DoF are achieved; due to the symmetry of the simulated system where $K_u = K_d$ and $b_d = b_u$, this amounts to 6 DoF each for UL and DL. Focusing on the results for $\alpha = 0.75, \beta = 10$ it can noticed that both DL and UL results have a high SNR slope that corresponds to 4 DoF, while for $\alpha = 0, \beta = 0.1$ the slopes corresponds to 0 DoF. In both cases the achieved DoF are equivalent to $\alpha G K_d b_d$ for the DL and $\alpha G K_u b_u$ for the UL, which confirms our expectations from Theorem 6.2.

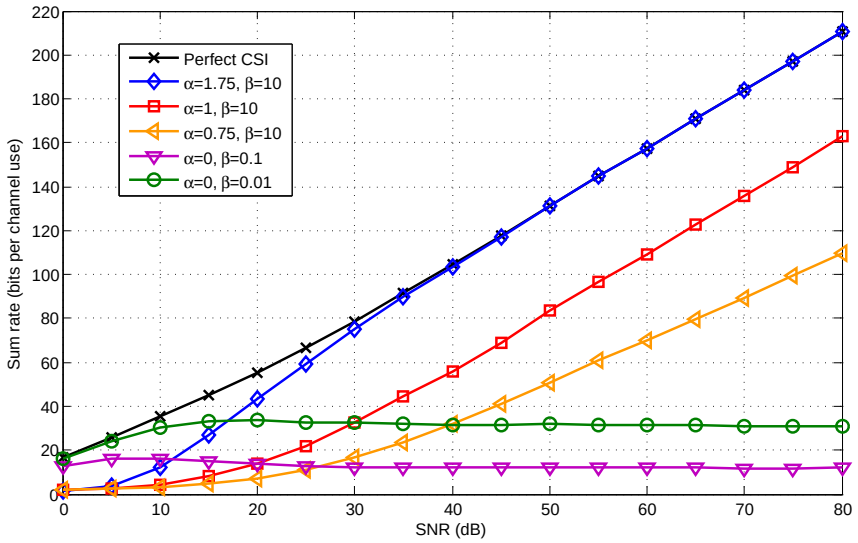


Figure 6.4: Average sum rates achieved by both FD-Max-SINR-Naive and FD-MMSE-Naive algorithms under various imperfect CSI conditions for system with $G = 1$, $K_d = K_u = 4$, $b_d = b_u = 1$, $M_B = 4$ and $N_d = N_u = 3$.

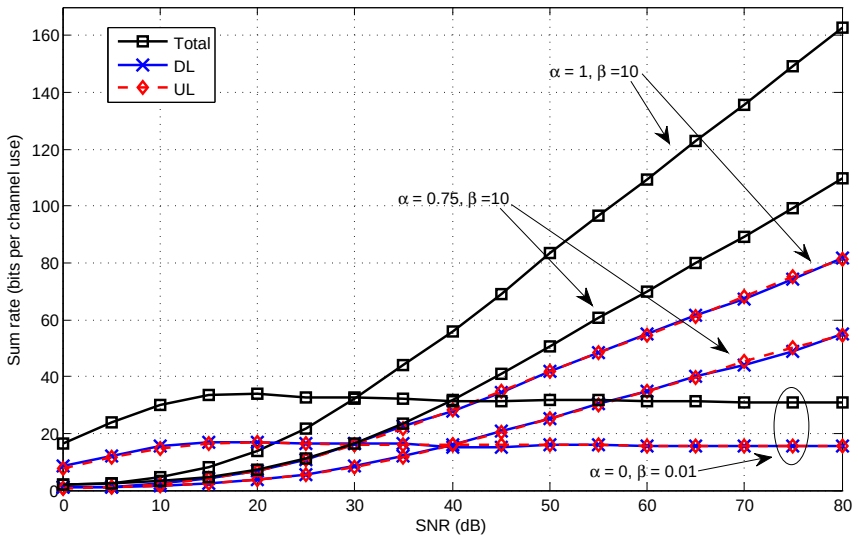


Figure 6.5: Average total, DL and UL rates achieved by both FD-Max-SINR-Naive and FD-MMSE-Naive algorithms under various imperfect CSI conditions for system with $G = 1$, $K_d = K_u = 4$, $b_d = b_u = 1$, $M_B = 4$ and $N_d = N_u = 3$.

Similar rate and DoF behaviour can also be noticed for the system configuration considered in Fig. 6.4. The result for $\alpha = 1.75$ overlaps the perfect CSI one for asymptotically high SNR, indicating full DoF are achieved as expected from Theorem 6.2, and also confirming that there is no rate loss as expected from Theorem 6.1. The latter theorem also predicts that for $\alpha = 1$ the asymptotic

sum rate loss is upper bounded by $\Omega \approx 49.20$ bits per channel use; measuring the actual gap from Fig. 6.4 we obtain 47.9 bits per channel use, confirming that the bound is not excessively loose. Additionally, for $\alpha \leq 1$ Theorem 6.2 predicts achievable DoF of $\alpha(GK_d b_d + GK_u b_u)$, as can be confirmed by the slopes of the $\alpha = 0.75$ and $\alpha = 0$ results that obtain 6 DoF and 0 DoF respectively. The even distribution of this DoF loss between UL and DL can be verified from Fig. 6.5.

6.6.2 Results for SKCE algorithms

We use a system having $G = 1$, $K_d = K_u = 3$, $b_d = b_u = 2$ and $M_B = M_d = M_u = 6$, which is known to be feasible [8], to obtain Fig. 6.6 and Fig. 6.7. As can be seen from both figures while the SKCE versions of the algorithms produce results that are very close, the curves don't overlap completely in the manner that results for the naive versions do. Such behavior is expected since the Max-SINR and MMSE equivalence established in Section 6.4.3 holds only for cases where η is set to 0 for beamformer calculation.

As seen from Fig. 6.6 and Fig. 6.7, the SKCE versions of the algorithms outperform the naive versions both in terms of sum rate and BER. For example for $\alpha = 1$, $\beta = 10$ at an SNR of 40 dB, MMSE-SKCE has a sum rate improvement of 12.3 bits per channel use, while Max-SINR-SKCE has a gain of 12.1 bits per channel. For the same α and β combination, MMMSE-SKCE achieves a BER of 1×10^{-2} at around 21.9 dB and Max-SINR-SKCE achieves it at 22.1 dB, while the naive version requires approximately 23.6 dB to obtain the same performance. Analogously, for $\alpha = 0.75$, $\beta = 10$ we have a rate gain of 14.2 bits per channel use for MMSE-SKCE and 14.1 bits per channel use for Max-SINR-SKCE. In term of BER for $\alpha = 0.75$, $\beta = 10$, MMSE-SKCE requires approximately 8.8 dB less than the naive version to reach a BER level of 1×10^{-2}, whilst Max-SINR-SKCE requires around 8.5 dB less than Max-SINR-Naive.

Considering the results for MMSE-Naive and Max-SINR-Naive with $\alpha = 0$, $\beta = 0.1$ in Fig. 6.6 and Fig. 6.7 it can be noticed that performance initially improves in the region of -10 dB up to around 10 dB, and then starts to degrade until it eventually settles to a steady state value for SNR $\geq$ 45 dB. For this specification of α and β, the CSI error is quite significant and independent of SNR. In the range of -10 db to 10 dB the power of the leakage is reasonably small, since the power levels we are dealing with are low; this allows for performance

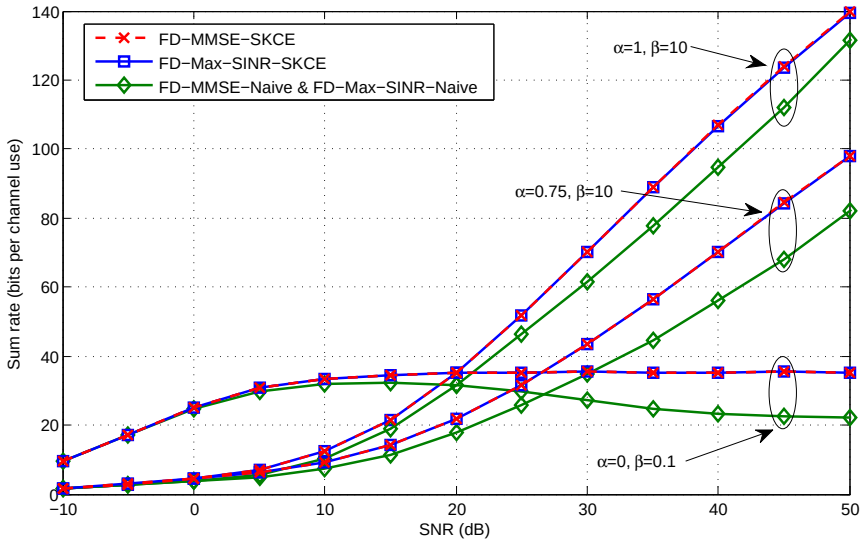


Figure 6.6: Average sum rates achieved for system with $G = 1$, $K_d = K_u = 3$, $b_d = b_u = 2$ and $M_B = N_d = N_u = 6$ under various imperfect CSI scenarios.

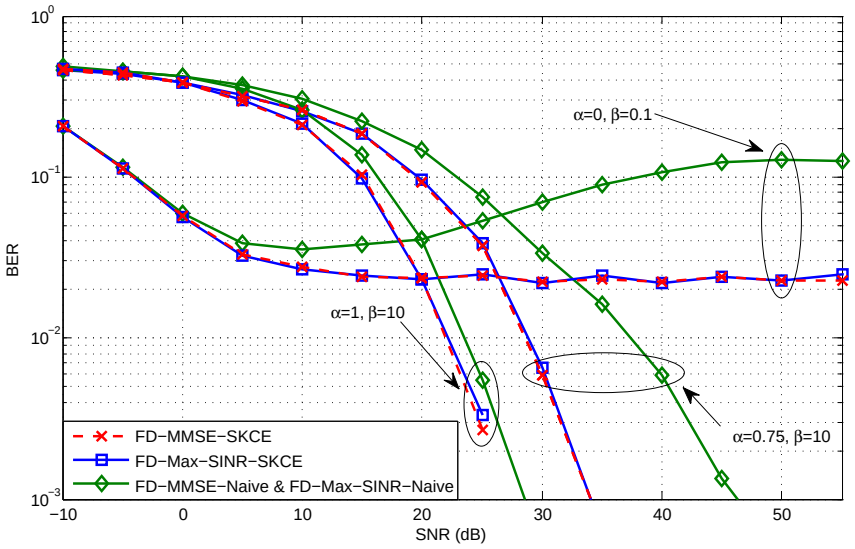


Figure 6.7: BER achieved for system with $G = 1$, $K_d = K_u = 3$, $b_d = b_u = 2$ and $M_B = N_d = N_u = 6$ under various imperfect CSI scenarios, using QPSK modulation.

improvement across the region. However, once SNR increases beyond 10 dB the interference leakage starts to become more significant, resulting in an interference limited system; this leads to a degradation in performance that eventually settles to a steady state value. Such behavior is avoided by the SKCE version of the algorithms, which also improve the overall performance. In fact for $\alpha = 0$, $\beta = 0.1$ the SKCE algorithms settle at approximately 13.0 bits per channel use above

their naive counterparts. Additionally, in terms of BER, MMSE-SKCE and Max-SINR-SKCE both settle at around 2.3×10^{-2}, while the naive versions settle at 1.3×10^{-1}.

6.6.3 Determining IA feasibility in multi-cell systems

Next we focus on how the proposed algorithms can be used to give an indication of IA feasibility for FD multi-cell systems with HD users. For example, consider a system having $G = 2$ and $K_d = K_u = 2$ with each user requiring 2 streams, i.e. $b_d = b_u = 2$. We want to determine the antennas required at the BS, M_B, and at the users, $N = N_d = N_u$, to ensure that full DoF equal to $G(K_d b_d + K_u b_u) = 16$ are achievable.

If all nodes have 16 antennas, i.e. $\{M_B = N = 16\}$, the desired number of streams can easily be delivered, however from an achievable DoF perspective this leads to an unnecessarily large number of antennas; with IA we should achieve the same DoF with less antennas. For a HD system, with M_B BS antennas and N user antennas, to deliver 2 streams each to $K = 4$ users per cell across two cells (i.e achieve total DoF of 16), we need $M_B \geq 2(4 + p)$ and $N \geq 2(5 - p)$ where $p \in \{1, 2, 3, 4\}$ [7]. With $p = 1$ this evaluates to $M_B \geq 10$ and $N \geq 8$, implying that $\{M_B = 10, N = 8\}$ is the minimum number of antennas required to achieve 16 DoF in the HD system.

Moving on to our FD system, we use the proposed multi-cell algorithms from Section 6.5 with perfect CSI to obtain the results in Fig. 6.8. As can be seen results for $\{M_B = N = 16\}$, $\{M_B = 10, N = 8\}$, $\{M_B = 10, N = 7\}$ and $\{M_B = 9, N = 8\}$ have the same slope and achieve full DoF. However for $\{M_B = 9, N = 7\}$, $\{M_B = 10, N = 6\}$, $\{M_B = 9, N = 6\}$, $\{M_B = 10, N = 5\}$, the sum rate flattens out as SNR increases, indicating that IA is infeasible. Table 6.1 relates the feasibility of the various system configurations simulated in Fig. 6.8 with the properness of the system according to (6.49). As can be seen systems marked as improper are always infeasible, however systems marked as proper are not necessarily feasible. In fact for $\{M_B = 9, N = 7\}$ and $\{M_B = 10, N = 6\}$, where the properness condition is met with equality, the resulting scenario is proper but infeasible.

Moreover, it can be noticed that results for $\{M_B = 10, N = 8\}$, $\{M_B = 10, N = 7\}$ and $\{M_B = 9, N = 8\}$ obtain very similar rates with a marginal

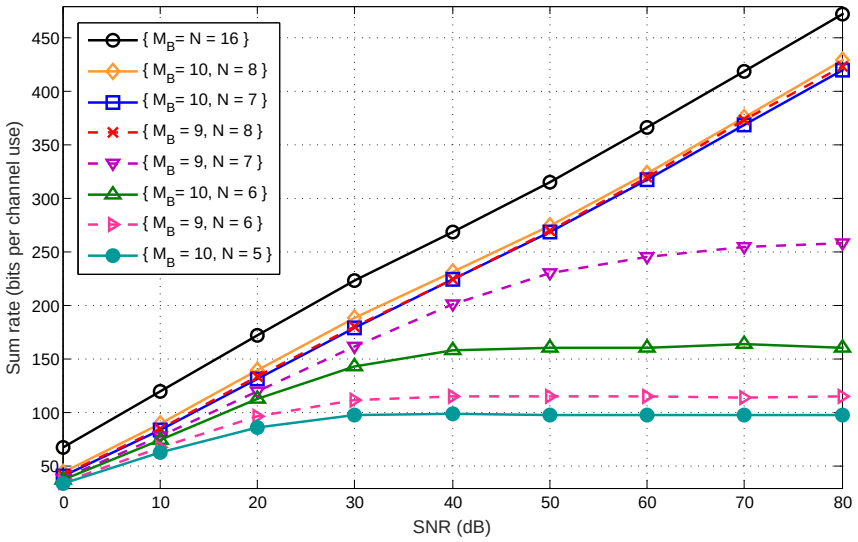


Figure 6.8: Sum rate achieved using both FD-Max-SINR-Naive and FD-MMSE-Naive under perfect CSI conditions for system with $G = 2$, $K_d = K_u = 2$, $b_d = b_u = 2$ and varying antenna numbers.

Table 6.1: Properness and IA feasibility for systems simulated in Fig. 6.8.

M_B	N	*Properness of system*	*IA Feasibility*
16	16	Proper with $N_v > N_e$	Feasible
10	8	Proper with $N_v > N_e$	Feasible
10	7	Proper with $N_v > N_e$	Feasible
9	8	Proper with $N_v > N_e$	Feasible
9	7	Proper with $N_v = N_e$	Infeasible
10	6	Proper with $N_v = N_e$	Infeasible
9	6	Improper with $N_v < N_e$	Infeasible
10	5	Improper with $N_v < N_e$	Infeasible

increase for an increasing number of antennas. The rate for $\{M_B = N = 16\}$ is the highest across the whole SNR range; however this rate advantage comes from having a significantly larger number of antennas compared to the other configurations where IA is also feasible.

6.6.4 Convergence results

Fig. 6.9 shows the convergence behaviour of the designed IA algorithms. For each scenario plotted we consider an SNR of 10 dB and average the results over 200 channel realisations under perfect CSI. As can be seen for all scenarios the proposed algorithms do indeed converge to a constant value.

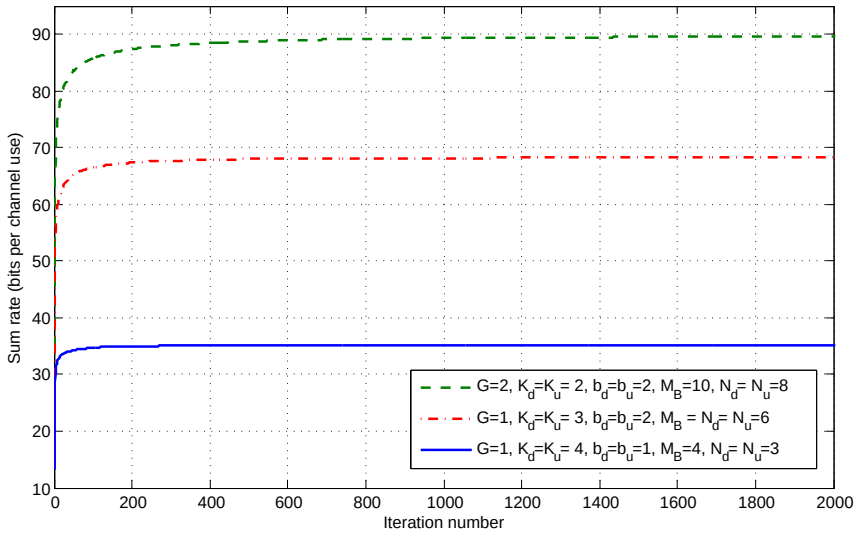


Figure 6.9: Sum rate convergence trend averaged over 200 channel realisations for both Max-SINR and MMSE based algorithm designs at an SNR of 10 db, under perfect CSI.

6.7 Conclusion

The combination of FD technology and IA provides a promising solution to tackle the ever increasing resource demand problem in wireless networks. While the performance benefits are clear under perfect CSI, it is important to consider imperfect CSI scenarios to obtain a more practical characterisation of the system's behaviour. In this chapter, we considered the use of linear IA in a multi-cell system with FD BSs and legacy HD users, and characterised the performance by deriving a bound on the loss in sum rate and quantifying the DoF loss incurred due to imperfect CSI. Results show that the rate loss is bounded by a derived value when the error is exactly inversely proportional to SNR, it goes to zero when the error scales with SNR to the power of a proper fraction, and is otherwise unbounded. Additionally, depending on how the error scales with SNR, full DoF are still achievable under imperfect CSI. We also proposed two novel IA algorithms based on MMSE and Max-SINR, referred to as FD-MMSE-SKCE and FD-Max-SINR-SKCE respectively, that are applicable to an FD multi-cell system with HD users. Our designs exploit statistical knowledge with respect to the CSI error and produce unitary beamformers. They are shown to be equivalent for cases where η is set to 0, and improve performance over the naive designs both in terms of sum rate and BER. Moreover, for the multi-cell case where the feasibility of linear IA

is not yet explored in literature, they can help discern system feasibility.

Appendix 6.A Useful Lemmas

Lemma 6.1. $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}|^2\}$ *and* $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m}|^2\}$ *are both equal to* $\eta/(1+\eta)\ \forall\ k,g,n,i,j,m.$

Proof. Let us first focus on $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}|^2\}$. From the definition of the error model in Section 6.2.3 we know that $\widehat{\mathbf{H}}$ and $\boldsymbol{\Upsilon}$ are independent of each other. Since $\widehat{\mathbf{u}}_{k_g^d}^{n}$ and $\widehat{\mathbf{v}}_{i_j^d}^{m}$ are calculated on $\widehat{\mathbf{H}}$, this makes both beamformers automatically independent of $\boldsymbol{\Upsilon}$. Additionally, $\boldsymbol{\Upsilon}$ is Gaussian and bi-unitarily invariant [74], thus for unitary beamformers the product $\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}\quad\forall\ k,g,n,i,\ j,m$ is a Gaussian random variable with zero mean and variance $\eta/(1+\eta)$. Finally, using central absolute moments we can evaluate $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\widehat{\mathbf{v}}_{i_j^d}^{m}|^2\}$ which is equal to $\eta/(1+\eta)$. A similar argument based on $\widehat{\mathbf{u}}_{k_g^d}^{n}$ and $\widehat{\mathbf{v}}_{i_j^u}^{m}$ can be used to prove that $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\widehat{\mathbf{v}}_{i_j^u}^{m}|^2\}$ is equal to $\eta/(1+\eta)$. □

Lemma 6.2. $\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\widehat{\mathbf{H}}_{g,i^u}\widehat{\mathbf{v}}_{i^u}^{m}\widehat{\mathbf{v}}_{i^u}^{m\,H}\boldsymbol{\Upsilon}_{g,i^u}^{H}\}=\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\boldsymbol{\Upsilon}_{g,i^u}\widehat{\mathbf{v}}_{i^u}^{m}\widehat{\mathbf{v}}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{g,i^u}^{H}\}=0\ \forall\ m,i.$

Proof. Beamforming elements are calculated using $\widehat{\mathbf{H}}$, thus they are automatically independent of $\boldsymbol{\Upsilon}$ from the definition of the imperfect CSI model in Section 6.2.3. □

Lemma 6.3. *[116] For matrix* $\mathbf{A}\in\mathbb{C}^{M\times M}$ *and vector* $\mathbf{b}\in\mathbb{C}^{M\times 1}$

$$(\mathbf{A}-\mathbf{b}\mathbf{b}^{H})^{-1}\mathbf{b}=\frac{\mathbf{A}^{-1}\mathbf{b}}{1-\mathbf{b}^{H}\mathbf{A}^{-1}\mathbf{b}}\ .$$

Chapter 7

Conclusion

7.1 Summary of contributions

The demand for mobile wireless network resources is constantly on the rise, pushing for new communication technologies that are able to support unprecedented rates. In this thesis we have addressed the issue by studying interference management solutions in order to exploit the available RF spectrum more efficiently under relaxed CSI conditions. The relaxed CSI contexts considered range from the availability of only global topological information at the transmitters in Chapter 3, to the availability of global imperfect CSI in Chapters 4, 5 and 6. Additionally, while the studies in Chapters 3 and 4 focused on current HD technology, the latter half of the thesis in Chapters 5 and 6 considered FD capability at the BSs as a way to further boost spectral efficiency.

Starting with the TIM framework, in Chapter 3 we studied the DoF of a two-cell two-user-per-cell IBC with alternating inter-cell connectivity and global topological information. We derived DoF outerbounds for both the no spatial multiplexing case (SISO) and the case where spatial multiplexing can be applied (MISO with local CSIT and MIMO without local CSIT). After proposing novel transmission schemes based on joint coding across states, we also obtained achievable DoF expressions for both scenarios. The derived bounds are shown to achievable under certain conditions, and for the best case scenario provide a two fold increase in achievable DoF. Results from this chapter clearly show that significant DoF gains can be obtained when transmitters are provided with global topological information, indicating that even such a minimal level of global CSIT

is still highly beneficial.

Moving on to the availability of global imperfect CSI, we considered the application of linear IA in multi-user multi-cell systems, both for a HD downlink scenario in Chapter 4, and also for systems where FD BSs communicate with DL and UL users simultaneously in Chapter 6. Our goal in these chapters was to obtain a more practical understanding of the impact of imperfect CSI on IA performance. The CSI error model used allowed us to treat the error either as a function of the SNR (ρ) or as independent of it, by representing the error variance as $\eta = \beta\rho^{-\alpha}$. Based on this error model, we derived outer bounds on the asymptotic mean loss in sum rate and characterised the DoF loss due to CSI mismatch for the HD system in Chapter 4 and for the FD one in Chapter 6. Results show that the SNR exponent is highly important in determining the performance loss behaviour for both the HD and the FD systems. When the error is exactly inversely proportional to SNR, the rate loss is upper bounded by a derived value and no DoF losses are incurred. Additionally, both metrics go to zero when the error scales with SNR to the power of a proper fraction. However, when they scale to the power of an improper fraction, the rate loss is unbounded and achievable DoF go to zero.

Moreover, in Chapters 4 and 6 we also designed a number of linear IA algorithms applicable to the corresponding system mode. These take into account statistical knowledge of the CSI error for added robustness and provide performance improvements over their naive counterparts both in terms of sum rate and BER, without incurring any additional computational costs. For the HD system we proposed a novel version of the Max-SINR algorithm, called Max-SINR-SKCE, in Chapter 4. For the FD case we proposed two algorithms in Chapter 6, an MMSE based one called FD-MMSE-SKCE, and a Max-SINR based one called FD-Max-SINR-SKCE. These FD algorithms do not only provide numerical values for the beamformers, but can be used in conjunction with the derived proper condition to discern IA feasibility for different system configurations, since such theoretical knowledge is not yet available in literature so far.

When it comes to FD systems, we also considered filter design for WSR maximisation in multi-user multi-cell MIMO networks with FD BSs and legacy HD users in Chapter 5. Since WSR problems are known to be non-convex, we established a relationship between rate and MSE for both DL and UL communication, and proposed WMMSE alternating optimisation algorithms that are proven to

converge. These problems are solved both for perfect CSI, and imperfect CSI under two types of error models. Furthermore, we proposed an additional design that maximises the total DL rate subject to a per UL user target rate, suitable for situations where it is important to ensure that each UL user is served in every time slot. Simulation results for small cell scenarios show that replacing standard HD BSs with FD ones within this context can indeed increase achievable sum rate for low to intermediate transmitter and/or distortion levels, with gains of up to 1.92 for the best case scenario. They also confirm that these advantages are still present under imperfect CSI conditions.

Therefore, across Chapters 3 to 6 of this thesis we have contributed to further the understanding of how the RF spectrum can be used more efficiently, and we have shown that this is indeed possible, both via the exploitation of any available CSI (even if it is of limited quality), and also via the introduction of FD enabled nodes.

7.2 Future work

The work presented in this thesis opens a number of interesting areas for future research, some of the more promising directions are listed below.

7.2.1 Topological interference management

- The TIM study in Chapter 3 shows that even such a limited amount of global CSIT can result in significant DoF gains. However, our work is restricted to the two-cell two-user-per-cell scenario, and for the MIMO setting is only applicable to $M \times 2$ and $2 \times N$ configurations. Analysing a more general G-cell K-user-per-cell $M \times N$ system would be of great interest due to its wider applicability. Such studies can be carried out not only within the alternating connectivity setting considered in Chapter 3, but also for other types of partially connected networks, including ones that use FD nodes.

- The TIM problem can be expanded to cater for cases where the topological knowledge is not uniform across all nodes, for example: (a) *partial topological information* - where the status of some links is unknown, and (b)

mismatched topological information - where different nodes have different knowledge with respect to the network's structure.

- The TIM framework requires the use of a pre-established noise floor in order to determine whether a link is classified as weak or strong. If the chosen threshold value is too low, then the network can mistakenly be classified as fully connected. However, if it is too high then strong interfering links are classified as noise, affecting the operating SNR significantly. Therefore, looking into ways on how to determine the optimal value for the noise floor is an important research direction for the practical realisation of TIM solutions.

7.2.2 Interference alignment

- The IA studies in Chapters 4 and 6 design IA beamformers that are applicable to perfect and imperfect CSI scenarios; however, it would also be of interest to study novel transceiver design methods that can achieve IA when only topological information is available at the transmitters, i.e. topological interference alignment (TIA) solutions [117]. To implement TIA, instead of satisfying the standard IA conditions that take into account the actual channel value, in addition to the precoder and the receiver, we need to observe a new set of conditions that are independent of the channel itself. For example for the IC, the TIA the feasibility conditions are expressed as

$$\begin{aligned} \mathbf{U}_i\mathbf{V}_j &= \mathbf{0} \;\; \forall\, i \neq j \\ |\mathbf{U}_i\mathbf{V}_i| &> \mathbf{0} \;\; \forall\, i\,. \end{aligned} \tag{7.1}$$

This is contrast to the standard IA feasibility conditions for the IC given by

$$\begin{aligned} \mathbf{U}_i\mathbf{H}_{ij}\mathbf{V}_j &= \mathbf{0} \;\; \forall\, i \neq j \\ |\mathbf{U}_i\mathbf{H}_{ii}\mathbf{V}_i| &> \mathbf{0} \;\; \forall\, i\,. \end{aligned}$$

From (7.1), it can be noticed that TIA is only applicable to cases where the channel $\mathbf{H}_{ij}$ can be expressed as $h_{ij}\mathbf{I}$. For example this can occur in a slow fading scenario with time extensions, resulting in an asymptotic IA scheme.

- The IA study in Chapter 6 proposes algorithms that are applicable to multi-cell networks with FD BSs and HD users, and also derives the proper condition for multi-cell systems. While these provide an insight into the possible feasibility of various system configurations, they do not provide a complete set of expressions that can be directly used to determine whether any given system is feasible or not. Thus, a theoretical study that looks into the derivation of all the necessary and sufficient conditions to achieve IA for the system under consideration would be highly complementary to the study in Chapter 6.

7.2.3 FD enabled networks

- The WSR problems in Chapter 5 consider per node sum power constraints for transmission, however, it also possible to consider per antenna power constraints. Such constraints are more practical since they take into account the possibility that the transmitter may be unable to allocate power arbitrarily amongst its own antennas. Relevant scenarios include distributed MIMO systems where antennas are not co-located, and also non-distributed ones due to limitations in the individual RF chains. Considering the perfect CSI WSR problem with sum power constraints in (5.15), we can adapt it as follows

$$\begin{aligned} \max_{\mathbf{V}} \quad & \sum_{g=1}^{G}\sum_{k=1}^{K_g^d} \mu_{k_g^d} R_{k_g^d} + \sum_{g=1}^{G}\sum_{k=1}^{K_g^u} \mu_{k_g^u} R_{k_g^u} \\ \text{s.t.} \quad & \operatorname{diag}(\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H) \leq \mathbf{P}_{k_g^u} \ \ \forall\, k, g \\ & \sum_{k=1}^{K_g^d} \operatorname{diag}(\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H) \leq \mathbf{P}_g \ \ \forall\, g\,. \end{aligned} \tag{7.2}$$

  to consider per antenna power constraints. Here, $\mathbf{P}_{k_g^u}$ is an all zero matrix except for elements along the diagonal which consist of $P_{k_g^u}^{[1]}, P_{k_g^u}^{[2]}, \ldots, P_{k_g^u}^{[M_U]}$ consecutively, with $P_{k_g^u}^{[i]}$ being the maximum transmit power at the ith antenna of user k_g^u, and $\mathbf{P}_g$ is an all zero matrix except for elements along the diagonal which consist of $P_g^{[1]}, P_g^{[2]}, \ldots, P_g^{[M_B]}$ consecutively, with $P_g^{[i]}$ being the maximum transmit power at the ith antenna of BS g. The remaining variables are defined as in (5.15). Similar per antenna power constraint adjustments can be made for all of the problems considered in Chapter 5.

- For the DL rate maximisation problem with a per UL user target rate, we only consider perfect CSI in Section 5.6, however this problem can also be solved for imperfect CSI scenarios. Considering the norm-bounded error model defined in (5.5), we need to solve

$$\begin{aligned}
\max_{\mathbf{V}} \min_{\Delta} \quad & \sum_{g=1}^{G} \sum_{k=1}^{K_g^d} \mu_{k_g^d} R_{k_g^d} \\
\text{s.t.} \quad & R_{k_g^u} \geq R_{UL} \ \ \forall\, k, g \\
& \mathrm{Tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \leq P_U \ \ \forall\, k, g \\
& \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) \leq P_B \ \ \forall\, g \\
& \{\Delta_{k_g^d, i_j^u} : \|\Delta_{k_g^d, i_j^u}\|_F \leq \varepsilon_{k_g^d, i_j^u}\} \ \ \forall\, k, g, i, j \\
& \{\Delta_{k_g^d, j} : \|\Delta_{k_g^d, j}\|_F \leq \varepsilon_{k_g^d, j}\} \ \ \forall\, k, g, j \\
& \{\Delta_{g, i_j^u} : \|\Delta_{g, i_j^u}\|_F \leq \varepsilon_{g, i_j^u}\} \ \ \forall\, g, i, j \\
& \{\Delta_{g,j} : \|\Delta_{g,j}\|_F \leq \varepsilon_{g,j}\} \ \ \forall\, g, j
\end{aligned} \tag{7.3}$$

  where variable definitions follow those originally specified for (5.25). For a stochastic error model as defined in (5.6), we need to consider

$$\begin{aligned}
\max_{\mathbf{V}} \quad & \sum_{g=1}^{G} \sum_{k=1}^{K_g^d} \mu_{k_g^d} R_{k_g^d}^S \\
\text{s.t.} \quad & R_{k_g^u}^S \geq R_{UL} \ \ \forall\, k, g \\
& \mathrm{Tr}(\mathbf{V}_{k_g^u} \mathbf{V}_{k_g^u}^H) \leq P_U \ \ \forall\, k, g \\
& \sum_{k=1}^{K_g^d} \mathrm{Tr}(\mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H) \leq P_B \ \ \forall\, g
\end{aligned} \tag{7.4}$$

  where the variable definitions follow those specified for (5.32).

- The FD scenario analysed in this thesis (Chapters 5 and 6) focuses on a system with FD BSs and legacy HD users. This setting is chosen since it is more foreseeably realisable in the near future; however, once the current HD infrastructure is replaced by FD enabled devices, FD technology is expected to proliferate to user devices as well. Therefore, studying interference management solutions for multi-cell multi-user networks equipped

with FD nodes at both BS and user ends is of great interest for the more distant future. Work in this direction can focus both on a variety of optimisation problems, and also the application and theory of IA and TIM techniques.

7.2.4 Relaxed CSI conditions

With respect to relaxed CSI conditions, this thesis has considered either the availability of global topological information (Chapter 3) or the availability of imperfect CSI (Chapters 4, 5 and 6); however, CSI may have a number of additional impairments. Some issues considered in BC and IC literature include:

- *delayed CSI* - where at time (t) transmitters only have perfect knowledge of the CSI up to time $(t-1)$ [16];
- *mixed delayed and current CSI* - where at time (t) transmitters have perfect knowledge of the CSI up to time $(t-1)$ and also imperfect knowledge of the CSI at time (t) [118];
- *partial CSI* - where transmitters have access to only a subset of the global CSI [19], this may either be different for each transmitter or the same across all.

The aforementioned CSI contexts are as of yet relatively unexplored for the IBC/IMAC and FD cellular systems. Since such network models provide a better representation of both present day (IBC/IMAC) and future (FD enabled networks) practical communication systems, extending the study of such relaxed CSI settings to these more complex system models is highly relevant.

List of Publications

- P. Aquilina, A. C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks", *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1590-1608, Apr. 2017
- P. Aquilina and T. Ratnarajah, "On the degrees of freedom of interference broadcast channels with topological interference management," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1477-1489, Apr. 2016.
- P. Aquilina and T. Ratnarajah, "Performance analysis of IA techniques in the MIMO IBC with imperfect CSI," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1259-1270, Apr. 2015.
- P. Aquilina and T. Ratnarajah, "Beamformer design for interference alignment in full-duplex cellular networks with imperfect CSI," in *Proc. IEEE ICC*, May 2017.
- P. Aquilina and T. Ratnarajah, "Topological interference management for interference broadcast channels with alternating connectivity," in *Proc. IEEE ISIT*, Jun. 2015, pp. 2426-2430.
- P. Aquilina and T. Ratnarajah, "Topological interference management for two cell interference broadcast channels with alternating connectivity," in *Proc. IEEE ICASSP*, Apr. 2015, pp. 3043-3047.

Bibliography

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020 white paper," 2016. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] O. E. Ayach, A. Lozano, and R. W. Heath, "On the overhead of interference alignment: Training, feedback, and cooperation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4192–4203, Nov. 2012.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.

[4] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[5] S. A. Jafar, "Interference alignment - A new look at signal dimensions in a communication network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, pp. 1–134, 2011. [Online]. Available: http://dx.doi.org/10.1561/0100000047

[6] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.

[7] T. Liu and C. Yang, "On the feasibility of linear interference alignment for MIMO interference broadcast channels with constant coefficients," *IEEE Trans. Signal Process*, vol. 61, no. 9, pp. 2178–2191, May 2013.

[8] K. Kim, S. W. Jeon, and D. K. Kim, "The feasibility of interference alignment for full-duplex MIMO cellular networks," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1500–1503, Sep. 2015.

[9] R. Tresch, M. Guillaud, and E. Riegler, "On the achievability of interference alignment in the K-user constant MIMO interference channel," in *Proc. 15th IEEE/SP Workshop on Statistical Signal Process.*, Aug. 2009, pp. 277–280.

[10] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[11] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *Proc. 43rd Asilomar Conf. on Signals, Systems and Computers*, Jan. 2009, pp. 1106–1110.

[12] S. A. Jafar and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[13] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6355–6371, Oct. 2012.

[14] S. A. Jafar, "Blind interference alignment," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 3, pp. 216–227, Jun. 2012.

[15] T. Gou, C. Wang, and S. A. Jafar, "Aiming perfectly in the dark-blind interference alignment through staggered antenna switching," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2734–2744, Jun. 2011.

[16] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.

[17] T. Gou and S. A. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MIMO BC with mixed CSIT," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1084–1087, Jul. 2012.

[18] Z. Wang, M. Xiao, C. Wang, and M. Skoglund, "Degrees of freedom of two-hop MISO broadcast networks with mixed CSIT," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6982–6995, Dec. 2014.

[19] P. de Kerret and D. Gesbert, "Interference alignment with incomplete CSIT sharing," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2563–2573, May 2014.

[20] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

[21] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[22] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[23] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM*, Aug. 2014, pp. 375–386.

[24] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[25] M. E. Knox, "Single antenna full duplex communications using a common carrier," in *Proc. IEEE WAMICON*, Apr. 2012, pp. 1–6.

[26] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. MOBICOM*, 2010, pp. 1–12.

[27] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[28] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. MOBICOM*, Sep. 2011, pp. 301–312.

[29] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. K. Ramakrishnan, C. W. Rice, and N. K. Shankaranarayanan, "Design and characterization of a full-duplex multiantenna system for WiFi networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1160–1177, Mar. 2014.

[30] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: Self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 128–137, May 2015.

[31] S. Goyal, P. Liu, S. Panwar, R. A. DiFazio, R. Yang, and E. Bala, "Full duplex cellular systems: Will doubling interference prevent doubling capacity?" *IEEE Commun. Magazine*, vol. 53, no. 5, pp. 121–127, May 2015.

[32] P. Aquilina and T. Ratnarajah, "On the degrees of freedom of interference broadcast channels with topological interference management," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1477–1489, Apr. 2016.

[33] ——, "Topological interference management for two cell interference broadcast channels with alternating connectivity," in *Proc. IEEE ICASSP*, Apr. 2015, pp. 3043–3047.

[34] ——, "Topological interference management for interference broadcast channels with alternating connectivity," in *Proc. IEEE ISIT*, Jun. 2015, pp. 2426–2430.

[35] T. Gou, S. A. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3286–3308, Jun. 2011.

[36] H. Maleki, S. A. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 3, pp. 228–240, Jun. 2012.

[37] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.

[38] R. Tandon, S. A. Jafar, S. Shamai, and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106–4128, Jul. 2013.

[39] V. Aggarwal, A. S. Avestimehr, and A. Sabharwal, "On achieving local view capacity via maximal independent graph scheduling," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2711–2729, May 2011.

[40] X. Yi and D. Gesbert, "Topological interference management with transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6107–6130, Nov. 2015.

[41] Y. Gao, G. Wang, and S. A. Jafar, "Topological interference management for hexagonal cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2368–2376, May 2015.

[42] H. Sun, C. Geng, and S. A. Jafar, "Topological interference management with alternating connectivity," in *Proc. IEEE ISIT*, Jul. 2013, pp. 399–403.

[43] S. Gherekhloo, A. Chaaban, and A. Sezgin, "Topological interference management with alternating connectivity: The Wyner-type three user interference channel." 2013. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr1310.html#GherekhlooCS13a

[44] ——, "Resolving entanglements in topological interference management with alternating connectivity," in *Proc. IEEE ISIT*, Jun. 2014, pp. 1772–1776.

[45] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[46] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 849–857, Feb. 2012.

[47] I. Kontoyiannis and M. Madiman, "Sumset and inverse sumset inequalities for differential entropy and mutual information," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4503–4514, Aug. 2014.

[48] P. Aquilina and T. Ratnarajah, "Performance analysis of IA techniques in the MIMO IBC with imperfect CSI," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1259–1270, Apr. 2015.

[49] J. Thukral and H. Bolcskei, "Interference alignment with limited feedback," in *Proc. IEEE ISIT*, Jun. 2009, pp. 1759–1763.

[50] R. T. Krishnamachari and M. K. Varanasi, "Interference alignment under limited feedback for MIMO interference channels," *IEEE Trans. Signal Process*, vol. 61, no. 15, pp. 3908–3917, Aug. 2013.

[51] B. Nosrat-Makouei, J. G. Andrews, and R. W. Heath, "MIMO interference alignment over correlated channels with imperfect CSI," *IEEE Trans. Signal Process*, vol. 59, no. 6, pp. 2783–2794, Jun. 2011.

[52] O. E. Ayach and R. W. Heath, "Interference alignment with analog channel state feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 626–636, Feb. 2012.

[53] R. Tresch and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. IEEE ICC Workshops*, Jun. 2009, pp. 1–5.

[54] C. Suh and D. Tse, "Interference alignment for cellular networks," in *Proc. 46th Annual Allerton Conf. on Commun., Control, and Computing*, Sep. 2008, pp. 1037–1044.

[55] Y. Ma, J. Li, R. Chen, and Q. Liu, "On feasibility of interference alignment for L-cell constant cellular interfering networks," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 714–716, May 2012.

[56] J. Kim, S. H. Park, H. Sung, and I. Lee, "Spatial multiplexing gain for two interfering MIMO broadcast channels based on linear transceiver," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3012–3017, Oct. 2010.

[57] G. Sridharan and W. Yu, "Degrees of freedom of MIMO cellular networks with two cells and two users per cell," in *Proc. IEEE ISIT*, Jul. 2013, pp. 1774–1778.

[58] T. Kim, D. J. Love, B. Clerckx, and D. Hwang, "Spatial degrees of freedom of the multicell MIMO multiple access channel," in *Proc. IEEE GLOBECOM*, Dec. 2011, pp. 1–5.

[59] T. Liu and C. Yang, "Degrees of freedom of general symmetric MIMO interference broadcast channels," in *Proc. IEEE ICASSP*, May 2013, pp. 4364–4368.

[60] G. Sridharan and W. Yu, "Degrees of freedom of MIMO cellular networks: Decomposition and linear beamforming design," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3339–3364, Jun. 2015.

[61] T. Liu and C. Yang, "On the degrees of freedom of asymmetric MIMO interference broadcast channels," in *Proc. IEEE ICC*, Jun. 2014, pp. 1971–1976.

[62] B. Zhuang, R. A. Berry, and M. L. Honig, "Interference alignment in MIMO cellular networks," in *Proc. IEEE ICASSP*, May 2011, pp. 3356–3359.

[63] J. Schreck and G. Wunder, "Distributed interference alignment in cellular systems: Analysis and algorithms," in *Proc. European Wireless Conf.*, Apr. 2011, pp. 1–8.

[64] R. K. Mungara, A. Tölli, and M. Juntti, "Degrees of freedom and interference mitigation for MIMO interfering broadcast channels," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2011, pp. 441–446.

[65] J. Park, Y. Sung, and H. V. Poor, "On beamformer design for multiuser MIMO interference channels," 2010. [Online]. Available: https://arxiv.org/abs/1011.6121

[66] I. Santamaria, O. Gonzalez, R. W. Heath, and S. W. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–6.

[67] J. Maurer, J. Jaldén, D. Seethaler, and G. Matz, "Vector perturbation precoding revisited," *IEEE Trans. Signal Process*, vol. 59, no. 1, pp. 315–328, Jan. 2011.

[68] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[69] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[70] J. Baltersee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 12, pp. 2358–2368, Dec. 2001.

[71] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008.

[72] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[73] S. M. Razavi and T. Ratnarajah, "Performance analysis of interference alignment under CSI mismatch," *IEEE Trans. Veh. Tech.*, vol. 63, no. 9, pp. 4740–4748, Nov. 2014.

[74] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004. [Online]. Available: http://dx.doi.org/10.1561/0100000001

[75] P. Aquilina, A. C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1590–1608, Apr. 2017.

[76] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3933–3938, Aug. 2009.

[77] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.

[78] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2498–2511, May 2013.

[79] A. C. Cirik, R. Wang, Y. Rong, and Y. Hua, "MSE-based transceiver designs for full-duplex MIMO cognitive radios," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2056–2070, Jun. 2015.

[80] D. Nguyen, L. N. Tran, P. Pirinen, and M. Latva-aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896–4910, Sep. 2014.

[81] S. Goyal, P. Liu, S. Panwar, R. A. DiFazio, R. Yang, J. Li, and E. Bala, "Improving small cell capacity with common-carrier full duplex radios," in *Proc. IEEE ICC*, Jun. 2014, pp. 4987–4993.

[82] S. Goyal, P. Liu, and S. Panwar, "User selection and power allocation in full duplex multi-cell networks," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–1, 2016.

[83] R. K. Mungara and A. Lozano, "Interference surge in full-duplex wireless systems," in *Proc. 49th Asilomar Conf. Signals, Systems and Computers*, Nov. 2015, pp. 25–29.

[84] S. Goyal, P. Liu, S. Hua, and S. Panwar, "Analyzing a full-duplex cellular system," in *Proc. CISS*, Mar. 2013, pp. 1–6.

[85] S. Wang, V. Venkateswaran, and X. Zhang, "Exploring full-duplex gains in multi-cell wireless networks: A spatial stochastic framework," in *Proc. INFOCOM*, Apr. 2015, pp. 855–863.

[86] 3GPP TR 36.828, "Further enhancements to LTE TDD for DL-UL interference management and traffic adaptation," Release 11, Jun. 2012.

[87] Y. S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.

[88] P. Tehrani, F. Lahouti, and M. Zorzi, "Resource allocation in OFDMA networks with half-duplex and imperfect full-duplex users," in *Proc. IEEE ICC*, May 2016, pp. 1–6.

[89] D. Nguyen, L. N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.

[90] A. C. Cirik, R. Wang, Y. Hua, and M. Latva-aho, "Weighted sum-rate maximization for full-duplex MIMO interference channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 801–815, Mar. 2015.

[91] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," in *Proc. IEEE ICC*, Jun. 2009, pp. 1–6.

[92] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[93] J. Shin and J. Moon, "Weighted-sum-rate-maximizing linear transceiver filters for the K-user MIMO interference channel," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2776–2783, Oct. 2012.

[94] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Multi-objective optimization for robust power efficient and secure full-duplex wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5511–5526, Aug. 2016.

[95] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5473–5484, Nov. 2011.

[96] J. Jose, N. Prasad, M. Khojastepour, and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *Proc. IEEE ICC*, Jun. 2011, pp. 1–6.

[97] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.

[98] J. Wang, M. Bengtsson, B. Ottersten, and D. P. Palomar, "Robust MIMO precoding for several classes of channel uncertainty," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3056–3070, Jun. 2013.

[99] J. Wang and D. P. Palomar, "Robust MMSE precoding in MIMO channels with pre-fixed receivers," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5802–5818, Nov. 2010.

[100] L. Vandenberghe, S. Boyd, and S. P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 19, no. 2, pp. 499–533, Apr. 1998. [Online]. Available: http://dx.doi.org/10.1137/S0895479896303430

[101] R. Hunger, "Floating point operations in matrix-vector calculus," 2007. [Online]. Available: https://mediatum.ub.tum.de/doc/625604/625604.pdf

[102] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, no. 1, pp. 193 – 228, 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0024379598100320

[103] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 1931–1946, Jul. 2004.

[104] P. Aquilina and T. Ratnarajah, "Beamformer design for interference alignment in full-duplex cellular networks with imperfect CSI," in *Proc. IEEE ICC*, May. 2017.

[105] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," in *Proc. IEEE ITW*, Sep. 2013, pp. 1–5.

[106] J. Bai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of multi-user MIMO full-duplex network," in *Proc. IEEE ISIT*, Jun. 2015, pp. 864–868.

[107] S. H. Chae and S. H. Lim, "Degrees of freedom of cellular networks: Gain from full-duplex operation at a base station," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 4048–4053.

[108] S. W. Jeon, S. H. Chae, and S. H. Lim, "Degrees of freedom of full-duplex multiantenna cellular networks," in *Proc. IEEE ISIT*, Jun. 2015, pp. 869–873.

[109] M. A. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Farajzadeh-Tehrani, "Scaling wireless full-duplex in multi-cell networks," in *Proc. INFOCOM*, Apr. 2015, pp. 1751–1759.

[110] R. de Francisco and D. T. M. Slock, "An optimized unitary beamforming technique for MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 990–1000, Mar. 2010.

[111] S. M. Razavi and T. Ratnarajah, "Adaptive LS- and MMSE-based beamformer design for multiuser MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 132–144, Jan. 2016.

[112] M. Yang, S. W. Jeon, and D. K. Kim, "Degrees of freedom of full-duplex cellular networks with reconfigurable antennas at base station," 2016. [Online]. Available: https://arxiv.org/abs/1604.07957

[113] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 206–218, Jan. 2011.

[114] C. Wilson and V. Veeravalli, "A convergent version of the max sinr algorithm for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2952–2961, Jun. 2013.

[115] S. H. Park and I. Lee, "Degrees of freedom of multiple broadcast channels in the presence of inter-cell interference," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1481–1487, May 2011.

[116] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices," *SIAM Review*, vol. 23, no. 1, pp. 53–60, Jan. 1981.

[117] Y. Shi, J. Zhang, and K. B. Letaief, "Low-rank matrix completion for topological interference management by Riemannian pursuit," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4703–4717, Jul. 2016.

[118] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8336–8356, Dec. 2013.