SAPHYRE

Contract No. FP7-ICT-248001

Interference and Utility Modelling (final)
D2.3b

Contractual date:	M24
Actual date:	M24
Authors:	Leonardo Badia, Eleftherios Karipidis, Martin Schubert, Jan Sýkora, Lei Yu, Jianshu Zhang
Participants:	CFR, LiU, FhG, CTU, TUIL
Work package:	WP2
Security:	Public
Nature:	Report
Version:	1.1
Number of pages:	69

Abstract

This final report describes the results of Task 2.3 (Interference and Utility Modelling). In the first part different interference models are discussed. Fundamental properties like convexity and monotonicity are explored and the problem of interference balancing is studied for a coupled multi-user system. Also, results on the multi-operator two-way relay channel are presented. In the second part of the deliverable, the resulting utility sets are discussed and a framework for joint scheduling and power control is derived. Also, game-theoretic strategies are studied. Finally, an approach based on wireless network coding is proposed.

Keywords

Quality-of-Service, interference, wireless network coding, relaying, game theory, resource allocation.

Contents

Notations

Mathematical Notations

$\mathbb{C}$	set of complex numbers
$\mathbb{R}$	set of real numbers
$\mathbb{Z}$	set of integers
$\mathrm{E}\{\cdot\}$	expectation
$\mathrm{Tr}\{\cdot\}$	trace of matrix
$\{\cdot\}^{\mathrm{T}}$	transpose
$\{\cdot\}^{\mathrm{H}}$	Hermitian transpose
$\{\cdot\}^{+}$	Moore-Penrose pseudo inverse
$\boldsymbol{I}_m$	m-by-m identity matrix
$\mathbf{0}_{m\times n}$	m-by-n matrix with all zero elements
$\|\cdot\|$	Euclidean norm of a vector
$\boldsymbol{x} \geq \boldsymbol{y}$	component-wise inequality of two vectors $\boldsymbol{x}$, $\boldsymbol{y}$
$\boldsymbol{x} \geq 0$	component-wise greater or equal than zero
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm of a matrix

Abbreviations

2-SRN	2-Source Relay Network
AF	Amplify and Forward
AWGN	Additive White Gaussian Noise
BC	Broadcast
BS	Base Station
CDMA	Code Division Multiple Access
C-SI	Complementary Side-Information
DF	Decode and Forward
DoF	Degree of Freedom
HDF	Hierarchical Decode and Forward
HI	Hierarchical Information
HXC	Hierarchical eXclusive Code
IC	Interference Channel
IFC	Interference
IRC	Interference Relay Channel
JDF	Joint DF
LTE	Long Term Evolution
LP	Linear Programming
MAC	Medium Access Control

MILP	Mixed-Integer Linear Programming
MIMO	Multiple-Input Multiple-Output
NC	Network Coding
NP	Non-deterministic Polynomial-time
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PHY	Physical Layer
PLNC	Physical Layer Network Coding
QoS	Quality of Service
RRA	Radio Resource Allocator
SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio
TD	Time Division
TDD	Time Division Duplex
UT	User Terminal
WNC	Wireless Network Coding
ZMCSCG	Zero-Mean Circularly Symmetric Complex Gaussian

Executive Summary

SAPHYRE aims at demonstrating how spectrum / infrastructure sharing in wireless networks improves spectral efficiency, enhances coverage, increases user satisfaction, maintains Quality-of-Service (QoS) performance, leads to increased revenue for operators, and decreases capital and operating expenditures.

Conventional wireless system design is based on the concept of interference-free communication links. Then the system becomes a collection of quasi-independent communication links. In the past, this practical approach has greatly simplified the analysis and optimisation of such systems.

However, assigning each user a separate resource is not always an efficient way of organising the system. If the number of users is high, then each user only gets a small fraction of the overall resource. Shortages are likely to occur when users have high capacity requirements. This will become even more problematic for future wireless networks, which are expected to provide high-rate services for densely populated user environments. The system then might be better utilised by allowing users to share resources, as investigated in SAPHYRE [1].

This development drives the demand for new sharing principles based on the dynamic reuse of the system resources *frequency*, *power*, and *space* (i.e. the distribution and usage of transmitting and receiving antennas over the service area). The classical design paradigm of independent point-to-point communication links is gradually being replaced by a new network-centric point of view, where users avoid or mitigate interference in a flexible way by dynamically adjusting the resources allocated to each user. In this context, interference emerges as the key performance-limiting factor [2]. Interference affects all layers of the communication system, thus new cross-layer approaches are needed for the optimisation and system level evaluation.

Sharing, and especially non-orthogonal sharing, is inherently cross-layer oriented, since it requires a joint optimisation of all the network functionalities that depend on interference. This mainly involves the physical layer and the medium access control layer, but also higher layers. Resource and infrastructure sharing are only possible if the level of mutual interference remains below a certain threshold. In order to develop efficient sharing strategies, we need interference models that provide the following degrees of freedom:

- Allocation of carrier frequencies / time slots,
- Control and allocation of transmission powers,
- Incorporation of coding schemes,
- MIMO processing schemes,
- Interoperability distance (spatial separation).

Interference determines how many users/terminals per area can be served at a certain data rate. Assigning each user a separate resource is not always an efficient way of organising the system. If the number of users is high then each user only gets a small fraction of the overall resource. Shortages occur when many users have high capacity requirements. Hence, the classical design paradigm of independent point-to-point communication links is gradually being replaced by a new network-centric point of view, where users interact and compete for the limited system resources.

MIMO offers another degrees of freedom to avoid and mitigate interference. The multi-user case is quite different from the single-user case, because the users compete for the available resources. This typically means that there is no single optimum strategy. The performance of some users can be increased at the cost of decreasing the performance of other users. The chosen operating point is often a compromise between fairness and efficiency.

All these aspects should be taken into account when modelling the interactions between different users and operators within SAPHYRE.

This report aims at providing a basis for the research carried out in the following tasks of work packages WP2, WP3 and WP4:

- T2.1 Basic Limits for System Design,
- T2.2 Applied Game Theory,
- T3.1 Applied Signal Processing,
- T3.2 Network, Resource and Interference Aware Coding and Decoding,
- T4.1 Joint PHY/MAC Optimisation and Self-Organisation.

1 Introduction

1.1 Main Contributions

In the first part of the report it is shown that interference has a certain mathematical structure that can be used as a basis for the development of efficient iterative algorithms, which are guaranteed to converge in a reasonable time. Specifically, the following results are shown:

- Many interference-coupled systems can be modelled as special cases of the framework of interference functions. This means that they behave monotonic with respect to the underlying resources. This is an important property which deserves more attention. Interesting analytical opportunities arise from interference functions which are both monotonic and convex. Then certain resource allocation problems have an equivalent convex form. That is, even if the original problem is non-convex, it is possible to derive an equivalent convex problem that yields the solution of the original problem.

- The interference framework was extended to multicarrier systems and power constraints

- The problem of interference balancing (max-min SINR balancing) is fundamental, since it provides an indicator for the congestion of an interference coupled system. In [3] sufficient conditions for existence and uniqueness of an optimiser are derived.

- A semi-distributed fixed point iteration [4] was derived for SIR balancing (see Subsection 3.1.2). The algorithm is distributed in a sense that only measured interference powers need to be exchanged between base stations that possibly belong to different operators.

- A centralized throughput maximization strategy was developed (see Subsection 3.1.3). The algorithm, which is based on a branch&bound stratgy provides the global optimum and can serve as a benchmark for the achievable performance in a shared system.

- In Section 2.4 the two-way relaying technique is extended to the resource sharing scenarios that are of interest to SAPHYRE.

The second part of the report is focused on utility models and the consequences for the optimisation of resources in the system.

- A framework for joint scheduling, power control and beamforming is derived for SINR-based QoS targets in Section 3.2. It is shown how this framework can enable non-orthogonal spectrum sharing between operators. The resource allocation and transmit design problem is jointly tackled by novel constrained optimization formulations.

- In Section 3.3, an algorithm is proposed that reaches a Pareto efficient point, which trades throughput for fairness in an efficient and tunable manner. This involves the interaction between multiple base stations, possibly belonging to different operators.

- In Section 3.4 a conceptually different approach is taken, based on ideas from network information theory. While in the classical design paradigm interference is treated as something that has to be avoided, the framework for Wireless Network Coding (WCN) proposed in Section 3.4 exploits the information-theoretic insight that in a network, "interference" can carry information. Also strong interference can be cancelled, so under certain circumstances it can be desirable to have strong interference rather than weak interference. The WNC technique introduces shared relays and the relevant coding, modulation and processing technique required to achieve the sharing gain. The PHY resource sharing management layer supported by a proper utility metric and corresponding resource allocation strategies, e.g. game theory, can provide an *additional* sharing gain on top of WNC itself.

1.2 Structure of this Report

The report is organised in two parts. In the first part, an overview on interference models used in SAPHYRE will be presented. In the second part, QoS models will be studied.

First Part (Chapter 2)

Section 2.1 derives models for the MIMO interference channel. The interference channel is an important reference scenario for the work within SAPHYRE. It serves as a basic model for the interaction between two operators.

Section 2.2 presents a fundamental mathematical approach, which models interference as a function of the transmission powers (so-called interference functions). The behaviour of the resulting system is determined by the properties of the interference functions. Typical properties include (but are not limited to) different assumptions on convexity or concavity. This approach has the advantage of being general and being amenable to mathematical analysis.

SAPHYRE is aiming at a system-wide approach for sharing. A full-blown modelling of all interdependencies in such a system is usually prohibitive and can only be managed by means of numerical simulations. Thus, having an interference model with reduced complexity is an important prerequisite for achieving the ambitious project goals. By abstracting away from real-world system details, we are able to handle certain interference-coupled systems in an analytical manner. This is expected to provide valuable insight and intuition for the development of sharing algorithms and game-theoretic approaches.

Section 2.3 analyses the fundamental issue of interference coupling. The motivation for this theoretical work stems from the lack of understanding of the max-min SIR balancing problem, which is at the core of certain resource allocation problems. By deriving conditions for the existence and uniqueness of an optimizer we provide a basis for the development of iterative distributed algorithms for balancing resources in an interference-coupled network.

Section 2.4 addresses another important interference scenario, resulting from the use of relays shared by multiple operators (infrastructure and spectrum sharing).

Second Part (Chapter 3)

Section 3.1 is on SINR-based QoS models. Many performance measures can be expressed as monotonic functions of the SINR. This provides an abstract framework for the interference models (the "physical layer") and higher layer utility models. Based on this framework, two different algorithms are proposed. A semi-distributed fixed point iteration [4] was derived for SIR balancing (see Subsection 3.1.2), and a centralized throughput maximization strategy was developed (see Subsection 3.1.3).

Section 3.2 derives a QoS framework for joint scheduling, power control and beamforming. The aim is to allocate the resources non-orthogonally, in such a way that the interference between the operators is tolerable and that QoS is guaranteed.

Section 3.3 explores the link between game theory and interference/utility models.

Section 3.4 addresses a conceptually new approach for modelling and handling interference. Instead of treating interference as noise, which should be avoided, an information-theoretic approach is explored, where signals are combined at nodes in the network, and interference-cancellation is used.

2 Interference Models

The first part of the deliverable is strongly influenced by ideas originating from resource allocation and power control theory. This is crucial, since the efficacy of sharing depends on the way resources are allocated and interference is avoided. In the existing literature, there is a great deal of research focusing on systems where the users are coupled by power cross-talk. The design philosophy behind this approach is based on the assumption that the transmission power used by one operator does interfere with the communication links of another operator (Figure 2.1). Thus, the basic idea is to avoid interference by cleverly allocating the resources or by reducing interference by signal processing, either at the transmitter or receiver. This should ideally be done under full exploitation of the available degrees of freedom listed at the beginning of the introduction.

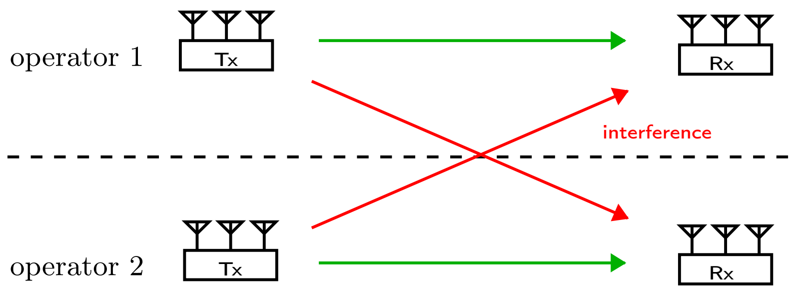


Figure 2.1: Interference between operators must be avoided in order to facilitate resource sharing. This can be achieved by orthogonal allocation of resources (the conventional case). But the goal of this section is to derive interference model that is able to describe the more interesting case of interference-coupled networks. This enables the development of non-orthogonal sharing techniques that tolerate a certain amount of interference.

In order to assess the impact of interference, we consider the following interference models known from the literature:

1) Linear functions, based on the notion of a *link gain matrix*, which characterizes the power coupling between users,

2) Interference resulting from adaptive beamforming or other receiver designs. This includes such diverse areas as beamforming [4–10], CDMA [11, 12], base station assignment [13] and axiomatic interference modesl [14]

3) Standard interference functions introduced in [15],

4) Homogeneous (scale-invariant) interference functions proposed in [4].

In recent work [16], it was shown that 2) and 3) are closely connected. In fact, every standard interference can be modelled by using the framework 3). Also, every linear interference function and interference resulting from adaptive interference mitigation strategies can be regarded as a special case of the model 3). Thus, the focus of this deliverable will be on the framework 3).

Later, in Section 3.4 we will take a different look at interference, which is motivated by recent ideas from network information theory. In this context, interference is not always considered as harmful. It is known from information theory that the optimal approach of dealing with strong interference is to cancel it, e.g. successive interference cancellation. Also, it is possible to combine signals at network nodes (Network Coding or Physical Layer Network Coding), which means that other signals are not necessarily treated as interference. This is a relatively new area of research which is still less well understood.

Consider a system of K non-orthogonal communication links that can belong to different operators. Independent signals are assumed. Interference among communication links is modelled by interference functions $\mathcal{I}_1(\boldsymbol{p}), \ldots, \mathcal{I}_K(\boldsymbol{p})$ as explained later.

The performance of each link is measured by the SINR

$$\mathrm{SINR}_k(\boldsymbol{p}) = \frac{p_k}{\mathcal{I}_k(\boldsymbol{p})} \tag{2.1}$$

The SINR and resulting performance measures depend on the following system components.

- *Power Control/Assignment.* By reducing the transmission power, less interference is caused to other links. This enables more efficient sharing strategies and the reuse of resources.
- *Signal processing.* By optimising the signal processing (e.g. MIMO), interference is avoided and multiple communication links (operators) can utilise the same resource.
- *Resource allocation.* By dynamically allocating links to frequency carriers, more bandwidth is available for each operator.

The joint optimisation of all three components involves different layers of the communication model, hence a cross-layer approach is required.

2.1 MIMO Interference Channel

In this section we discuss a special case of non-linear interference functions, resulting from the assumption of adaptive spatial filtering at the receiver. The framework can be extended to adaptive linear pre-processing at the transmitter.

The MIMO case is of special importance for SAPHYRE, since a large portion of the work will be focused on this scenario.

2.1.1 Interference Structure

Consider an uplink system with K communication links (users) $\mathcal{K} = \{1, 2, \dots, K\}$ and an M-element antenna array at the receiver, as illustrated in Figure 2.2. The

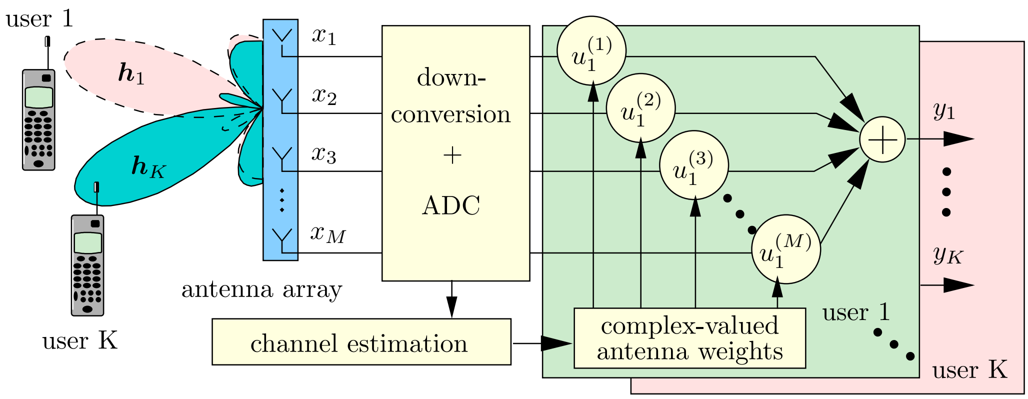


Figure 2.2: Spatial filtering at the receiver: superimposed signals are separated by a bank of spatial filters (beamformers).

following discussion is confined to the case where each user transmits a single data stream. But the framework can easily be extended to the case where each user has multiple data streams (multiplexing). For example, data can be multiplexed over different beams (spatial multiplexing) or carriers (OFDM). This adds additional degrees of freedom but it also complicates the optimisation of the communication links.

The K signals are modelled as random variables $S_1, \dots, S_K$ with $p_k = \mathrm{E}[|S_k|^2]$. Let $\boldsymbol{t}_l$, with $\|\boldsymbol{t}_l\|_2 = 1$, be a complex-valued vector whose length is equal to the number of transmit antennas of user l. The vector $\boldsymbol{t}_l$ maps the signal S_1 to the transmit antennas (Figure 2.3). The resulting signal $\boldsymbol{t}_l S_1$ is transmitted over the

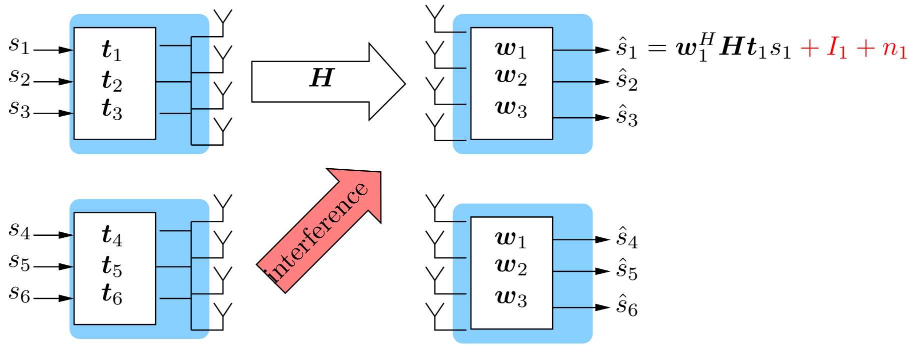


Figure 2.3: MIMO interference channel.

propagation channel, which is characterised by a matrix $\boldsymbol{H}_l$, which is assumed to be a random variable. User l is transmitting over an *effective channel* $\boldsymbol{h}_l = \boldsymbol{H}_l \boldsymbol{t}_l$, which is random as well.

The matrix $\boldsymbol{H}_l$ contains the channel coefficients between all transmit and receive antennas of user l. The number of columns equals the number of transmit antennas and the number of rows equals the number of receive antennas. Such a multiplicative channel model can be used if the signal bandwidth is relatively small in comparison to the coherence bandwidth of the propagation channel. This is the case, for example, in an OFDM system with narrowband carriers.

At the receiver, the resulting output of the antennas is $\boldsymbol{h}_l S_l$. The overall array output vector $\boldsymbol{x}$ is the superposition of the signals of all K users plus a random noise signal $\boldsymbol{n}$.

$$\boldsymbol{x} = \sum_{l\in\mathcal{K}} \boldsymbol{h}_l S_l + \boldsymbol{n} \ .$$

Assuming M receive antennas, we have $\boldsymbol{x} \in \mathbb{C}^M$. The transmitted signals $S_1, \ldots, S_K$ can be recovered by a bank of linear filters $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K \in \mathbb{C}^M$ ("beamformers"). The output of the kth filter is

$$y_k = \boldsymbol{w}_k^H \big(\sum_{l\in\mathcal{K}} \boldsymbol{h}_l S_l + \boldsymbol{n}\big) = \underbrace{\boldsymbol{w}_k^H \boldsymbol{h}_k S_k}_{\text{desired signal}} + \underbrace{\sum_{l\in\mathcal{K};\, l\neq k} \boldsymbol{w}_k^H \boldsymbol{h}_l S_l}_{\text{interference}} + \underbrace{\boldsymbol{w}_k^H \boldsymbol{n}}_{\text{effective noise}} \ . \tag{2.2}$$

The desired signal is corrupted by interference and noise. An important performance measure is the *signal-to-interference-plus-noise ratio* (SINR). Many other performance measures, like bit error rate or capacity can be linked to the SINR.

2.1.2 SINR Model

The SINR of each user depends on the transmission powers $\boldsymbol{p} = [p_1, \ldots, p_K]^T$ and the noise variance σ_n^2, which can be collected in a single variable

$$\underline{\boldsymbol{p}} = [p_1, \ldots, p_K, \sigma_n^2]^T \ . \tag{2.3}$$

Now, we introduce $\boldsymbol{R}_k = \mathrm{E}[\boldsymbol{h}_k \boldsymbol{h}_k^H]$, which is the spatial covariance matrix of the channel of user l. Exploiting that the signals are uncorrelated and $\mathrm{E}[\boldsymbol{n}\boldsymbol{n}^H] = \sigma^2 \boldsymbol{I}$, the SINR of user k is

$$\mathrm{SINR}_k(\underline{\boldsymbol{p}}, \boldsymbol{w}_k) = \frac{\mathrm{E}[|\boldsymbol{w}_k^H \boldsymbol{h}_k S_k|^2]}{\mathrm{E}[|\sum_{l\neq k} \boldsymbol{w}_k^H \boldsymbol{h}_l S_l + \boldsymbol{w}_k^H \boldsymbol{n}|^2]} = \frac{p_k \boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k}{\boldsymbol{w}_k^H \big(\sum_{l\neq k} p_l \boldsymbol{R}_l + \sigma_n^2 \boldsymbol{I}\big) \boldsymbol{w}_k} \ . \tag{2.4}$$

The beamformer that maximizers the SINR can be found efficiently via eigenvalue decomposition [17]. The SINR is maximised by any $\boldsymbol{w}_k$ fulfilling

$$\Big(\sum_{l\neq k} p_l \boldsymbol{R}_l + \sigma_n^2 \boldsymbol{I}\Big)^{-1} \boldsymbol{R}_k \cdot \boldsymbol{w}_k = \lambda_{\max} \cdot \boldsymbol{w}_k \tag{2.5}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the matrix $\left(\sum_{l \neq k} p_l \boldsymbol{R}_l + \sigma_n^2 \boldsymbol{I}\right)^{-1} \boldsymbol{R}_k$. The resulting interference-plus-noise power of user k can be written as

$$\begin{aligned}\mathcal{I}_k(\underline{\boldsymbol{p}}) &= \frac{p_k}{\max_{\|\boldsymbol{w}_k\|=1} \mathrm{SINR}_k(\underline{\boldsymbol{p}}, \boldsymbol{w}_k)} = \min_{\|\boldsymbol{w}_k\|=1} \frac{p_k}{\mathrm{SINR}_k(\underline{\boldsymbol{p}}, \boldsymbol{w}_k)} \\ &= \min_{\|\boldsymbol{w}_k\|_2=1} \frac{\sum_{l \neq k} p_l \boldsymbol{w}_k^H \boldsymbol{R}_l \boldsymbol{w}_k + \|\boldsymbol{w}_k\|^2 \sigma_n^2}{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k} = \min_{\|\boldsymbol{w}_k\|_2=1} \underline{\boldsymbol{p}}^T \boldsymbol{v}_k \end{aligned} \tag{2.6}$$

where $\boldsymbol{v}_k$ is a vector of coupling coefficients, whose lth element is defined as follows.

$$[\boldsymbol{v}_k]_l = \begin{cases} \frac{\boldsymbol{w}_k^H \boldsymbol{R}_l \boldsymbol{w}_k}{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k} & 1 \leq l \leq K, \ l \neq k \\ \frac{\|\boldsymbol{w}_k\|^2}{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k} & l = K+1, \\ 0 & l = k \ . \end{cases} \tag{2.7}$$

The notation in (2.6) might appear unnecessarily complicated, but it turns out to be useful because it shows that the interference function (2.6) is a special case of a more general framework A1-A3, which will be discussed later in Section 2.2.

The vector $\boldsymbol{v}_k(\boldsymbol{w}_k)$ determines how user k is interfered by other users. For any given $\boldsymbol{p} > 0$, the beamformer $\boldsymbol{w}_k$ can adapt to the current interference situation. One possible way of choosing $\boldsymbol{w}_k$ is to enforce $\boldsymbol{w}_k^H \boldsymbol{R}_l \boldsymbol{w}_k = 0$. This approach, known as *zero-forcing*, is suboptimal in the presence of power constraints since it neglects the noise enhancement factor $\|\boldsymbol{w}_k\|^2$ in (2.6). Another strategy is to minimise $\|\boldsymbol{w}_k\|^2$. This maximizers the signal-to-noise ratio (SNR) and is known as the *spatial matched filter*. Clearly, this approach has the disadvantage of not completely eliminating the interference. The aforementioned SINR maximisation strategy (2.6) provides a compromise between eliminating interference and minimising $\|\boldsymbol{w}_k\|^2$.

A special case occurs if the channels $\boldsymbol{h}_k$ are deterministic, then $\boldsymbol{R}_k = \boldsymbol{h}_k \boldsymbol{h}_k^H$. Such a deterministic model is usually assumed if the channel $\boldsymbol{h}_k$ is constant within the time scale of interest, so that it can be estimated. Replacing $\boldsymbol{R}_k$ in (2.5) by $\boldsymbol{h}_k \boldsymbol{h}_k^H$, it can be observed that the SINR is maximised by scalar multiples of the vector

$$\boldsymbol{w}_k^{\text{opt}} = \left(\sum_{l \neq k} p_l \boldsymbol{R}_l + \sigma_n^2 \boldsymbol{I}\right)^{-1} \boldsymbol{h}_k \ . \tag{2.8}$$

Inserting $\boldsymbol{w}_k^{\text{opt}}$ in the SINR (2.4), we obtain a closed-form expression of the interference.

$$\mathcal{I}_k(\underline{\boldsymbol{p}}) = \frac{1}{\boldsymbol{h}_k^H \left(\sigma_n^2 \boldsymbol{I} + \sum_{l \neq k} p_l \boldsymbol{R}_l\right)^{-1} \boldsymbol{h}_k} \ . \tag{2.9}$$

Note, that the function (2.9) is concave in $\underline{\boldsymbol{p}}$. This is a consequence of (2.9) being a special case of (2.6). Recall that the minimum of linear functions is concave.

Within the project we will study how concavity can be exploited for the design of resource allocation algorithms.

2.2 Interference Calculus – An Axiomatic Approach

In this section we introduce an abstract mathematical interference model [4]. The framework is based on a few basic properties (axioms). This simplicity allows us to handle certain interference-coupled systems in an analytical manner. Such an analytical approach is highly desirable. Most interference-coupled systems can only be analysed by means of numerical simulations due to their inherent complexity. By using a clean-slate analytical approach, we can obtain valuable insight about the fundamental behaviour of interference-coupled systems. This might provide conceptually new long-term solutions, and a road map for the development of new resource allocation strategies.

2.2.1 Interference Functions

In this section we model the transmission powers of all L communication links in the system as an L-dimensional non-negative vector $\boldsymbol{p} = [p_1, p_2, \ldots, p_L]^T \in \mathbb{R}_+^L$. Here, L is the overall number of resources in the system and $\boldsymbol{p}$ is the optimisation variable, that determines how transmission powers are allocated to the system resources. We also use the index set $\mathcal{K}$ to denote the set of all communication link indices of the system.

Note that the vector $\boldsymbol{p}$ can contain the powers of different operators. Also, this model can include the allocation of communication streams to different OFDM carriers or other orthogonal resources. Different links can belong to the same user, e.g. when data streams are multiplexed over frequencies, or spatial beams. Inactive links correspond to zero entries of the vector $\boldsymbol{p}$.

The design goal is to choose $\boldsymbol{p} \geq 0$ in such a way that excess interference between users is avoided. This approach is based on the assumption that we have L independent data streams that possibly interfere with each other.

In order to assess the power cross-talk between the data streams, we use the theory of *interference functions* [4, 15]. Let $\mathcal{I} : \mathbb{R}_+^L \mapsto \mathbb{R}_+$. We say that $\mathcal{I}$ is a general interference function (or simply *interference function*) if the following axioms are fulfilled:

A1	(positivity)	There exists an $\boldsymbol{p} > 0$ such that $\mathcal{I}(\boldsymbol{p}) > 0$,
A2	(scale invariance)	$\mathcal{I}(\alpha\boldsymbol{p}) = \alpha\mathcal{I}(\boldsymbol{p})$ for all $\alpha > 0$,
A3	(monotonicity)	$\mathcal{I}(\boldsymbol{p}) \geq \mathcal{I}(\boldsymbol{p}')$ if $\boldsymbol{p} \geq \boldsymbol{p}'$.

Throughout this deliverable, all vector inequalities are element-wise, i.e. $\boldsymbol{p} \geq \boldsymbol{p}'$ means that $p_l \geq p'_l$ for all elements $l = 1, 2, \ldots, L$.

Property A3 is quite intuitive. If we increase the amount of expended power resources $\boldsymbol{p}$, then the resulting value $\mathcal{I}$ will increase (or at least not decrease). An immediate example is power control, where $\boldsymbol{p}$ is a vector of transmission powers and $\mathcal{I}(\boldsymbol{p})$ is the resulting interference at some other user. Hence, the name "interference function".

The axioms A1, A2, A3 were proposed and studied in [4], with extensions in [18–20].

Interference functions were originally proposed for power control problems. Consider K communication links with powers $\boldsymbol{p} \in \mathbb{R}^K_{++}$. Here, K can stand for the number of users, but we can also model the case where each user has multiple links. Optimisation strategies are mostly based on the SIR or the SINR, depending on whether the model includes noise or not.

The case with no noise corresponds to a system without power constraints. If the transmission powers can be arbitrarily large, then the impact of noise is negligible. This case is interesting from a theoretical point of view, since it allows to focus on the impact of interference coupling. Noiseless scenarios have been addressed in the early literature, e.g. [21], and later in the context of beamforming [22]. More recent extensions of these results can be found in [20, 23, 24]. Studying the noiseless case helps to obtain a thorough understanding of the underlying structure of power control problems. Two mathematical theories have been proved useful in this context. Firstly, the Perron-Frobenius theory of non-negative matrices. Secondly, the theory of interference functions [4, 15].

In the following we will discuss a system with receiver noise power $\sigma_n^2 > 0$. In order to incorporate noise in the framework A1, A2, A3, we introduce an extended power resource vector

$$\underline{\boldsymbol{p}} = \begin{bmatrix} \boldsymbol{p} \\ \sigma_n^2 \end{bmatrix} = [p_1, \ldots, p_K, \sigma_n^2]^T \ . \tag{2.10}$$

Then $\mathcal{I}(\underline{\boldsymbol{p}})$ stands for interference plus noise, as in the previous example (2.6). While the impact of noise is evident in (2.6), it is not so obvious when defining the interference via the axioms A1,A2,A3 alone. In order for the noise to have any impact, we need the following additional property.

A4 (strict monotonicity) $\mathcal{I}(\underline{\boldsymbol{p}}) > \mathcal{I}(\underline{\boldsymbol{p}}')$ if $\underline{\boldsymbol{p}} \geq \underline{\boldsymbol{p}}'$ and $\underline{p}_{K+1} > \underline{p}'_{K+1}$.

If $\underline{p}_{K+1} > 0$, then A4 ensures that $\mathcal{I}(\underline{\boldsymbol{p}}) > 0$ for arbitrary $\boldsymbol{p} \geq 0$. This can be easily shown by contradiction: Suppose that $\mathcal{I}(\underline{\boldsymbol{p}}) = 0$, then for any α with $0 < \alpha < 1$ we have

$$0 = \mathcal{I}(\underline{\boldsymbol{p}}) > \mathcal{I}(\alpha\underline{\boldsymbol{p}}) = \alpha\mathcal{I}(\underline{\boldsymbol{p}}) \ ,$$

which would lead to the contradiction $0 > \lim_{\alpha\to 0} \alpha\mathcal{I}(\underline{\boldsymbol{p}}) = 0$.

Axiom A4 connects the framework A1, A2, A3 with the framework of *standard interference functions* [15]. In [15] it was first proposed to model core properties of interference by an axiomatic approach. A function $Y(\boldsymbol{p})$ is called a standard interference function if the following axioms are fulfilled.

Y1	(positivity)	$Y(\boldsymbol{p}) > 0$ for all $\boldsymbol{p} \in \mathbb{R}^K_+$,
Y2	(scalability)	$\alpha Y(\boldsymbol{p}) > Y(\alpha\boldsymbol{p})$ for all $\alpha > 1$,
Y3	(monotonicity)	$Y(\boldsymbol{p}) \geq Y(\boldsymbol{p}')$ if $\boldsymbol{p} \geq \boldsymbol{p}'$.

Properties of this framework were studied in [11, 25, 26]. The connection between standard interference functions and the axiomatic framework A1, A2, A3 was studied in [16]. It was shown that any standard interference function Y can be expressed by a general interference function. To be precise, a function Y is a standard interference function if and only if the function

$$\mathcal{I}_Y(\underline{\boldsymbol{p}}) := \underline{p}_{K+1} \cdot Y\Big(\frac{p_1}{\underline{p}_{K+1}}, \dots, \frac{p_K}{\underline{p}_{K+1}}\Big) \tag{2.11}$$

fulfils A1, A2, A3 plus strict monotonicity A4. If we choose a constant value $\underline{p}_{K+1} = 1$, then we simply have $\mathcal{I}_Y(\underline{\boldsymbol{p}}) = Y(\boldsymbol{p})$. That is, $\mathcal{I}_Y$ is a standard interference function with respect to the first K components of its argument.

The discussion of the previous section shows that the axiomatic framework A1, A2, A3 is also suitable for a broad class of SINR-based power control problems, with interference that can be modelled by standard interference functions. Examples are beamforming [4–10], CDMA [11, 12], base station assignment [13, 14] and robust designs [27, 28].

2.2.2 Interference Balancing

The interference functions introduced in Section 2.2.1 are very general, they are only determined by axioms A1 and A2. Axioms A1 is of minor importance, it only serves the purpose of ruling out the trivial case $\mathcal{I}(\boldsymbol{p}) = 0$ for all $\boldsymbol{p} \geq 0$.

This model is indeed too general for the development of algorithmic solution. But is has been shown in first SAPHYRE publications [29–31] that this basic framework is useful for algorithm development if it is extended by additional properties.

One such property is the assumption of strict monotonicity with respect to noise, as discussed in Section 2.2.1. In this case, the framework is closely connected with standard interference functions, thus the results of [15] can be applied to all function falling within this case. In [15] a fixed point iteration was proposed for solving the problem of SINR-constrained power minimisation, also known as *SINR balancing.*

$$\min_{\boldsymbol{p}\in\mathcal{P}} \sum_{l\in\mathcal{K}} p_l \quad \text{s.t.} \quad \frac{p_k}{\mathcal{I}_k(\underline{\boldsymbol{p}})} \geq \gamma_k, \text{ for all } k \in \mathcal{K} . \tag{2.12}$$

Let's assume that the power set $\mathcal{P}$ and the targets $\boldsymbol{\gamma}$ are such that the constraints are feasible. It was shown in [15] that problem (2.12) has a unique optimiser, which is the exact point where all constraints are fulfilled with equality. This is the fixed point obtained by the iteration

$$p_k^{(n+1)} = \gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}}^{(n)}) , \quad \forall k \in \mathcal{K} , \quad \boldsymbol{p}^{(0)} = \boldsymbol{0} . \tag{2.13}$$

This iteration is component-wise monotonic convergent to the global optimiser $\boldsymbol{p}^*$. This was shown for standard interference functions in [15]. An alternative study based on the framework A1, A2, A3, A4 was presented in [4]. The iteration has linear convergence [25, 32], regardless of the actual choice of $\mathcal{I}_k$.

2.2.3 Exploiting Convexity and Monotonicity

Convexity is commonly considered as the dividing line between "easy" and "difficult" problems. Many examples can be found in the context of multiuser MIMO [7, 9, 10, 33, 34]. For example, equivalent convex reformulations exist for the well-known downlink beamforming problem, as observed in [7, 9, 10]. When investigating a problem, a common approach is to first look whether the problem is convex or not. Yates' pioneering work on interference functions [15] has shown that it is not just convexity that matters. The above problem (2.13) can be solved efficiently without exploiting convexity, by merely relying on monotonicity and scalability properties.

This underlines the importance of exploiting all available structure of the problem at hand. Standard approaches from convex optimisation theory do not make sufficient use of the special structure offered by interference functions. On the other hand, Yates' framework of standard interference functions [15] does not exploit convexity or concavity. As an example, consider the interference function (2.6), which results from optimum multi-antenna combining. It fulfils the axioms A1, A2, A3, and it is concave in addition. However, concavity is not exploited by the fixed point iteration (2.13).

In the remainder of this section we discuss how concavity (resp. convexity) and monotonicity can be exploited jointly. By only exploiting monotonicity and convexity, we show that the problem can be rewritten in an equivalent convex form.

To this end, assume a convex compact power set $\mathcal{P} \subseteq \mathbb{R}^K_{++}$. We rewrite (2.12) in an equivalent form

$$\min_{\boldsymbol{p} \in \mathcal{P}} \sum_{l \in \mathcal{K}} p_l \quad \text{s.t.} \quad \gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}}) - p_k \leq 0, \text{ for all } k \in \mathcal{K} \,. \tag{2.14}$$

If $\mathcal{I}_k$ is convex, then (2.14) is a convex optimisation problem. Property A4 ensures the existence of a non-trivial solution, provided that the targets γ_k are feasible.

Next, consider the case where $\mathcal{I}_k$ is strictly monotonic and *concave*. An example is the interference (2.6) resulting from beamforming, with either individual power constraints or a total power constraint. Then, problem (2.14) is non-convex because inequality constraint functions are *concave*, but not convex.

This observation is in line with the literature on multiuser beamforming [7, 9, 10], where it was observed that the corresponding problem is non-convex. Fortunately, the beamforming problem could be solved by deriving equivalent convex reformulations, enabled by the specific interference structure resulting from beamforming receivers [7, 9, 10].

An interesting question is: does an equivalent convex reformulation also exist for the more general problem (2.14), which is only based on the axiomatic framework?

Indeed, if $\mathcal{I}_1, \ldots, \mathcal{I}_K$ are concave and strictly monotonic interference functions, then problem (2.14) is equivalent to

$$\max_{\boldsymbol{p} \in \mathcal{P}} \sum_{l \in \mathcal{K}} p_l \quad \text{s.t.} \quad p_k - \gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}}) \leq 0, \quad \forall k \in \mathcal{K} \,. \tag{2.15}$$

Thanks to the monotonicity property of the interference functions, it can be shown that both problems (2.14) and (2.15) are solved by the same fixed point $\boldsymbol{p}^*$ [16]. In this sense they are equivalent. Problem (2.15) is convex and $\boldsymbol{p}^*$ can be found efficiently using standard algorithms from convex optimisation theory.

This result sheds some new light on the problem of multiuser beamforming, which is contained as a special case. It turns out that this problem has a generic convex reformulation (2.15). This insight possibly helps to better understand the convex reformulations observed in the beamforming literature [7, 10].

It should be emphasised that the reformulation (2.15) holds for arbitrary concave standard interference functions, not just the beamforming case. This also includes other receive strategies that aim at optimising the SINR. Examples are CDMA [11, 12] or base station assignment [13, 14]. Also, it is easy to incorporate additional constraints on the receivers, like the *shaping constraints* studied in [9].

2.3 Analysis of Interference-Coupled Systems

It is expected that the following min-max problem will play a fundamental role for the understanding of interference-coupled systems in SAPHYRE, and also for the development of distributed resource balancing algorithms.

$$C(\boldsymbol{\gamma}) = \inf_{\boldsymbol{p}\in\mathbb{R}^L_{++}} \max_{1\le k\le L} \frac{\gamma_k \mathcal{I}_k(\boldsymbol{p})}{p_k} . \tag{2.16}$$

Here, $\mathcal{I}_1, \ldots, \mathcal{I}_K$ are interference functions defined in Section 2.2.1 and $\boldsymbol{\gamma}$ is a vector of SIR targets.

The function $C(\boldsymbol{\gamma})$ is an indicator for the "congestion" of the system. If $C(\boldsymbol{\gamma}) \le 1$, then the SIR values $\boldsymbol{\gamma}$ can be jointly supported.

Within SAPHYRE we have studied problem (2.16). The results have been published in [3]. The main outcome of this research is a characterisation of conditions for the existence and uniqueness of an optimiser of problem (2.16).

This work has a very practical background. Existence and uniqueness of an optimiser is an important prerequisite for the convergence of algorithms. Thus, the results provide a basis for other work packages, where algorithms are developed. In particular, it is expected to be useful for the development of distributed fixed point iteration, similar to the approach of Yates [15].

2.3.1 Strongly-Connected System

Some basic properties of problem (2.16) were already studied in [4]. Here, we extend these results by assuming that the links are "strongly connected". This is defined by the following properties.

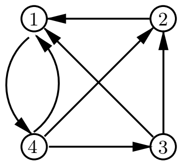
In order to model whether an interference function depends on some resource or not, we introduce the coupling matrix $\boldsymbol{A}_{\mathcal{I}}$, which characterises the interference coupling between the users. The *asymptotic coupling matrix* $\boldsymbol{A}_{\mathcal{I}}$ is defined as follows. Let $\boldsymbol{e}_l$ be the all-zero vector with the l-th component set to one. Then,

$$[\boldsymbol{A}_{\mathcal{I}}]_{kl} = \begin{cases} 1 & \text{if there exists a } \boldsymbol{p} > 0 \text{ such that} \\ & \lim_{\delta \to \infty} \mathcal{I}_k(\boldsymbol{p} + \delta \boldsymbol{e}_l) = +\infty, \\ 0 & \text{otherwise.} \end{cases} \tag{2.17}$$

If there is one $\boldsymbol{p}$ that fulfils the condition in (2.17), then this condition is fulfilled for *all* $\boldsymbol{p} > 0$.

We assume that $\boldsymbol{A}_{\mathcal{I}}$ is irreducible. A non-negative $L \times L$ matrix $\boldsymbol{A}_{\mathcal{I}}$ is said to be irreducible if and only if its directed graph $\mathcal{G}(\boldsymbol{A}_{\mathcal{I}})$ is strongly connected. This is illustrated by the following example.

$$\boldsymbol{A}_{\mathcal{I}} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \qquad \mathcal{G}(\boldsymbol{A}_{\mathcal{I}}) :$$



The graph $\mathcal{G}(\boldsymbol{A}_{\mathcal{I}})$ consists of $L = 4$ nodes $N_1, \dots, N_L$. A pair of nodes (N_i, N_j) is connected by a directed edge if $[\boldsymbol{A}_{\mathcal{I}}]_{ij} > 0$. A graph is called *strongly connected* if for each pair of nodes (N_i, N_j) there is a sequence of directed edges leading from N_i to N_j.

The dependency set of link k is

$$\mathsf{L}_k = \{ l \in [1, 2, \dots, K] : [\boldsymbol{A}_{\mathcal{I}}]_{kl} = 1 \} \, . \tag{2.18}$$

The set L_k is non-empty. Because of the assumed irreducibility there is at least one non-zero entry in each row and column.

$\mathcal{I}_k(\boldsymbol{p})$ is said to be *strictly positive* if for any $\boldsymbol{p} \geq 0$ with $p_l > 0$ for some $l \in \mathsf{L}_k$ we have $\mathcal{I}_k(\boldsymbol{p}) > 0$.

2.1 Definition. A system consisting of L interference functions is said to be strongly connected if the functions are strictly positive and $\boldsymbol{A}_{\mathcal{I}}$ is irreducible.

It was shown in [4, Thm. 2.7] that there always exists a $\boldsymbol{p}^* \gneqq 0$, such that

$$p_k^* \cdot C(\boldsymbol{\gamma}) = \gamma_k \mathcal{I}_k(\boldsymbol{p}^*), \quad 1 \leq k \leq K \, . \tag{2.19}$$

This result holds for all interference functions fulfilling A1-A3.

Under the assumption of a strongly connected system, the following result holds [3].

2.2 Theorem. *If the system is strongly connected then we have* $C(\boldsymbol{\gamma}) > 0$ *and any* $\boldsymbol{p}^* \geq 0$ *fulfilling* (2.19) *is strictly positive, i.e.* $\boldsymbol{p}^* > 0$*.*

Based on Theorem 2.2 we can derive the following result [3], which provides a sufficient condition for existence of an optimiser.

2.3 Corollary. *If the system is strongly connected then the min-max problem* (2.16) *has a solution* $\boldsymbol{p}^* > 0$.

Thus far, we have shown the existence of a solution $\boldsymbol{p}^* > 0$ but not uniqueness. There possibly are several solutions. Uniqueness will be discussed in the next section under the additional assumption of strict monotonicity.

2.3.2 Strict Monotonicity Implies a Unique Optimiser

We begin with a definition.

2.4 Definition (strict monotonicity)**.** $\mathcal{I}_k(\boldsymbol{p})$ is said to be strictly monotonic if for arbitrary $\boldsymbol{p}^{(1)}$, $\boldsymbol{p}^{(2)}$, the inequality $\boldsymbol{p}^{(1)} \geq \boldsymbol{p}^{(2)}$, with $p_l^{(1)} > p_l^{(2)}$ for some $l \in \mathsf{L}_k$, implies $\mathcal{I}_k(\boldsymbol{p}^{(1)}) > \mathcal{I}_k(\boldsymbol{p}^{(2)})$.

This leads to the next main result [3].

2.5 Theorem. *If the interference functions are strongly connected and strictly monotonic, then the set of equations* (2.19) *has a unique solution* $\boldsymbol{p}^* > 0$*, up to scalar multiples.*

The next result connects equation (2.19) and the balancing problem (2.16).

2.6 Theorem. *If the interference functions are strongly connected and strictly monotonic, then the SIR balancing problem* (2.16) *has an optimiser* $\boldsymbol{p}^* > 0$*, unique up to a scalar multiple, that balances all the ratios* SIR_k/γ_k *at the level* $C(\gamma)$*, i.e.*

$$C(\boldsymbol{\gamma}) = \min_{\boldsymbol{p} \in \mathcal{R}^L_{++}} \left(\max_{1 \leq l \leq L} \frac{\gamma_l \mathcal{I}_l(\boldsymbol{p})}{p_l} \right) \tag{2.20}$$

$$= \frac{\gamma_1 \mathcal{I}_1(\boldsymbol{p}^*)}{p_1^*} = \cdots = \frac{\gamma_L \mathcal{I}_L(\boldsymbol{p}^*)}{p_L^*} \; . \tag{2.21}$$

These results are very general, and thus far we haven't made any assumption on power constraints. This will be discussed in the next section.

2.3.3 Power-Constrained Interference Balancing

In this section we show how the results can be extended to include power constraints, which is important for most practical interference scenarios.

Thus far, we have analysed the SIR, which is invariant with respect to a scaling, i.e. $\mathrm{SIR}_l(\alpha\boldsymbol{p}) = \mathrm{SIR}_l(\boldsymbol{p})$ for all $\alpha > 0$. Thus, constraining the norm of $\boldsymbol{p}$ has no effect. This is typical for a system with no noise. In order to model power-constrained system with noise, the framework of *standard interference functions* is often used (cf. Section 2.2.1, page 15).

In the following we demonstrate that the power-constrained SINR balancing problem (2.12) can be understood as a special case of problem (2.16). Our approach is based on the introduction of an auxiliary interference function, which ensures that certain power constraints are fulfilled.

In order for a power constraint to have any effect, it is necessary to introduce noise. In this section we consider K user powers $p_1, \ldots, p_K$ and a noise power σ_n^2. These powers are stacked in the extended power vector $\underline{\boldsymbol{p}} \in \mathbb{R}_{++}^{K+1}$, which was already introduced and discussed in (2.10). In the following we normalise $\sigma_n^2 = 1$.

Consider a system with a total power constraint $\sum_{1 \le k \le K} \le P_{max}$, for some $P_{max} > 0$. In order to enforce this power constraint, we introduce an auxiliary interference function

$$\mathcal{I}_{K+1}(\underline{\boldsymbol{p}}) = \frac{1}{P_{max}} \cdot \sum_{k=1}^{K} \gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}}) \, . \tag{2.22}$$

It can be verified that $\mathcal{I}_{K+1}$ fulfils the axioms A1-A3, since the sum of interference functions is an interference function again.

Let $\boldsymbol{G}$ be the dependency matrix of the first K components. We assume that every user causes interference to at least one other user, thus each column of the matrix $\boldsymbol{G}$ has at least one non-zero entry. Also, every $\mathcal{I}_k(\underline{\boldsymbol{p}})$ depends on the noise component. Thus, the $K \times K+1$ asymptotic coupling matrix of the first K interference functions is $[\boldsymbol{G} \,|\, \mathbf{1}]$, where the last column models the dependency on the noise.

The interference function $\mathcal{I}_{K+1}$ depends on all powers, because of definition (2.22). Thus, the overall system coupling matrix becomes

$$\boldsymbol{A}_{\mathcal{I}} = \begin{bmatrix} \boldsymbol{G} & \mathbf{1} \\ \mathbf{1}^T & 1 \end{bmatrix} \, . \tag{2.23}$$

The matrix $\boldsymbol{A}_{\mathcal{I}}$ is irreducible because its last row and column are positive.

We further assume that the interference functions $\mathcal{I}_1, \ldots, \mathcal{I}_K$ are strictly monotonic. Since $\sigma_n^2 = 1$ is constant, we know that they are also strictly positive. This is a consequence of the properties A1-A3 and strict monotonicity, as observed in [4]. The function $\mathcal{I}_{K+1}$ is the sum of all other interference functions, thus positivity and strict monotonicity hold.

The interference functions $\mathcal{I}_1, \ldots, \mathcal{I}_K, \mathcal{I}_{K+1}$ constitute a fully connected, strictly monotonic system. The $K+1$ dimensional SIR balancing problem can be written as

$$\min_{\underline{\boldsymbol{p}} \in \mathcal{P}} \left(\max_{1 \le k \le K+1} \frac{\gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}})}{p_k} \right) \, . \tag{2.24}$$

Here, we optimise over the set

$$\mathcal{P} = \left\{ \underline{\boldsymbol{p}} \in \mathbb{R}_{++}^{K+1} : p_{K+1} = 1 \right\} \, . \tag{2.25}$$

As a consequence, $\underline{\boldsymbol{p}}$ always fulfils $p_{K+1} = 1$. Recall that the SIR is not affected by a simultaneous scaling of transmission powers and noise, so we could equivalently optimise over the unconstrained set $\mathbb{R}_{++}^{K+1}$. But the problem formulation (2.24) has the advantage of having a defined noise level.

The additional parameter γ_{K+1} can be set to one. Other values will just scale the available power budget.

From Theorem 2.6 we know that problem (2.24) has a unique optimiser $\underline{\boldsymbol{p}}^* > 0$ such that

$$\frac{p_1^*}{\gamma_1 \mathcal{I}_1(\underline{\boldsymbol{p}}^*)} = \cdots = \frac{p_K^*}{\gamma_K \mathcal{I}_K(\underline{\boldsymbol{p}}^*)} = \frac{P_{max}}{\sum_{k=1}^K \gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}}^*)}. \tag{2.26}$$

Taking the sum of the first K powers, we obtain the following identity.

$$\sum_{k=1}^K p_k^* = \sum_{k=1}^K \gamma_k \mathcal{I}_k(\underline{\boldsymbol{p}}^*) \cdot \frac{P_{max}}{\sum_{k=1}^K \gamma_l \mathcal{I}_l(\underline{\boldsymbol{p}}^*)} = P_{max}. \tag{2.27}$$

Hence, all user SIRs are balanced and the sum power constraint is fulfilled with equality.

Next, consider the functions $Y_k(\boldsymbol{p}) = \mathcal{I}_k(\underline{\boldsymbol{p}})$ for all $k = 1, 2, \ldots, K$. For constant noise $\underline{p}_{K+1}$, the function Y_k is a *standard interference function* [15], characterised by monotonicity and scalability, i.e. $Y(\alpha\boldsymbol{p}) < \alpha Y(\boldsymbol{p})$. Moreover, any standard interference function can be written as $\mathcal{I}_k(\underline{\boldsymbol{p}})$ with $p_{K+1} = 1$. Thus, both frameworks can be used interchangeably for modelling interference plus noise in communication system. For a detailed comparison between the axiomatic framework A1-A3 and the framework of standard interference functions, the reader is referred to [16].

With standard interference functions, the power-constrained SINR balancing problem can be written as follows.

$$\max_{\boldsymbol{p}>0:\sum_{k=1}^K p_k \le P_{max}} \min_{k\in\mathcal{K}} \frac{p_k}{\gamma_k Y_k(\boldsymbol{p})} \ . \tag{2.28}$$

Equivalently, we can focus on the problem

$$C(\boldsymbol{\gamma}, P_{max}) = \min_{\boldsymbol{p}>0:\sum_{k=1}^K p_k \le P_{max}} \max_{k\in\mathcal{K}} \frac{\gamma_k Y_k(\boldsymbol{p})}{p_k} \ . \tag{2.29}$$

Similar as in [8], we can show that

- Problem (2.29) has a unique optimiser $\boldsymbol{p}^* > 0$ that balances all SINR at a level $C(\boldsymbol{\gamma}, P_{max}) > 0$, i.e.

$$C(\boldsymbol{\gamma}, P_{max}) = \frac{\gamma_k Y_k(\boldsymbol{p}^*)}{p_k^*} \quad \text{for all } k = 1, 2, \ldots, K \tag{2.30}$$

- At the optimum, the power constraint is fulfilled with equality, i.e. $\sum_{k=1}^K p_k^* = P_{max}$.

- $C(\boldsymbol{\gamma}, P_{max})$ is strictly monotonic decreasing in P_{max}.

From the latter property it follows that there cannot be another balanced level. If we find powers $\boldsymbol{p}^* > 0$ achieving a balanced level and if this solution fulfils $\sum_{k=1}^{K} p_k^* = P_{max}$, then this is the unique optimiser of the SINR balancing problem (2.29).

Hence, the solution (2.26) obtained by the SIR balancing approach with an auxiliary interference function yields the unique global optimum of the SINR balancing problem.

These results contribute to a better understanding of interference balancing in multiuser systems. In Section 2.3.3 it has been demonstrated that the interference model A1,A2,A3 can be applied to practical scenarios with power constraint, so it provides an alternative approach to the well-known framework of standard interference functions.

One could argue that standard interference functions are an established model, so the value of using a different model A1,A2,A3 might not be clear at first sight. But is has been shown in a series of recent publications [4, 18–20] that the framework A1,A2,A3 has an interesting structure which is amenable to mathematical analysis. In a sense, it can be seen as an extension of the Perron-Frobenius theory, which has proved to be useful in many areas of research that involve coupled systems.

2.3.4 Multicarrier Systems

For resource sharing, it is important to extend the interference framework to a system, where M users from the set $\mathcal{M} = \{1, 2, \ldots, M\}$ can possibly transmit over N "resources" $\mathcal{N} = \{1, 2, \ldots, N\}$. Here, "resources" are used as an abstract term referring to any signaling dimension (e.g., frequency carriers, time slots, sectorization patterns). The user-resource pairs define $L = M \times N$ transmission links. We will assume that the resources can be accessed freely. In other words, each user can use multiple resources simultaneously, and each resource can be shared by multiple users. It should be emphasized that signals transmitted over different resources can interfere with each other, e.g., inter-carrier interference or interference caused by non-perfect sectorization. The interference will be modeled by means of axiomatic interference functions that will be introduced later. On each communication link, we can employ interference mitigation strategies, referred to as "Tx/Rx strategies". For example, in an M-user, multicarrier, single-input, multiple-output (SIMO) system, all available subcarriers can be considered as the "resources", and transmit power allocations and receive beamformers as the "Tx/Rx strategies".

In order to avoid excessive interference between the users, we want to assign the resources dynamically, depending on the current interference situation. The users should be able to make use of the available resources and Tx/Rx strategies to fully exploit the potential of the system. The design goal is to determine an optimal allocation pattern that determines how the M users are mapped to the N resources. Each user-resource pair (transmission link) is characterized by its transmission power, which can be positive in case of an active link, or zero in case of an inactive link. It is also possible that there is no active link on some resource,

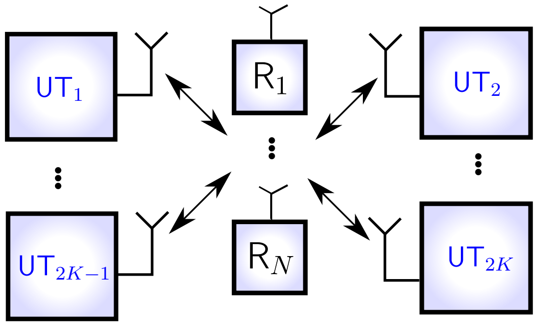


Figure 2.4: Multi-pair two-way relaying with multiple single-antenna amplify and forward relays.

which means this resource is left idle.

Let $P_{m,n} \geq 0$ be the transmission power used by user m on resource n. The case $P_{m,n} = 0$ corresponds to an unused (inactive) resource. By denoting the user-resource pair (m, n) with the link index l, all powers $P_{m,n}$ can be stacked into a power vector: $\boldsymbol{p} \triangleq [P_{1,1}, ...P_{M,1}, ..., P_{1,N}, ..., P_{M,N}]^T \triangleq [p_1, ..., p_L]^T$, where $l = 1, ..., m + (n-1)M, ..., L$.

We further assume that $\boldsymbol{p}$ is constrained to a convex polytope $\mathcal{P}$, which in particular means that the available power is finite. This includes the most common power constraints, like sum-power constraint or per-user power constraints.

Based on this framework, we will analyze the achievable QoS region in Chapter 3.

2.4 Interference Model in Two-Way Relay Channel

In Deliverable D2.3a we have described the relay assisted resource sharing scenarios where pairs of single-antenna user terminals (UTs) belonging to different operators communicate with each other via the help of a two-way MIMO amplify and forward (AF) relay. In this scenario (Figure 2.4), both the spectrum and the relay are shared among difference operators. Although such a scenario has practical application and is of interest to SAPHYRE, the optimization problem in this scenario is more complex compared to one-way relaying protocol or regenerative relaying strategy such as decode and forward (DF) relaying. This is due to the fact that the optimization problem in the two-way relaying (TWR) requires the joint optimization over two transmission phases, i.e., the multiple access phase and the broadcasting phase. Nevertheless, in this section we study the sum rate maximization problem of this scenario. Although the original problem is non-convex, we demonstrate that the sum rate maximization problem subject to a total power constraint at the relay can be solved using monotonic optimization and the power method.

To illustrate the generality of our problem, we start with a similar scenario, i.e., multi-pair TWR with multiple single-antenna AF relays. We find optimal relay transmit coefficients which maximize the sum rate of the system subject to a total power constraint in the relay network. Then we show that the same methods can be applied to the multi-operator TWR channel in D2.3a. The scenario under investigation is shown in Fig. 2.4. K pairs of single-antenna users would like to communicate with each other via the help of N single-antenna relays. We assume perfect synchronization and the channel is frequency flat and quasi-static block fading. The vector channel from the $(2k-1)$th user (on the left-hand side of Fig. 2.4) to the relays is denoted as $\boldsymbol{f}_{2k-1} = [f_{2k-1,1}, f_{2k-1,2}, \ldots, f_{2k-1,N}]^{\mathrm{T}} \in \mathbb{C}^N$, while the channel from the $2k$th user (on the right-hand side of Fig. 2.4) to the relay is denoted as $\boldsymbol{g}_{2k} = [g_{2k,1}, g_{2k,2}, \ldots, g_{2k,N}]^{\mathrm{T}} \in \mathbb{C}^N$, for $k \in \{1, 2, \cdots, K\}$. For notational simplicity, we assume an ideal time-division duplex (TDD) system, i.e., the channels are *reciprocal.* The transmission takes two time slots. In the first time slot, the signal received at all relays can be combined in a vector as

$$\boldsymbol{r} = \sum_{k=1}^{K} (\boldsymbol{f}_{2k-1} s_{2k-1} + \boldsymbol{g}_{2k} s_{2k}) + \boldsymbol{n}_{\mathrm{R}} \in \mathbb{C}^N \tag{2.31}$$

where s_{2k-1} and s_{2k} are i.i.d. symbols with zero mean and unit power. The vector $\boldsymbol{n}_{\mathrm{R}}$ denotes the zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise and $\mathbb{E}\{\boldsymbol{n}_{\mathrm{R}} \boldsymbol{n}_{\mathrm{R}}^{\mathrm{H}}\} = \sigma_{\mathrm{R}}^2 \boldsymbol{I}_N$.

Afterwards, the AF relays broadcast the weighted signal as

$$\bar{\boldsymbol{r}} = \boldsymbol{W} \cdot \boldsymbol{r} \tag{2.32}$$

where $\boldsymbol{W} = \mathrm{diag}\{\boldsymbol{w}^*\}$ and $\boldsymbol{w} = [w_1, w_2, \ldots, w_N]^{\mathrm{T}}$ is the vector which consists of the N complex weights of all the relays.

In the second time slot, the received signal at the $(2k-1)$th user (on the left-hand side of Fig. 2.4) is expressed as

$$\begin{aligned} y_{2k-1} \;&=\; \underbrace{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{F}_{2k-1} \boldsymbol{g}_{2k} s_{2k}}_{\text{desired signal}} + \underbrace{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{F}_{2k-1} \boldsymbol{f}_{2k-1} s_{2k-1}}_{\text{self-interfernce}} \\ &+\; \underbrace{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{F}_{2k-1} \sum_{\substack{\ell \neq k \\ \ell=1}}^{K} (\boldsymbol{f}_{2\ell-1} s_{2\ell-1} + \boldsymbol{g}_{2\ell} s_{2\ell})}_{\text{inter-pair interference}} \\ &+\; \underbrace{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{F}_{2k-1} \boldsymbol{n}_{\mathrm{R}} + n_{2k-1}}_{\text{effective noise}} \end{aligned} \tag{2.33}$$

where $\boldsymbol{F}_{2k-1} = \mathrm{diag}\{\boldsymbol{f}_{2k-1}\}$ and n_{2k-1} is the ZMCSCG noise with variance σ_{2k-1}^2. Similar expressions can be obtained for the $2k$th user.

Assume that perfect channel knowledge can be obtained such that the self-interference terms can be canceled. Let P_{R} be the total transmit power consumed by the relays in the network. Our goal is to find the weight vector $\boldsymbol{w}$ such that the sum rate of the system is maximized subject to the sum power constraint.

The optimization problem can be formulated as

$$\begin{aligned} \max_{\boldsymbol{w}} \quad & \tfrac{1}{2}\sum_{m=1}^{2K} \log_2(1+\mathrm{SINR}_m) \\ \text{subject to} \quad & \mathbb{E}\{\|\bar{\boldsymbol{r}}\|^2\} \leq P_{\mathrm{R}}, \end{aligned} \tag{2.34}$$

where the factor 1/2 is due to the two channel uses (half duplex). When $m = 2k-1$, from the expression (2.33), the SINR of the mth user can be calculated as

$$\mathrm{SINR}_{2k-1} = \frac{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{B}_{2k-1}\boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}}(\boldsymbol{D}_{2k-1}+\boldsymbol{E}_{2k-1})\boldsymbol{w}+\sigma_{2k-1}^2} \tag{2.35}$$

where $\boldsymbol{D}_{2k-1} = \sum_{\substack{\ell \neq k \\ \ell=1}} (\tilde{\boldsymbol{h}}^{(\mathrm{o})}_{2k-1,\ell}\tilde{\boldsymbol{h}}^{(\mathrm{o})^{\mathrm{H}}}_{2k-1,\ell} + \tilde{\boldsymbol{h}}^{(\mathrm{e})}_{2k-1,\ell}\tilde{\boldsymbol{h}}^{(\mathrm{e})^{\mathrm{H}}}_{2k-1,\ell})$ and $\boldsymbol{B}_{2k-1} = \boldsymbol{h}_{2k-1}\boldsymbol{h}^{\mathrm{H}}_{2k-1}$ are $N \times N$ positive semidefinite Hermitian matrices. Matrices $\boldsymbol{D}_{2k-1}$ and $\boldsymbol{B}_{2k-1}$ are related to the interference power and the desired signal power, respectively, ($\boldsymbol{h}_{2k-1} = \boldsymbol{f}_{2k-1} \odot \boldsymbol{g}_{2k}$, $\tilde{\boldsymbol{h}}^{(\mathrm{o})}_{2k-1,\ell} = \boldsymbol{f}_{2k-1} \odot \boldsymbol{f}_{2\ell-1}$ and $\tilde{\boldsymbol{h}}^{(\mathrm{e})}_{2k-1,\ell} = \boldsymbol{f}_{2k-1} \odot \boldsymbol{g}_{2\ell}$). The term which is related to the forwarded noise from the relay is denoted by an $N \times N$ full rank diagonal matrix $\boldsymbol{E}_{2k-1} = \sigma_{\mathrm{R}}^2 \boldsymbol{F}_{2k-1}\boldsymbol{F}^{\mathrm{H}}_{2k-1}$. Similar SINR expression can be obtained when $m = 2k$. Furthermore, the total transmit power is given by $\mathbb{E}\{\|\bar{\boldsymbol{r}}\|^2\} = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{\Gamma}\boldsymbol{w}$ with

$$\boldsymbol{\Gamma} = \sum_{k=1}^{K} (\boldsymbol{F}_{2k-1}\boldsymbol{F}^{\mathrm{H}}_{2k-1} + \boldsymbol{G}_{2k}\boldsymbol{G}^{\mathrm{H}}_{2k}) + \sigma_{\mathrm{R}}^2 \boldsymbol{I}_N. \tag{2.36}$$

To simplify the optimization problem we note that the inequality constraint in (2.34) has to be satisfied with equality at optimality. Otherwise, the optimal $\boldsymbol{w}$ can be scaled up to satisfy the constraint with equality while increasing the objective function, which contradicts the optimality. Inserting the constraint into the objective function in (2.34), the original problem can be reformulated as an *unconstrained* optimization problem

$$\max_{\boldsymbol{w}} \prod_{m=1}^{2K} \frac{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{A}_m\boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{C}_m\boldsymbol{w}} \tag{2.37}$$

where $\boldsymbol{C}_m = \boldsymbol{D}_m + \boldsymbol{E}_m + \frac{\sigma_m^2}{P_{\mathrm{R}}}\boldsymbol{\Gamma}$ and $\boldsymbol{A}_m = \boldsymbol{B}_m + \boldsymbol{C}_m$ are positive definite. Problem (2.37) is equivalent to (2.34) since the objective function is homogeneous and any scaling in $\boldsymbol{w}$ does not change the optimality. Nevertheless, if $\bar{\boldsymbol{w}}$ is the solution to (2.37), it should be scaled to fullfill the power constraint, i.e., the optimal solution to (2.34) is given by

$$\boldsymbol{w} = \sqrt{\frac{P_{\mathrm{R}}}{\bar{\boldsymbol{w}}^{\mathrm{H}}\boldsymbol{\Gamma}\bar{\boldsymbol{w}}}}\,\bar{\boldsymbol{w}}. \tag{2.38}$$

Problem (2.37) is non-convex and in general NP-hard.

Generalized Polyblock Algorithm

Monotonic optimization (see [35], [36]) deals with the maximization or minimization of an increasing function over an intersection of normal and reverse normal sets. The polyblock approximation approach is a unified algorithm to find the global optimum of the monotonic optimization problem. Prior work that used this approach in the area of wireless communications can be found in [37], [38]. We show that the problem (2.37) is a monotonic optimization problem and then propose a version of the polyblock algorithm to solve it.

1 Proposition. Problem (2.37) is a monotonic optimization problem.

Proof. Problem (2.37) is equivalent to the following problem

$$\max_{\boldsymbol{y}} \{\Phi(\boldsymbol{y})|\boldsymbol{y} \in \mathbb{D}\} \tag{2.39}$$

where $\Phi(\boldsymbol{y}) = \prod_{m=1}^{2K} y_m$ and $\mathbb{D} = \mathbb{G} \cap \mathbb{L}$. The sets $\mathbb{G} = \{\boldsymbol{y} \in \mathbb{R}_+^{2K} | y_m \leq \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A}_m \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{C}_m \boldsymbol{w}}, \boldsymbol{w} \in \mathbb{C}^N\}$ and $\mathbb{L} = \{\boldsymbol{y} \in \mathbb{R}_+^{2K} | y_m \geq \min_{\boldsymbol{w}} \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A}_m \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{C}_m \boldsymbol{w}}\}$ are normal set and reverse normal set, respectively. Function $\Phi(\boldsymbol{y})$ is an increasing function since $\Phi(\bar{\boldsymbol{y}}) \geq \Phi(\tilde{\boldsymbol{y}})$ for $\bar{\boldsymbol{y}} \succeq \tilde{\boldsymbol{y}}$. Then the proof of the equivalence follows similar steps as in [36]. Thus, problem (2.37) is a monotonic optimization problem. The definitions of increasing function, normal set, and reverse normal set are the same as in [36]. □

A polyblock $\mathbb{P}$ with vertex set $\mathbb{T} \subset \mathbb{R}_+^{2K}$ is defined as the finite union of all the boxes $[\boldsymbol{0}, \boldsymbol{z}]$, $\boldsymbol{z} \in \mathbb{T}$. It is dominated by its proper vertices. A vertex $\boldsymbol{z}$ is proper if there is no $\bar{\boldsymbol{z}} \neq \boldsymbol{z}$ and $\bar{\boldsymbol{z}} \succeq \boldsymbol{z}$ for $\bar{\boldsymbol{z}} \in \mathbb{T}$.

According to Proposition 2 in [36], the global maximum of the problem (2.39), if exists, is attained on $\partial^+\mathbb{D}$, i.e., the upper boundary of $\mathbb{D}$. The main idea of the polyblock approximation algorithm for solving (2.39) is to approximate $\partial^+\mathbb{D}$ by polyblocks, i.e., construct a nested sequence of polyblocks which approximate $\mathbb{D}$ from above, that is,

$$\mathbb{P}_1 \supset \mathbb{P}_2 \supset \cdots \supset \mathbb{D} \text{ s.t. } \max_{\boldsymbol{y} \in \mathbb{P}_k} \Phi(\boldsymbol{y}) \to \max_{\boldsymbol{y} \in \mathbb{D}} \Phi(\boldsymbol{y}) \tag{2.40}$$

when $k \to \infty$ and $\boldsymbol{y}_k \succeq \boldsymbol{y}_\ell$ for all $\ell \geq k$.

Now we outline how to construct the subset $\mathbb{P}_k$ in our case, which is clearly the critical step of a polyblock approximation. Let $\mathbb{T}_k$ be the proper vertex set of $\mathbb{P}_k$ and define the maximizer at iteration k as

$$\bar{\boldsymbol{y}}_k \in \arg\max_{\bar{\boldsymbol{y}}} \{\Phi(\bar{\boldsymbol{y}})|\bar{\boldsymbol{y}} \in \mathbb{T}_k\}. \tag{2.41}$$

Compute the unique intersection point of $\partial^+\mathbb{D}$ and $\bar{\boldsymbol{y}}_k$ as $\hat{\boldsymbol{y}}_k = \alpha_k \bar{\boldsymbol{y}}_k$ with $\alpha_k \in [0, 1]$. Then the proper vertex set $\mathbb{T}_{k+1}$ of $\mathbb{P}_{k+1}$ in step $k+1$ is the set obtained by substituting $\bar{\boldsymbol{y}}_k$ in $\mathbb{T}_k$ with the new vertices $\{\bar{\boldsymbol{y}}_k^1, \cdots, \bar{\boldsymbol{y}}_k^{2K}\}$ defined by

$$\bar{\boldsymbol{y}}_k^m = \bar{\boldsymbol{y}}_k - (\bar{y}_{k,m} - \hat{y}_{k,m})\boldsymbol{e}_m, \; m = 1, \cdots, 2K \tag{2.42}$$

and removing all the improper vertices as well as the vertices not belonging to $\mathbb{L}$. The scalar $\bar{y}_{k,m}$ is the mth element of $\bar{\boldsymbol{y}}_k$ and $\boldsymbol{e}_m \in \mathbb{R}_+^{2K}$ is the mth unit vector. The factor α_k is calculated as [36]

$$\alpha_k = \max_{\boldsymbol{w}} \min_{m} \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A}_m \boldsymbol{w}}{\bar{y}_{k,m} \boldsymbol{w}^{\mathrm{H}} \boldsymbol{C}_m \boldsymbol{w}}. \tag{2.43}$$

Although (2.43) is non-convex, it is an easier sub-problem which can be solved approximately ($\eta-$optimality) using the algorithm in [39]. Finally, the proposed (ϵ, η)-optimal solution using the polyblock algorithm is described in Table 2.1. The proof of the global convergence follows similar steps as in [36].

Table 2.1: (ϵ, η)-optimal polyblock algorithm for solving (2.37)

Initialization step: set initial vertex set $\mathbb{T}_0 = \{\boldsymbol{b}\}$,[1] maximum iteration number $N_{\max}$, and the threshold values ϵ, η.

Main step:
for $k = 1$ to $N_{\max}$ **do**
Solve (2.41) and (2.43) finding $\bar{\boldsymbol{y}}_k$ and $\eta-$optimal α_k.
Construct a smaller polyblock $\mathbb{P}_k$ using $\bar{\boldsymbol{y}}_k$ and α_k.
if $\max_m\{(\bar{y}_{k,m} - \hat{y}_{k,m})/\bar{y}_{k,m}\} \leq \epsilon$ **then**
break
end if
end for

Extended GPI Algorithm

The problem (2.37) can also be solved using the the general power iterative (GPI) algorithm which is introduced in [40]. However, the condition for applying GPI is not explicitly given in [40] and it is not trivial.

Let us briefly review the GPI method in [40]. According to the optimality condition, all the local maximizers for the problem (2.37) should satisfy

$$\left.\frac{\partial \lambda(\boldsymbol{w})}{\partial \boldsymbol{w}}\right|_{\boldsymbol{w}=\bar{\boldsymbol{w}}} = 0 \tag{2.44}$$

where $\lambda(\boldsymbol{w}) = \prod_{m=1}^{2K} \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A}_m \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{C}_m \boldsymbol{w}}$. After differentiation and some algebraic manipulation, the condition in (2.44) can be converted into

$$\boldsymbol{V}(\bar{\boldsymbol{w}})\bar{\boldsymbol{w}} = \lambda(\bar{\boldsymbol{w}})\boldsymbol{Q}(\bar{\boldsymbol{w}})\bar{\boldsymbol{w}} \tag{2.45}$$

[1]Here $\boldsymbol{b} \in \mathbb{R}_+^{2K}$ satisfies $\boldsymbol{b}_m = \max_{\boldsymbol{w}} \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{A}_m \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{C}_m \boldsymbol{w}}$, $m = 1, \cdots, 2K$.

where $\boldsymbol{V}(\bar{\boldsymbol{w}}) = \sum_{m=1}^{2K}(\prod_{i\neq m} \bar{\boldsymbol{w}}^{\mathrm{H}}\boldsymbol{A}_i\bar{\boldsymbol{w}})\boldsymbol{A}_m$ and $\boldsymbol{Q}(\bar{\boldsymbol{w}}) = \sum_{m=1}^{2K}(\prod_{i\neq m} \bar{\boldsymbol{w}}^{\mathrm{H}}\boldsymbol{C}_i\bar{\boldsymbol{w}})\boldsymbol{C}_m$. Equation (2.45) is a generalized eigenvalue problem and $\lambda(\bar{\boldsymbol{w}})$ can be thought as the generalized eigenvalue of matrices $\boldsymbol{V}(\bar{\boldsymbol{w}})$ and $\boldsymbol{Q}(\bar{\boldsymbol{w}})$. Thus, the maximum generalized eigenvalue $\lambda_{\max}(\bar{\boldsymbol{w}})$ is the maximum of the problem (2.37). Since both matrices are functions of $\bar{\boldsymbol{w}}$, a closed-form solution is not possible. Therefore, the authors in [40] apply the recursive power method of [41] to obtain the solution. It is also numerically shown that the GPI algorithm converges in 30 iterations. However, this is not true in general. In [41], it is shown that the power method converges only if the largest eigenvalue is dominant and the convergence speed depends on the ratio between the largest and the second largest eigenvalues. Although we can only demonstrate this via numerical simulations, we claim that GPI should have similar features as the original power method. Thus, the following conjecture is given.

1 Conjecture. The GPI algorithm converges if there is a dominant eigenvalue. The convergence behavior depends on the dispersion of the eigenvalue profiles of the matrices of $\boldsymbol{A}_m$ and $\boldsymbol{C}_m$.

Nevertheless, the GPI algorithm can be applied to our scenario especially when $\boldsymbol{D}_m$ is rank deficient ($N > 2(K-1)$ since $\mathrm{rank}\{\boldsymbol{D}_m\} = \min\{N, 2K-2\}$), i.e., there will be a dominant eigenvalue when SNR $\rightarrow \infty$. Moreover, it converges faster in the high SNR regime with a given error tolerance factor. For a detailed implementation one can be referred to [40].

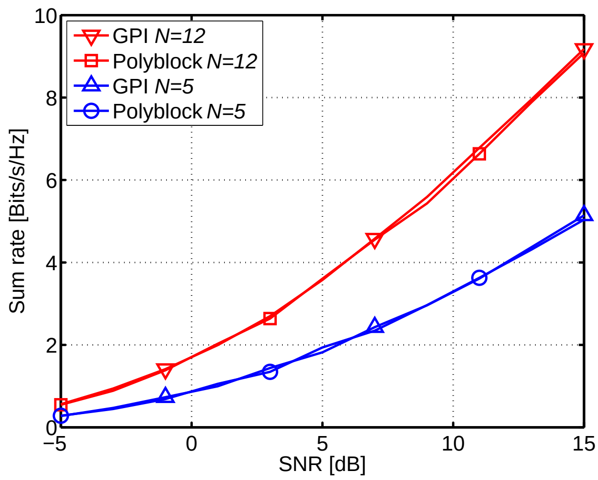


Figure 2.5: Sum rate comparison. There are $L = 2$ pairs of UTs.

Fig. 2.5 shows the comparison of different algorithms with $N = 5$ relays and $N = 12$ relays in the network. In general, the two algorithms converge. However, the polyblock algorithm performs slightly worse than the GPI algorithm. This is due to the (ϵ, η)-optimality.

Remark. Although all the proposed algorithms do not have any requirements on N, to cancel the interference completely $N > 2K(K-1)$ is required since the rank of the sum of the interference terms is equal to $\mathrm{rank}\{\sum_{m=1}^{2K} \boldsymbol{D}_m\} = 2K(K-1)$ [42]. If $N \leq 2K(K-1)$, the results will be unfair for some users since they will suffer from extremely lower throughputs compared to the other users.

Extension to the Multi-Operator Two-Way Relay Channel

Now we show that the sum rate maximization problem in the multi-operator TWR scenario can be also solved using the polyblock framework or the GPI method. For the multi-operator TWR scenario we compute the optimal $\boldsymbol{G}$ which maximizes the sum rate of the system subject to a transmit power constraint at the relay, i.e.,

$$\begin{aligned} \max_{\boldsymbol{G}} \quad & \tfrac{1}{2}\sum_{\ell=1}^{L}\sum_{k=1}^{2}\log_2(1+\eta_k^{(\ell)}) \\ \text{subject to} \quad & \mathbb{E}\{\|\bar{\boldsymbol{r}}\|^2\} \leq P_{\mathrm{R}}. \end{aligned} \tag{2.46}$$

Recalling from D2.3b, the SINR $\eta_k^{(\ell)}$ of each UT is expressed as

$$\eta_k^{(\ell)} = \frac{\mathbb{E}\{|\boldsymbol{h}_k^{(\ell)^{\mathrm{T}}}\boldsymbol{G}\boldsymbol{h}_{3-k}^{(\ell)}x_{3-k}^{(\ell)}|^2\}}{\mathbb{E}\{|\sum_{\bar{k},\bar{\ell}\neq\ell}\boldsymbol{h}_k^{(\ell)^{\mathrm{T}}}\boldsymbol{G}\boldsymbol{h}_{\bar{k}}^{(\bar{\ell})}x_{\bar{k}}^{(\bar{\ell})}|^2\} + \mathbb{E}\{\|\boldsymbol{h}_k^{(\ell)^{\mathrm{T}}}\boldsymbol{G}\boldsymbol{n}_{\mathrm{R}}\|^2\} + \sigma_k^{(\ell)^2}} \tag{2.47}$$

To derive the optimal $\boldsymbol{G}$, further algebraic manipulations are required. The transmit power at the relay can be expanded as

$$\begin{aligned} & \mathbb{E}\{\|\bar{\boldsymbol{r}}\|^2\} \\ = \; & \mathbb{E}\{\mathrm{Tr}\{\boldsymbol{G}\boldsymbol{r}(\boldsymbol{G}\boldsymbol{r})^{\mathrm{H}}\}\} \\ = \; & \mathrm{Tr}\left\{\boldsymbol{G}\left(\sum_{k,\ell}P_k^{(\ell)}\boldsymbol{h}_k^{(\ell)}\boldsymbol{h}_k^{(\ell)^{\mathrm{H}}} + \sigma_{\mathrm{R}}^2\boldsymbol{I}_{M_{\mathrm{R}}}\right)\boldsymbol{G}^{\mathrm{H}}\right\} \\ = \; & \sum_{k,\ell}\mathrm{Tr}\left\{P_k^{(\ell)}\boldsymbol{G}\boldsymbol{h}_k^{(\ell)}\boldsymbol{h}_k^{(\ell)^{\mathrm{H}}}\boldsymbol{G}^{\mathrm{H}}\right\} + \mathrm{Tr}\left\{\sigma_{\mathrm{R}}^2\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\right\} \\ = \; & \sum_{k,\ell}P_k^{(\ell)}(\boldsymbol{G}\boldsymbol{h}_k^{(\ell)})^{\mathrm{H}}\boldsymbol{G}\boldsymbol{h}_k^{(\ell)} + \sigma_{\mathrm{R}}^2\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g} \\ = \; & \boldsymbol{g}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{g} \end{aligned} \tag{2.48}$$

where $\boldsymbol{g} = \mathrm{vec}\{\boldsymbol{G}\}$. The fact $\mathrm{Tr}\{\boldsymbol{\Gamma}\boldsymbol{\zeta}\} = \mathrm{Tr}\{\boldsymbol{\zeta}\boldsymbol{\Gamma}\}$ and $\mathrm{vec}\{\boldsymbol{\Gamma}\boldsymbol{X}\boldsymbol{\zeta}\} = (\boldsymbol{\zeta}^{\mathrm{T}}\otimes\boldsymbol{\Gamma})\mathrm{vec}\{\boldsymbol{X}\}$ is used in the derivation. Moreover, $\boldsymbol{\Lambda}$ is a positive definite Hermitian matrix which is defined as

$$\boldsymbol{\Lambda} = \sum_{k,\ell}P_k^{(\ell)}((\boldsymbol{h}_k^{(\ell)}\boldsymbol{h}_k^{(\ell)^{\mathrm{H}}})^{\mathrm{T}}\otimes\boldsymbol{I}_{M_{\mathrm{R}}}) + \sigma_{\mathrm{R}}^2\boldsymbol{I}_{M_{\mathrm{R}}^2}. \tag{2.49}$$

Following a similar procedure, the SINR $\eta_k^{(\ell)}$ can be rewritten as

$$\eta_k^{(\ell)} = \frac{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{\Phi}_k^{(\ell)}\boldsymbol{g}}{\boldsymbol{g}^{\mathrm{H}}(\boldsymbol{\Upsilon}_k^{(\ell)} + \boldsymbol{\Delta}_k^{(\ell)})\boldsymbol{g} + \sigma_k^{(\ell)^2}} \tag{2.50}$$

where $\boldsymbol{\Phi}_k^{(\ell)}$, $\boldsymbol{\Upsilon}_k^{(\ell)}$, and $\boldsymbol{\Delta}_k^{(\ell)}$ are defined as

$$\begin{aligned}
\boldsymbol{\Phi}_k^{(\ell)} &= P_k^{(\ell)}(\boldsymbol{h}_{3-k}^{(\ell)\mathrm{T}} \otimes \boldsymbol{h}_k^{(\ell)\mathrm{T}})^{\mathrm{H}}(\boldsymbol{h}_{3-k}^{(\ell)\mathrm{T}} \otimes \boldsymbol{h}_k^{(\ell)\mathrm{T}}) \\
\boldsymbol{\Upsilon}_k^{(\ell)} &= \sum_{\substack{\bar{k}=1,2 \\ \tilde{\ell}\neq\ell}} P_{\bar{k}}^{(\tilde{\ell})}(\boldsymbol{h}_{\bar{k}}^{(\tilde{\ell})\mathrm{T}} \otimes \boldsymbol{h}_k^{(\ell)\mathrm{T}})^{\mathrm{H}}(\boldsymbol{h}_{\bar{k}}^{(\tilde{\ell})\mathrm{T}} \otimes \boldsymbol{h}_k^{(\ell)\mathrm{T}}) \\
\boldsymbol{\Delta}_k^{(\ell)} &= \sigma_{\mathrm{R}}^2(\boldsymbol{I}_{M_{\mathrm{R}}} \otimes (\boldsymbol{h}_k^{(\ell)}\boldsymbol{h}_k^{(\ell)\mathrm{H}})^{\mathrm{T}}).
\end{aligned} \tag{2.51}$$

Inserting (2.50) and (2.48) into (2.46), the original problem can be reformulated as

$$\begin{aligned}
&\max_{\boldsymbol{g}} \quad \tfrac{1}{2}\sum_{k,\ell}\log_2\left(1+\frac{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{\Phi}_k^{(\ell)}\boldsymbol{g}}{\boldsymbol{g}^{\mathrm{H}}(\boldsymbol{\Upsilon}_k^{(\ell)}+\boldsymbol{\Delta}_k^{(\ell)})\boldsymbol{g}+\sigma_k^{(\ell)2}}\right) \\
&\text{subject to} \quad \boldsymbol{g}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{g} \leq P_{\mathrm{R}}.
\end{aligned} \tag{2.52}$$

Clearly, after using the fact that the constraint in (2.52) is active at the optimality, problem (2.52) can be reformulated into a similar form as (2.37), i.e.,

$$\max_{\boldsymbol{g}} \ \lambda(\boldsymbol{g}) = \prod_{\ell=1}^{L}\prod_{k=1}^{2}\frac{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}_k^{(\ell)}\boldsymbol{g}}{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{C}_k^{(\ell)}\boldsymbol{g}} \tag{2.53}$$

where $\boldsymbol{C}_k^{(\ell)} = \boldsymbol{\Upsilon}_k^{(\ell)} + \boldsymbol{\Delta}_k^{(\ell)} + \frac{\sigma_k^{(\ell)2}}{P_{\mathrm{R}}}\boldsymbol{\Lambda}$ and $\boldsymbol{A}_k^{(\ell)} = \boldsymbol{C}_k^{(\ell)} + \boldsymbol{\Phi}_k^{(\ell)}$ are positive definite Hermitian matrices. The solution to problem (2.53) differs from the solution to (2.46) only in scaling and reshaping, i.e., if $\bar{\boldsymbol{g}}$ is the solution to (2.53), the optimal solution to (2.46) is given by

$$\boldsymbol{G} = \mathrm{unvec}_{M_{\mathrm{R}}\times M_{\mathrm{R}}}\left\{\bar{\boldsymbol{g}}\sqrt{\frac{P_{\mathrm{R}}}{\bar{\boldsymbol{g}}^{\mathrm{H}}\boldsymbol{\Lambda}\bar{\boldsymbol{g}}}}\right\}. \tag{2.54}$$

3 Utility Models

In the following we discuss utility models. The results partly depend on the previous section, where interference models were derived.

We start be reviewing SINR-based utility models. Then we discuss how this can be exploited for a framework for joint scheduling and power control. We end with the introduction of a model for WNC based sharing in Section 3.4.

3.1 SINR-Based QoS Models and Regions

In this section, QoS stands for an arbitrary performance measure, which depends on the SINR by a strictly monotone and continuous function ϕ defined on $\mathbb{R}_+$. The QoS of user k is

$$q_k(\boldsymbol{p}) = \phi_k\big(\mathrm{SINR}_k(\boldsymbol{p})\big) \ , \quad k \in \mathcal{K} \ . \tag{3.1}$$

Many performance measures depend on the SINR in this way. Examples are SINR, logarithmic SINR, MMSE, bit error rate, etc.

Note that $\boldsymbol{p}$ can be regarded as the resource vector, which determines how the transmit powers are allocated to the different resources, as explained in Subsection 2.3.4. Thus, the user index k denotes a particular data stream, and multiple data streams can be transmitted by every user. Thus, SINR_k is the SINR of the kth communication link.

3.1.1 QoS Region

The theoretical limits can be described in terms of the QoS region, i.e., the set of all jointly achievable QoS, as illustrated in Fig. 3.1. In order to fully exploit the available resources, it is desirable to achieve an operating point on the boundary of the region. Common optimisation goals are to maximise the sum efficiency (e.g. sum throughput), to achieve max-min fairness, or some compromise between fairness and efficiency.

Let γ_k be the inverse function of ϕ_k, then $\gamma_k(q_k)$ is the minimum SINR level needed by the kth communication link to satisfy some QoS target q_k. Assume that the QoS is defined on some domain $\mathbb{Q}$ and the K-dimensional domain is denoted by $\mathbb{Q}^K$. Let $\boldsymbol{q} \in \mathbb{Q}^K$ be a vector of QoS values, then the corresponding SINR vector is

$$\boldsymbol{\gamma}(\boldsymbol{q}) = [\gamma_1(q_1), \ldots, \gamma_K(q_K)]^T \ . \tag{3.2}$$

That is, $\boldsymbol{q}$ can be achieved if and only if the SINR targets $\boldsymbol{\gamma} := \boldsymbol{\gamma}(\boldsymbol{q})$ can be achieved. Thus, in many properties of the QoS region, like Pareto optimality, have

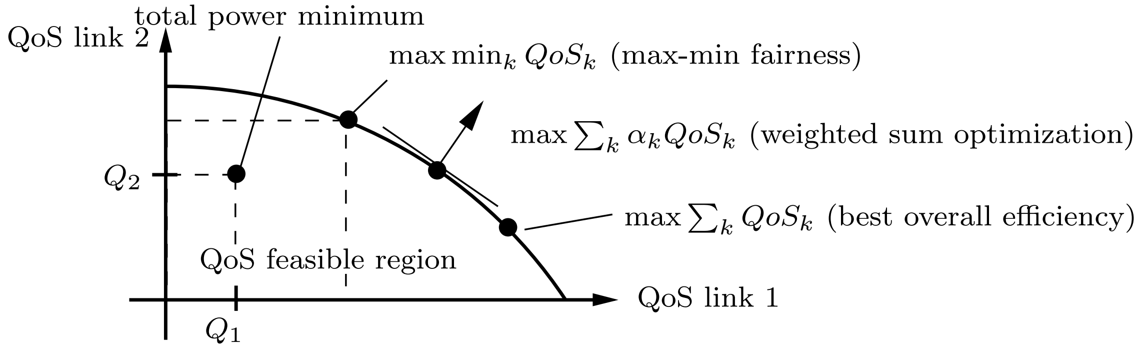


Figure 3.1: Theoretical limits of a shared system. There is a trade-off between the performances of the individual communication links. Increasing the performance of one link generally comes at the cost of another links. The goal of the system optimisation is to find an operating point on the boundary of the region.

direct correspondence in the SINR region. Hence, the importance of thoroughly understanding the SINR region.

We say that $\boldsymbol{\gamma}$ is feasible if for any $\epsilon > 0$ there exists a $\boldsymbol{p}_\epsilon > 0$ such that

$$\mathrm{SINR}_k(\boldsymbol{p}_\epsilon) \geq \boldsymbol{\gamma}_k - \epsilon, \quad \forall k \in \mathcal{K} \; . \tag{3.3}$$

If this is fulfilled for $\epsilon = 0$, then the targets $\boldsymbol{\gamma}$ can actually be attained, otherwise they are achieved in an asymptotic sense. For practical scenarios, the former case is mostly ensured by noise and power constraints. A necessary and sufficient condition for feasibility is $C(\boldsymbol{\gamma}) \leq 1$, where $C(\boldsymbol{\gamma})$ is defined by (2.16). Thus, the SINR region is

$$\mathcal{S} = \{\boldsymbol{\gamma} \geq 0 : C(\boldsymbol{\gamma}) \leq 1\} \; . \tag{3.4}$$

This definition of an SINR (resp. QoS) region is quite general and not tied to any particular channel or interference mitigation strategy. The definition (3.4) contains other known SINR regions as special cases.

The indicator function $C(\boldsymbol{\gamma})$ fulfils the axioms A1, A2, A3. Thus, $\mathcal{S}$ is a subset of an interference function. This has some interesting consequences: Certain properties of $\mathcal{S}$ directly correspond with the properties of $C(\boldsymbol{\gamma})$. By analysing $C(\boldsymbol{\gamma})$ we obtain valuable information about the structure of QoS regions [18].

For example, the property A3 translates into *comprehensiveness.* A set $\mathcal{V} \subset \mathbb{R}^K_{++}$ (strictly positive reals) is said to be *upward-comprehensive* if for all $\boldsymbol{w} \in \mathcal{V}$ and $\boldsymbol{w}' \in \mathbb{R}^K_{++}$, the inequality $\boldsymbol{w}' \geq \boldsymbol{w}$ implies $\boldsymbol{w}' \in \mathcal{V}$. If the inequality is reversed, then $\mathcal{V}$ is said to be *downward-comprehensive.* In the context of game theory, comprehensiveness is interpreted as *free disposability of utility* [43]. If a user can achieve a certain utility, then it can freely dispose of all utilities below.

It was shown in [18] that any set from $\mathbb{R}^K_{++}$ is closed downward-comprehensive if and only if it is a sublevel set of an interference function. Likewise, it is closed upward-comprehensive if and only if it is a superlevel set of an interference function.

This shows a one-to-one correspondence between interference functions and comprehensive sets. Moreover, it was shown that such a comprehensive set is convex if and only if the corresponding interference function is convex.

Convexity of the QoS region is a desirable property which often facilitates efficient algorithmic solutions. However, many QoS regions are comprehensive but not convex. A standard approach is to "convexify" the utility set by randomisation techniques, e.g. [44, 45], or by resource sharing. However, such a strategy may not always be possible or relevant in all situations.

Another more elegant way is to exploit "*hidden convexity*", if available. Sometimes, a given problem is not convex but there exists an equivalent convex problem formulation. That is, the original non-convex problem can be solved indirectly by solving the equivalent problem instead.

An example is *logarithmic convexity* (log-convexity). A function $f(x)$ is said to be log-convex on its domain if $\log f(x)$ is convex. Log-convexity was already exploited in the context of power control [46–52]. Assume that $\text{SINR}_k(\boldsymbol{p}) = p_k/\mathcal{I}_k(\boldsymbol{p})$, and $\mathcal{I}_1, \dots, \mathcal{I}_K$ are log-convex interference functions after a change of variable $\boldsymbol{s} = \log \boldsymbol{p}$. That is, $\mathcal{I}(\exp(\boldsymbol{s}))$ is log-convex with respect to $\boldsymbol{s}$. This is fulfilled, for example, for all linear interference functions and also for certain worst-case designs [20].

Under the assumption of SINRs defined by log-convex interference functions, the resulting indicator function $C(\boldsymbol{\gamma})$ itself is a log-convex interference function after another change of variable $q_k = \log \gamma_k$ [20]. That is, $C(\exp(\boldsymbol{q}))$ is log-convex with respect to $\boldsymbol{q} = (q_1, \dots, q_K)$. Every log-convex function is convex, and a sub-level set of a convex function is convex. Thus, the set $\mathcal{S}$ is convex on a logarithmic scale. This can be exploited for the development of algorithms that operate on the boundary of the region [50].

3.1.2 The Boundary and Max-Min Interference Balancing

The boundary of the QoS region is fully characterized by the SINR balancing function

$$C(\boldsymbol{\gamma}) = \max_{\boldsymbol{p} \in \mathcal{P}} \min_k \frac{\text{SINR}_k(\boldsymbol{p})}{\gamma_k} \,. \tag{3.5}$$

This a special case of (2.16), which was already discussed in Section 2.3. Here, we have the specific case that the scenario includes noise and power constraints. Note that SINR_k can be the SINR that results from adaptive beamforming, which can be implicitly included in the definition of the interference model.

Within SAPHYRE, the following iterative fixed point algorithm was developed [31]

$$\boldsymbol{p}^{(n+1)} = \boldsymbol{\mathcal{I}}(\boldsymbol{p}^{(n)})/\|\boldsymbol{\mathcal{I}}(\boldsymbol{p}^{(n)})\|, \quad \boldsymbol{\mathcal{I}}(\boldsymbol{p}) \triangleq [\mathcal{I}_1(\boldsymbol{p}), \dots, \mathcal{I}_L(\boldsymbol{p})]^T.$$

The algorithm measures the interference for each communication link. In each iteration step n the transmit powers $\boldsymbol{p}^{(n)}$ are updated. In each step, the vector is rescaled by the norm $\|\boldsymbol{\mathcal{I}}(\boldsymbol{p}^{(n)})\|$.

Note that the power updates can be performed locally, provided that the interference values are known. the only common value is the the norm $\|\boldsymbol{\mathcal{I}}(\boldsymbol{p}^{(n)})\|$, which needs to be exchanged in the network. This can be achieved by every user broadcasting its measured interference value.

The iteration balances the relative SIR values of the communication link. This optimisation strategy aims at achieving fairness. Another goal is discussed in the following subsection.

3.1.3 Throughput Maximization

Another way of optimizing the interference in a shared network is to maximize the sum rate $\sum_{l=1}^{L} \log(1 + \text{SINR})$. In Fig. 3.1 this operating point corresponds to the maximum efficiency. This strategy maximizes the total network throughput. Individual user priorities can be included by adding weighting factors w_l.

In SAPHYRE we study the weighted sum-rate maximization problem

$$R = \max_{\boldsymbol{p} \in \mathcal{P}} \sum_{l=1}^{L} w_l \log_2 \left(1 + \frac{p_l}{Y_l(\boldsymbol{p})}\right) . \tag{3.6}$$

Here, Y_l is a concave standard interference function, as explained in Subsection 2.3.3. This includes many signal processing techniques that aim at minimizing interference, e.g. beamforming or linear precoding.

Problem (3.6) is NP hard. Thus, it it is not possible to find an equivalent convex reformulation. Convex reformulations only exist for special cases or approximations. Such a special case is the SIMO channel with dirty paper coding, or the MISO channel with perfect successive interference cancellation. An example of a convex approximation follows from replacing $\log(1 + \text{SINR})$ by $\log(\text{SINR})$. In this case, problem (3.6) becomes logarithmically convex [19].

Fortunately, we can exploit that any concave standard interference function $Y_l(\boldsymbol{p})$ has a representation [16]

$$Y_l(\boldsymbol{p}) = \min_{z_l \in \mathcal{Z}_l} \left(\boldsymbol{p}^T \boldsymbol{\psi}_l(z_l) + \sigma^2_{n_l}(z_l)\right) , \quad l = 1, \dots, L, \tag{3.7}$$

where vectors $\boldsymbol{\psi}_l(z_l)$, $\forall l$, model the cross-power coupling between the links. The coupling coefficients depend on parameters $z = (z_1, \dots, z_L)$, which are chosen from compact sets $\mathcal{Z}_1, \dots, \mathcal{Z}_L$, respectively. The vector $\sigma^2_{\boldsymbol{n}}(z) \triangleq [\sigma^2_{n_1}(z_1), \dots, \sigma^2_{n_L}(z_L)]^T$ is associated with the effective noise powers, which are also influenced by the chosen receive strategies.

The minimization of (3.7) corresponds to the maximization of the SINR, thus we refer to z as "receive strategies". We can exploit that the vector $\boldsymbol{\psi}_l(z_l)$ depends only on z_l. As explained in [30], the problem (3.6) can be solved by a branch&bound algorithm, which computes the global optimum up to a chosen ϵ accuracy. The algorithm provides a benchmark for the achievable throughput in a shared system.

3.2 QoS Framework for Joint Scheduling, Power Control, and Beamforming

In this section, a framework is proposed for joint optimization of bandwidth allocation and transmit strategies. A wireless network is considered comprising K single-antenna users and a spectrum band divided in $N < K$ orthogonal resources (frequency channels). The spectrum sharing scenario of interest to SAPHYRE is when the users belong to (at least) two operators, but they are served in the *same* spectrum band. The definition of service is that a predetermined QoS target is met for each user allocated a resource. The assumption is that at most one resource can be allocated to each user, but this allocation is not exclusive; that is, the same resource can be allocated to multiple users. Hence, the performance of each user is subject to the interference caused by the co-channel transmission, which have to be appropriately designed in order to simultaneously meet the QoS targets of all served users.

The resource allocation and transmit design problem is jointly tackled by novel constrained optimization formulations. Two cases are considered for the transmit strategies. First, the transmitters are assumed to have a single antenna each, so that only their transmit powers can be optimized. In this case, the proposed formulation is a mixed-integer linear-programming (MILP) problem in standard form. Second, the transmitters are assumed to have multi-antenna capability, so that transmit beamforming optimization can be performed in addition to power control. In this case, the proposed formulation is a mixed-integer second-order-cone programming (MI-SOCP) problem in standard form. In both cases, the joint problem of interest is NP-hard due to the scheduling component, which is of combinatorial nature and is captured by the integer part of the joint formulations. However, the proposed framework, due to its specific structure (linear or conic, respectively), enables solving the problem *exactly* and relatively efficiently for the vast majority of instances, using off-the-shelf algorithms. This framework can be used for benchmarking and for motivating the development of low-complexity coordinated algorithms.

The motivation for this work is the increased sophistication of wireless communication systems envisioned for the future. For example, we witness a number of emerging applications, such as streaming video and interactive gaming, which bring high requirements on the QoS. This puts additional demands on the resource allocation algorithms used in the network. The SAPHYRE vision is that links that belong to different infrastructure (operators) may coexist in the same spectrum. Collectively, the envisioned technology that will be able to offer such coexistence is called *non-orthogonal spectrum sharing*. Spectrum sharing will require a wide range of new enabling technology and algorithms, which includes, among others, algorithms for optimally solving scheduling, power control, and transmit beamforming problems.

3.2.1 QoS Power Control

As discussed in Section 2.2, the fundamental aspects of power allocation in wireless networks can be understood by considering a generic model comprising K transmitter-receiver pairs. Here, the transmissions on the K links under study take place concurrently in the same channel. Due to the broadcast nature of the wireless medium there is coupling between the transmitter-receiver pairs. The effect is that each receiver listens to a superposition of the desired signal and of all the other $K-1$ transmitted signals, which constitute interference. The communication quality of the kth link can be quantified via the received SINR

$$\mathrm{SINR}_k \triangleq \frac{G_{kk} p_k}{\sum\limits_{\ell \neq k} G_{\ell k} p_\ell + \sigma_n^2}. \tag{3.8}$$

In (3.8), p_k is the transmit power used on the kth link, $G_{\ell k}$ is the gain of the channel between the ℓth transmitter and the kth receiver, and σ_n^2 is the variance of the AWGN noise. The channel gains include the effects of propagation loss, shadowing and fading.

The receivers are assumed to treat the interference as noise. Hence, the maximum achievable rate, that a link can support, is dictated by the SINR. In practice, a rate requirement, that is imposed by an application, can be directly translated to an SINR requirement, for a given modulation/coding scheme and a given target bit-error rate. Thus, QoS can be guaranteed to the kth link when the SINR_k exceeds a predetermined threshold γ_k. In order to ensure this condition, the physical-layer protocol may adjust the transmit power p_k up to a bound P, which is typically determined by regulatory and/or hardware constraints. However, while boosting p_k increases SINR_k, it reduces at the same time SINR_ℓ $\forall \ell \neq k$. Hence, the transmission powers need to be determined jointly.

The SINR-constrained power control problem (2.12) is rewritten as

$$\min_{\{p_k \in [0,P]\}_{k=1}^K} \sum_{k=1}^K p_k \tag{3.9}$$

$$\text{s.t.} \quad \mathrm{SINR}_k \geq \gamma_k \quad \forall k \in \mathcal{K}, \tag{3.10}$$

where SINR_k is defined in (3.8) and $\mathcal{K} \triangleq \{1, \ldots, K\}$ is the set of all direct links. The objective function in (3.9) strives to minimize the total transmission power subject to a QoS constraint (3.10) for each link. This minimizes the overall interference emitted by the network, and at the same time prolongs the operating lifetime of energy-starved transmitters.

Problem (3.9)–(3.10) has been extensively studied in the past, e.g. [4, 53]. From an optimization viewpoint, it is a convex linear programming (LP) problem. Hence, when an optimum solution exists, it can be found very efficiently. The solution is found by a central controller which knows all the channel gains $\{G_{\ell k}\}$ and QoS requirements $\{\gamma_k\}$. Owing to the linearity of the problem, even in the absence of a central controller efficient algorithms have been proposed to find the optimum solution in a distributed manner [15, 54].

However, problem (3.9)–(3.10) may be infeasible. This typically happens when the requested SINR thresholds $\{\gamma_k\}$ are large or when the coupling channel gains $\{G_{\ell k}\}_{\ell \neq k}$ are large relative to the corresponding direct link gains $\{G_{kk}\}$. If the QoS constraints for all K links cannot be simultaneously satisfied by power control, the different links need to be scheduled to more than one orthogonal resources (frequency channels). Algorithms available to date take such access-control decisions based mostly on the channel gains and disregard valuable insights gained by the attempt to solve the QoS power control problem in the physical layer. Significant improvements are expected by cross-layer approaches to the scheduling and power control problem. Previous, related results in this direction include the joint power and admission control[1] problems considered in [55, 56]. Therein, modifications to the distributed power control algorithm of [54] are proposed to determine whether another link can be served in the same resource without yielding the problem infeasible.

A different, more disciplined, approach is proposed in [57–59]. This line of work follows a common methodology. Initially, the two-layer problem of interest is formulated as a joint optimization problem, with the introduction of auxiliary binary variables that model the scheduling decisions. The joint problem is inherently NP-hard and in a form that cannot be directly solved. In a second step, the joint optimization problem is relaxed to its convex counterpart. Finally, heuristic algorithms are proposed which iteratively solve different instances of this convex problem. These algorithms are solved centrally and provide high-quality suboptimal solutions with polynomially-bounded worst-case complexity. The case of admission control (one resource) is treated in [57] and [58], jointly with beamforming and power control, respectively. The joint formulations are relaxed to semidefinite programming problems, which are convex. The general case of scheduling (many resources) and power control is treated in [59]. Therein, the joint formulation is relaxed to a geometric programming problem, which admits a convex reformulation.

The problem considered in [59] is revisited in [60], for a slightly different scenario, and an alternative formulation is proposed. The key difference is that the novel formulation maintains the linearity of the original QoS power control problem (3.9)–(3.10). Specifically, the joint scheduling and QoS power control problem is modeled as a so-called MILP problem. MILP problems are NP-hard in general, but due to their linearity they can be solved rather efficiently, in most instances, by means of branch-and-bound techniques. Hence, contrary to the approach in [57–59], the proposed formulation in [60] enables finding the optimal solution, yet at the cost of occasional high complexity.

[1] Some authors refer to this problem as scheduling. Herein, the term scheduling is reserved for the explicit allocation of links to resources.

3.2.2 Joint Scheduling and QoS Power Control

The fundamental question is how to *jointly* allocate resources and power *optimally*, in order to minimize the overall interference and maximize the number of links that can be served. To state this problem formally, we assume that there are $N < K$ available orthogonal resources and denote the set of their indexes as $\mathcal{N} \triangleq \{1, \ldots, N\}$. This is a valid assumption when the goal is to maximize the spectral efficiency. If it was $N > K$ the solution of the problem would have been to trivially schedule one link per resource. The other main assumption in the spectrum-sharing paradigm, that is of interest to SAPHYRE, is that the K links requesting service can be potentially scheduled to any resource. We denote the transmission powers and the channel gains in the nth resource as $\{p_k^n\}_{k=1}^K$ and $\{G_{\ell k}^n\}_{\ell,k=1}^K$, respectively. The SINR that the kth receiver experiences when tuned to the nth resource is then equal to

$$\mathrm{SINR}_k^n \triangleq \frac{G_{kk}^n p_k^n}{\sum\limits_{\ell \neq k} G_{\ell k}^n p_\ell^n + \sigma_n^2}. \tag{3.11}$$

We say that the kth link is *assigned* to the nth resource when there exist feasible powers $\{p_k^n\}_k$ such that $\mathrm{SINR}_k^n \geq \gamma_k$. We call the kth link *served* or *admitted* when it is assigned to some resources. The problem is then to find the optimum (i) scheduling, i.e. assignment of links to resources; and (ii) transmission powers, that maximize the number of admitted receivers and minimize the total transmission power required to serve them.

In order to solve the joint QoS problem, we formulate it as a constrained optimization problem. To proceed, we introduce the auxiliary binary variables $\{s_k^n \in \{0,1\}\}_{k\in\mathcal{K}}^{n\in\mathcal{N}}$, one per resource and link. Each binary variable s_k^n models the following scheduling question: Can the kth link be assigned to the nth resource? The answer is "yes" when $s_k^n = 1$ and "no" otherwise. It is evident that the primary goal of the optimization should be to maximize the number of positive answers. Hence, using the auxiliary variables $\{s_k^n\}$, we formulate the problem as

$$\max_{\left\{\substack{p_k^n \in [0,P] \\ s_k^n \in \{0,1\}}\right\}_{k\in\mathcal{K}}^{n\in\mathcal{N}}} \quad \sum_{k=1}^K \sum_{n=1}^N s_k^n - W \sum_{k=1}^K \sum_{n=1}^N p_k^n \tag{3.12}$$

subject to

$$\frac{G_{kk}^n p_k^n + M(1 - s_k^n)}{\sum\limits_{\ell \neq k} G_{\ell k}^n p_\ell^n + \sigma_n^2} \geq \gamma_k \qquad \forall k \in \mathcal{K},\ \forall n \in \mathcal{N}, \tag{3.13a}$$

$$p_k^n - P s_k^n \leq 0 \qquad \forall k \in \mathcal{K},\ \forall n \in \mathcal{N}, \tag{3.13b}$$

$$\sum_{n=1}^N s_k^n \leq 1 \qquad \forall k \in \mathcal{K}. \tag{3.13c}$$

In what follows, we will explain the role of each equation in (3.12)–(3.13), starting the discussion with the conditions (3.13) and concluding it with the motivation of the objective function (3.12). In the process, we will also elaborate on the operational meaning of the (yet undefined) scalar positive-real parameters W and M that appear in (3.12) and (3.13a).

Equation (3.13a) defines N SINR constraints, one per available resource, for each of the K links. Effectively, the binary variable s_k^n acts as an "if statement" that determines whether the (k, n)th inequality of (3.13a) is active in the power control problem. When $s_k^n = 1$, the (k, n)th inequality falls back to defining a standard SINR constraint. When $s_k^n = 0$, the (k, n)th inequality does not impose any constraint on $\{p_k^n\}_k$, provided that M is large enough to satisfy the inequality for all feasible values of $\{p_k^n\}_k$.

Hence, we need to choose the parameter M so that all KN inequalities (3.13a) are fulfilled when $\{s_k^n = 0\}$, irrespective of the values for $\{p_k^n\}$. It suffices to consider the worst-case scenario, where all the interfering transmitters use full power, whereas the transmitter in the direct link is silent. Setting $\{p_\ell^n = P\}_{\ell \neq k}$ and $p_k^n = 0$ in each inequality in (3.13a), and selecting the maximum resulting lower bound we have

$$M \geq \max_{k,n} \ \gamma_k \sum_{\ell \neq k} G_{\ell k}^n P + \gamma_k \sigma_n^2. \tag{3.14}$$

Note that with $s_k^n = 0$, the optimization (3.12)–(3.13) will yield $p_k^n = 0$. This is so because then p_k^n appears only as interference in the denominator of the "active" SINR constraints in (3.13a), while the second term of the objective function (3.12) seeks to minimize the total transmission power. Hence, there is no need to explicitly account for the links that do not transmit in the denominator of SINR_k^n.

Equation (3.13b) is a technical condition that is, strictly speaking, redundant in the sense that the optimization problem will yield the same solution without this constraint. When $s_k^n = 1$, (3.13b) basically restates the power constraint, and when $s_k^n = 0$, it effectively sets the optimal p_k^n to zero. This will however be the result even without requiring it explicitly in a separate constraint, as discussed above. While the constraint (3.13b) does not affect the optimal solution, including it may speed up the algorithms used to find a numerical solution.

Equation (3.13c) makes sure that each link is assigned to at most one resource. Since QoS is already guaranteed to the link when just one out of N respective constraints (3.13a) is active, multiple assignments would solely increase the interference in the wireless network. If there are extra resources, i.e. other feasible assignments, the system would rather utilize them to serve more links or to decrease the total transmission power.

Note that by letting the sum in (3.13c) take on values smaller than 1 (actually 0), admission control functionality is included in the joint optimization problem (3.12)–(3.13). This means that when there are not enough resources, the system may deny service to some links in order to ensure service to the remaining ones. For a denied link k, we get $\{s_k^n = 0\}^n$. This insight leads to the following result.

1 Claim. Optimization problem (3.12)–(3.13) is always feasible.

Proof. A trivial feasible solution is always $\{s_k^n = 0\}_{k \in \mathcal{K}}^{n \in \mathcal{N}}$ for any $\{p_k^n \in [0, P]\}_{k \in \mathcal{K}}^{n \in \mathcal{N}}$. However, this is the worst possible solution from QoS perspective, since it corresponds to the case that none of the K links is served. Another trivial solution is to serve one link per resource. □

If the inequalities (3.13c) were replaced with equalities, the resulting problem would be a restriction of (3.12)–(3.13), since the set of feasible solutions would be a subset of the original one. This restricted problem would become infeasible when it is impossible to admit all links.

The objective function in (3.12) is a sum of two terms and the second term is scaled with a weight $W \geq 0$. The first term effectively counts the number of links served; hence, it rewards solutions that provide service to many users. The second term represents the total amount of power spent. Due to the minus sign, it penalizes solutions that are power-inefficient. By choosing W, we can trade off between the two conflicting objectives of saving power, and serving many links. In the special case that $W = 0$, we obtain the solution that maximizes the number of served users, subject to the power constraints.

Somehow remarkably, exploiting the special properties of the two terms, we can simultaneously achieve the best of both objectives, by fine-tuning the parameter W. This is possible because the first term is discrete (with step size 1) and the second term is bounded by KP (when all K transmitters use maximum power P). Hence, adapting the ruler analogy argument of [57, p. 2684], we can ensure that the objectives do not overlap when W is chosen such that $1 > WKP$. The interpretation of this choice is that the scheduling objective is prioritized over the power minimization, since the reward for serving one link is higher that the maximum potential power saving. Thus, we have the following result

2 Claim. If W is chosen according to

$$0 < W < \frac{1}{KP} \tag{3.15}$$

then the solution has the following properties: (i) The maximum possible number of links will be served, as if $W = 0$; and (ii) No other solution that serves this set of links can operate with less power.

The proposed formulation of the joint power control and scheduling problem in (3.12)–(3.13) is nonconvex, owing to the binary variables $\{s_k^n\}$. In fact, it can be shown that (3.12)–(3.13) is NP-hard. The computationally intensive components of the problem are the scheduling and admission control. Due to its combinatorial nature, the optimization with respect to the variables $\{s_k^n\}_{k \in \mathcal{K}}^{n \in \mathcal{N}}$ has a worst-case complexity that is exponential in the number KN of binary optimization variables. In what follows, we show that (3.12)–(3.13) admits a MILP problem representation. This is no surprise, cause we have *a priori* designed it so that it does. The key point is that the binary variables were introduced so that constraints (3.13a) are actually linear inequalities. Since the denominator of the fraction is positive, (3.13a) can be equivalently rewritten as

$$\begin{aligned}
&G_{kk}^n p_k^n + M(1 - s_k^n) \geq \gamma_k \sum_{\ell \neq k} G_{\ell k}^n p_\ell^n + \gamma_k \sigma_n^2 \Leftrightarrow \\
&\gamma_k \sum_{\ell \neq k} G_{\ell k}^n p_\ell^n - G_{kk}^n p_k^n + M s_k^n \leq M - \gamma_k \sigma_n^2 \Leftrightarrow \\
&\sum_{\ell=1}^{K} A_{\ell k}^n p_\ell^n + M s_k^n \leq B_k,
\end{aligned} \tag{3.16}$$

where we have defined $B_k \triangleq M - \gamma_k \sigma_n^2$ and

$$A_{\ell k}^n \triangleq \begin{cases} -G_{kk}^n & \text{if } \ell = k, \\ \gamma_k G_{\ell k}^n & \text{if } \ell \neq k. \end{cases} \tag{3.17}$$

Replacing (3.13a) with (3.16), the optimization (3.12)–(3.13) is equivalently recast as the following MILP problem in standard form

$$\max_{\left\{ \substack{p_k^n \in [0,P] \\ s_k^n \in \{0,1\}} \right\}_{k \in \mathcal{K}}^{n \in \mathcal{N}}} \quad \sum_{k=1}^{K} \sum_{n=1}^{N} s_k^n - W \sum_{k=1}^{K} \sum_{n=1}^{N} p_k^n \tag{3.18}$$

subject to

$$\sum_{\ell=1}^{K} A_{\ell k}^n p_\ell^n + M s_k^n \leq B_k \qquad \forall k \in \mathcal{K},\ \forall n \in \mathcal{N}, \tag{3.19a}$$

$$p_k^n - P s_k^n \leq 0 \qquad \forall k \in \mathcal{K},\ \forall n \in \mathcal{N}, \tag{3.19b}$$

$$\sum_{n=1}^{N} s_k^n \leq 1 \qquad \forall k \in \mathcal{K}. \tag{3.19c}$$

MILP are LP problems that comprise both integers and continuous variables. Even though they are in general NP-hard, there are many algorithms and (free or commercial) software packages (e.g. GLPK or IBM ILOG CPLEX, respectively), that find the global solution, very efficiently in the vast majority of the instances. The proposed framework can be directly used in practice for small-scale problems, where "small-scale" is of course relative to the computational capacity of the resource allocation unit. In addition, as the MILP reformulation allows us to solve the problem exactly and much more efficiently than a brute-force search, it also provides a benchmark for performance evaluation of competing, suboptimal algorithms in offline simulations. This is particularly important for the SAPHYRE simulations that are being performed in WP4.

A basic prerequisite is that there is a central controller that is aware of all the channel gains and requested SINR thresholds. This renders the solution directly applicable in systems that have such a controller. The proposed method can also be used to provide fundamental limits for decentralized solutions, assuming that the signaling overhead is negligible and that the optimal feedback can be fed back under tight latency constraints (which is NP-hard in practice).

3.2.3 Joint Scheduling and QoS Beamforming

In this section, we consider a generic wireless network consisting of multiple base stations (BSs) and multiple mobile users (MSs) within each cell. The interest of SAPHYRE is on topologies arising when the networks of two (or more) operators are overlayed. Frequency reuse one is assumed, i.e. all the BSs share non-orthogonally the same frequency resources. The BSs are equipped with multiple antennas and the MSs with single antenna. In every cell, transmit beamforming is employed to spatially multiplex the streams intended to MSs scheduled in the same resource. The assumption is that the BSs coordinate to perform scheduling and beamforming design jointly.

We assume that the network under discussion has L BSs and K MSs in each cell. The lth ($l \in \mathcal{L} \triangleq \{1, \cdots, L\}$) BS is denoted as BS_l and the kth ($k \in \mathcal{K} \triangleq \{1, \cdots, K\}$) MS in the lth cell is denoted as $\mathrm{MS}_{l,k}$. There are N resources available and they are shared by all BSs. We assume the total number of MSs in the network exceeds the number of resources, i.e., $LK > N$. Each BS has N_t transmit antennas and each MS has a single receive antenna. Hence, the downlink can be modeled as a MISO interference channel. We denote the channel vector between BS_j and $\mathrm{MS}_{l,k}$ in the nth ($n \in \mathcal{N} \triangleq \{1, \cdots, N\}$) resource as $\mathbf{h}^n_{j,l,k} \in \mathbb{C}^{Nt\times 1}$, and the transmit beamforming vector for $\mathrm{MS}_{l,k}$ in the nth resource as $\mathbf{w}^n_{l,k} \in \mathbb{C}^{Nt\times 1}$. The SINR for $\mathrm{MS}_{l,k}$ in the nth resource can then be expressed as

$$\mathrm{SINR}^n_{l,k} = \frac{|(\mathbf{w}^n_{l,k})^H \mathbf{h}^n_{l,l,k}|^2}{\sum\limits_{b\neq k} |(\mathbf{w}^n_{l,b})^H \mathbf{h}^n_{l,l,k}|^2 + \sum\limits_{j\neq l}\sum\limits_{b} |(\mathbf{w}^n_{j,b})^H \mathbf{h}^n_{j,l,k}|^2 + (\sigma^n_{l,k})^2} \tag{3.20}$$

where $(\sigma^n_{l,k})^2$ is the variance of background noise. In order to achieve an acceptable QoS for each served MS, $\mathrm{SINR}^n_{l,k}$ is required to be higher than a threshold $\gamma_{l,k}$, i.e., $\mathrm{SINR}^n_{l,k} \geq \gamma_{l,k}$.

We assume that each MS can only be served in one resource. Therefore, if $\mathrm{MS}_{l,k}$ is served in the nth resource, we have a none-zero beamforming vector for $\mathrm{MS}_{l,k}$ in the nth resource, i.e., $\mathbf{w}^n_{l,k} \neq \mathbf{0}$, while in all of the other resources we have $\mathbf{w}^m_{l,k} = \mathbf{0}$ for $m \in \mathcal{N}$, $m \neq n$. Moreover, we assume the aggregate transmit power at each BS_l is upper-bounded by a power P, i.e., $\sum_{k=1}^{K}\sum_{n=1}^{N} \|\mathbf{w}^n_{l,k}\|^2 \leq P$.

With the above set up, our goal is to improve the overall performance of the network by coordinated resource allocation, user selection, and beamforming design of multiple BSs and multiple MSs within each cell. We use the total number of served MSs and the total amount of transmit power as performance metrics. We have two objectives: the first objective is to serve as many MSs as possible by using the available resources while satisfying both the SINR constraints of the MSs and the power constraints of the BSs; the second objective is to find the optimal beamforming vectors for those served MSs that minimize the total transmit power. This optimization problem is a combinatorial problem. The two objectives can be achieved simultaneously in a joint formulation by using auxiliary binary schedule variables $s^n_{l,k} \in \{0, 1\}$. If $s^n_{l,k} = 1$, $\mathrm{MS}_{l,k}$ is served in the nth resource, and $s^n_{l,k} = 0$

otherwise. The combinatorial problem can be expressed in a joint formulation as

$$\max_{\left\{\substack{\mathbf{w}_{l,k}^n \in \mathbb{C}^{Nt\times 1} \\ s_{l,k}^n \in \{0,1\}}\right\}} \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{N} s_{l,k}^n - W \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{N} \|\mathbf{w}_{l,k}^n\|^2 \tag{3.21}$$

subject to

$$\frac{|(\mathbf{w}_{l,k}^n)^H \mathbf{h}_{l,l,k}^n|^2 + M_{l,k}^n(1-s_{l,k}^n)}{\sum\limits_{b\neq k} |(\mathbf{w}_{l,b}^n)^H \mathbf{h}_{l,l,k}^n|^2 + \sum\limits_{j\neq l}\sum\limits_{b} |(\mathbf{w}_{j,b}^n)^H \mathbf{h}_{j,l,k}^n|^2 + (\sigma_{l,k}^n)^2} \geq \gamma_{l,k} \quad \forall l \in \mathcal{L}, \forall k \in \mathcal{K}, \forall n \in \mathcal{N}, \tag{3.22a}$$

$$\sum_{k=1}^{K}\sum_{n=1}^{N} \|\mathbf{w}_{l,k}^n\|^2 \leq P \qquad \forall l \in \mathcal{L}, \tag{3.22b}$$

$$\sum_{n=1}^{N} s_{l,k}^n \leq 1 \qquad \forall l \in \mathcal{L}, \forall k \in \mathcal{K}. \tag{3.22c}$$

The objective function in (3.21) is a sum of two terms and the second term is scaled with a weight $W \geq 0$. The first term effectively counts the number of MSs served; hence, it rewards solutions that provide service to many users. The second term represents the total amount of power spent. Due to the minus sign, it penalizes solutions that are power-inefficient. By choosing W, we can trade off between the two conflicting objectives of serving many MSs and saving power. In the special case that $W = 0$, we obtain the solution that maximizes the number of served users, subject to the power constraints. Because the first term is discrete with step size -1 and the second term is bounded by LP, we can ensure that the two terms do not overlap when W is chosen as $0 < W < 1/(LP)$. This choice of W implies that the maximum possible number of MSs will be served and no other solution that serves the same number of MSs can operate with less power.

Equation (3.22a) defines LKN SINR constraints, one per MS and resource. When $\text{MS}_{l,k}$ is served in the nth resource, i.e., $s_{l,k}^n = 1$, (3.22a) is a standard SINR constraint. On the other hand when $s_{l,k}^n = 0$, we have $\mathbf{w}_{l,k}^n = \mathbf{0}$ at optimum due to the power minimization in the second term of (3.21), provided that $M_{l,k}^n$ is large enough to satisfy the inequality for all feasible values of $\mathbf{w}_{l,k}^n$. We need to choose the parameter $M_{l,k}^n$ so that all LKN inequalities (3.22a) are fullfilled when $s_{l,k}^n = 0$, irrespective of the values for $\mathbf{w}_{l,k}^n$. We choose $M_{l,k}^n$ as

$$M_{l,k}^n \geq \gamma_{l,k} P \sum_{l=1}^{L} \|\mathbf{h}_{l,l,k}^n\|^2 + \gamma_{l,k}(\sigma_{l,k}^n)^2. \tag{3.23}$$

Equation (3.22c) makes sure that each MS is served in at most one resource. With this constraint, admission control functionality is included in the joint optimization problem (3.21)–(3.22). This means that when there are not enough resources, the system may deny service to some MSs in order to ensure service to the remaining ones. For a denied $\mathrm{MS}_{l,k}$, we get $\{s_{l,k}^n = 0\}_{n\in\mathcal{N}}$.

It is well known [61] that an arbitrary phase rotation can be added to the beamforming vectors $\mathbf{w}_{l,k}^n$ without affecting the objective (3.21) and constraints (3.22a) and (3.22b). Thus, the term $(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n$ in the numerator of (3.22a) can be constrained to be real and nonnegative without the loss of generality. We define a vector $\mathbf{u} \in \mathbb{C}^{LK\times 1}$ as $\mathbf{u} \triangleq [u_{1,1}\ldots u_{1,K},\ldots,u_{L,1}\ldots u_{L,K}]^T$, where each element is $u_{j,b} = (\mathbf{w}_{j,b}^n)^H\mathbf{h}_{j,l,k}^n$ with $j \in \mathcal{L}, b \in \mathcal{K}$. By using $\mathbf{u}$, we can write the SINR constraints (3.22a) in second order cone formulation as,

$$\left\|[\mathbf{u}^T,\ \sigma_{l,k}^n]\right\| \le \sqrt{(1+1/\gamma_{l,k})}(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n + \sqrt{M_{l,k}^n/\gamma_{l,k}}(1-s_{l,k}^n), \tag{3.24a}$$

$$(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n \ge 0, \tag{3.24b}$$

$$\mathrm{Im}\left\{(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n\right\} = 0. \tag{3.24c}$$

Next, we define new vectors $\mathbf{v} \in \mathbb{C}^{N_tLKN\times 1}$ and $\mathbf{v}_l \in \mathbb{C}^{N_tKN\times 1}$ which, respectively, stack all the LKN beamforming vectors of the system, and the KN beamforming vectors of BS_l. They are expressed as $\mathbf{v} = \left[\mathbf{v}_1^T,\ldots,\mathbf{v}_L^T\right]^T$ and $\mathbf{v}_l = \left[\mathbf{w}_{l,1}^T,\ldots,\mathbf{w}_{l,K}^T\right]^T$, where $\mathbf{w}_{l,k}$ stacks the beamforming vectors of $\mathrm{MS}_{l,k}$ in all the N resources, i.e., $\mathbf{w}_{l,k} = \left[(\mathbf{w}_{l,k}^1)^T\cdots(\mathbf{w}_{l,k}^N)^T\right]^T$. By using $\mathbf{v}$ and $\mathbf{v}_l$, the transmit power of the system and the transmit power of BS_l can be represented as $\|\mathbf{v}\|^2$ and $\|\mathbf{v}_l\|^2$ respectively.

Then by changing the maximization in (3.21) into a minimization we can reformulate (3.21)–(3.22) into a mixed integer second order cone program (MI-SOCP) problem as,

$$\min_{\left\{\substack{\mathbf{w}_{l,k}^n\in\mathbb{C}^{Nt\times 1}\\ s_{l,k}^n\in\{0,1\}}\right\}} \quad \sqrt{W}\|\mathbf{v}\| - \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{N} s_{l,k}^n \tag{3.25}$$

subject to

$$\left\|[\mathbf{u}^T,\ \sigma_{l,k}^n]\right\| \le \sqrt{(1+1/\gamma_{l,k})}(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n + \sqrt{M_{l,k}^n/\gamma_{l,k}}(1-s_{l,k}^n) \quad \forall l\in\mathcal{L},\ \forall k\in\mathcal{K},\ \forall n\in\mathcal{N}, \tag{3.26a}$$

$$(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n \ge 0 \quad \forall l\in\mathcal{L},\ \forall k\in\mathcal{K},\ \forall n\in\mathcal{N}, \tag{3.26b}$$

$$\mathrm{Im}\left\{(\mathbf{w}_{l,k}^n)^H\mathbf{h}_{l,l,k}^n\right\} = 0 \quad \forall l\in\mathcal{L},\ \forall k\in\mathcal{K},\ \forall n\in\mathcal{N}, \tag{3.26c}$$

$$\|\mathbf{v}_l\| \le \sqrt{P} \quad \forall l\in\mathcal{L}, \tag{3.26d}$$

$$\sum_{n=1}^{N} s_{l,k}^n \le 1 \quad \forall l\in\mathcal{L},\ \forall k\in\mathcal{K}. \tag{3.26e}$$

This MI-SOCP problem (3.25)–(3.26) can be solved by using branch-and-bound based solvers such as IBM ILOG CPLEX.

3.3 Game-Theoretic Strategies

Radio resource allocation on wireless channels is known to involve several design choices from the algorithmic point of view. One key issue involves the definition of the main allocation objective: since the channel quality is perceived differently by different users, it may be thought of allocating most of the resources to the users who see the best channel conditions, which however lead to unfairness from the users' perspective. Alternatively, some form of fairness may be sought, which imposes sometimes not to allocate the users with the best channel conditions (and thereby potentially decreasing the allocation efficiency).

Generally speaking, this trade-off is often solved with a priori choices. However, these solutions are not easy to set up, nor they can be dynamically adapted. Conversely, game theory can be used to solve this problem in a more efficient manner, by putting the burden of the trade-off resolution on some preference/utility definitions, which are much easier to define from the operator's standpoint, and also enable dynamic adaptation of the allocation.

A sample challenge of this kind arises in multiple access schemes using Orthogonal Frequency Division Multiple Access (OFDMA), such as the downlink of Long Term Evolution (LTE) systems. Assume that several users need to be served by allocating packets belonging to their requested flows on the OFDMA frame. Since their perceived channel quality is different (and, additionally, varies also according to the subcarrier of choice) the problem becomes a complex task, in view of the high number of possible allocations among which to choose. Additionally, the aforementioned trade-off between maximizing the throughput and achieving fairness (at least in a long term perspective) further complicates the problem.

To study the problem through a game-theoretic approach, we follow the model proposed in [62]. Here, a modular representation is introduced, where the radio resource management is split between two functional entities, i.e. a credit-based scheduler and the actual resource allocator. The scheduler determines which packets, taken from different flows, are candidates to be served in the next allocation round. The resource allocator associates the packets with groups of OFDMA subcarriers, also accessed in a Time Division (TD) fashion, so that the resources to allocate are time/frequency resource blocks. In this choice, the resource allocator exploits a degree of freedom, represented by the number of packets selected by the scheduler (larger than the number of slots).

The resulting allocation can be regulated according to a trade-off between two contrasting objectives, i.e. that of throughput maximization, which is achieved by selecting the packets only according to a channel quality rationale, and fairness among the flows, which requires to pursue equity among the achieved rates. Indeed, this trade-off is reflected by the number of packets selected by the scheduler: when it is minimum, i.e. only the packets that fit the OFDMA frame are selected, all packets are mandatorily allocated and the resource allocator has no choice. Here the allocation is only determined by the credit-based scheduler, which ensures fairness (the users with higher credits are allocated). Conversely, if the number of selected packets is high, the resource allocator can restrict the selection to the packets of the users with the best quality, entirely neglecting any fairness among flows.

Within this framework, we propose a game-theoretic approach [63] to resolve this trade-off between contrasting objectives. In fact, the idea is to establish two virtual players, one representing the scheduler needs, i.e. to ensure fairness to the users, and the other reproducing the resource allocation perspective, i.e. to select those users which are experiencing better channel quality. A coordination game is established between these two players, which leads to the derivation of a simple yet effective algorithm to identify a Pareto-efficient trade-off point.

3.3.1 System Model

The downlink transmission on a LTE system, using OFDMA multiplexing, requires coordination of multiple flows directed to the users to be coordinated, so that a number of packets are selected for possible transmission from each flow. In the following, this operation will be referred to as *scheduling*. However, actual transmission also requires to match the selected packets to a given resource block in a channel-aware fashion. Thus, it is necessary to eventually select which resources to utilize for the selected packets. Such an operation will be referred to as *resource allocation*.

In the LTE standards, the design of policies for resource management is intentionally left open to allow developers to implement their own strategy of choice. However, in the following we adopt a two-fold model where scheduling and resource allocation are managed by two different modules: a scheduler, operating at the transport layer (thereby possibly distinguishing among different kinds of traffic) and a resource allocator, which actually implements the Medium Access Control (MAC) sublayer. The scheduler determines which packets must be passed to the allocator and their order according to an internal scheduling policy. The allocator selects for transmission a subset of them with the aim of maximizing the advantages of multiuser diversity. In this case only a loose cross-layer is introduced, guaranteeing a certain modularity between scheduler and radio resource allocator (RRA).

In particular, we call L the number of resource blocks that the resource allocator is entitled to assign. This is subject to a constraint $L \leq L_{\max}$, where $L_{\max}$ is a maximum value which corresponds to assigning every resource block. For simplicity, we consider that, to limit the interference caused to the neighboring cells, L is set to a fixed value which is less than or equal to $L_{\max}$. The value assigned to L is communicated to the scheduler by the resource allocator. Actually, this represents a form of cross-layer interaction among the modules, which is intentionally kept to a minimum level, thereby promoting modularity and tunability of the approach.

Upon knowing L, the scheduler determines a number D of packets to send to the resource allocator, where in general $D \geq L$. The exact choice of D influences the entire allocation. As a matter of fact, if $D = L$, the resource allocator has no degree of freedom as to which packets to allocate (while, obviously, it must allocate the packets to the best channels as perceived by the users). By increasing D, the

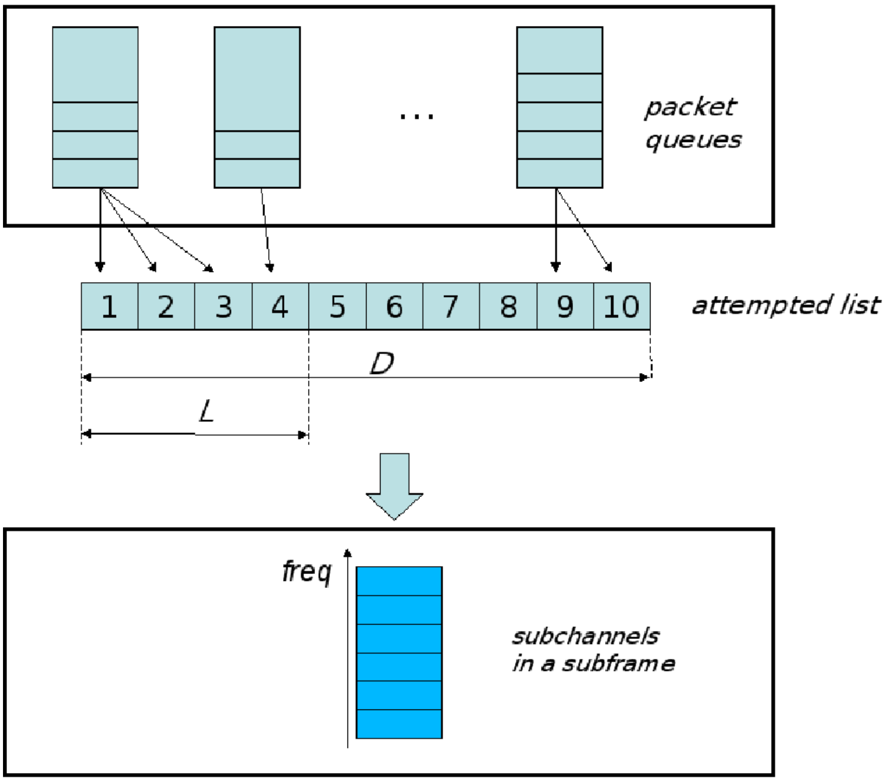


Figure 3.2: System model

resource allocator can achieve a higher throughput by selecting only L packets out of D, according to a channel-aware policy, although at the price of a possibly decreased fairness. A graphical representation of the system is given in Figure 3.2.

3.3.2 The Proposed Game-Theoretic Approach

The choice of D determines a trade-off between the possible objectives of throughput and fairness. We now present a game-theoretic approach to set D; our proposed methodology enables a dynamic setup of D without any need for a preliminary evaluation, e.g. where D is set to some arbitrary value, of the possible equilibria of the system, nor it is required to re-compute the system equilibria if the network and channel conditions change. Instead, the choice of D is directly derived from the definitions of the contrasting utilities between which a trade-off is sought (specifically, throughput and fairness). Together with the separation of the resource management process into two functional entities (scheduler and RRA), this is key to achieve a computationally efficient online allocation strategy.

In our formulation, the scheduler (player 1) and the RRA (player 2) are represented as players of a game whose aim is the decision of the value for D. Both players make a proposal s_j, with $j = 1, 2$, respectively. The game is inspired by *coordination games*, where two players get non-zero payoff only if they converge on

Resource Allocator

		L	$L{+}1$	$\dots$	$2L$
	L	$1, T_{\text{min}}$	$0,0$	$0,0$	$0,0$
Scheduler	$L+1$	$0,0$	$\dots$	$0,0$	$0,0$
	$\dots$	$0,0$	$0,0$	$\dots$	$0,0$
	$2L$	$0,0$	$0,0$	$0,0$	$\frac{1}{2}, T_{\text{max}}$

Figure 3.3: Bi-matrix representation of the game.

a common agreement. In our case, if proposals s_1 and s_2 coincide, D is selected as their common value. However, the choice of s_1 and s_2 is also done according to the utility of the proposer, i.e. the fairness for the scheduler and the throughput for the RRA, respectively.

In the following, we introduce some assumptions for the sake of simplicity in the exposition. We consider a network scenario with only two users; this is not to be confused with the two "virtual" players of the game, i.e. the scheduler and the resource allocator. Besides, this assumption is just made to ease the presentation, but can be relaxed quite naturally to scenarios with $N > 2$ users. We model the system as a static game in normal form, as follows:

- The players are the scheduler and the RRA.
- Their action spaces are the set of values of D that can be proposed, i.e. $S_1 = S_2 = \{L, L+1, \dots, 2L\}$.
- Both payoffs are 0 if the proposals s_1 and s_2 do not coincide, i.e. there is no agreement on the value of D. This assumption is drawn from the more general theory about coordination games.
- When $s_1 = s_2$, the payoffs are assigned to fairness $F(s_1, s_2)$ for the scheduler, measured using Jain's index, and the throughput $T(s_1, s_2)$ for the RRA.

The last point is arbitrary, as other definitions can be used; the important requirement is that $F(s, s)$ and $T(s, s)$ are decreasing and increasing in s, respectively. The resulting bi-matrix representation of the game is given in Figure 3.3. The fairness is a decreasing function of D: its maximum value is 1 while the minimum is 1/2, i.e. $1/N$ where N is the number of flows. On the other hand, the throughput is an increasing function of D varying in the range $[T_{\text{min}}, T_{\text{max}}]$, where T_{min} is achieved when no degree of freedom is given to the allocator, while T_{max} is obtained when the RRA has enough freedom to allocate only the best L resources. Both maximum throughput and minimum fairness are reached for $D = 2L$, under the assumption that there are always at least L packets available for selection by the scheduler from each queue. All the strategies along the diagonal are Pareto efficient Nash equilibria. This means that improving the payoff of one player results in worsening the other's outcome. Thus, once the value of L is fixed, there is no unique evolution of the game and, in any case, a trade-off is encountered.

To determine a trade-off point, we propose an algorithm which tries to automatically estimate an efficient value of D for each frame. The value is chosen considering the entire history of the game, thus the model we propose is a *repeated game with perfect information*. The aim is to reach an acceptable level for both payoffs after a number of repetitions. Note that this proposed algorithm is just an example and can be replaced by other analogous procedures.

1) Both scheduler and RRA randomly pick a value for D.

2) If the choices coincide, D is set and the game ends, otherwise a bargaining phase goes on until a common point is chosen. Every time the players disagree, both get zero payoff.

3) The goal of each round of the loop is moving towards the diagonal of the bimatrix step-by-step. Each player decides whether or not to change its previous proposal based on its level of satisfaction, i.e. the ratio between the value actually achieved and the maximum achievable. The higher the satisfaction, the higher the probability that a player changes its choice with a value more convenient for the other. If S_D and RRA_D are the proposals for D made by the scheduler and the allocator, respectively, and S_s and RRA_s the respective levels of satisfaction when the game is played, we select the changes as follows.

– If $S\_D > RRA\_D$, we are in the lower triangle of the matrix. We can move towards the diagonal by going up (decrement of S_D), or right (increment of RRA_D), or in both directions. For both players, these options lead to higher values in their own utility function to the detriment of the other's, thus the willingness to change should be a decreasing function of the respective satisfaction level. Thus, we select

$$Prob\{S\_D \; up\} = 1 - S\_s \tag{3.27}$$

$$Prob\{RRA\_D \; right\} = 1 - RRA\_s \tag{3.28}$$

– If $S\_D < RRA\_D$, we are in the upper triangle of the matrix. The diagonal can be reached by going down (S_D increment), or left (RRA_D decrement), or in both directions. The situation is now reversed, as a deviation in its own action implies a reduction in the payoff of each player in favor of the other's. Therefore, the probability of moving must be an increasing function of the respective satisfaction, which is obtained for example by choosing

$$Prob\{S\_D \; down\} = S\_s \tag{3.29}$$

$$Prob\{RRA\_D \; left\} = RRA\_s \tag{3.30}$$

In this manner, we define an algorithm whose goal is to lead the choice of D towards an intermediate value which offers both good throughput and satisfactory fairness.

3.3.3 Applications of the Proposed Approach in a Multi-Agent Context

The algorithm proposed above can be shown to reach a Pareto efficient point, which trades throughput for fairness in an efficient and tunable manner. However, this should not be seen just as a way to set the equilibrium between contrasting needs. In fact, a direct extension may be identified to cases where the multiple players of the game are not just *virtual* agents representing different layers of the same entity, e.g. the radio resource management procedure at one base station. Rather, the game may be extended to a wider population of actors, where still coordination is sought but among different classes of actors.

There are at least two such extensions which are relevant to the scope of the SAPHYRE project. A first extension of this game-theoretic setup involves the interaction between multiple base stations, possibly belonging to different operators (this may also partially applied to the case where the operator is the same, but the exchange of management policies among the base stations is made difficult by some externality).

This case, which can be analyzed in a practical scenario similar to those studied by [64] can be framed in the context of *games with incomplete information*. A multitude of games may be used to represent the different base stations, each one of them using two virtual players to represent the contrasting needs of throughput and fairness. In other words, the game discussed above, as well as some specific procedure to solve it, is played several times at the same time. Under the assumption of perfectly rational players, it may still be assumed that a Pareto efficient equilibrium point is sought by all base stations. However, due to interference caused by neighboring cells, a repeated version of this game may also be considered in order to reach a further equilibrium among the games played locally.

A further extension may involve the management of contrasting objectives among different operators. In this case, the game agents are not only virtual players which are assumed to blindly pursue the task of finding an efficient trade-off between throughput and fairness (or any other objective). Rather, the operators also try to drive the whole allocation of the system toward a favorable allocation for them, which is possibly reflected by the virtual players at one base station trying to influence the outcomes of other games. The extension of such games, which involves further utility modeling and possibly extension to Bayesian games, is a possible direction for further study.

3.4 Utility Model for WNC-Based Sharing

3.4.1 WNC-Based Sharing

Basic Concept

Wireless Network Coding (WNC) (a.k.a. Physical Layer Network Coding (PLNC)) [65], [66], [67], [68], [69] is a novel paradigm of the *network structure aware* modulation and coding. It follows ideas similar to the Network Coding (NC) [70]. Unlike to NC, the WNC operates directly in the *signal space* domain, i.e. directly with the continuous signal waveform representation. Major distinguishing features of WNC are: (1) the fact that it actively uses knowledge of the network topology and (2) the fact that the information is not routed through a single network path but it is rather "flooded" over the whole network and uses all possible paths between the source and the final destination. All this being performed at PHY layer.

The above stated paradigm enables the PHY layer to actively utilise all incoming received signals. Apart of the useful signal carrying the desired data, we also define so called Complementary Side-Information (C-SI) signals [67] which carry the information contents that does not directly depend on the desired data but contains the other source data that also influence the useful signal. The C-SI can be viewed as the "friendly interference". The WNC strategy works with three types of the signals: (1) useful signal, e.g. signal with hierarchical data for HDF strategy, (2) signal with C-SI, (3) classical interference. See [67], [68], [69] for details.

From the perspective of the utility model, the WNC PHY layer strategy introduces several new aspects. First of all, we cannot define the utility function purely in terms of the interference levels. The classical interference can however be present. But the most important input entity is the form and amount of the C-SI available at given node. This requires a very different set (cf. with SINR) of the input parameters to be used in the utility function. This topic is not currently covered, to the best of our knowledge, by any existing research works.

We formulated an initial simplified form of the core utility input elements. It can serve for the investigation of the mutual relation between WNC related performance parameters under various strategies and game theory utility functions. One of the desired outputs is the identification (colouring) of the geographical regions according to the optimality under different WNC strategies – PHY relay sharing schemes (HDF (Hierarchical Decode and Forward), JDF (Joint DF), AF (Amplify and Forward), no relay sharing). This optimality map and corresponding relay operation is likely to be different for 2 different operators at one geographical region and therefore a conflict on a required relay operation must be solved by the game theory.

The current initial stages of the work concentrate on identifying proper quantitative description and elements suitably reflecting the strategy, the topology and the core performance parameters.

Relevance to the SAPHYRE Project Goals

The SAPHYRE project goals, the spectrum and the infrastructure sharing, are directly addressed by the WNC technique on its own. The WNC technique introduces shared relays and the relevant coding, modulation and processing technique required to achieve the sharing gain. The PHY resource sharing management layer supported by the a proper utility metric and corresponding resource allocation strategies, e.g. the game theory, can provide an *additional* sharing gain on top of WNC itself.

3.4.2 Core Utility Model Input Elements

Signal Types

The WNC works with three *types* of received signals. Their definitions follow ($I(\cdot;\cdot)$ denotes the mutual information).

- Hierarchical Information (HI) signal
  - The signal x is HI signal on desired data a iff $I(a;x) > 0$.

  The signal x carries information about desired data.

- Complementary Side-Information (C-SI)
  - Let x be HI signal on desired data a then the signal z is C-SI for complementary data b iff $I(b;x) > 0$ and $I(b;z) > 0$.

  The HI signal of the desired data a is influenced by other data b and C-SI signal provides the information about those b data. Thus, the C-SI is *friendly interference* – helps removing influence of the complementary data b. For example, in 2-Way Relay Channel, the signal forwarded from the relay to destination A is HI; the information from the other source B is C-SI. The C-SI should *not be confused* with Side-Information signal, where y is SI signal if $I(a;y) > 0$, i.e. additional observation of the desired data.

- Interference (IFC) (classical, harmful) signal r w.r.t. $x(a)$
  - It must hold $I(a;r) = 0$, $I(b;r) = 0$ for any other complementary data ($I(x;b) > 0$).

Relaying Strategy

The WNC can use various relaying strategies (HDF, JDF, AF, no sharing). Each relaying strategy is further described by the particular PHY layer technique used by the relay. Particularly, the HDF strategy can use various relay hierarchical codebook classes (minimal, extended, full/joint mapping).

Channel State

The channel state, including its parametrisation, is fully defined by the stochastic input-output description, i.e. the observation likelihood function conditioned by the (hierarchical) data and channel parameters. It can be conveniently equivalently described by the set of the parameters for a given channel, e.g. the SNR in the case of AWGN linear channel.

Figure 3.4: 2-Source Relay Network (2-SRN) with Partial C-SI.

3.4.3 Core Utility Performance Metric

This section defines the utility function model. In the initial simple model, we work with simple utility μ which is the achievable kth user rate[2] $\mu = R_i^{(k)}$ in the ith stage of the network (typically imposed by the half-duplex constraint). The utility μ is generally a function of (1) the signal type which can be HI, C-SI or IFC $T \in \{H, C, I\}$, (2) the relaying strategy S including particular codebook class $\mathcal{C}$, e.g. HDF with Minimal-HXC, (3) the index $i \in \mathbb{N}$ of the network stage, (4) the channel state represented by a proper parameter $\mathbf{\Gamma}$ (typically a vector of SNRs for all signal types and all incoming signals)

$$\mu_i^{(k)} = \mu\left(S_i^{(k)}(\mathcal{C}), \mathbf{\Gamma}_{H,i}^{(k)}, \mathbf{\Gamma}_{C,i}^{(k)}, \mathbf{\Gamma}_{I,i}^{(k)}\right). \tag{3.31}$$

3.4.4 2-SRN HDF Example

The 2-Source Relay Network (2-SRN) is defined in Figure 3.4. The 2-SRN is a two-stage network due to the half-duplex constraint at the relay. We use the HDF strategy. The first stage is the Hierarchical MAC and the second stage is the Hierarchical BC. The utility functions for users $k \in \{A, B\}$

$$R_1^{(k)} = R_{\text{HMAC}(\mathcal{C})}\left(\mathbf{\Gamma}_{H,1}^{(k)}\right), \tag{3.32}$$

$$R_2^{(k)} = R_{\text{HBC}(\mathcal{C})}\left(\mathbf{\Gamma}_{H,2}^{(k)}, \mathbf{\Gamma}_{C,2}^{(k)}\right) \tag{3.33}$$

where $\mathbf{\Gamma}_{H,1}^{(k)} = \gamma_{x,k}$, $\mathbf{\Gamma}_{H,2}^{(k)} = \gamma_{y,k}$, $\mathbf{\Gamma}_{C,2}^{(k)} = \gamma_{z,k}$.

The resulting functions can be generally complicated functions of the channel states. However, for a particular case of 2-SRN using HDF strategy with *minimal*

[2]At later stages, it can be generalised for other utility metric, e.g. sum-rate, service outage probability, etc.

HXC (Hierarchical eXclusive Code) with variety of constellations, we have shown [68] that the following approximation holds with high fidelity

$$R_{\mathrm{HBC}}(\gamma_y, \gamma_z) \approx \frac{R_{\mathrm{HBC}}^{\mathrm{perfCSI}}(\gamma_y) R_{\mathrm{CSI}}(\gamma_z)}{\lg |\mathcal{A}|}. \tag{3.34}$$

This approximation simplifies the evaluation of the utility.

Similar approximations can be found also for the case of HDF strategy with *extended cardinality HXC*[3]. It can be shown that the utility function R_{HBC} can be approximated by

$$R_{\mathrm{HBC}}\left(\gamma_y, \gamma_z\right) \approx R'_{\mathrm{HBC}}\left(\gamma_y, \gamma_z \to -\infty\right) + R''_{\mathrm{HBC}}\left(\gamma_z\right), \tag{3.35}$$

where $R'_{\mathrm{HBC}} \leq \lg\left(\frac{|\mathcal{A}^R|}{|\mathcal{A}|}\right)$. As it is obvious from the examples in Figs. 3.5, 3.6, 3.7, 3.8, the particular form of $R''{}_{\mathrm{HBC}}\left(\gamma_z\right)$ depends on the cardinality of the relay output alphabet, the choice of the eXclusive hierarchical mapper at the relay ($\mathcal{X}\left(\cdot,\cdot\right)$) and even on the choice of source alphabet indexing – compare two examples in Figs. 3.8, 3.9, where a different behaviour is observed for 8PSK source alphabet with two different output symbol mappings.

Apparently, (3.35) is valid also for a minimal cardinality HXC ($R'_{\mathrm{HBC}} = 0$) and full cardinality HXC ($R''_{\mathrm{HBC}} = 0$). Our goal in future project stages is to a derive a general form of approximation of $R''_{\mathrm{HBC}}\left(\gamma_z\right)$ and to find similar approximations for the more complicated networks and relaying strategies.

[3]For the extended cardinality HXC, the relay output alphabet cardinality ($|\mathcal{A}^R|$) is larger than the source alphabet cardinality ($|\mathcal{A}|$), resulting in a lower requirements on the C-SI link quality.

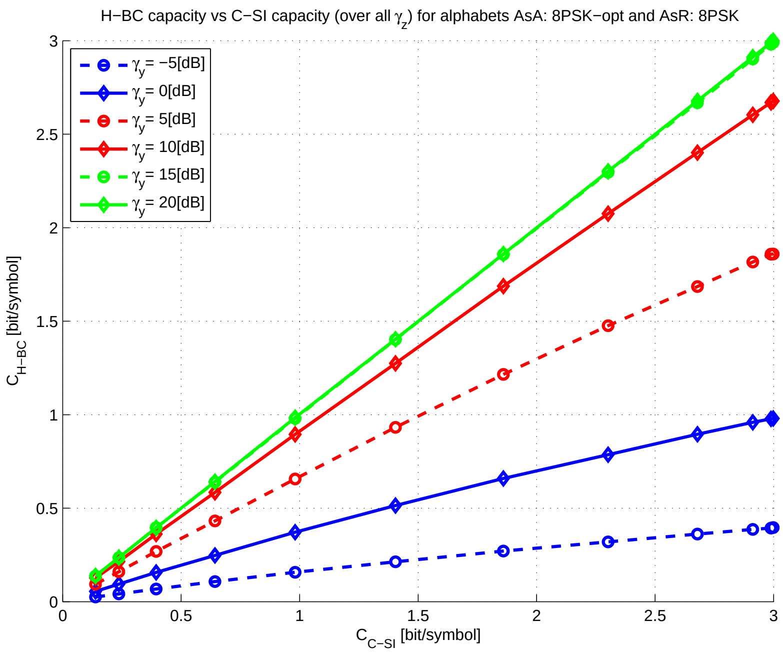


Figure 3.5: 8PSK alphabet constrained capacity of H-BC versus capacity of C-SI link as a parametric plot over all γ_z - minimal HXC.

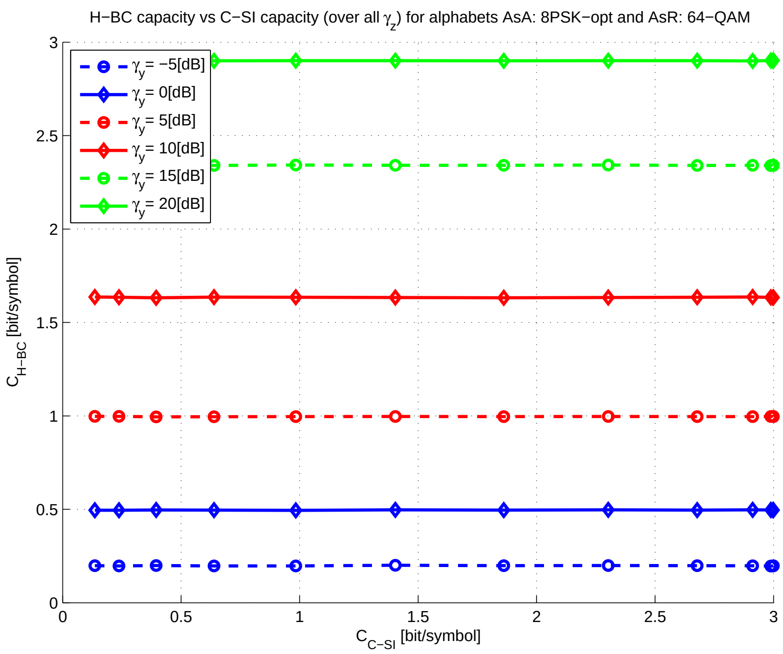


Figure 3.6: 8PSK alphabet constrained capacity of H-BC versus capacity of C-SI link as a parametric plot over all γ_z - full HXC.

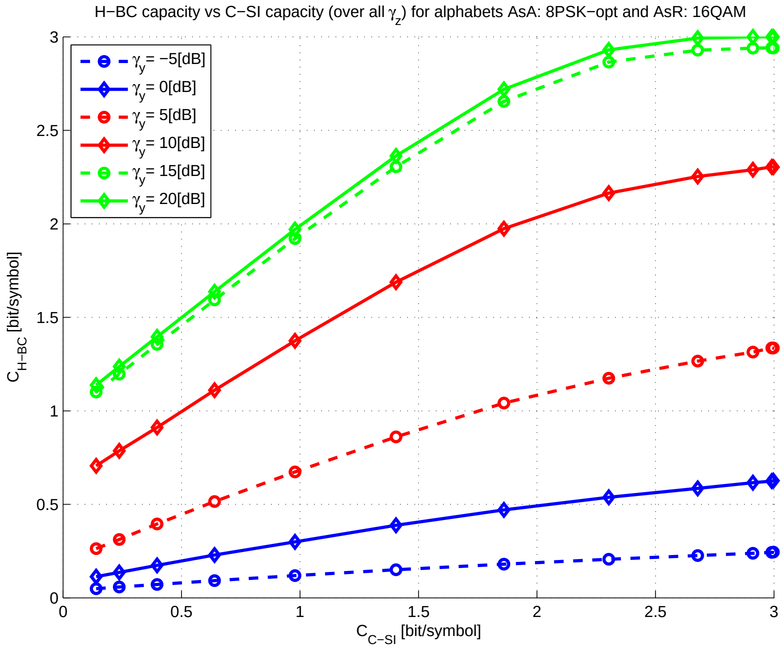


Figure 3.7: 8PSK alphabet constrained capacity of H-BC versus capacity of C-SI link as a parametric plot over all γ_z - extended HXC ($|\mathcal{A}^R|$=16).

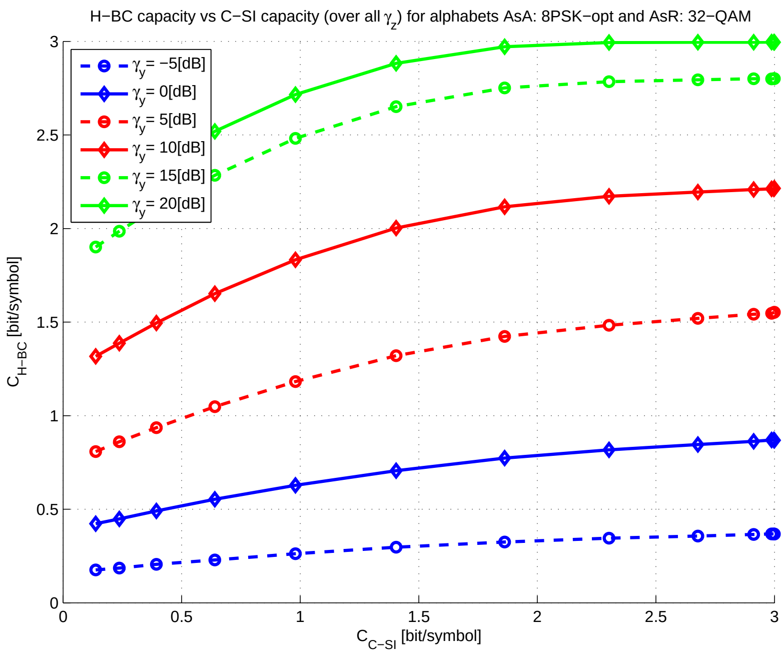


Figure 3.8: 8PSK alphabet constrained capacity of H-BC versus capacity of C-SI link as a parametric plot over all γ_z - extended HXC ($|\mathcal{A}^R|$=32).

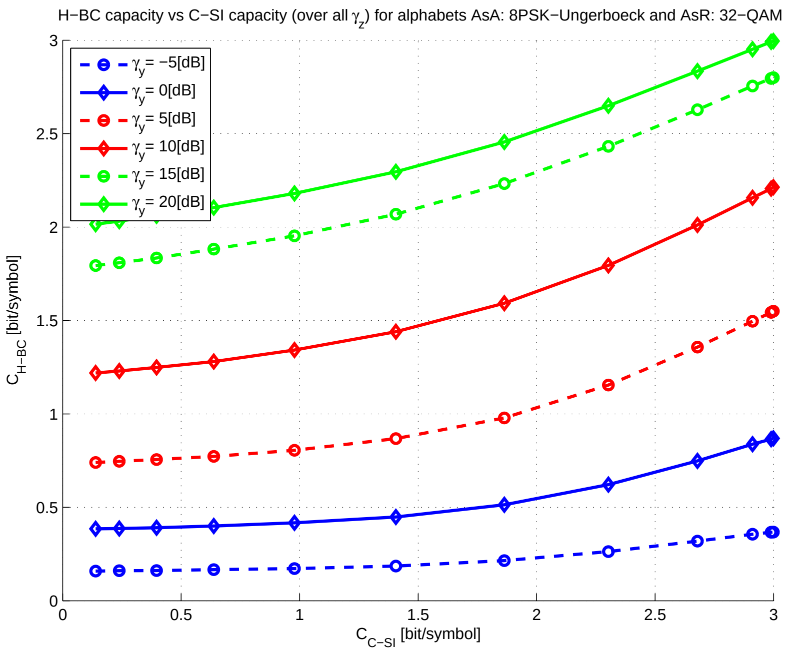


Figure 3.9: 8PSK alphabet (with Ungerboeck mapping to constellation symbols) constrained capacity of H-BC versus capacity of C-SI link as a parametric plot over all γ_z - extended HXC ($\left|\mathcal{A}^R\right|$=32).

4 Concluding Remarks

This report aims to provide a basis for interference and utility models for the SAPHYRE project as a whole. We have reviewed existing models from the literature and discussed new approaches that have already been developed in recent work of the SAPHYRE partners. Part of these results were developed in a cooperation between the partners FhG and LiU. The results are published in [29,30]. Other SAPHYRE-publications have been submitted by FhG [3, 31] and by CFR [63]. A joint SAPHYRE work of TUIL and TUD can be found in [71]. CTU's publications [67], [68], [69] present an alternative approach where some form of the interference in WNC systems can help decoding the data at the final destination. This type of strategy requires different form of the utility metric. Results of [68] provide a simple form of such utility to be used by other partners in related project tasks.

This deliverable has summarised various approaches to interference and utility modelling, involving different areas like power control theory, game theory and information theory. This report provides the basis for sharing strategies developed in other work packages, especially WP4 where the sharing gain will be evaluated by numerical simulations.

Bibliography

[1] "FP7 SAPHYRE Sharing Physical Resources - Mechanisms and Implementations for Wireless Networks. project (FP7-ICT-248001)," 2011. [Online]. Available: http://www.saphyre.eu

[2] D. Gesbert, M. Kountouris, R. W. Heath Jr., C.-B. Chae, and T. Sälzer, "Shifting the MIMO paradigm," *IEEE Signal Processing Magazine*, pp. 36–46, Sep. 2007.

[3] M. Schubert, N. Vucic, and H. Boche, "SIR balancing for strongly connected interference networks – existence and uniqueness of a solution," in *Proc. The Seventh International Symposium on Wireless Communication Systems (ISWCS)*, York, United Kingdom, Sep. 2010.

[4] M. Schubert and H. Boche, *QoS-Based Resource Allocation and Transceiver Optimization*. Foundations and Trends in Communications and Information Theory, 2005/2006, vol. 2, no. 6.

[5] C. Farsakh and J. A. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1497–1506, Nov. 1998.

[6] F. Rashid-Farrokhi, K. J. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1437–1449, Oct. 1998.

[7] M. Bengtsson and B. Ottersten, *Handbook of Antennas in Wireless Communications*. CRC press, Aug. 2001, ch. 18: Optimal and Suboptimal Transmit Beamforming.

[8] M. Schubert and H. Boche, "Solution of the multi-user downlink beamforming problem with individual SINR constraints," *IEEE Trans. on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[9] D. Hammarwall, M. Bengtsson, and B. Ottersten, "On downlink beamforming with indefinite shaping constraints," *IEEE Trans. Signal Proc.*, vol. 54, no. 9, pp. 3566–3580, Sep. 2006.

[10] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Proc.*, vol. 54, no. 1, pp. 161–176, 2006.

[11] P. Kumar, R. Yates, and J. Holtzman, “Power control based on bit error (BER) measurements,” in *Proc. IEEE Military Communications Conference MILCOM 95*, McLean, VA, Nov. 1995, pp. 617–620.

[12] S. Ulukus and R. Yates, “Adaptive power control and MMSE interference suppression,” *ACM Wireless Networks*, vol. 4, no. 6, pp. 489–496, 1998.

[13] S. Hanly, “An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity,” *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1332–1340, Sep. 1995.

[14] R. Yates and H. Ching-Yao, “Integrated power control and base station assignment,” *IEEE Trans. on Vehicular Technology*, vol. 44, no. 3, pp. 638–644, Aug. 1995.

[15] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1341–1348, Sep. 1995.

[16] H. Boche and M. Schubert, “A unifying approach to interference modeling for wireless networks,” *IEEE Trans. Signal Proc.*, vol. 58, no. 6, pp. 3282–8297, Jun. 2010.

[17] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays.* Wiley, New York, 1980.

[18] H. Boche and M. Schubert, “The structure of general interference functions and applications,” *IEEE Trans. Inform. Theory*, vol. 54, no. 11, pp. 4980–4990, Nov. 2008.

[19] ——, “Concave and convex interference functions – general characterizations and applications,” *IEEE Trans. Signal Proc.*, vol. 56, no. 10, pp. 4951–4965, Oct. 2008.

[20] ——, “A calculus for log-convex interference functions,” *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5469–5490, Dec. 2008.

[21] J. Zander, “Performance of optimum transmitter power control in cellular radio systems,” *IEEE Trans. on Vehicular Technology*, vol. 41, no. 1, pp. 57–62, Feb. 1992.

[22] D. Gerlach and A. Paulraj, “Base station transmitting antenna arrays for multipath environments,” *Signal Processing (Elsevier Science)*, vol. 54, pp. 59–73, 1996.

[23] H. Boche and M. Schubert, “Perron-root minimization for interference-coupled systems with adaptive receive strategies,” *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3173–3164, Oct. 2009.

[24] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of Resource Allocation in Wireless Networks: Theory and Algorithms*, second expanded edition ed., ser. Foundations in Signal Processing, Communications and Networking. Springer, 2008, vol. 3.

[25] C. Huang and R. Yates, "Rate of convergence for minimum power assignment algorithms in cellular radio systems," *Baltzer/ACM Wireless Networks*, vol. 4, pp. 223–231, 1998.

[26] K. K. Leung, C. W. Sung, W. S. Wong, and T. Lok, "Convergence theorem for a general class of power-control algorithms," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1566–1574, Sep. 2004.

[27] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust downlink power control in wireless cellular systems," *EURASIP Journal on Wireless Communications and Networking*, no. 2, pp. 261–272, 2004.

[28] M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization," *IEEE Selec. Areas in Commun.*, vol. 25, no. 7, pp. 1390–1401, Sep. 2007.

[29] N. Vucic, S. Shi, and M. Schubert, "DC programming approach for resource allocation in wireless networks," in *Proc. Rawnet*, Jun. 2010.

[30] K. Eriksson, S. Shi, N. Vucic, M. Schubert, and E. Larsson, "Globally optimal resource allocation for achieving maximum weighted sum rate," in *Proc. Globecom*, Miami, USA, Dec. 2010.

[31] N. Vucic and M. Schubert, "Fixed point iteration for max-min sir balancing with general interference functions," in *Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, Prague, Czech Republic, May 2011.

[32] H. Boche and M. Schubert, "A superlinearly and globally convergent algorithm for power control and resource allocation with general interference functions," *IEEE/ACM Trans. on Networking*, vol. 16, no. 2, pp. 383–395, Apr. 2008.

[33] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Proc.*, vol. 52, no. 4, pp. 1037–1052, Apr. 2004.

[34] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1426–1438, Aug. 2006.

[35] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM J. Optim.*, vol. 11, 2000.

[36] N. T. H. Phuong and H. Tuy, "A unified monotonic approach to generalized linear fractional programming," *Journal of Global Optimization*, vol. 26, 2003.

[37] L. P. Qian, Y. J. Zhang, and J. Huang, "Mapel: Achieving global optimality for a non-convex wireless power control problem," *IEEE Transactions on Wireless Communications*, vol. 8, Mar. 2009.

[38] E. A. Jorswieck and E. G. Larsson, "Monotonic optimization framework for the two-user MISO interference channel," *IEEE Transactions on Communications*, vol. 58, Jul. 2010.

[39] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization based beamforming," *Signal Processing Magazine*, May 2010.

[40] N. Lee, H. J. Yang, and J. Chun, "Achievable sum-rate maximizing af relay beamforming scheme in two-way relay channels," in *Proc. IEEE Int. Conf. on Communications (ICC 2008)*, Beijing, China, May 2008.

[41] G. H. Golub and C. F. V. Loan, *Matrix Computations.* The Johns Hopkins University Press, 1996.

[42] C. Wang, H. Chen, Q. Yin, A. Feng, and A. F. Molisch, "Multi-user two-way relay networks with distributed beamforming," *IEEE Transactions on Wireless Communications*, 2011, accepted for publication.

[43] H. J. M. Peters, *Axiomatic Bargaining Game Theory.* Kluwer Academic Publishers, Dordrecht, 1992.

[44] L. Zhou, "The Nash bargaining theory with non-convex problems," *Econometrica, Econometric Soc.*, vol. 3, no. 65, pp. 681–686, May 1997.

[45] J. E. Suris, L. A. DaSilva, Z. Han, and A. B. MacKenzie, "Cooperative game theory for distributed spectrum sharing," in *Proc. IEEE Int. Conf. on Comm. (ICC), Glasgow, Scotland*, Jun. 2007.

[46] C. W. Sung, "Log-convexity property of the feasible SIR region in power-controlled cellular systems," *IEEE Communications Letters*, vol. 6, no. 6, pp. 248–249, Jun. 2002.

[47] D. Catrein, L. Imhof, and R. Mathar, "Power control, capacity, and duality of up- and downlink in cellular CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1777–1785, 2004.

[48] H. Boche and S. Stanczak, "Log-convexity of the minimum total power in CDMA systems with certain quality-of-service guaranteed," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 374–381, Jan. 2005.

[49] ——, "Convexity of some feasible QoS regions and asymptotic behavior of the minimum total power in CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2190–2197, Dec. 2004.

[50] S. Stanczak, M. Wiczanowski, and H. Boche, *Theory and Algorithms for Resource Allocation in Wireless Networks*, ser. Lecture Notes in Computer Science (LNCS). Springer-Verlag, 2006.

[51] M. Chiang, C. W. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[52] C. W. Tan, D. P. Palomar, and M. Chiang, "Exploiting hidden convexity for flexible and robust resource allocation in cellular networks," in *IEEE Infocom*, May 2007, pp. 964–972.

[53] M. Chiang, P. Hande, T. Lan, and C. W. Tan, *Power Control in Wireless Cellular Networks*, ser. Foundation and Trends in Networking. now Publishers, Jul. 2008, vol. 2.

[54] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 641–646, Nov. 1993.

[55] N. Bambos, C. Chen, and G. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," *IEEE/ACM Trans. Networking*, vol. 8, pp. 583–597, Oct. 2000.

[56] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 74–85, Jan. 2004.

[57] E. Matskani, N. D. Sidiropoulos, and L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2682–2693, Jul. 2008.

[58] I. Mitliagkas, N. D. Sidiropoulos, and A. Swami, "Convex approximation-based joint power and admission control for cognitive underlay networks," in *IEEE Int. Wireless Commun. and Mobile Computing Conf. (IWCMC)*, Crete Island, Greece, Aug. 2008, pp. 28–32.

[59] E. Karipidis, N. D. Sidiropoulos, and L. Tassiulas, "Joint QoS multicast power/admission control and base station assignment: A geometric programming approach," in *IEEE Workshop on Sensor Array and Multi-Channel Signal Proces. (SAM)*, Darmstadt, Germany, Jul. 2008, pp. 155–159.

[60] E. Karipidis, E. G. Larsson, and K. Holmberg, "Optimal scheduling and QoS power control for cognitive underlay networks," in *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Aruba, Dec. 2009, pp. 5–8.

[61] M. Bengtsson and B. Ottersten, *Optimal and Suboptimal Transmit Beamforming*, ser. Handbook of Antennas in Wireless Communications, Boca Raton, FL, 2001, ch. 18.

[62] L. Badia, A. Baiocchi, A. Todini, S. Merlin, S. Pupolin, A. Zanella, and M. Zorzi, "On the impact of physical layer awareness on scheduling and resource allocation in broadband multicellular IEEE 802.16 systems communication system," *IEEE Wireless Communications Magazine*, vol. 14, no. 1, pp. 36–43, Feb. 2007.

[63] L. Anchora, L. Canzian, L. Badia, and M. Zorzi, "On the impact of physical layer awareness on scheduling and resource allocation in broadband multicellular IEEE 802.16 systems communication system," in *Proc. IEEE International Workshop on Computer Aided Modeling Analysis and Design of Communication Links and Networks (CAMAD)*, Fort Lauderdale, FL, Dec. 2010.

[64] S. Merlin, S. Begnini, A. Zanella, L. Badia, and M. Zorzi, "QoS-aware distributed resource allocation for hybrid FDMA / TDMA multicellular networks," in *Proc. International Symposium on Wireless Personal Multimedia Communications (WPMC)*, San Diego, CA, Sep. 2006.

[65] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, Jun. 2009.

[66] ——, "Denoising maps and constellations for wireless network coding in two-way relaying systems," in *Proc. IEEE Global Telecommun. Conf. (GlobeCom)*, 2008.

[67] J. Sykora and A. Burr, "Hierarchical alphabet and parametric channel constrained capacity regions for HDF strategy in parametric wireless 2-WRC," in *Proc. IEEE Wireless Commun. Network. Conf. (WCNC)*, Sydney, Australia, Apr. 2010, pp. 1–6.

[68] ——, "Network coded modulation with partial side-information and hierarchical decode and forward relay sharing in multi-source wireless network," in *Proc. European Wireless Conf. (EW)*, Lucca, Italy, Apr. 2010, pp. 1–7.

[69] ——, "Wireless network coding: The network aware PHY layer," in *Proc. Int. Symp. of Wireless Communication Systems (ISWCS)*, York, United Kingdom, Sep. 2010, tutorial.

[70] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, *Network Coding Theory.* now Publishers, 2006.

[71] F. Roemer, J. Zhang, M. Haardt, and E. Jorswieck, "Spectrum and infrastructure sharing in wireless networks: A case study with relay-assisted communications," in *Proc. Future Network and Mobile Summit 2010*, Florence, Italy, Jun. 2010.